ERSA Conference
August 2003
Títle

Estimation of Gender Wage Discrimination in the Portuguese Labour Market

Amélia Bastos
Instituto Superior de Economia Gestão
R. Quelhas nº 6
1200-781 LISBOA
tel: 351 21 3925831
fax: 351 21 3922781
abastos@iseg.utl.pt

Graça Leão Fernandes
Instituto Superior de Economia e Gestão
R. Quelhas nº 6
1200-781 LISBOA
tel: 213925870
fax: 351 21 3922781
gleao@iseg.utl.pt

José Passos
Instituto Superior de Economia e Gestão
R. Quelhas nº 6
1200-781 LISBOA
tel: 213925870
fax: 351 21 3922781
passos@iseg.utl.pt

Abstract:

Gender wage discrimination is a reality in the Portuguese labour markets although no study has been until now carried on to measure its dimension. We think that economists should contribute to the knowledge of the dimension and significance of this phenomenon by giving orientation for the definition of political measures towards its reduction.

In this paper we measure the size of gender wage discrimination in the Portuguese labour market. Furthermore, we evaluate this measure for the two Portuguese main cities, Lisboa and Porto.

In recent literature about the measurement of gender wage discrimination the Oaxaca (1973) decomposition and its developments is a commonly used approach. We extend this approach using bootstrap techniques for statistical inference purposes.

I. Introduction

This paper intends to measure the size of gender wage discrimination in the Portuguese labour market .

In recent literature about this issue the Oaxaca (1973) decomposition is the most widely used technique.

From descriptive statistics as simple as the ratio between male/female average wages (0.30 in our data) it is clear that gender wage discrimination is a reality in the Portuguese labour market. However little is known about how wide this phenomenon is or it´s spread in Portruguese labour market. This analysis, as long as we keep in mind its limitations, helps the recognition of gender wage discrimination and gives elements that help the definition of political measures towards its reduction.

In this paper we use the Oaxaca; Ramson (1994) decomposition to measure gender wage discrimination in and Porto, the two biggest regional labour markets in Portugal. Furthermore, we use bootstrap techniques to find the distribution of the wage gender gap due to discrimination in each one of the regions.

The paper is divided in four sections. This first section introduces the problem of gender wage discrimination in Portugal. In section II we briefly review the Oaxaca, Ramson (1994) decomposition. Section III describes the data used for the wage equation estimation and the methodology employed. Section IV introduces the results and presents the relevant conclusions.

II. The gender wage gap decomposition

The Oaxaca (1973) decomposition to estimate a measure of wage discrimination decomposes the average wage gap between two groups of workers in two components, one explained by productive differences due to workers skill differences and the other, not explained by individual characteristics thus considered discrimination.

Mincerian wage equations[1] are estimated for each of the groups (male and female). Let $\overline{W}_m$ and $\overline{W}_f$ refer to the mean of the actual wage received by men (group m) and women (group f) . The average wage gap is calculated from:

$$\begin{aligned}\Delta\overline{W} &= \ln\overline{W}_m - \ln\overline{W}_f \\ &= \left(\hat{\beta}_m\overline{X}_m\right) - \left(\hat{\beta}_f\overline{X}_f\right) \\ &= \left(\hat{\beta}_m - b\right)\overline{X}_m + \left(b - \hat{\beta}_f\right)\overline{X}_f + \left(\overline{X}_m - \overline{X}_f\right)b \qquad (1)\end{aligned}$$

where: b - represents the non-discriminatory wage structure;
$\hat{\beta}_i$ $(i = m, f)$ - the estimated coefficients of the wage equations;
$\overline{X}_i$ $(i = m, f)$ - vectors of average individual skill endowments.

The first term in the RHS of equation 1 measures the discrimination in favour of male, the second reports to the discrimination against female and the third is the gap due to differences on individual skills endowment.

Several authors[2] discuss the non-discriminatory wage structure definition that leads to different algorithms to estimate b. In this paper we follow Oaxaca, Ramson (1994) approach where b are the coefficients from the pooled (male/female) regression.

III. Data and Methodology

To estimate the wage equations we use data from “Quadros de Pessoal” a data gathered by the Department of Statistics of the Ministry of Employment and Social Security for 1997.

This data is the most extensive, complete and reliable microdata set available for the study of the Portuguese labour market. It is collected annually through a compulsory questionaire to firms employing salaried workers[3].

Our data has 2 227 717 workers (1 334 687 male and 933 030 female), individual characteristics (age, schooling levels, skill levels) as well as their firms location, sector and business volume.

From this data we selected all full time salaried workers in firms located in mainland Portugal for whom there are no missing values for the variables included in the wage equation. So the working data included 1 884 843 individuals (1 090 844 male and 793 999 female). From these 157271, 59281 male and 112223, 37667 female work in firms located respectively in and Porto are used to estimate the earming functions.

Our wage equation is not a typical mincerian one since it includes variables characterising the location, sector and business volume of the firms where the individuals are employed[4].

Our wage equation is then specified as:

$$\ln W = \alpha + \beta X + \gamma Z + u \qquad (2)$$

where: W - hourly wage rate before tax including base wage plus all regularly paid subsidies;
X - vector of workers characteristics (school level, experience, skill, level, time in current job);
Z - vector of firm characteristics (business volume, sector).

We don´t have a tenure variable because our data only has information about the time in current job.

We include sectoral dummies given the evidence that women are concentrated in some sectors which points to the fact that the distribution of male/female across sectors could, itself, be a result of discrimination.

The bootstrap was introduced by Efron (1979) as a computer-based method for estimating the variance of an estimator. Freedman (1981) extended this method to the regression framework.

Basically the bootstrap treats the data as if they were the population for the purpose of evaluating the quantity of interest. The method has been shown very useful in situations where the asymptotic distribution of an estimator is difficult to derive. Moreover it is often more accurate in finite samples than first-order asymptotic approximations.

In the present paper our purpose is twofold: i) testing the null hypothesis of non-discrimination which requires computing the standard deviation of the male-female wage differential due to discrimination and ii) testing the equality of this wage discrimination between the two major cities of Portugal (Lisboa e Porto). Due to the mathematical difficulty in obtaining the exact distribution of the wage discrimination estimator the bootstrap methodology is adopted.

In general, bootstrapping regression models can be carried on in two different ways: i) through resampling errors or ii) resampling cases. The main difference between these two approaches is related to the hypothesis underlying the regression model. With resampling cases the regression model still applies with no assumption on the random error other than independence, being robust to departures from the homoscedasticity assumption which is a typical problem in cross-sectional models.

In the following, resampling cases is applied to derive the distribution of the estimator defined as the proportion of male-female wage differential due to discrimination,

$$D = \frac{\left(\hat{\beta}_m - \beta^*\right)\overline{X}_m + \left(\beta^* - \hat{\beta}_f\right)\overline{X}_f}{\Delta\overline{W}} \qquad (3)$$

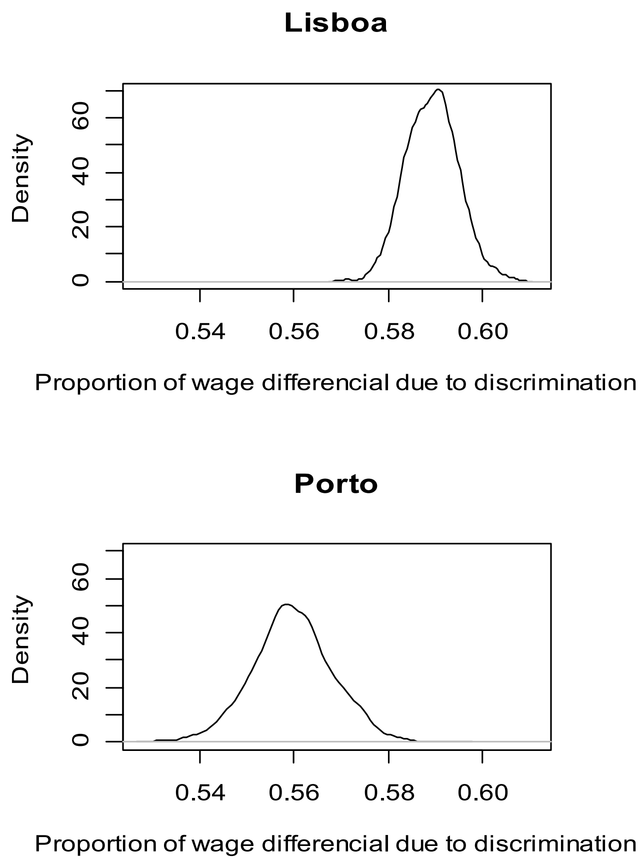
Having in mind the wage models for Lisboa and Porto, we draw 1000 bootstrap samples. From each bootstrap sample and for each city least squares regression is applied given estimates $\hat{\beta}^*_m$, $\hat{\beta}^*_f$ and D^* by (3). Table 1 summarizes the main results.

Table 1 - Bootstrap Estimates of D

	Min	1st Quart	Mean	3rd Quart	Max	Stdv
Lisboa	0.571	0.586	0.589	0.593	0.607	0.00545
Porto	0.532	0.555	0.560	0.565	0.593	0.00828

The standard deviation of the bootstrap estimator of D is very low compared to the mean and therefore the null hypothesis of non-discrimination is easily rejected. Moreover Lisboa and Porto exhibit different patterns of discrimination with Porto having a larger variance and a lower mean value of discrimination [see Figure 1].

Figure 1 - Distribution of male-female wage discrimination

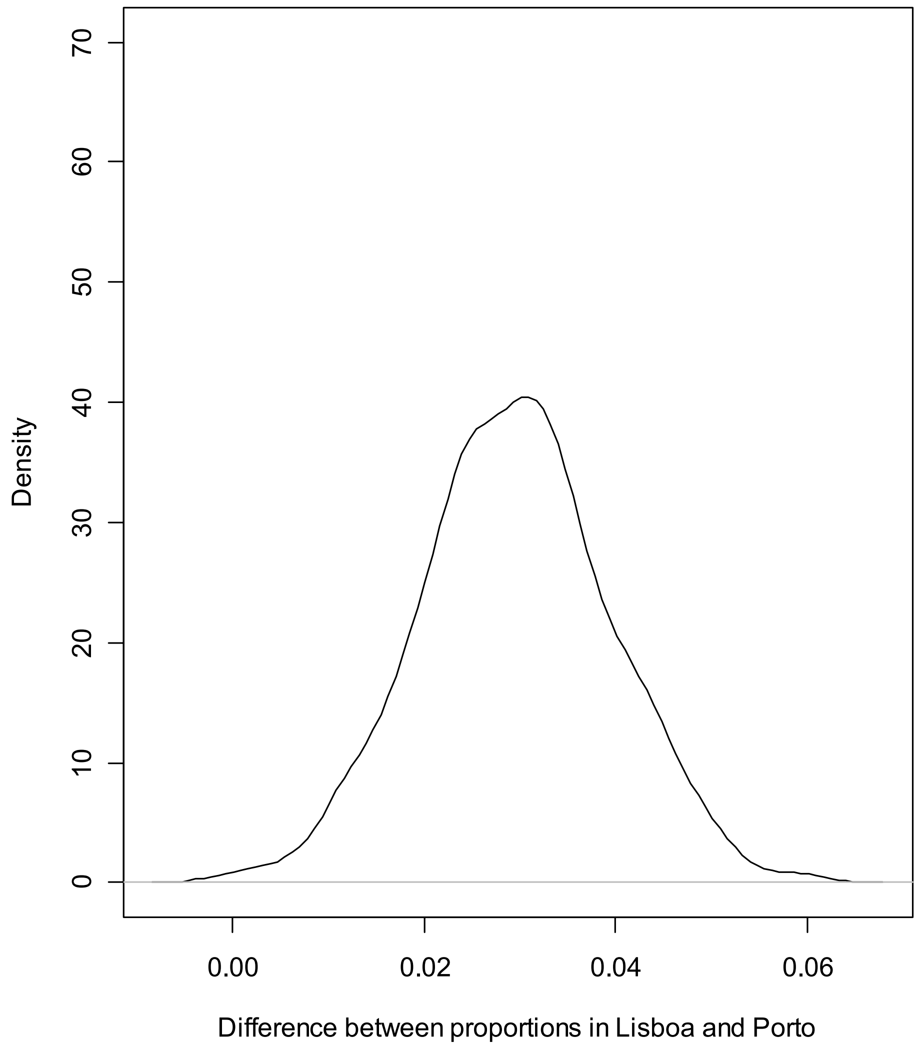


From the bootstrap estimates of D for Lisboa and Porto the distribution of the difference is also computed [see Figure 2] and the main results presented in Table 2. The null hypothesis of equal discrimination between these two cities is easily rejected.

Table 2 - Bootstrap Estimates of the difference between proportions in Lisboa and Porto

Min	1st Quart	Mean	3rd Quart	Max	Stdv
-0.002	0.023	0.030	0.036	0.061	0.00993

Figure 2 - Distribution of the difference between proportions in Lisboa and Porto



IV. Results and main conclusions

Table A_1 in appendix gives the means of some of the variables included in the earning functions, separately for men and for women and for and Porto. In this two cities men are, on average, older than women, have more experience and present higher skill levels. Nevertheless, they achieved less levels in school, compared to women. The similar pattern of this variables in Lisboa and Porto has, although, some differentiations. In fact, in the later, workers seem to be youngest on average and that's perhaps the reason why they did not achieved school levels as higher as those from Lisboa. In Porto women register more divergence in skills and schooling comparing to men, than in . In fact, in Porto women are less qualified than men in spite of the highest level on school that they achieved, on average. This evidence will certainly be associated with discrimination.

The results of the empirical estimation of the earning functions are presented in Tables A_2 and A_3 of the appendix. All estimates are roughly significant for a 0.001 significance level and have the magnitude[5] and expected signals. Standard errors are robust to heterokedasticity using White (1980) type estimation. The estimated models are all robust to alternative specifications. A comparative analysis between men and women, on one instance, and between the two cities studied, on the other instance, permits the following conclusions:

- The rates of return on schooling do not vary with gender or regional localisation;
- An extra year of experience benefits more men than women, in the two cities;
- Gains from skill up-grading are higher the higher the skill level for both men and women. This impact seems to be stronger for women, in terms of gender, and for Porto, in terms of regional localisation;
- Firms business volume have a positive and increasing impact on wages. This fact applies especially to women in and to men in Porto;
- There are differences between men and women concerning the impact of job sector. Industry and services have a higher negative impact on women wages, especially in Porto;
- Finally, the time in actual job variable introduced in the model affects positively wages, particularly for women.

Let us now analyse the decomposition of the wage difference between men and women, according to the regions considered.

The log wage differential is similar in Lisboa and Porto. On average, the proportion of wage difference due to discrimination is more significant in Lisboa than in Porto (Tables A_4 and A_5 in the Appendix). The later registers a higher effect of the differential wage productivity component and a lower pure discrimination component, both on average. This fact is perhaps related with the higher experience that workers from Lisboa seem to have on average compared with those from Porto. This variable showed empirically a weaker effect on wages in Porto, thus reducing the impact of wage discrimination due to productivity in this region.

The application of the bootstrap technique enables the construction of the discrimination factor distribution. On the one hand and according to the graphs presented earlier, Lisboa seems to register a higher component on discrimination, fact already referred earlier regarding the original sample. On the other hand, in Porto gender wage discrimination is expected to be lower than in Lisboa but presents a higher dispersion.

These results are robust to other wage discrimination decomposition procedures as Silber and Weber (1999) also pointed out.

One possible explanation for this pattern is related to cultural elements. In addition, the unequal distribution of men and women by sector in the two regions, result itself of discrimination, influences the value assumed by the discrimination coefficient. In fact the data shows that in Porto women are preferably concentrated on industry, which do not happen with men. This situation contributes to the variability of the discrimination factor. In , on the contrary, women have positions in sectors where the proportion of men is also significant, as services. This fact favours the intensity of discrimination. Nevertheless these hypothetical explanations need further research.

Finally, it should be mentioned that the estimated value for discrimination ranges between numbers that are common in this kind of studies, such as Silber and Weber (1999) and Neumark (1988).

Footnotes:

[1] Mincer (1974)

[2] Reimers (1983), Neumark (1988), Cotton (1988), Oaxaca, Ramson (1994)

[3] The data doesn´t include public administration and non-market services. The agriculture sector is poorly covered.

[4] Fernandes (1992) shows that such variables significantly explain wages in the Portuguese labour market.

[5] Fernandes (1992).

References:

- Cotton, J. (1988), On the decomposition of wage differentials, *Review of Economics and Statistics*, **70**, 236-243.
- Efron, B. (1979), Bootstrap methods: another look at the jackknife, *Annual of Statistics*, **7**, 1-26.
- Freedman, O.A. (1981), Bootstrapping regression models, *Annual of Statistics*, **9**, 1218-1228.
- Fernandes, G. L. (1992), *Wage determination in the Portuguese Labour market*, PhD thesis, .
- Mincer,J. (1974), *Schooling, Experience and Earnings*, (New York: National Bureau of Economic Research).
- Neumark, D. (1998),Employer´s discriminatory behaviour and the estimation of wage discrimination, *Journal of Human Ressources*, **23**, 279-295.
- Oaxaca, R. (1973), Male-female wage differentials in urban labor markets, *International Economic Review*, **9**, 693-709.
- Oaxaca, R. and Ramson, M. (1994), On discrimination and the decomposition of wage differentials, *Journal of Econometrics*, **61**, 5-22.

- Reimers, C. (1983), Labor market discrimination against Hispanic and black men, *Review of Economics and Statistics*, **65**, 570-579.
- Silber, J. and Weber, M. (1999), Labour market discrimination: are there significant differences between the various decomposition procedures ?, *Applied Economics*, **31**, 359-365.
- White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, **48**, 817-838.

Appendix

Table A_1: Mean value of the variables

Variable	Lisboa		Porto	
	Men	Women	Men	Women
Age	39.23	36.76	38.55	35.62
TCJ	9.25	8.01	10.29	8.23
Exper	24.57	21.46	24.98	21.3
School	8.66	9.30	7.57	8.31
SL1	.18	.12	.11	.08
SL2	.17	.15	.14	.11
SL3	.39	.38	.46	.40
SL4	.08	.14	.09	.19

TCJ - time in current job

Exper - experience

School - highest level of school achieved

SL_i (i=1,2,3,4) - skill level. From 1= highest skilled to 5=non skilled

Table A_2: Earning functions for men and women from Lisboa (dependent variable: logarithm of hourly wage)

Variable	Coeficients of the earning functions in Lisboa for			
	Men		Women	
C	5.37207	(686.718)	5.33265	(758.616)
BV2	.081548	(23.1189)	.094258	(30.4746)
BV3	.218683	(61.6843)	.224532	(68.9603)
BV4	.329691	(86.7804)	.269388	(69.1264)
TCJ	.011945	(28.9494)	.018723	(40.3758)
TCJ2	-.169730E-03	(- 14.048)	-.279324E-03	(- 18.455)
School	.062123	(136.691)	.059934	(118.941)
Exper	.033623	(87.1858)	.023829	(61.6381)
Exper2	-.428525E-03	(- 68.109)	-.303558E-03	(- 45.611)
Manuf	.088389	(17.9524)	-.771063E-03	(- .15982)
Constr	-.024347	(- 5.4528)	-.026699	(- 3.2491)
Ser	.049974	(12.1175)	-.046696	(- 14.469)
Transp	.135479	(33.7649)	.151951	(34.3382)
Financ	.378605	(80.1573)	.427586	(92.4851)
Immov	.076839	(15.2626)	.057136	(14.8820)
SL1	.541855	(102.962)	.549714	(91.8917)
SL2	.248974	(57.6662)	.351974	(74.7927)
SL3	-.492661E-02	(- 1.4434)	.074659	(22.7447)
SL4	-.083994	(- 20.155)	-.06596	(- 18.929)

t-values in parenthesis

Manuf - manufactoring

Constr - construction

Ser - services

Transp - transports

Financ - finance

Immov - immovable

Table A_3: Earning functions for men and women from Porto (dependent variable: logarithm of hourly wage)

Variable	Coeficients of the earning functions in Porto for			
	Men		Women	
C	5.36375	(426.687)	5.37757	(485.622)
BV2	.111779	(24.2211)	.102276	(25.0720)
BV3	.275860	(56.0028)	.249387	(47.5109)
BV4	.324904	(59.2414)	.308565	(43.2862)
TCJ	.011103	(19.5148)	.016474	(24.8727)
TCJ2	-.156061E-03	(- 9.9368)	-.287227E-03	(- .13068)
School	.056260	(78.6009)	.052734	(66.8344)
Exper	.030145	(51.9779)	.019721	(32.6494)
Exper2	-.387149E-03	(- 40.195)	-.238789E-03	(- 22.358)
Manuf	-.045063	(- 6.6934)	-.120297	(- 22.964)
Constr	- .09477	(- 13.095)	-. 05179	(- 3.5869)
Ser	- .01787	(- 2.7509)	-. 06289	(- 12.965)
Transp	.109937	(16.8902)	.144711	(15.6932)
Financ	.413139	(52.7856)	.467129	(54.5545)
Immov	- .05329	(- 6.4976)	. -02269	(- 3.5679)
SL1	.560473	(64.8093)	.585272	(56.7227)
SL2	.279327	(43.0500)	.371114	(46.7696)
SL3	.055978	(11.0056)	.075562	(16.3008)
SL4	- .02139	(- 3.5673)	- .03719	(- 7.3537)

t-values in parenthesis

Table A_4: Descriptive statistics for discrimination in Lisboa

	Mean	Std Dev	Minimum	Maximum
LWM	6.87840	0.0017005	6.87249	6.88308
LWF	6.64326	0.0018893	6.63751	6.64928
DW1	0.096579	0.0019346	0.090316	0.10227
DW2	0.13857	0.0015205	0.13359	0.14352
PDW2	0.58930	0.0054544	0.57110	0.60681

	Sum	Variance	Skewness	Kurtosis
LWM	6878.40227	2.89167D-06	-0.0061559	-0.12789
LWF	6643.25840	3.56951D-06	0.079905	-0.16702
DW1	96.57887	3.74253D-06	-0.080842	0.012573
DW2	138.56501	2.31201D-06	0.024278	0.0091486
PDW2	589.30061	0.000029750	0.074536	0.0067597

LWM - log wage for men
LWF - log wage for women
DW1 - wage differentiation due to productivity
DW2 - discrimination
PW2 - proportion of the wage differentiation due to discrimination

Table A_5: Descriptive statistics for discrimination in Porto

	Mean	Std Dev	Minimum	Maximum
LWM	6.65327	0.0025411	6.64492	6.66227
LWF	6.42250	0.0028049	6.41163	6.43130
DW1	0.10160	0.0029988	0.091029	0.11153
DW2	0.12917	0.0022720	0.12246	0.13735
PDW2	0.55980	0.0082830	0.53175	0.59254

	Sum	Variance	Skewness	Kurtosis
LWM	6653.26642	6.45715D-06	0.054233	0.18460
LWF	6422.49887	7.86719D-06	-0.080492	0.034687
DW1	101.59706	8.99271D-06	-0.0022146	0.12711
DW2	129.17049	5.16190D-06	0.091479	-0.022006
PDW2	559.79552	0.000068607	0.043958	0.25340