CEFAGE-UE Working Paper
2011/25

Inter-Regional Wage Differentials in Portugal: An Analysis Across the Wage Distribution

João Pereira , Aurora Galego

University of Évora, Department of Economics and CEFAGE-UE

CEFAGE-UE, Universidade de Évora, Palácio do Vimioso, Lg. Marquês de Marialva, 8, 7000-809 Évora, Portugal
Telf: +351 266 706 581 - E-mail: cefage@uevora.pt - Web: www.cefage.uevora.pt

INTER-REGIONAL WAGE DIFFERENTIALS IN PORTUGAL: AN ANALYSIS ACROSS THE WAGE DISTRIBUTION

João Pereira and Aurora Galego

University of Évora , Department of Economics and CEFAGE-UE

ABSTRACT

Typically, studies on regional wage differentials are based on OLS estimates and use Blinder (1973) and Oaxaca (1973) decomposition. Quantile regression is an alternative approach which allows for studying these differences across the whole wage distribution. In this study, the quantile regression framework is considered for the analysis of regional wage differences in Portugal. Our findings reveal significant differences in wage equations coefficients between regions for the various quantiles. Furthermore, we conclude that the regional wage differentials and the components explained by differences in endowments and differences in returns increase across the whole wage distribution.

Key-words: regions, wage differentials, quantile regression, quantile-based decompositions

JEL: J31; J38; C21

1. INTRODUCTION

The issue of regional wage differentials is relevant both for policy proposes and general public discussion. A sound knowledge of the distribution of wage inequalities and their causes is essential for defining policy measures for reducing spatial income inequalities. A range of empirical studies have analysed regional wage differentials for a number of countries (Blackaby and Manning, 1990; Blackaby and Murphy, 1995; Duranton and Monastiriotis, 2002; García and Molina, 2002). Typically, these studies are based on OLS estimates and the decomposition method devised by Blinder (1973) and Oaxaca (1973), which focuses on the analysis of wages differences at the mean of the conditional earnings distribution. This approach provides a reasonable description of wage distributions when they are unimodal, symmetric and have similar variances (Butcher and Dinardo, 2002). However, in general, these conditions may be not fulfilled. Therefore, wage differentials should be analysed along the entire wage distribution.

The quantile regression model (Koenker and Basset, 1978, 1982) offers a more complete and flexible approach than the usual OLS estimations and therefore provides a better framework for analysing regional wage differentials. In fact, by using this approach it is possible to study the effect of a covariate across the conditional distribution of the dependent variable (Melly, 2005a) and obtain more detailed and reliable insights as regards regional wage differentials. However, very few studies have considered the quantile regression model in the context of regional wage inequality. In particular, only Motellón et al. (2011) analyse inter-regional wage differentials following the methodology devised by Dinardo et al. (1996) and Butcher and Dinardo (2002), which is essentially non-parametric. The major drawback of this methodology is that it does not allow for carrying out significance tests for decomposition effects and therefore it is not possible to gauge whether the results reported with regard to wage decomposition are statistically significant or not.

This paper seeks to build on previous research in a number of different ways. Firstly, unlike most previous studies, we estimate regional wage equations by quantile regression in order to analyse the effect of covariates at several points on the wage distribution. Secondly, we apply the quantile-based decomposition method suggested by Machado and Mata (2005) and Melly (2005a, 2006) to decompose regional wage differentials at several points of the wage distribution into one component based on differences in observed characteristics and another based on the differences in rewards for these characteristics. This method is of a semi-parametric nature, which allows for the estimation of significance tests and confidence intervals of wage decomposition effects (characteristics and returns). This marks a clear difference in relation to the non-parametric method suggested by Dinardo *et al.* (1996) and Butcher and Dinardo (2002) and applied by Motellón et al. (2011), which does not allow such significance tests to be performed. To the best of our knowledge, this is the first application of the methods proposed by Machado and Mata (2005) and Melly (2005a, 2006) in the context of regional wage differentials, although they have been applied in other contexts. Other applications include the study of the public-private sector wage gaps (Melly, 2005b; Lucifora and Meurs, 2006; Cai and Liu, forthcoming), gender discrimination (Albrecht et al., 2003), and union wage premium (Cai and Liu, 2008).

We consider the case of Portugal, a small country with significant and quite stable regional wage differentials (Vieira et al., 2006; Pereira and Galego, 2011). In the empirical analysis, we take a sample from the Portuguese Ministry of Employment data set – Quadros *de Pessoal* - for the last available year – 2008. Our findings reveal that coefficients estimates along the wage distribution for each region and between the various regions are not stable. In most cases, differences in the returns to characteristics increase across the wage distribution, for both males and females. Moreover, these findings confirm previous evidence as to the existence of significant regional wage differences between the Lisboa region and the other regions, and also reveal increasing differentials across the wage distribution. Finally, with regard to the regional wage decomposition, we conclude that both the part relating to differences in characteristics and the part relating to differences in returns to these characteristics are in general statistically significant and increase across the entire wage distribution.

The paper is organised as follows. In Section 2 there is a summary presentation of the literature on spatial wage differentials. In Section 3, the methodology used in this study is presented. Section 4 provides a preliminary analysis of the data. In Sections 5 and 6 our findings are presented and discussed. In Section 7 we present our conclusions.

2. REVIEW OF LITERATURE ON SPATIAL WAGE DIFFERENCES

In a homogenous space - without amenity differences – in which labour and capital can move around freely, and where information is perfect and transportation costs are modest, neoclassical economic theory predicts a long-run economic equilibrium in which factor prices are equalised (Goldfarb and Yezer, 1976). Price differentials may, however, arise in this context if relevant differences in amenities are present, such as extreme climatic conditions or pollution. Such price (or wage) differentials are required in order to attract people to less amenable areas (Roback, 1982) and thus in order to equalise workers' utility throughout the space.

Temporary shifts in demand and supply may disturb the economic equilibrium and cause wage differentials in addition to those explained by amenities (Blackaby and Manning, 1990). Labour market inefficiencies, such as a non-competitive housing market (Henley, 1998), tend to exacerbate these disequilibrium situations.

There are, however, other contexts in which spatial wage differentials may arise. For example, human capital concentration in cities or regions may provide a source of important knowledge spillovers (Lucas, 1988), which increase economic efficiency and leads for higher wages levels. In fact, people who live in areas where human capital is highly concentrated have the opportunity to learn from others and thus improve their own productivity (Glaeser et al., 1992; Lucas, 1988). In the case of industrial concentration in cities or regions, external economies may also take place (Marshall, 1890; Porter, 1990; Romer, 1986).

The new economic geography literature (see, for example, Fujita et al., 1999; Krugman, 1991), on the other hand, highlights the role of scale economies and transportation costs in the creation spatial demand linkages that contribute to economic agglomeration. These models generally explain nominal wage differentials assuming real wage equalisation: they predict that nominal wages will be higher in regions that have easy access to economic centres due to stronger demand linkages (Fujita et al., 1999; Krugman, 1991). This approach is related to Harris' (1954) market-potential function, which states that the demand for goods produced in a location is the sum of the purchasing power in other locations, weighted by transportation costs.

Another topic in the literature on spatial wage differentials focuses on wage differences between urban and non-urban areas. Empirical evidence has consistently demonstrated that workers in densely-populated urban areas earn more than their non-urban counterparts (Glaeser and Maré, 2001; Yankow, 2006; Addario and Patacchini, 2008). Economic theory puts forward several explanations for the urban wage premium emphasising, in general, that the higher costs of living in these areas are not enough to account for all the differentials between urban and non-urban areas (Glaeser and Maré, 2001; Yankow, 2006). One possible explanation for this is that urban areas attract the most productive workers: the ability bias hypothesis. Another is that firms may experience productivity advantages arising from external economies when they are located in densely-populated urban areas. Also, sorting effects, as a consequence of urban agglomeration, may produce more efficient and productive matches between workers and firms (Wheeler, 2001; Combes et al., 2008).

Most microeconometric studies on regional wage differentials base their analysis on the estimation of human capital wage equations and on the classical Blinder (1973) and Oaxaca (1973) decomposition estimated at the mean of the conditional wage distribution (Blackaby and Manning, 1990; Blackaby and Murphy, 1995; García and Molina, 2002; Duranton and Monastiriotis, 2002; Simón et al., 2006). Wage differentials are explained either by differences in regional characteristics (endowments) or by the fact that these characteristics (endowments) are rewarded differently in different locations. Wage differentials explained by differences in both human capital and industry related characteristics are compatible with the neoclassical view. However, if the same productivity related characteristics are not rewarded at the same price throughout the space, we might have a temporary situation of disequilibrium, agglomeration economies or sorting effects.

The empirical evidence provided by these studies on regional wage differentials varies from country to country. For instance, Blackaby and Murphy (1995) found that wage differentials between the North and the South of Britain are relatively small. The results show that the wage differential that can be explained by differences in rewards to workers with the same level of skills is about 2.4%, in favour of the South, and therefore, the situation is not too far from the neoclassical equilibrium. On the other hand, for Spain, García and Molina (2002) found important wage differences between Madrid and the other Spanish regions. The reasons for these differences are mixed and depend on the specific case, but differences in both characteristics and in their rewards play an important role in explaining the regional wage gap in Spain.

For Portugal, few studies on regional wage differentials have been carried out. Pereira and Galego (2011) using information from Quadros de Pessoal for 1995 and 2002 and considering level tow of regional aggregation (NUTS 2), found important and stable regional wage differentials, mainly between Lisboa and the other regions. The estimated differentials range from about 20% to 30%. Both the characteristics and the returns effect play an important role in the explanation of these wage differentials. Vieira et al. (2006) examined this issue at level of a more disaggregated administrative division (distritos). Although, there are some differences in the explanatory variables and in the methodology used, the results concerning to the estimated wage differentials between Lisboa and other regions of the country are qualitatively similar to those of Pereira and Galego (2011). However, none of these studies analyses regional wage differences across the wage distribution.

Recently, two other studies have applied quantile regression techniques to the issue of regional wage differentials and regional inequality. Dickey (2007) analysed regional wage inequality in the UK but her focus is on wage inequality within regions and on the forces shaping rather than on inter-regional wage inequality. Motelon et al. (2011) also applied the quantile regression model for studying inter-regional wage differentials in Spain, concluding that there are increasing wage differentials across the wage distribution. They used the approach devised by Dinardo et al. (1996) and Butcher and Dinardo (2002). This methodology is essentially non-parametric, which has the advantage of imposing restrictions neither on covariate effects nor on density shapes.

However, if there are too many variables, counterfactual distributions cannot be estimated non-parametrically (Melly, 2005a, 2006). Moreover, by applying this methodology, no standard errors are estimated and therefore neither can significance tests carried out nor can confidence bands be calculated for decomposition effects. Hence, it is not possible to find out whether the results reported by Motelon et al. (2011) concerning wage gap decomposition are statistically significant or not.

3. METHODOLOGY

Let y_i be the log hourly wage of worker i and x_i a vector of covariates representing individual and workplace characteristics. Assuming a linear relationship between y_i and x_i, the θth quantile of the conditional distribution of y_i given x_i is given by:

$$Q_\theta(y_i \mid x_i) = x_i\beta_\theta,\ \theta \in (0,1) \qquad (1)$$

Koenker and Basset (1978, 1982) showed that the quantile regression estimator of β_θ is the solution of the following minimization problem:

$$\hat{\beta}_\theta = \arg\min_{\beta}\left[\sum_{i:y_i \geq x_i\beta} \theta\left|y_i - x_i\beta\right| + \sum_{i:y_i < x_i\beta} (1-\theta)\left|y_i - x_i\beta\right|\right]$$

Blinder (1973) and Oaxaca (1973) decompose the difference in average earnings between two groups of workers in two components: one is attributable to the difference in the average values of explanatory variables (characteristics effect) and the other, which is unexplained, is due to differences in the estimated coefficients (price effect). This decomposition can be used for analysing wage differentials at the mean of earnings distributions and not over the wage distribution as a whole, which may be limitative and potentially hide important aspects of earnings distributions.

Machado e Mata (2005) extended this decomposition to the quantile regression model. Their procedure is based on randomly drawing θ and x and estimating the whole conditional distribution by quantile regression and then integrating the conditional distribution over the range of regressors in order to obtain an estimate of the unconditional distribution (Machado e Mata, 2005; Melly, 2005a; Cai and Liu, forthcoming). This estimator is, however, time-consuming as it combines quantile regression and bootstrapping. In this paper, we use a simplified but asymptotically equivalent estimator proposed by Melly (2005a, 2006). This procedure is implemented by following the next three steps:

1: Estimate the quantile regression coefficients for the *j* different quantiles: $\beta^{\alpha}\left(\theta_j\right)$ and $\beta^{\gamma}\left(\theta_j\right)$, for $\theta_j \in \left(0,1\right)$ and $j = 1,........J$, using workers from region α and region γ, respectively. Based in these estimates, both the predicted wage density $\left\{\left\{x_i^{\gamma}\beta^{\gamma}\left(\theta_j\right)\right\}_{j=1}^{J}\right\}_{i=1}^{N\gamma}$ and the counterfactual wage density for region γ workers $\left\{\left\{x_i^{\gamma}\beta^{\alpha}\left(\theta_j\right)\right\}_{j=1}^{J}\right\}_{i=1}^{N_\gamma}$ are then calculated; where x_i^{γ} stands for the observed characteristics of region γ workers and x_i^{α} stands for the observed characteristics of region α workers. Likewise, N_γ refers to the number of workers in region γ, while N_α refers to the number of workers in region α. The counterfactual wage density of region γ workers represents the wage density that would exist if they maintained their characteristics but were rewarded as region α workers.

2: Estimate the θth quantile of the sample $\left\{\left\{x_i^{\gamma}\beta^{\alpha}\left(\theta_j\right)\right\}_{j=1}^{J}\right\}_{i=1}^{N_\gamma}$, represented by $Q_\theta\left(x^{\gamma},\beta^{\alpha}\left(\tau\right)\right)$, and of the sample $\left\{\left\{x_i^{\gamma}\beta^{\gamma}\left(\theta_j\right)\right\}_{j=1}^{J}\right\}_{i=1}^{N_\gamma}$, represented by $Q_\theta\left(x^{\gamma},\beta^{\gamma}\left(\tau\right)\right)$.

3: Decompose the unconditional wage distribution difference at each quantile into two components: one component explained by differences in the regions´ observed characteristics and the other explained by differences in the returns to these characteristics. The full expression for the wage decomposition at θ quantile of the wage distribution is then given by:

$$\underbrace{\mathrm{Q}_\theta\left(x^{\alpha},\beta^{\alpha}\left(\tau\right)\right)\text{-}\mathrm{Q}_\theta\left(x^{\gamma},\beta^{\gamma}\left(\tau\right)\right)}_{\substack{\textit{estimated wage differential}\\ \textit{at } \theta \textit{ quantile}}}=\underbrace{\mathrm{Q}_\theta\left(x^{\alpha},\beta^{\alpha}\left(\tau\right)\right)\text{-}\mathrm{Q}_\theta\left(x^{\gamma},\beta^{\alpha}\left(\tau\right)\right)}_{\textit{diferences in the observed characteristics}}+$$
$$+\underbrace{\mathrm{Q}_\theta\left(x^{\gamma},\beta^{\alpha}\left(\tau\right)\right)\text{-}\mathrm{Q}_\theta\left(x^{\gamma},\beta^{\gamma}\left(\tau\right)\right)}_{\textit{returns to observed characteristics}}$$

The bootstrap method was used for estimating standard errors and confidence intervals. In this study, 100 replications were carried out in order to estimate the confidence intervals.

4. DATA DESCRIPTION AND PRELIMINARY ANALYSIS

In this study, we use individual data from *Quadros de Pessoal* for 2008, a matched employer-employee dataset produced by the Portuguese Ministry of Employment, which includes information about all private firms in Portugal; the survey does not provide information about unemployed, those employed in the field of public administration, the self-employed or the armed forces. The available data contains information on both workers and firms, including earnings, hours of work, age, education, tenure, firm size, industry affiliation, occupation and also information about the region where the firms and establishments are located. Given the amount of data available in *Quadros de Pessoal* – more than 2 millions of individuals – and in view of the timing-consuming methods used in this study, we randomly-selected a sample of 5% individuals per region from the raw data.

In our final sample, we considered only workers between 16 and 65 years of age and excluded those working in the agriculture and fisheries sectors, as well as unpaid family workers and apprentices. Individuals working in the Madeira and Açores regions were also not considered[1]. Wage outliers were dropped, namely wages above 20 times the 99th percentile and wages below half of the 1st percentile. The final sample comprises 68,322 males and 56,245 females.

Figure 1 displays kernel density estimates (Gaussian kernel) for hourly wage distributions by region. Clearly, the main differences in the shape of these distributions seem to be between the Lisboa region and other regions. Density for Lisboa lies somewhat to the right of that of other regions, has broader lower tails, and displays a higher area of probability in the upper tail and on the whole of the right side. In addition, density for Lisboa suggests higher wage dispersion. In the case of men, Lisboa's mode is clearly located to the right of that of other regions.

Figure 1 around here

Overall, this suggests that average wages are higher in the Lisboa region and that the main wage differences are those between Lisboa and all other regions, which bears out the findings of previous studies (Pereira and Galego, 2011; Vieira et al., 2006). In addition, this analysis also reveals the fact that wage differentials probably increase over the wage distribution. Indeed, the

[1] These regions are made up of islands and therefore present a quite different situation to those located in mainland Portugal.

test proposed by Kruscal and Wallis (1952, 1953) for independent samples provides confirmation that regional earnings distributions are significantly different from one another and consequently should be analysed separately[2].

table 1 around here

Table 1 displays descriptive statistics for the main variables used in our analysis. It may be concluded that there are important differences in the distribution of human capital throughout the country, especially with regard to the percentage of males and females with university degrees and, to a lesser extent, the percentage of individuals who have attended secondary education. It is in the Lisboa region that we find the highest percentage of individuals with highest level of education. Lisboa is also the region which displays the highest average firm size. Regional differences in terms of levels of *experience* and *tenure* are not so apparent.

There are differences between regions in terms of industrial and occupational structure, as can be seen in Appendix B. For instance, in the Lisboa region one of the most important industries is Administration and services, whereas in Algarve tourism and related activities are of crucial significance. As regards occupation, Lisboa has the largest number of managers, professional and associated professional staff, while in other regions craft workers, plant and machine operators and unskilled workers are predominant.

[2] For both males and females the null hypothesis that the populations are the same is rejected at any significance level below 1%.

5. RESULTS AND DISCUSSION

5.1 WAGE EQUATIONS ESTIMATES

Let us firstly analyse coefficients estimates for the regional wage equations and for both genders. Our analysis is based on Mincer-type wage equations estimated for each region (and gender) by OLS and by quantile regression (for selected quantiles). We consider as explanatory variables worker experience, tenure, 10 control dummies for industry affiliation, 9 occupational dummies, dummies for education, and the logarithm of firm size[3]. In order to take into account regional differences in the cost of living, wages were deflated by the Instituto Nacional de Estatística (INE) regional consumer price index and are at 2006 prices.

Selectivity bias is a possible problem when estimating wage equations which might have some impact on our results. In this case, however, it cannot be controlled as there is not enough information in the data set: *Quadros de Pessoal* does not include unemployed and non-active individuals. Nevertheless, as stated in Pereira and Galego (2011), we believe that the results contained in this paper are not markedly influenced by sample selectivity. Firstly, this is a particularly important issue when comparing male and female wage equations. In our analysis we are not comparing men with women in different regions, but men (and women) in the Lisboa region with men (and women) in other regions. Secondly, as those working in the agriculture and fisheries sectors are excluded from the sample[4] and as the Portuguese population is heavily concentrated in urban and coastal areas, we think that the results are not significantly influenced

---

2 A definition of variables is given in Appendix A.

4 As in Pereira and Galego (2011).

by the possibility that individuals from urban and rural areas may make different participation decisions. Thirdly, our wage equations include detailed controls for occupation, which may capture some unobserved ability components and therefore partially correct for possible spatial selection biases (Duranton and Monastiriotis, 2002). Finally, previous empirical work for Portugal using the *European Community Household Panel* for 1995 (Pereira, 2003), did not reveal statistically significant sample selection effects in regional wage equations estimates for either men or women.

Table 2 around here

OLS estimates of the regional wage equations are presented in table 2. Coefficient estimates are all significant and reveal the expected effects. In particular, *experience* and *tenure* positively affect both women's and men´s wages and a higher level of educational achievement is associated with a higher wage return. The size of the individual's place of work also has a positive and statistically significant impact for both genders and for all regions, which suggests the existence of efficiency wages effects in the Portuguese labour market. Yet, there are important differences in coefficients estimates between regions. For example, for both genders, returns to education are highest for the Lisboa region; Algarve, by contrast, displays the lowest returns to education. Lisboa also shows the highest values for coefficients of *experience* and *tenure*. By contrast, the firm size coefficient is in general lower for Lisboa than for other regions. Finally, we may conclude that returns to occupation for higher-skilled occupations are, in most cases, greatest for Lisboa.

[figure 2 and figure 3, about here]

With regard to the quantile regression, coefficients estimates for selected variables are displayed in figure 2 for males and in figure 3 for females[5]. With the exception of three occupational dummies identifying *senior officials and managers*, *professionals, and technicians and associate professionals*, we do not present the results concerning industry and occupational controls. Quantile analysis provides an understanding of the distribution of coefficients estimates across the wage distribution for each region and among the several regions. Estimates reveal that, in most cases, returns to characteristics increase across the wage distribution for both males and females. Exceptions to this pattern are returns to tenure (*tenure* and *tenure2*) and the coefficient (elasticity) of the firm size variable (*lsfize*). Moreover, returns to characteristics are typically highest for the Lisboa region for most of the wage distribution or even along the entire wage distribution or some variables (*exp* and exp2, *senior officials and managers*). Once again, the elasticity of wages to the firm size provides the exception to this general pattern.

We also performed inter-quantile tests for the hypothesis of equal coefficients for each explanatory variable. We considered several inter-quantile differences (90^{th}-10^{th}; 90^{th}-50^{th}; 50^{th}-10^{th}; 75^{th}-25^{th}) and, in general, all the differences were statistically significant[6]. The statistical significance of the inter-regional difference between coefficients and for a given quantile is addressed in section 5.2 with an analysis of the statistical significance of the price effect.

[tables 3 and table 4 about here]

Analysis of coefficient dispersion for the five regions, measured by the standard deviation of the coefficient estimates across the distribution, is also enlightening. For men, dispersion is typically higher at the top end of the distribution (table 3), whereas for women the pattern is less clear

[5] Due to the large number of coefficients for the five regions and for the several quantiles, we choose to display the covariates effects by graphics for selected variables.
[6] Due to the large amount of results involved these are not present. However, the results are available upon request.

(table 4). In the case of women, some variables, like education dummies and experience, display a higher dispersion at the top end of the distribution, while for others the dispersion is higher in the middle (*tenure, tenure2, Professionals*) or at the bottom of the distribution (*Senior officials and managers*). Therefore, we may conclude that the differences among the regions as regards the effects of wage explanatory variables are dissimilar across the wage distribution. All in all, these results point to different returns to characteristics across the wage distribution for each region and between regions, suggesting an uneven regional wage differential distribution.

5.2. DECOMPOSITION OF REGIONAL WAGE GAPS

In this sub-section, the decomposition of regional wage differentials in Portugal is analysed by considering the difference between the Lisboa region and other regions. Lisboa is taken as the reference region as it is that which displays the highest wages. In order to decompose the regional wage gap into the contribution of endowments and returns by quantiles, we apply both the Blinder (1973) and Oaxaca (1973) decomposition method and the Melly (2005a, 2006) decomposition method. The first step of the Melly´s estimator requires the estimation of a number J of quantile regression models for each region and gender. We estimate 200 equally spaced between 0 and 1 quantile regression models. The initial number suggested by Melly (2005a, 2006) is 100[7]. A higher number J increases the precision of the estimates but is time-consuming. Our estimations with J=100 and J=200 are quite stable, in spite of this, we choose J=200. In order to estimate the standard errors and confidence intervals of the decomposition affects, the bootstrap method was used and 100 replications were carried out.

[7] Albrecht et al. (2003) carried out their calculations using J=100 in a slightly different version of the Machado and Mata estimator.

Estimated regional wage differentials for each region relative to Lisboa along the entire wage distribution are displayed in figure 4 for men and in figure 5 for women. These figures also display the standard wage gap computed at the mean of the conditional wage distribution using the Blinder (1973) and Oaxaca (1973) decomposition and OLS estimates. At the mean of the conditional wage distribution, the regional wage gaps between Lisboa and each of the other regions range from 21.9% (Alentejo and Centro) to 27% (Norte) in the case of men; and in the case of women from 18.6% (Algarve) to 24.4% (Norte).

[figure 4 and figure 5, about here]

The quantile approach reveals rather different wage differentials across the wage distribution. In the case of men (figure 4), the estimated wage differential increases almost uniformly and linearly over the wage distribution. For women (figure 5), this pattern it is not so evident at the top of the wage distribution as there are some cases in where the wage differential decreases slightly (Norte) or where the slope of the wage differential decreases (Norte and Centro). For example, in the case of men (see also table 5), at the 10th percentile, compared to Lisboa, wages are 3% lower for the Centro region, 6.4% for Norte and about 4% for Alentejo and Algarve; at the top of the wage distribution (the 90th percentile) this differential is 46% for Centro, 48.5% for Norte, 43.2% for Alentejo and 51.5% for Algarve; finally, at the median of the conditional wage distribution, the estimated wage differential in relation to Lisboa ranges from 18.8% for the the Centro region to 26.6% for Norte. From these findings it is clear that the conclusions usually drawn using the standard Blinder and Oaxaca decomposition and OLS estimates do not reveal the whole picture and may produce inaccurate or only partially valid conclusions. This pattern of increasing wage differentials across the wage distribution was also found by Mottellón et al. (2011) for Spain, however, as we have stated, it is not possible to gauge whether these findings are statistically significant or not.

The decomposition of the wage differentials on characteristics (endowments) and coefficients (returns) effects at selected quantiles (10th, 20th, ……, 90th) using the Melly (2005a, 2006) estimator is displayed, together with the OLS estimates, in Table 5 for males and Table 6 for females. Analysis of results obtained at the selected quantiles shows that, in general, both effects (endowments and returns) are statistically significant and increase monotonically over the wage distribution for both genders.

[tables 5 and table 6 about here]

As regards the characteristics effect, Lisboa is the region with the largest endowed workforce. This effect is particularly evident at the top end of the wage distribution but is also a significant and important effect for lower quantiles. For example, for males at the 90th percentile, assuming that all workers are rewarded according to the Lisboa wage function, there is an estimated wage differential with direct input to differences in the level of observable characteristics (workers skills and firm characteristics) ranging from 21.2% for the Centro region to 32.4% for Algarve; at the 10th percentile this differential ranges from 3.1% to 10.7% for the same regions; at median position of the conditional wage distribution it ranges from 8.7% for the Centro region to 21.1% for Algarve. Finally, OLS estimates of the characteristics effect range from 10.6% for Norte to 23.5% for Algarve. Therefore, OLS estimates are a long way from representing the wage distribution as a whole.

The decomposition devised by Melly (2005a, 2006) and Machado and Mata (2005) does not allow for identifying which economic variables in the wage decomposition explain this effect or the coefficients effect. However, the existing evidence (Pereira and Galego, 2011) using the Blinder (1973) and Oaxaca (1973) decomposition method indicates that the most important factors explaining this effect are the higher percentage of large firms, more highly educated workers and more highly paid occupations for Lisboa as compared with other regions of the country. It is likely that the same variables provide an increasing contribution at upper quantiles of the wage distribution.

A similar pattern occurs for both men and women with regard to the coefficients effect. There are significant increasing wage differentials for workers with the same level of observable skills over the wage distribution. In fact, in general, a worker in the Lisboa region earns more than his or her counterparts in other regions across the entire wage distribution. For example, a Lisboa male worker with the same skills earns 21.2% more at the 90th percentile than his counterpart in the Norte; at the 10th percentile this difference is 2.7%, while at the median of the conditional wage distribution it is about 15%. In the case of Algarve, however, for workers at or below the 40th percentile, wages are higher for Algarve than Lisboa. The OLS estimate of this differential, computed at the mean of the conditional wage distribution, indicates an estimated difference of about 15% for the Norte region, which is not representative of the differentials for the wage distribution as a whole.

As the characteristics effect is significant and positive, policies for improving regional human capital in Portugal will potentially have a significant effect in the reduction of inter-regional wage inequalities. However, these policies seem to be more effective for lower quantiles of the wage distribution (low-skill and low-wage-earning workers), as for upper quantiles (highly-qualified and high-wage-earning workers) wage differentials are also highly explained by differences in returns to characteristics. Therefore, even equalizing workers' characteristics across the entire wage distribution, will not automatically lead to regional wage equalization.

The key question is why workers with the same level of observable skills earn more in Lisboa and why this differential increases typically across the wage distribution. This may be related with the fact that Lisboa is, to a great extent, a large urban area. There is a vast amount of empirical evidence (Glaeser and Maré, 2001; Yankow, 2006; Addario and Patacchini, 2008) reporting significant wage premiums for workers in large urban areas. Also, the evidence reported by Yankow (2006) for the USA supports the hypothesis that cities attract workers with higher unmeasured skills (ability bias), produce more efficient and productive matches between workers and firms (sorting effects) and create conditions for agglomeration economies. The fact that returns to characteristics are typically higher in Lisboa across the entire wage distribution may be related to agglomeration economies. These economies may cause a general increase in the productivity and wages levels. Furthermore, the inter-regional differential of returns to characteristics (price effect) typically increases across the wage distribution, which might be related to the ability bias hypothesis and/or sorting effects. Indeed, some job types are typically located in cities and especially in the capital city: senior administrative staff, directors of large enterprises, etc. Consequently, it would be very difficult, if not impossible, for workers filling those positions to earn the same wage in other locations. Hence, the great urban area of Lisboa may produce more efficient matches between workers and firms – mainly for those at top of the wage distribution - than other regions. Similar effects associated with agglomeration and urban development may be present in Algarve as it is a touristic region and densely populated. This may explain why in Algarve the part of wage differentials explained by differences in returns to characteristics is smaller than for other regions. Further investigation is, however, needed to confirm these suggestions.

Other explanations for the wage differential between Lisboa and the other regions are difficult to accept. First of all, it is quite unlikely that the large estimated wage differentials could be explained by a temporary disequilibrium situation, as other studies for Portugal (Pereira and Galego, 2011; Vieira et al., 2006), and for very different years (1995, 1996, 2000, 2002), report wage differentials of the same magnitude for workers with the same level of observable characteristics (workers' skills and firm characteristics). In fact, is difficult to believe that such disequilibrium could persist for so long. It is also very unlikely that compensating differentials related to crime, pollution or amenities might be responsible for these differentials, as there is no significant disadvantage of the Lisboa region at this level.

6. CONCLUSIONS

Previous studies carried out on regional wage differentials have typically analysed the issue using OLS estimates of wage equations and the Blinder (1973) and Oaxaca (1973) decomposition method. This approach may allow for a reasonable description of wage distributions when they are unimodal, symmetric and have similar variances. In practice, however, these conditions are unlikely to hold. Therefore, wage differentials provided by Blinder and Oaxaca decomposition, at the mean of the conditional wage distributions, may not be representative of the wage distribution as a whole.

The quantile regression model (Koenker and Basset, 1978, 1982) allows for analysis of the effect of the covariates across the wage distribution. Moreover, the extension of the Blinder and Oaxaca decomposition to the quantile regression model (Machado e Mata, 2005) enables the estimation of wage differentials at different points on the wage distribution and, consequently, a better understanding of the wage distribution. In this study, we use the quantile regression model and the estimator proposed by Melly (2005a, 2006) - asymptotically equivalent to Machado and Mata

(2005) -, which has not been previously used in this context, for estimating regional wage differentials in Portugal at selected quantiles.

Our results for regional wage equations show that coefficient estimates for covariates are quite different across the wage distribution and between the various regions. Furthermore, in the majority of cases, differences between coefficients estimates increase across the wage distribution for both males and females. Using the Melly (2005a, 2006) decomposition method, we also find that estimates of regional wage differentials at the mean of the conditional wage distribution do not provide sufficient information about the wage distribution as a whole. Clearly, regional wage differentials in Portugal increase in an almost linear fashion over the wage distribution. This pattern is similar for both the characteristics effect and the returns effect.

Our findings also suggest that public policy measures for reducing inter-regional human capital inequalities alone will not be sufficient for eliminating the inter-regional wage gap. In addition, such measures may be more efficient for low-skill and low-wage-earning workers than for highly-qualified and high-wage-earning workers. In fact, a growing and significant part of the estimated wage differential is explained by differences in returns for workers with same level of observable skills. It is quite likely that agglomeration economies, sorting effects and other mechanisms associated with urban development are the source of these differentials. Further investigation is required to provide a better understanding of these matters.

REFERENCES

Addario S and Patacchini E (2008) Wages and the city. Evidence from Italy. *Labour Economics* 15: 1040–1061

Albrecht J, Bjorklund A and Vroman S (2003) Is There a Glass Ceiling in Sweden? Journal of Labor Economics 21 (1): 145-177

Blinder A (1973) Wage discrimination: reduced forms and structural estimates. *The Journal of Human Resources* 8 (4): 436-55

Blackaby D and Manning D (1990) The north-south divide: questions of existence and stability? *The Economic Journal* 100: 510-527

Blackaby D and Murphy P (1995) Earnings, unemployment and Britain's north-south divide: real or imaginary? *Oxford Bulletin of Economics and Statistics* 57(4): 487-512

Bucther and Dinardo, (2002) The immigrant and native-born wage distributions: evidence from united states censuses. *Industrial and Labor Relations Review* 56 (1): 97-121

Cai L and Liu A (2008) Union wage effects in Australia: is there variation along the distribution?. *The Economic Record* 84: 496-510

Cai L and Liu A (forthcoming) Public–private sector wage gap in Australia: variation along the Distribution. *British Journal of Industrial Relations*

Combes P, Duranton G and Gobillon L (2008) Spatial wage disparities: sorting matters! *Journal of Urban Economics* 63: 723-742

Dickey H (2007) Regional earnings inequality in Great Britain: evidence from quantile regressions. *Journal of Regional Science* 47: 775-806

Dinardo J, Fortin N and Lemieux T (1996) Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica* 64(5): 1001-1044

Duranton G and Monastiriotis V (2002) Mind the gaps: the evolution of regional inequalities in the UK, 1982-1997. *Journal of Regional Science* 42(2): 219-256

Fujita M, Krugman P and Venables J (1999) The Spatial Economy: Cities, Regions, and International Trade. Cambridge, Mass

García I and Molina J (2002) Inter-regional wage differentials in Spain. *Applied Economic Letters* 9: 209-215

Glaeser E, Kallal H, Scheinkman J and Shleifer A (1992) Growth in cities. *Journal of Political Economy* 100(6): 1126-1152

Glaeser E and Maré D (2001) Cities and skills. *Journal of Labour Economics* 19 (2): 316-342

Goldfarb R and Yezer M (1976) Evaluating alternative theories of intercity and interregional wage differentials. *Journal of Regional Science* 16(3): 345-363

Harris C (1954) The market as a factor in the localization of industry in the United States. *Annals of Association of American Geographers* 44: 315-348

Henley A (1998) Residential mobility, housing equity and the labour market. *The Economic Journal* 108: 414-427

Lucas R (1988) On the mechanisms of economic development. *Journal of Monetary Economics* 22: 3-42

Lucifora C and Meurs D (2006) The public sector pay gap in France, Great Britain and Italy. *Review of Income and Wealth* 52 (1): 43-59.

Koenker R and Basset G (1978) Regression quantiles. *Econometrica* 46: 33-50

Koenker R and Basset G (1982) Robust tests for heteroscedasticity based on regression quantiles. *Econometrica* 50: 43-61

Krugman P (1991) Increasing returns and economic geography. *Journal of Political Economy* 99(3): 483-499

Kruskal, H and Wallis W (1952) Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47: 583–621

Kruskal, H and Wallis W (1953) Errata: Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 48: 907–911

Machado J and Mata J (2005) Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics* 20: 445-465

Marshall A (1890) *Principles of Economics*. Macmillan

Motellón E, Bazo E and Attar M (fortcoming) Regional heterogeneity in wage distributions: evidence from spain, *Journal of Regional Science*

Melly B (2005a) Decomposition of differences in distribution using quantile regression. *Labour Economics* 12: 577-590

Melly B (2005b) Public-private sector wage differentials in Germany: evidence from quantile regression. Empirical Economics 30: 505–520

Melly B (2006). “Estimation of Counterfactual Distributions using Quantile Regression”, m*imeo*, University of St. Gallen, http://www.alexandria.unisg.ch/Publikationen/22644

Oaxaca R (1973) Male-female wage differentials in urban labour markets. *International Economic Review* 14*:* 693-709

OECD (2000) *Employment Outlook*. June, Paris

Pereira J (2007) Mobilidade Geográfica e Distância da Deslocação em Portugal. *Notas Económicas* 25: 42-58

Pereira J and Galego A (2011) Regional wage differentials in Portugal: static and dynamic approaches, *Papers in Regional Science* 90(3): 529-548.

Porter M (1990) *The Competitive Advantage of Nations*. Free Press, New York

Roback J (1982) Wages, rents, and the quality of life. *Journal of Political Economy* 90(6): 1257-1278

Romer P (1986) Increasing returns and long run growth. *Journal of Political Economy* 98: S71-S101

Vieira J, Couto J and Tiago M (2006) Inter-regional wage dispersion in Portugal. *Regional and Sectoral Economic Studies 6*. URL: http://www.usc.es~economet/reviews/eers616.pdf

Wheeler C (2001) Search, sorting, and urban agglomeration. *Journal of Labor Economics* 19(4): 879-899

Yankow J (2006) Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics* 60: 139-161

APPENDIX A. Definition of variables

ln hourly wage	Logarithm of the hourly wage rate (wage rate includes base remuneration, other regularly paid components and payment for overtime; the hours of work includes the normal duration of work and overtime hours). Wages were deflated by the regional consumer price index from the INE and are at 2006 prices.
exp	number of potential years of experience in the labour market = (age - years of education - 6)
exp2	$exp^2/100$
tenure	number of years of tenure in the current job
tenure2	$tenure^2/100$
< secondary education	dummy variable; equals one if individual has less than secondary education (twelve years).
secondary education	dummy variable; equals one if individual has a secondary education (twelve years).
university degree	dummy variable; equals one if individual has a university degree.
lfsize	The logarithm of the firm size
occupational dummies	The estimations were carried out using dummies identifying occupations at one digit level of aggregation of the Portuguese occupational classification.
industry dummies	The estimations were carried out using dummies at one digit level of aggregation identifying the economic sector where the employee works.

APPENDIX B. Descriptive statistics

Table B. Distribution of occupations and industry (%)

	Norte		Centro		Lisboa		Alentejo		Algarve	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
occupations										
Senior officials and Managers	0.047	0.026	0.044	0.024	0.064	0.038	0.046	0.024	0.049	0.030
Professionals	0.049	0.072	0.049	0.068	0.090	0.104	0.033	0.054	0.037	0.052
Technicians and Associate professionals	0.102	0.088	0.103	0.084	0.163	0.142	0.091	0.088	0.088	0.088
Clerks	0.100	0.170	0.084	0.186	0.127	0.258	0.101	0.190	0.090	0.197
Service workers and shop and market sales workers	0.084	0.231	0.078	0.285	0.109	0.257	0.097	0.320	0.188	0.362
Skilled agricultural and fishery workers	0.004	0.003	0.006	0.002	0.004	0.002	0.014	0.012	0.022	0.004
Craft and related trades workers	0.349	0.222	0.322	0.109	0.200	0.023	0.323	0.076	0.255	0.023
Plant and machine operators and assemblers	0.157	0.054	0.198	0.069	0.109	0.021	0.176	0.051	0.111	0.013
Elementary occupations	0.108	0.134	0.117	0.173	0.133	0.156	0.120	0.185	0.160	0.231
industry										
Mining	0.007	0.001	0.011	0.002	0.001	0.001	0.024	0.006	0.007	0.001
Manufacture	0.321	0.364	0.320	0.271	0.130	0.073	0.246	0.192	0.058	0.035
Electricity, gas, water supply; sewerage, waste management and remediation activities	0.011	0.004	0.015	0.005	0.013	0.004	0.020	0.006	0.018	0.005
Construction	0.228	0.024	0.213	0.025	0.149	0.024	0.226	0.022	0.263	0.042
Wholesale and retail Trade	0.178	0.181	0.179	0.209	0.180	0.206	0.199	0.220	0.191	0.224
Transport and storage	0.055	0.012	0.087	0.016	0.092	0.033	0.073	0.018	0.065	0.015
Hotels and restaurants	0.031	0.061	0.029	0.077	0.057	0.095	0.041	0.094	0.191	0.301
Information and communication	0.014	0.009	0.009	0.006	0.054	0.040	0.006	0.002	0.009	0.004
Financial intermediation	0.024	0.023	0.022	0.021	0.051	0.059	0.030	0.022	0.022	0.022
Real estate, renting and business activities	0.006	0.006	0.003	0.005	0.008	0.011	0.006	0.008	0.022	0.036
Professional, scientific and technical activities	0.022	0.030	0.016	0.030	0.051	0.073	0.018	0.032	0.019	0.030
Administrative and support service activities	0.053	0.064	0.037	0.055	0.155	0.173	0.048	0.064	0.068	0.056
Public administration and defence, compulsory social security	0.009	0.017	0.012	0.018	0.009	0.012	0.023	0.025	0.020	0.026
Education	0.012	0.045	0.014	0.043	0.012	0.043	0.007	0.028	0.007	0.031
Human health and social work activities	0.014	0.116	0.016	0.165	0.017	0.108	0.018	0.213	0.014	0.122
Other services	0.017	0.043	0.018	0.053	0.022	0.050	0.017	0.049	0.026	0.049

Table 1 Sample Averages – selected variables

	Norte		Centro		Lisboa		Alentejo		Algarve	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
ln hourly wage	1.485	1.324	1.531	1.321	1.771	1.593	1.542	1.340	1.492	1.374
<Secondary education	0.746	0.650	0.730	0.625	0.559	0.465	0.726	0.627	0.721	0.632
Secondary education	0.159	0.201	0.174	0.227	0.263	0.300	0.198	0.241	0.210	0.251
University degree	0.095	0.149	0.096	0.148	0.178	0.235	0.076	0.132	0.069	0.117
Exp	22.666	20.616	23.075	21.344	21.842	20.523	23.361	22.367	22.516	21.942
Tenure	7.550	7.074	7.402	6.895	6.767	6.433	6.910	6.359	6.505	4.966
Lfsize	3.416	3.427	3.255	3.272	4.091	3.956	3.129	3.179	2.960	2.958

Table 2 Wage equations – OLS estimates

	Norte		Centro		Lisboa		Alentejo		Algarve	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Constant	0.655* (0.028)	0.698* (0.016)	0.720* (0.035)	0.725* (0.020)	0.697* (0.031)	0.677* (0.019)	0.721* (0.051)	0.756* (0.036)	0.775* (0.047)	0.847* (0.045)
Secondary education	0.168* (0.008)	0.175* (0.007)	0.116* (0.010)	0.130* (0.099)	0.171* (0.008)	0.171* (0.008)	0.128* (0.019)	0.124* (0.015)	0.112* (0.019)	0.101* (0.017)
University degree	0.507* (0.018)	0.528* (0.014)	0.447* (0.0214)	0.462* (0.017)	0.568* (0.015)	0.526* (0.012)	0.556* (0.044)	0.519* (0.033)	0.394* (0.052)	0.393* (0.034)
Exp	0.021* (0.001)	0.016* (0.001)	0.020* (0.001)	0.011* (0.001)	0.028* (0.001)	0.021* (0.001)	0.022* (0.002)	0.010* (0.002)	0.014* (0.002)	0.012* (0.002)
Exp2	-0.034* (0.002)	-0.025* (0 .002)	-0.035* (0.002)	-0.017* (0.002)	-0.045* (0.002)	-0.036* (0.002)	-0.035* (0.004)	-0.016* (0.004)	-0.024* (0.005)	-0.022* (0.004)
Tenure	0.012* (0.001)	0.011* (0.001)	0.014* (0.001)	0.013* (0.001)	0.024* (0.001)	0.022* (0.001)	0.016* (0.002)	0.014* (0.002)	0.019* (0.002)	0.014* (0.001)
Tenure2	-0.012* (0.001)	-0.014* (0.002)	-0.016* (0.003)	-0.017* (0.002)	-0.033* (0.004)	-0.025* (0.003)	-0.013* (0.007)	-0.004* (0.007)	-0.023* (0.003)	-0.014* (0.003)
Lfsize	0.080* (0.002)	0.042* (0.001)	0.071* (0.002)	0.045* (0.002)	0.069* (0.002)	0.051* (0.002)	0.075* (0.005)	0.036* (0.003)	0.088* (0.005)	0.054* (0.004)
Senior officials and Managers	0.556* (0.020)	0.515* (0.025)	0.501* (0.026)	0.431* (0.034)	0.775* (0.023)	0.684* (0.027)	0.384* (0.051)	0.343* (0.060)	0.351* (0.051)	0.443* (0.059)
Professionals	0.537* (0.022)	0.617* (0.017)	0.466* (0.027)	0.554* (0.022)	0.560* (0.018)	0.589* (0.016)	0.325* (0.057)	0.499* (0.040)	0.482* (0.064)	0.576* (0.044)
Technicians and Associate professionals	0.440* (0.012)	0.387* (0.013)	0.404* (0.015)	0.365* (0.016)	0.470* (0.013)	0.448* (0.013)	0.314* (0.033)	0.326* (0.028)	0.327* (0.033)	0.413* (0.034)
Residual Standard deviation	0.366	0.298	0.363	0.2955	0.427	0.370	0.375	0.2817	0.369	0.320
R2	0.444	0.584	0.401	0.5229	0.552	0.585	0.430	0.531	0.348	0.432
N	25156	20191	14256	11755	21790	18307	3820	3154	3300	2838

Notes: Robust standard errors in parenthesis. Industry dummies and other 5 professional dummies were included but not reported.
(*) significant at 1% level

Table 3: Wage equation coefficients estimates dispersion for the 5 regions, men

variable / percentile	Standard deviation								
	10	20	30	40	50	60	70	80	90
constant	0.057	0.073	0.071	0.076	0.057	0.031	0.043	0.071	0.066
Secondary education	0.018	0.025	0.028	0.025	0.027	0.026	0.028	0.032	0.034
University degree	0.05	0.09	0.09	0.08	0.08	0.06	0.07	0.08	0.08
lfsize	0.012	0.012	0.010	0.010	0.010	0.012	0.015	0.018	0.017
exp	0.002	0.004	0.005	0.004	0.004	0.004	0.004	0.006	0.008
exp2	0.003	0.007	0.008	0.007	0.007	0.006	0.006	0.008	0.012
tenure	0.007	0.007	0.007	0.006	0.006	0.006	0.005	0.005	0.008
tenure2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03
Senior officials and Managers	0.06	0.12	0.16	0.17	0.19	0.18	0.17	0.18	0.16
Professionals	0.07	0.08	0.08	0.10	0.11	0.09	0.07	0.07	0.13
Technicians and Associate professionals	0.04	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.14

Table 4:Wage equation coefficients estimates dispersion for the 5 regions, women

variable / percentile	Standard deviation								
	10	20	30	40	50	60	70	80	90
Constant	0.06	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.06
Secondary education	0.008	0.013	0.019	0.024	0.032	0.032	0.035	0.039	0.06
University degree	0.03	0.04	0.05	0.07	0.07	0.08	0.08	0.04	0.08
lfsize	0.009	0.008	0.007	0.009	0.009	0.009	0.006	0.004	0.009
exp	0.002	0.004	0.004	0.004	0.004	0.004	0.004	0.06	0.008
exp2	0.006	0.006	0.005	0.006	0.006	0.006	0.008	0.009	0.015
tenure	0.005	0.005	0.005	0.006	0.005	0.005	0.005	0.005	0.003
tenure2	0.006	0.007	0.010	0.009	0.008	0.007	0.010	0.009	0.003
Senior officials and Managers	0.30	0.23	0.27	0.23	0.23	0.14	0.14	0.10	0.17
Professionals	0.06	0.04	0.03	0.05	0.06	0.07	0.04	0.06	0.05
Technicians and Associate professionals	0.04	0.04	0.03	0.04	0.04	0.06	0.06	0.08	0.10

Table 5: Oaxaca's decomposition of the Regional Wage Gap- Males

Component	Norte		Centro		Alentejo		Algarve	
	Estimate	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**
OLS								
Raw difference	0.270*		0.219*		0.219*		0.254*	
Characteristics	0.106*	(0.115,0.096)	0.107*	(0.117, 0.097)	0.150*	(0.164, 0.135)	0.235*	(0.250,0.220)
Coefficients	0.165*	(0.173,0.156)	0.112*	(0.122, 0.102)	0.069*	(0.084, 0.055)	0.019**	(0.035,0.003)
Quantile .1								
Raw difference	0.064*		0.031*		0.044*		0.041*	
Characteristics	0.037*	(0.031,0.043)	0.031*	(0.022,0.039)	0.046*	(0.030,0.062)	0.107*	(0.093,0.120)
Coefficients	0.027*	(0.023,0.032)	0.001	(-0.006,0.007)	-0.002	(-0.015,0.010)	-0.066*	(-0.079, -0.053)
Quantile .2								
Raw difference	0.125*		0.071*		0.087*		0.088*	
Characteristics	0.057*	(0.051,0.063)	0.044*	(0.035,0 .053)	0.062*	(0.045,0.078)	0.135*	(0.121,0.148)
Coefficients	0.068*	(0.064,0.072)	0.027*	(0.021,0.033)	0.025*	(0.012,0.037)	-0.047*	(-0.058,-0.035)
Quantile .3								
Raw difference	0.173*		0.107*		0.126*		0.134*	
Characteristics	0.076*	(0.069,0.082)	0.056	(0.047,0.065)	0.081*	(0.065,0.098)	0.161*	(0.146,0.177)
Coefficients	0.097*	(0.093,0.102)	0.051*	(0.045,0.057)	0.045*	(0.032,0.057)	-0.027*	(-0.040,-0.015)
Quantile .4								
Raw difference	0.219*		0.145*		0.161*		0.179*	
Characteristics	0.095*	(0.088,0.103)	0.070*	(0.060,0.079)	0.1004*	(0.083,0.118)	0.187*	(0.169,0.205)
Coefficients	0.124*	(0.119,0.128)	0.075*	(0.069,0.082)	0.061*	(0.048,0.074)	-0.007	(-0.021,0.006)
Quantile .5								
Raw difference	0.266*		0.188*		0.198*		0.226*	
Characteristics	0.118*	(0.109,0.126)	0.087*	(0.076,0.098)	0.121*	(0.102,0.140)	0.211*	(0.191,0.232)
Coefficients	0.149*	(0.143,0.154)	0.101*	(0.094,0.108)	0.077*	(0.063,0.091)	0.015***	(-0.001,0.030)
Quantile .6								
Raw difference	0.317*		0.238*		0.240*		0.278*	
Characteristics	0.146*	(0.135,0.156)	0.109*	(0.096,0.121)	0.143*	(0.122,0.165)	0.237*	(0.214,0.261)
Coefficients	0.172*	(0.165,0.178)	0.129*	(0.121,0.137)	0.097*	(0.081,0.112)	0.041*	(0.023,0.059)
Quantile .7								
Raw difference	0.373*		0.296*		0.289*		0.340*	
Characteristics	0.183*	(0.171,0.196)	0.136*	(0.121,0.151)	0.170*	(0.143,0.196)	0.264*	(0.237,0.291)
Coefficients	0.189*	(0.181,0.198)	0.160*	(0.149,0.170)	0.119*	(0.101,0.137)	0.075*	(0.054,0.096)
Quantile .8								
Raw difference	0.431*		0.367*		0.352*		0.416*	
Characteristics	0.229*	(0.214,0.245)	0.171*	(0.153,0.190)	0.214*	(0.179,0.250)	0.293*	(0.260,0.327)
Coefficients	0.202*	(0.190,0.213)	0.196*	(0.183,0.209)	0.138*	(0.115,0.160)	0.123*	(0.097,0.149)
Quantile .9								
Raw difference	0.485*		0.460*		0.432*		0.515*	
Characteristics	0.273*	(0.253,0.292)	0.212*	(0.186,0.237)	0.268*	(0.216,0.321)	0.324*	(0.277,0.370)
Coefficients	0.212*	(0.197,0 .228)	0.249*	(0.231,0.267)	0.164*	(0.131, 0.197)	0.192*	(0.156,0.227)

Notes: CI – Confidence interval; (*), (**),(***) - significant at 1%,5 % and 10% level, respectively

Table 6: Oaxaca's decomposition of the Regional Wage Gap- Females

Component	Norte		Centro		Alentejo		Algarve	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
OLS								
Raw difference	0.244*		0.243*		0.224*		0.186*	
Characteristics	0.108*	(0.120,0.096)	0.112*	(0.122,0.101)	0.136*	(0.151,0.122)	0.187*	(0.203,0.172)
Coefficients	0.136*	(0.147,0.125)	0.131*	(0.141,0.122)	0.087*	(0.100,0.075)	-0.001	(0.014,-0.015)
Quantile .1								
Raw difference	0.044*		0.034*		0.011*		-0.008*	
Characteristics	0.038*	(0.033,0.042)	0.034*	(0.028,0.040)	0.034*	(0.022,0.045)	0.069*	(0.056,0.082)
Coefficients	0.006*	(0.004,0.009)	-0.00003	(-0.004,0.004)	-0.023*	(-0.033, -0.013)	-0.077*	(-0.086,-0.067)
Quantile .2								
Raw difference	0.104*		0.082*		0.049*		0.025*	
Characteristics	0.061*	(0.056, 0.066)	0.049*	(0.043,0.055)	0.043*	(0.030,0.055)	0.087*	(0.074,0.100)
Coefficients	0.043*	(0.040,0.046)	0.033*	(0.029,0.038)	0.007	(-0.003,0.017)	-0.062*	(-0.073, -0.052)
Quantile .3								
Raw difference	0.157*		0.127*		0.093*		0.063*	
Characteristics	0.084*	(0.077,0.090)	0.062*	(0.055,0.069)	0.055*	(0.041,0.070)	0.103*	(0.089,0.117)
Coefficients	0.073*	(0.069,0.076)	0.065*	(0.060,0.070)	0.037*	(0.026,0.048)	-0.040*	(-0.051,-0.029)
Quantile .4								
Raw difference	0.209*		0.177*		0.143*		0.110*	
Characteristics	0.107*	(0.010,0.115)	0.077*	(0.069,0.085)	0.073*	(0.056,0.090)	0.121*	(0.105,0.137)
Coefficients	0.102*	(0.097,0.107)	0.100*	(0.094,0.105)	0.070*	(0.058,0.083)	-0.012*	(-0.024,0.001)
Quantile .5								
Raw difference	0.263*		0.232*		0.201*		0.163*	
Characteristics	0.134*	(0.125,0.142)	0.097*	(0.087,0.106)	0.095*	(0.075,0.114)	0.141*	(0.123,0.160)
Coefficients	0.130*	(0.124,0.136)	0.136*	(0.129,0.142)	0.107*	(0.093,0.121)	0.022*	(0.008,0.036)
Quantile .6								
Raw difference	0.321*		0.296*		0.269*		0.226*	
Characteristics	0.168*	(0.157,0.178)	0.123*	(0.111,0.135)	0.126*	(0.102,0.151)	0.166*	(0.144,0.189)
Coefficients	0.153*	(0.146,0.161)	0.172*	(0.165,0.180)	0.143*	(0.126,0.160)	0.060*	(0.043,0.077)
Quantile .7								
Raw difference	0.377*		0.367*		0.345*		0.297*	
Characteristics	0.208*	(0.195,0.221)	0.157*	(0.141,0.172)	0.174*	(0.140,0.208)	0.198*	(0.168,0.228)
Coefficients	0.169*	(0.160,0.179)	0.210*	(0.200,0.220)	0.171*	(0.150,0.193)	0.099*	(0.078,0.120)
Quantile .8								
Raw difference	0.415*		0.439*		0.418*		0.372*	
Characteristics	0.240*	(0.224,0.256)	0.200*	(0.181,0.220)	0.235*	(0.190,0.280)	0.233*	(0.193,0.272)
Coefficients	0.175*	(0.161,0.189)	0.238*	(0.225,0.252)	0.183*	(0.155,0.212)	0.140	(0.110,0.169)
Quantile .9								
Raw difference	0.396*		0.473*		0.456*		0.438*	
Characteristics	0.231*	(0.211,0.252)	0.225*	(0.199,0.250)	0.262*	(0.201,0.323)	0.272*	(0.209,0.336)
Coefficients	0.164*	(0.145,0.183)	0.248*	(0.229,0.268)	0.194*	(0.156,0.232)	0.165*	(0.124,0.207)

Notes: CI – Confidence interval; (*), (**),(***) - significant at 1%,5 % and 10% level, respectively

Figures

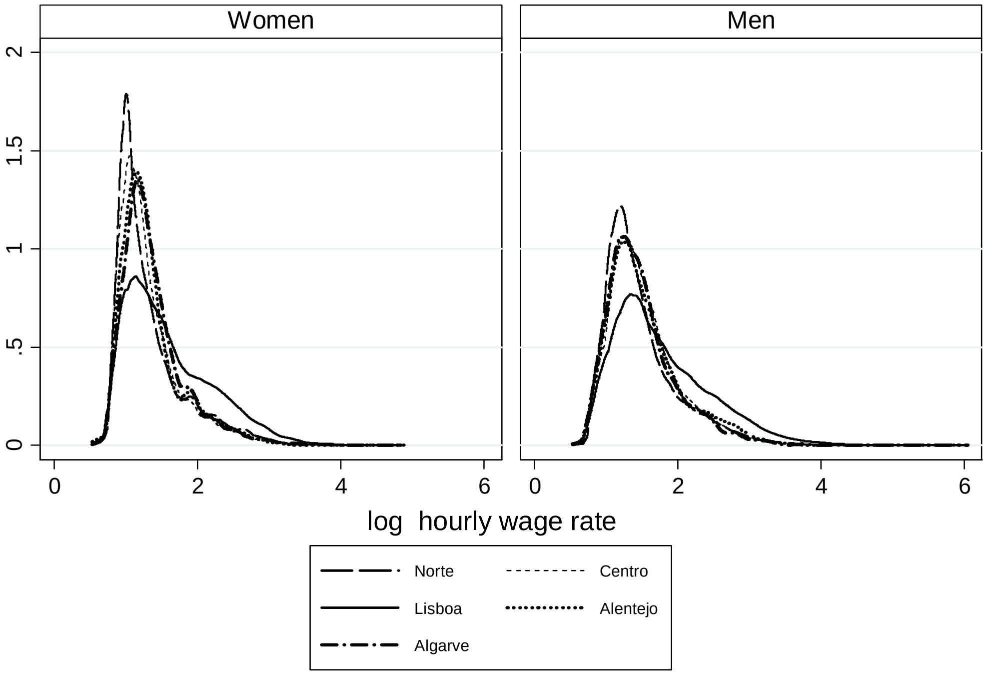

Figure 1: regional real hourly wage kernel density estimates

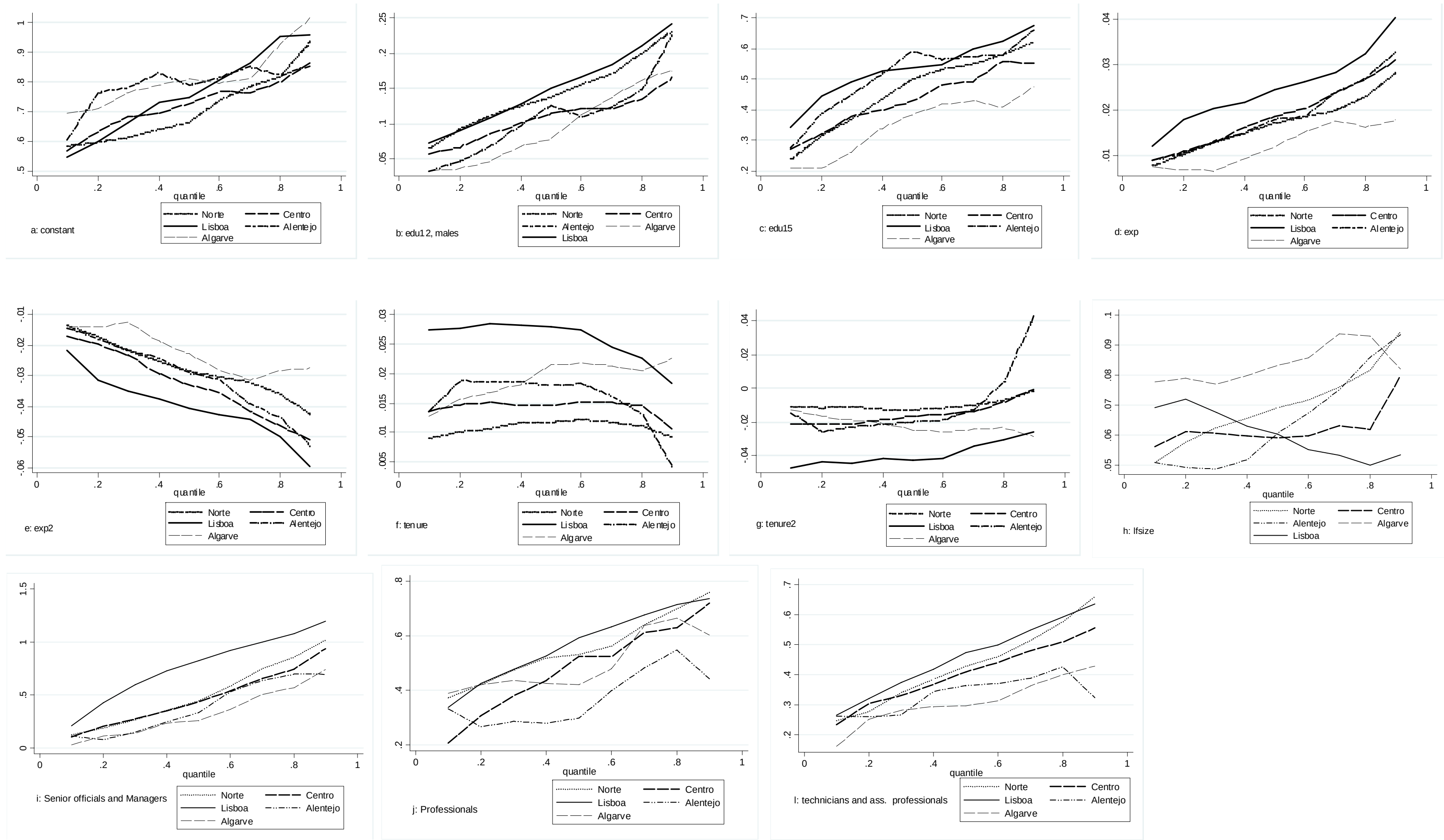

Figure 2: coeficientes estimates by quantile , males

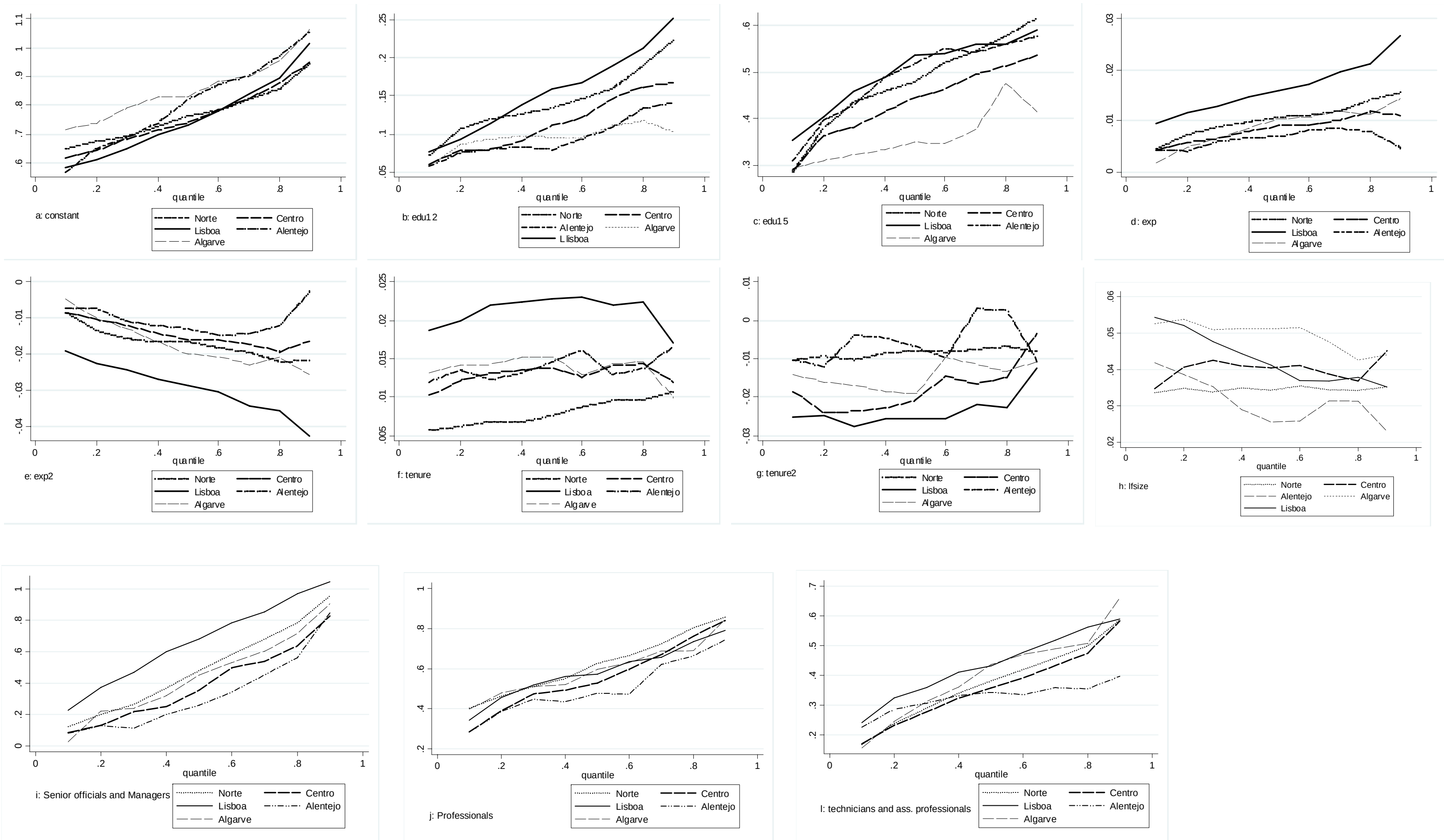

Figure 3: coeficientes estimates by quantile, females

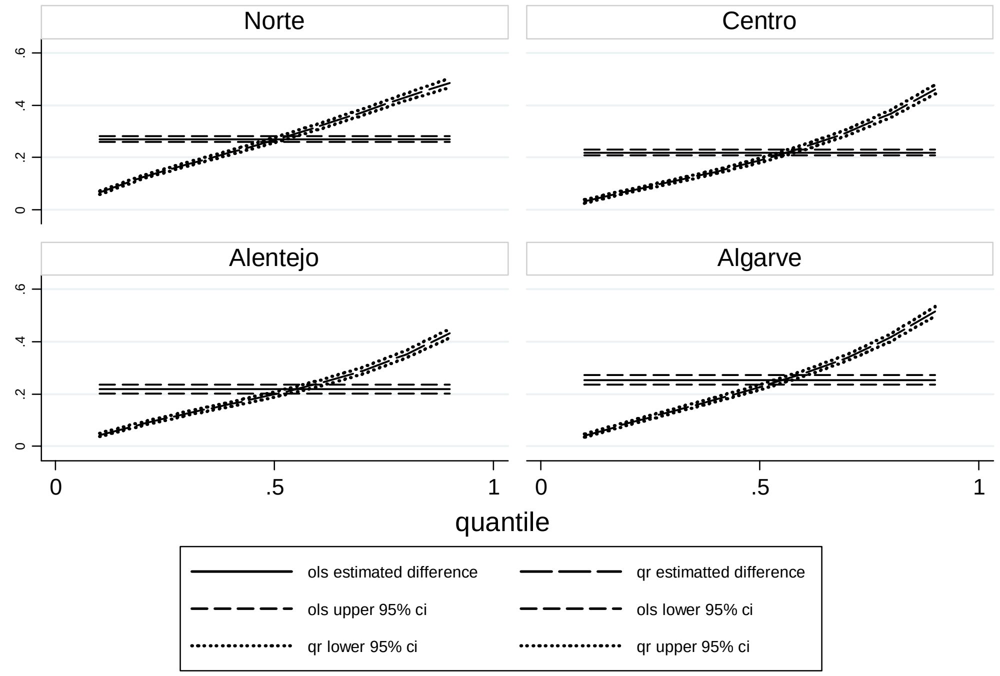

Figure 4: estimated regional wage differentials, males

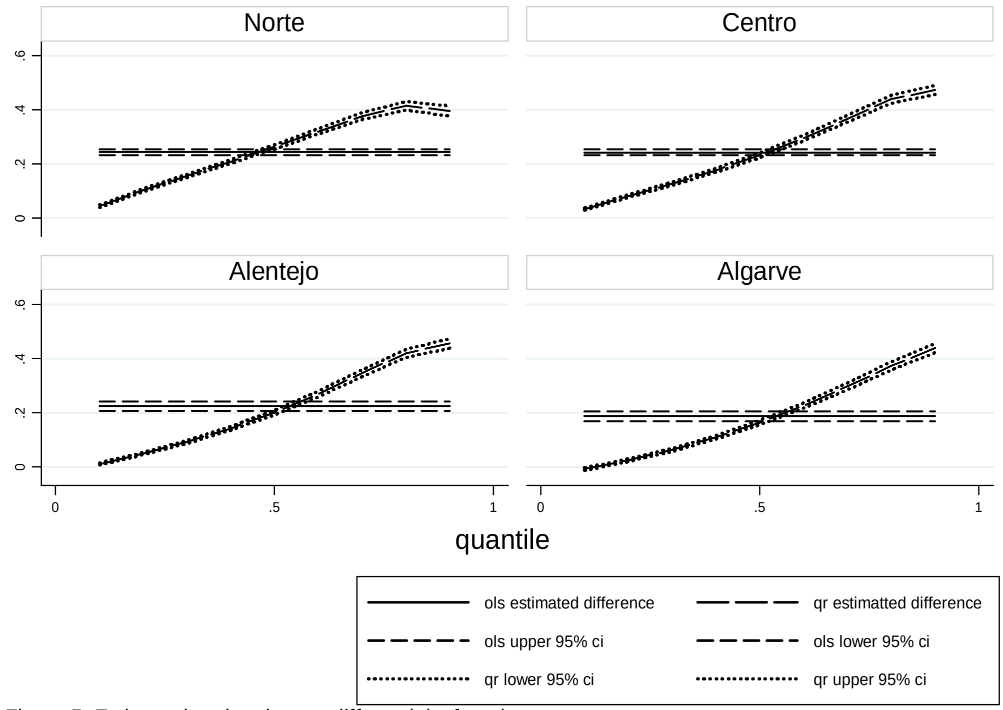

Figure 5: Estimated regional wage differentials, females