Department of Economics and Management

DEM Working Paper Series

The Analysis of the Gender Wage Gap in the Italian Public Sector: a Quantile Approach for Panel Data

Carolina Castagnetti
(Università di Pavia)

 # 109 (11-15)

November 2015

The Analysis of the Gender Wage Gap in the Italian Public Sector:

a Quantile Approach for Panel Data

Carolina Castagnetti[*]

Dipartimento di Scienze Economiche e Aziendali, Università di Pavia, Italy

November 4, 2015

Abstract

This paper analyzes the gender wage gap in the Italian public sector for the years $2005 - 2010$. We find a consistent level of gender wage gaps and an increasing path along the wage distribution. We use quantile regression methods to estimate and decompose the wage gap. The decomposition analysis supports the idea of a *glass ceiling* mechanism in action. However, the results change dramatically when we take into account the unobserved individual-specific heterogeneity by means of quantile technique for panel data. The evidence of a glass ceilings vanishes and the significant unexplained GWG is almost stable across the distribution.

Keywords: Gender wage gap, quantile regression for panel, public-private wage differential

JEL - Classification: J3, J45

1 Introduction

Although, there exist substantial literatures on the gender wage gap (henceforth GWG) in the private sector on the one hand and on the public-private wage premium by genders, on the other hand,[1] to our knowledge the analysis of the public sector gender wage gap throughout the wage distribution has been poorly analyzed.

---

[1]See Altonji and Blank (1999) for a survey on the gender gaps for the United States, Blau and Kahn (2003a) for international comparisons of gender wage gaps, and Giordano, Depalo, Pereira, Eugene, Papapetrou, Perez, Reiss, and Roter (2011) for an analysis of the public-private wage differentials in the Euro area.

One exception is represented by the work of Barón and Cobb-Clark (2010) and by the analysis of Blau and Kahn (2003b). Barón and Cobb-Clark (2010) investigate the GWG across public and private sector wage distribution for the Australia. They find that the gender wage gap among high-wage workers is largely unexplained in both the private and the public sector while is more than explained by differences in individual characteristics among low-paid workers. This finding suggests that glass ceilings rather than sticky floors may be prevalent in explaining the gender wage gap in the different sectors.

The results of Barón and Cobb-Clark (2010) are confirmed by the analysis of Blau and Kahn (2003b). They find that on average, discrimination on the basis of gender differences, as well as the differential between male and female income increase over the work life of an individual. Further, the unexplained gender gap in the public sector increases along the wage distribution and with respect to the private sector.

The focus of this paper is the analysis of the GWG in the Italian public sector and the decomposition in its determinants. To this task we first consider the public-private pay determination by gender; an higher level of the unexplained component for the female sample with respect to the male one, for instance, may be considered as a softening of a gender discrimination in the public sector, if any. Moreover, the quantile analysis may shed some light on the dynamic and the persistence of the aforementioned behaviors through the wage distribution.

In general, the analysis on the public-private pay differential has shown that the wage gap tends to be higher in the case of woman and typically narrows as one moves up the earning distributions. The difference in pay is usually lower in the public compared to the private sector (see for example Melly (2005b); Lucifora and Moeurs (2006), Giordano, Depalo, Pereira, Eugene, Papapetrou, Perez, Reiss, and Roter (2011)). In particular, Lucifora and Moeurs (2006) show that there are differences by gender, and returns to skill between the public and private sector. Giordano, Depalo, Pereira, Eugene, Papapetrou, Perez, Reiss, and Roter (2011) find the existence of different premium across the wage distribution and across genders; this premium is typically higher for female workers compared to their male counterparts. However, in countries with more pronounced wage compression, the premium across quantiles for women is flatter than in other group of countries.

Several are the theoretical interpretations of differences in pay among sectors; Gregory and Borland (1999), among others, argue that these differences in wage structure are not surprising given that wage setting in the public sector occurs in a political environment whereas private-

sector decision making occurs in a market environment. The empirical evidence on how the relative wage of men and women varies across different sectors has shown that the mean gender wage gap is typically considerably smaller in public-sector jobs (see Arulamplam, Booth, and Bryan (2006), Gregory and Borland (1999), Gunderson (1989)) while the distribution of wages varies dramatically across sectors (see Arulamplam, Booth, and Bryan (2006), Kee (2006)).

Because of the more standardized career path in the public sector with respect to the private one and the different hiring selection method (by competition in the public sector), the unexplained component of the gender wage gap, at least at the beginning of the career, should be lower with respect to those for the private one because the so called *discrimination* components should be counteracted. Hence, we are expecting a larger unexplained component for female public-private wage gap with respect to the male counterpart; at the beginning of the career because of the public wage premium and at the end because of non discriminatory forces operating.

However, the analysis does not completely support what we were expecting. While we find higher premia for female workers compared to their male counterparts, the public sector wage gap is accounted completely by differences in returns at the beginning of the income distribution while this component vanishes in the upper part of the distribution.

On the contrary, when we analyze the GWG within sectors, the results are in line with those highlighted by Barón and Cobb-Clark (2010) and Blau and Kahn (2003b). We find a lower level of the GWG in the public sector with respect to the private one, but still significant. More interesting, the GWG increases along the wage distribution in both sectors. The Oaxaca-Blinder decomposition shows that the unexplained component mostly exceeds the explained part and the distance grows as the wage increases. This pattern is much more evident in the public sector where we find evidence of a *glass ceiling* mechanism in action.

The results change dramatically when we take into account the unobserved individual-specific heterogeneity. To this task, we perform the analysis by considering the quantile approach for panel data proposed by Canay (2011). In order to assess how the GWG varies across the wage distribution we propose a two step procedure for computing the Oaxaca-Blinder decomposition. First, we estimate the Public GWG by means of Canay (2011) approach and then we run the Machado and Mata (2005) decomposition for quantile regression.

We find a sharply decrease of the public premium which suggests that there is a positive selection into the public sector, especially for females. Second, when we look at the GWG within sectors, the evidence of a glass ceilings in the public sector vanishes. However, in both sectors there is a

significant unexplained GWG almost stable across the distribution.

The paper is organized as follows. Section 2 presents the econometric approach. Section 3 describes the data. Section 4 reports and discusses the results about the Public sector premia and the GWG within sectors on the basis of standard cross section analysis. Section 5 extends the analysis to the longitudinal sample. Section 6 concludes.

2 Econometric Methodology

2.1 Estimation and Decomposition Method

We estimate the wage equations by means of quantile regression, as developed by Koenker and Bassett (1978). Following Buchinsky (1998) and assuming a linear specification, the model is defined as

$$q_\theta(y_i|x_i) = x_i'\beta_\theta \tag{1}$$

$$y_i = x_i'\beta_\theta + u_{\theta i} \tag{2}$$

where $q_\theta(y_i|x_i)$ defines the *conditional quantiles* of the dependent variable y (log wages), given the covariates x (individual characteristics). The distribution of the error term $u_{\theta i}$ is left unspecified and it is assumed that $q_\theta(u_{\theta i}|x_i) = 0$.

To investigate the gender wage gap in the public sector, we estimate this model for men and women separately at different quantiles, namely $\theta = \{0.10, 0.25, 0.50, 0.75, 0.90\}$. Results based on quantile regressions provide a complete view of how the wage gaps between and within sectors varies along the distribution. Moreover, as the quantile regression (QR) allows the regressors, i.e. individual observable characteristics, to have a different impact at different quantiles, we can control more deeply for differences between men and women's wages that depend on their characteristics.

2.2 Quantile Decomposition

To decompose the wage gap in explained and unexplained components, we make use of the procedure proposed by Machado and Mata (2005), that generalizes the Oaxaca-Blinder decomposition to a quantile regression framework. The advantage of the quantile decomposition is that we can estimate the unexplained component of the wage gap across the distribution of income, that is, at any quantile of the wage distribution.

While in the Oaxaca-Blinder setting, the wage gap is divided by means of a *counterfactual wage structure*, the Machado and Mata (2005) decomposition is based on the construction of a counterfactual distribution of y^f, i.e. a distribution of what would be female income, had the wage structure been the same as the male one.

Let $k \in \{m, f\}$ represent male and female observations, so that we have samples $\{(y_i^k, x_i^k) : i = 1, \ldots, n_k\}$ for all populations k, and we can estimate $q_\theta(y^k)$ separately for the two groups.

Formally, the Machado-Mata approach to estimate the counterfactual distribution of y^f can be summarized as follows:[2]

1. Draw a random sample θ_i^*, $i = 1, 2, \ldots, 5000$ from a uniform distribution $U[0, 1]$.

2. For each θ_i, estimate $\beta^m(\theta)$ and $\beta^f(\theta)$ as

$$\hat{\beta}^k(\theta_i^*) = \arg \min_{\beta \in \mathbb{R}^p} \sum_{j=1}^{n_k} \rho_{\theta_i^*}(y_j^k - x_j'^k \beta) \quad k = m, f.$$

   using the male and female dataset, respectively.

3. randomly draw $5,000$ women with replacement and use their characteristics (x^{*f}) to predict the wages using the estimated coefficients $\beta^m(\theta)$ generating a set of predicted wages, $\tilde{y}^f(\theta) = x'^{*f} \hat{\beta}^m(\theta)$. The empirical c.d.f. of these values is the estimated counterfactual distribution, namely what women would have earned if they were paid like men.

4. Then compare the counterfactual distribution with the empirical male and female distributions whose θquantiles are defined by $\hat{y}^m(\theta) = x'^m \beta^m(\theta)$ and $\hat{y}^f(\theta) = x'^f \beta^f(\theta)$, respectively.

As in the Oaxaca-Blinder decomposition for the mean differential, the pay gap between males and females can be divided in two parts; one representing the effect of different characteristics between the two groups; the other representing differences unexplained by the quantile regression model. The advantage of the quantile decomposition is that we can estimate the two components across the distribution of income, that is, at any θth quantile of the wage distribution.

More precisely, we can write

$$y^m(\theta) - y^f(\theta) = [\hat{y}^m(\theta) - \tilde{y}^f(\theta)] + [\tilde{y}^f(\theta) - \hat{y}^f(\theta)] + residual \tag{3}$$

---

[2]The decomposition proposed by Machado and Mata (2005) grounds on the probability integral transformation theorem from elementary statistics: if U is uniformly distributed on $[0, 1]$, then $F^{-1}(U)$ has distribution F. Thus, for a given x_i and a random $\theta \sim U[0, 1]$, $x_i'\beta(\theta)$ has the same distribution as $y_i | x_i$. If, instead of keeping x_i fixed, we draw a random x from the population, $x'\beta(\theta)$ as the same distribution of y.

where $y^k(\theta)$ denotes the observed log wages for $k = (male, female)$, $\hat{y}^k(\theta)$ denotes the estimator of the $k = (male, female)$ log wages based on the observed sample, and $\tilde{y}^f(\theta)$ denotes the estimated counterfactual log wages. By *counterfactual*, we mean the wage that females would get, if their abilities had been rewarded according to the male pays' schedule. The residual term captures the changes unaccounted for by the estimation method.

The first part of the income differential is the so-called *characteristics* effect, since it is the consequence of the different distribution of covariates for the two groups. The second addend in (3) represents the so-called *coefficient* effect, since it is obtained by evaluating female characteristics using two different conditional distributions.

In the following analysis we make use of the estimation procedure for standard errors proposed by Chernozhukov, Fernandez-Val, and Melly (2013). In fact, Machado and Mata (2005) proposed quantile regression-based estimators to evaluate distributional effects, but provided no econometric theory for these estimators. The asymptotic behavior of the estimators' error is studied by Chernozhukov, Fernandez-Val, and Melly (2013) who also show the validity of *exchangeable bootstrap* methods to obtain the asymptotic covariance matrix.

2.3 Quantile Regression for Panel

In order to take into account the unobserved individual heterogeneity in explaining the wage gap across the distribution, we extend our empirical analysis by exploiting the longitudinal structure of the data.[3] To this task, we consider the following quantile regression fixed effect model (hereafter FE-QR):

$$Q_\theta(y_{it}|x_{it}) = \alpha_i + x'_{it}\beta_\theta \qquad (4)$$

$$y_{it} = x'_{it}\beta_\theta + u_{\theta it} \qquad (5)$$

While estimation methods for cross-sectional conditional quantile regression models are well developed, corresponding methods for panel data (especially FE models) have received attention only recently. One problem associated with FE-QR is that, as it is the case with nonlinear panel data models, the method of differencing out the fixed effects used for the conditional linear mean model does not carry over to the conditional quantiles. Koenker (2004) proposes to treat each individual effect as a parameter to estimate[4] by means of a penalized estimation method. However,

---

[3] See Section 3 for the characteristics of the data when we rely on panel observations.

[4] The individual parameters are assumed to have a pure location shift effect on the conditional distribution of the

controlling fixed effects by directly estimating them is not without difficulty - known as incidental parameter problem (Neyman and Scott, 1948), which manifest itself in inconsistency of the common parameters when the number of individuals goes to infinity while the number of time period is fixed.[5]

A second problem arises because the objective function is not differentiable. The implication is that standard asymptotic analysis of panel data model is not directly applicable to QR. Kato and Galvao (2011) propose the smoothing of the objective function and study the properties of the estimator. They show that the estimator is asymptotically normally distributed and propose a bias correction for the estimator's mean. Flores, Flores-Lagunes, and Kapetanakis (2014) estimate a two-way fixed effects model where both effects vary over quantiles. Flores, Flores-Lagunes, and Kapetanakis (2014) account for the problem of *quantile crossing* adopting the method proposed by Chernozhukov, Fernandez-Val, and Galichon (2010) to transform the original estimated quantiles into monotonic ones. However, the objective function they consider is not smooth and they rely on Monte Carlo experiment to show the small bias in their estimates.

Alternative approaches that not consider the case of unobserved heterogeneity represented by the classical individual effects α_i are introduced by Harding and Lamarche (2014) who propose a quantile regression estimator for a model with a multifactor error structure and interactive effects potentially correlated with covariates.

In our application we follow the approach proposed by Canay (2011). Assuming that the individual parameters (fixed effects) have a pure location shift effect, Canay (2011) proposes an easy-to-use two-step estimator. In the first step, the individual effects α_i are estimated by traditional mean estimations (for instance estimation in first differences or by means of the within estimator), then corrected wages $\hat{y}_{it} = y_{it} - \hat{\alpha}_i$ are estimated on the other covariates by means of traditional quantile regression. Given $\hat{y}_{it}$ we estimate the wages by quantile regression and we rely on Machado and Mata (2005) method to decompose the wage gap in observed and unobserved components.

3 Data

To carry out our analysis we exploit individual data drawn from the 2005, 2006, 2008 and 2010 waves of the ISFOL-PLUS survey.[6] Since the first PLUS survey, in 2005, each consecutive year includes a proportion of panel interviews, taken from the previous sample. In the analysis we ex-

---

dependent variable.

[5]The analysis of an incidental parameter problem in FE-QR is described in Graham, Hahn, and Powell (2009) and Kato and Galvao (2011).

[6]Isfol: Istituto per lo Sviluppo della Formazione Professionale dei Lavoratori.

ploit both the cross-section dimension of the database, considering only the 2005 and 2010 samples, and the panel dimension, considering all the available years. The target population is composed of individuals between 15 and 64 years old,[7] and the sample size for each year is reported in Table 1.

Table 1: Isfol-PLUS Samples Size and Composition

Year		2005	2006	2008	2010
Observations	*Males*	16,292	16,825	15,277	17,817
	Females	24,094	20,688	18,653	20,858
	Total	40,386	37,513	33,930	38,675

The ISFOL-PLUS questionnaire is composed of specific sections for five sub-groups of the population: young individuals between 15 and 29 years old; women between 20 and 49 years old; elderly population between 50 and 64 years old; unemployed individuals; employed population. A rich set of information for each of these categories is included, ranging from family characteristics to individual skills and personal history. Even if the response rate for some of these sections were not very high, it is possible, for properly specified subsets of the sample, to provide a detailed explanation of some individual decisions (such as the choices of working, having children, studying) and to investigate which social environment factors may have influenced them.

Despite the fact that also self-employed and those with project-linked positions are present in the PLUS samples, for our analysis we have considered only salaried employees, which form by far the largest category. Indeed, since data on income is not harmonized between these three groups of workers, there are some difficulties in comparing them.

We have used log-hourly net income (adjusted to the 2010 level) as the dependent variable. Moreover, we have selected a group of around twenty independent variables, which include: years of schooling; a quadratic term for market experience; type of contract (part-time, short-term); family characteristics (civil status, presence of pre-schooling age children); sector, occupation and industry dummies; geographic variables (denoting people living in northern regions, and people living in urban areas); personal skills, as approximated by the university final mark (adjusted for years lost),[8] knowledge of English, and knowledge of how to use a computer for particular basic tasks.

As already mentioned, we will look at the evolution of the gender wage gap by analyzing two

---

[7]In 2005, there were $38,827,322$ individuals aged 15-64 in Italy, while the same population was $39,655,921$ in 2010.

[8]The use of this variable (here labelled S) has been suggested by Castagnetti, Chelli, and Rosti (2005) and is

cross-section samples of salaried employees, taken from the 2005 and 2010 ISFOL-PLUS data sets. The observations for which only an imputed wage was available have been excluded, together with those observations which showed a monthly net income lower than roughly 120 Euro.[9] Finally, we have randomly assigned all the observations for which a panel was available, by considering either 2005 or 2010 income with equal probability.

The samples obtained this way are composed of 8,807 individuals in 2005 (4,185 males), and of 8,546 individuals in 2010 (4,217 males).
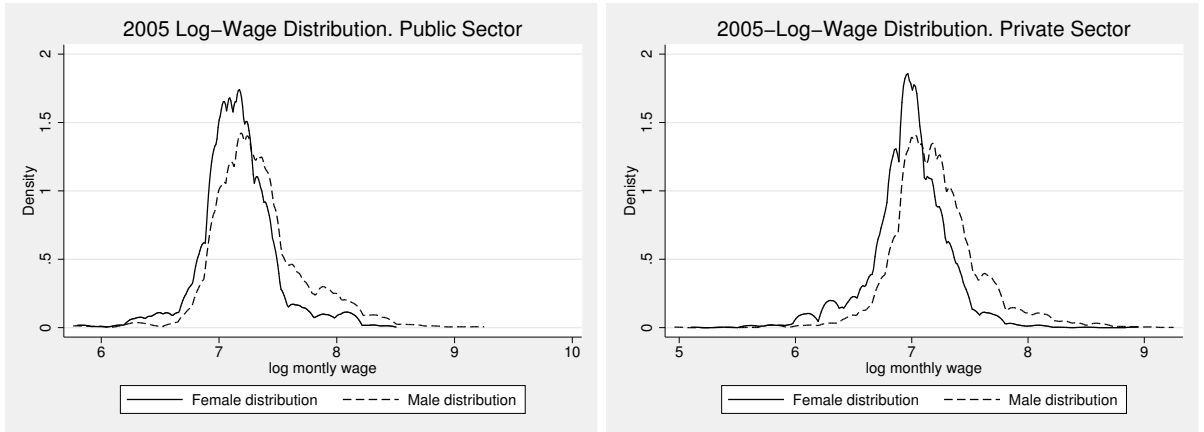

Figure 1: Kernel-density estimates of monthly net log-wages, divided by sector. The dotted line represents the male distribution.
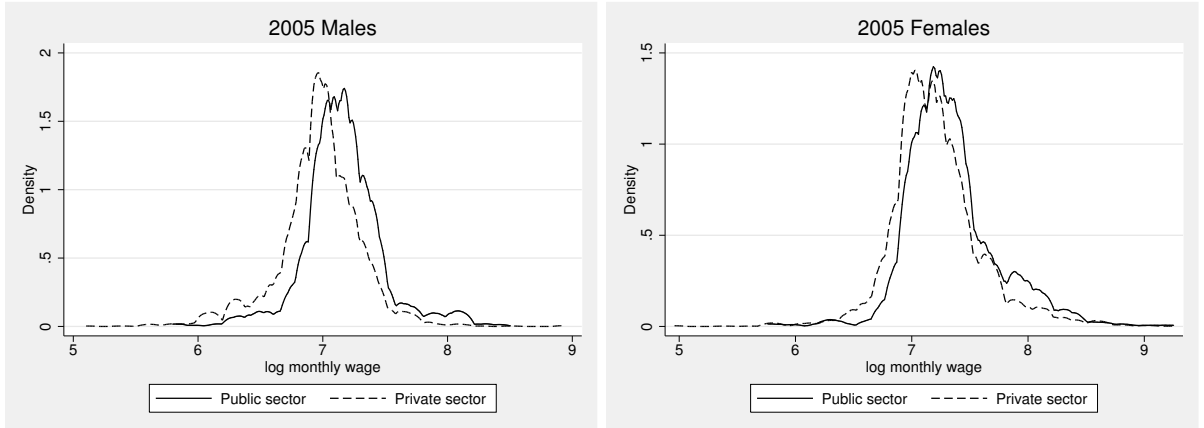

Figure 2: Kernel-density estimates of monthly net log-wages, divided by gender. The dotted line represents the private sector.

A first summary of the wage distribution across sectors, gender and within sectors is provided in Figures 1 and 2. The density functions were estimated using Epanechnikov kernel estimator. In Figure 1, we have plotted the kernel-density estimation of the log-wage distributions in the public

calculated as follows

$$S = \frac{\text{University Mark}}{1 + (0.05 \times \text{Years Lost})}$$

[9]The lower threshold has been 4.8 in log-wage terms

and the private sector, respectively. For both sectors, the female income distribution is shifted to the left with respect to the men's one, which gives us a preliminary evidence of a gender pay gap.

Figure 2 shows the wage distribution across sectors. It can be seen that for both genders, the public sector earnings distribution is characterized by a higher density function around the mode and a lower dispersion. The graph show that the public sector compresses the unconditional wage distribution for both genders. However, at this preliminary stage, we are considering only the unconditional wage distribution, without taking into account possible factors affecting it.

In Table 2 we have reported mean and standard deviation for some of the controls that will be included in our analysis. The selected data is described for the initial and final years of the period under investigation. As can be noticed, on average, men earn higher salaries in both sectors and have a longer working history. On the other hand, women have more years of schooling and show higher university performance. The number of years of schooling has been constructed from the available information on educational attainments, and thus has a relatively low variability. Table 2 shows that workers in the public sectors have more experience on average; they have more frequently achieved a university degree while employees in the private sector have more often reached an high school education only.

4   Cross Section Analysis

In this section we are going to discuss the main findings with respect to the cross-section dimension of our data sets. To this task we have considered the observations referring to 2005 and 2010, only. The principal technique employed in this context has been a quantile regression-based decomposition whose main strength lies in the fact that it allows to estimate productivity and discrimination wage differentials across the wage distribution. Since this analysis relies heavily on it, as a preliminary step we have estimated quantile regression at five quantiles of the income distribution, considering our sample separately by gender and by year. In this way, we were able to estimate the overall accuracy of our Mincerian wage equations, to test the significance of each of our proxies for productivity, and to appreciate any difference, among the time and the gender dimensions, in the shape taken by the wage structures. As preliminary step of the analysis, our specification includes in addition to standard individual and work-related characteristics, a dummy indicating public sector employment to identify the public sector wage premium. Results, for a selected number of controls are presented in Tables 3 and 4.[10]

---

[10]The full set of estimation results are available by the author upon request.

In reading the Tables, recall that each coefficient represents the effect on income, at a given quantile, of a shift in the corresponding covariate, keeping all else constant. The standard errors were computed, using the bootstrap method with 800 replications, a procedure that involves weaker assumptions with regard to the distributional form taken by the variables of interest, since it provides a consistent estimate even in the presence of heteroskedasticity.[11] Indeed, quantile regression is able to capture both the shape and the shift effect on the response variable, and bootstrap standard errors are reliable under both circumstances.

The overall specification of the model seems to fit well, and most of the reported coefficients have the predicted signs. It's worth noticing a gender variation in the impact of the variable denoting employment in the public sector. While this dummy is not always significant for males, the same is not true for females, which tend to have lower wages in the private sector. When we look at the dummy for being a public sector employee, the corresponding parameter represents the public-private wage gap. The higher rent gained by women working in the public sector suggests that the gender pay gap might be stronger in the private sector.

A second evidence coming out from this model is that the return on schooling tends to be stronger for women only at lower quantiles of the distribution. Further, the return on the *University Performance* seems higher for male. This findings is in contrast to what found by Castagnetti and Rosti (2009) for Italian graduate students where they show that females face a greater increase in labor market returns from signalling through academic performance.

Among the variables not included in the tables, occupational dummies tend to be more significant than sector dummies, a fact that is reflected by the high standard error measured for the variable denoting employment in the production-of-goods sector. The geographical variables included tend to be significant. In particular, there is a considerable gain among individuals living in Northern Italy, as compared to individuals living in central and southern regions. Finally, while the dummies denoting the presence of children are significant only for a small number of quantiles, the variables denoting civil status tends to be more significant in both years, especially among men.

In the following of the analysis, we analyze the public-private pay gap by means of the decomposition method provided by Machado and Mata (2005). The dummy-based approach whose results are presented in Table 3, has the important shortcoming of assuming that the return to individual and job characteristics are the same in both sectors. Moreover, when in Section 4.3 we

---

[11]Two good and short reviews on inference methods for quantile regression are given by Buchinsky (1995) and by Buchinsky (1998). For a more comprehensive treatment of the topic, see Koenker (2005).

conduct the analysis using a panel data model we have to take into account the low variability of the "public sector" dummy. In fact, we identify 75 changes from the public to the private sector and 153 switches in the opposite direction. Hence, the variation of the "public sector" variable is mostly cross-sectional and there is a little variation across time. In this case the fixed effect estimates of the dummy variable "public sector" may be biased.

4.1 Public-Private Sector Wage Differential

The results for the public-private wage gap are presented in Figure 3 for the 2005 cross section data and in Figure 4 for the pooled data[12] for females and males, respectively. We present results for the pooled data also because pooling data across waves makes results more robust to shocks affecting the labour markets in specific years. Moreover, increasing the number of observations improves the precision of the estimates.[13] For both genders, we observe that there is a positive wage gap for those working in the public sector, which shrinks as we move towards higher percentiles of the income distribution. These results confirm the findings of the the dummy-based approach shown in Table 3.

These results confirm the findings of Disney and Gosling (1998), Lucifora and Moeurs (2006) and Campos and Centeno (2012), among others. Disney (2007) considers several explanation for the observed differences in pay between the Public and the Private sectors. One justification often found in the literature for the public wage gap is that governments are less competitively-driven than the private sector and more inclined to equity and fairness in wage settlements, which translates into higher earnings than market levels at the bottom and moderate remuneration to top level managers. Another line of argument concerns the potential existence of compensating differentials, i.e. in-kind advantages and fringe benefits that would offset pay differences. However, this type of gain is to be found mostly in the public sector, where employees benefit from job protection and more advantageous pension plans. Other structural differences, due for instance to fundamentally different occupational compositions between sectors, should have been taken into account by the *Occupational dummies* included in the estimation.[14] Bargain and Melly (2008) explain the disparity in pay between the two sectors by means of the differences in workers' unobserved characteristics. Bargain and Melly (2008) show that when they control for unobserved heterogeneity by means of

---

[12]We pool all the observations for the waves 2005, 2006, 2008 and 2010

[13]See also Barón and Cobb-Clark (2010).

[14]At the end of Table 2 the dummies for the occupation we consider are listed. In our estimation, the reference category is the dummy *White-collars workers*.
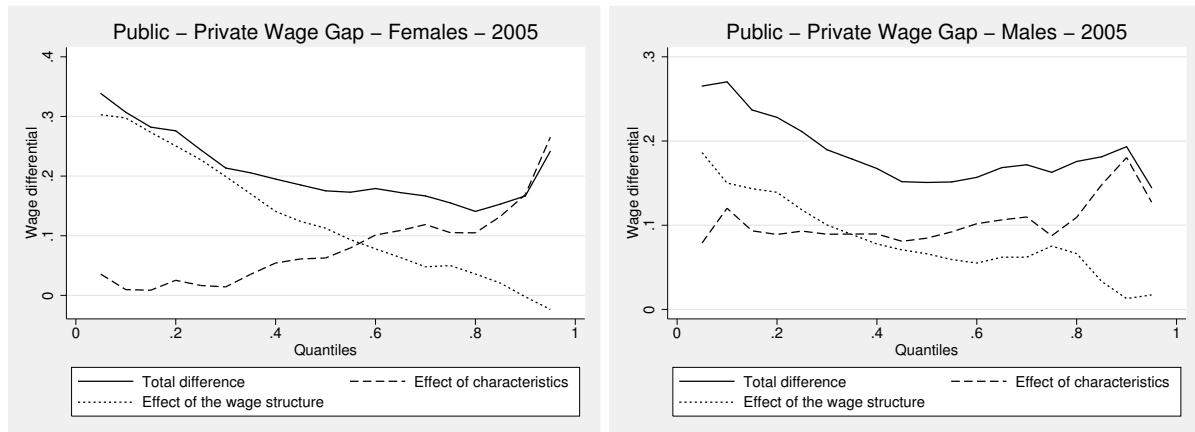
Figure 3: Public-Private sector wage gap decomposition, divided by gender.

standard panel techniques, the Public-Private wage premium vanishes.[15]

Results of the decomposition analysis of the public sector wage differential are reported in Figure 3 for the 2005 cross section data and in Figure 4 for the pooled data for females and males, respectively. We refer to Section 2.2 for the quantile decomposition in terms of measurable attributes (*Effect of characteristics*) and differences in the return to the same attributes (*Effect of the wage structure*) where the latter component is generally interpreted as the wage premium.

We observe that the estimated unexplained component of the public sector wage gap varies strongly with θ. Further, we observe that, for both genders, the portion of the public sector wage gap accounted by differences in returns to (observed) characteristics declines (almost) monotonically, to almost zero, from lower to upper quantiles. That is, a significant portion of the differential, in the lower part of the wage distribution, can be accounted for by differences in returns; while in the upper part there are almost no differences. For females, the sharp decline is more evident, and the estimated wage gap due to differences in returns becomes negative at top deciles, implying that there are significant differences in individual (observed) characteristics and occupations across sectors. For males, the contribution of the observed characteristics in explaining the sector wage gap is always more important with respect to what happens for the female sample.

4.2 Gender Wage Gap by Sector

The descriptive analysis of Section 3 has shown that the public sector compresses the pay dispersion especially for males and reduces the within-group pay inequality, i.e. the unobserved component of the public sector wage gap (see Figures 3 and 4). Blau and Kahn (2003a) show that wage structure has an important effect on the gender pay gap and wage compression reduces the gender

---

[15]See Section 5.0.1 for the panel analysis of the public-private wage premium.
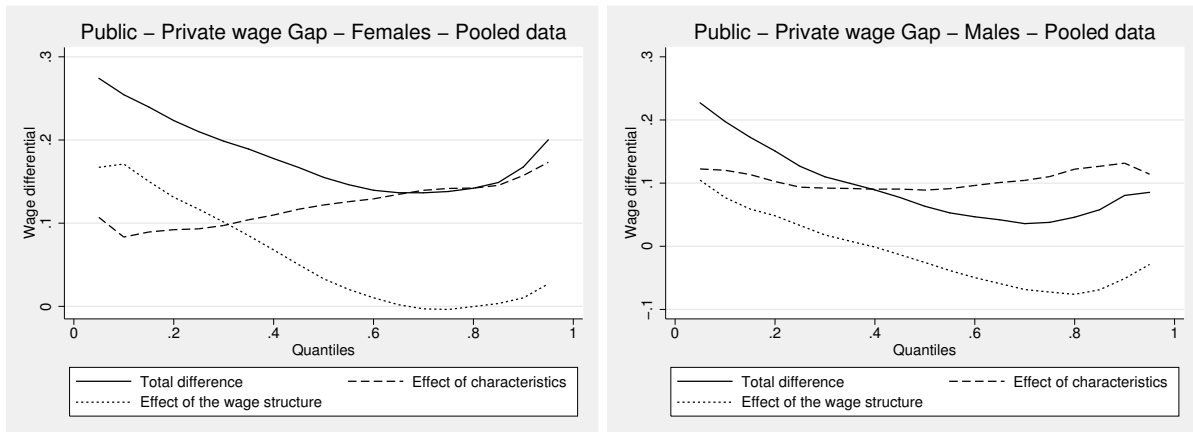
Figure 4: Public-Private sector wage gap decomposition, divided by gender. Pooled sample.

pay gap. Hence, next step is to verify if the previous findings imply a lower unexplained component of the gender wage gap in the public sector with respect to the private one.[16] Figure 5 shows the Machado-Mata decomposition by sector and by year.

The results presented in Figure 5 lead to several observation. First, in both sectors, the relative wages are increasing across the distribution and the GWG in the private sector is always bigger than those in the public kind. However the GWG for both sectors has been sharply decreased in 2010 with respect to 2005 and also the distance between the two sectors is narrowing. When we look at the decomposition of the GWG, we observe that a significant part of the gender wage gap remains unexplained in both sectors, after controlling for individual characteristics, education, job attributes and regional specific effects. Moreover, the weight of the unobserved component in explaining the gender wage gap is always bigger in the public sector with respect to the private sector. For both years considered and for the pooled sample, we observe that the coefficients component decreases along the wage distribution for the private sector while it increases for the public sector. Comparing the two sectors, we observe that among high wage workers, the wage gap faced by women is completely unexplained in the public sector while is mostly unexplained in the private sector. . In other words, the *discrimination component* looks much stronger for the public sector. More deeply, it appears that high-wage public-sector employees in Italy may face more employer discrimination (i.e., glass ceilings) than low wage workers (i.e., sticky floors). This result contrasts with the findings of Melly (2005a) for the Germany but confirms the findings of Barón and Cobb-Clark (2010). Further, Arulamplam, Booth, and Bryan (2006) and Kee (2006) find no evidence of sticky floors in public sector employment for Europe and Australia, respectively.

When we look at the private sector, instead, we observe that, unlike the public sector, the

---

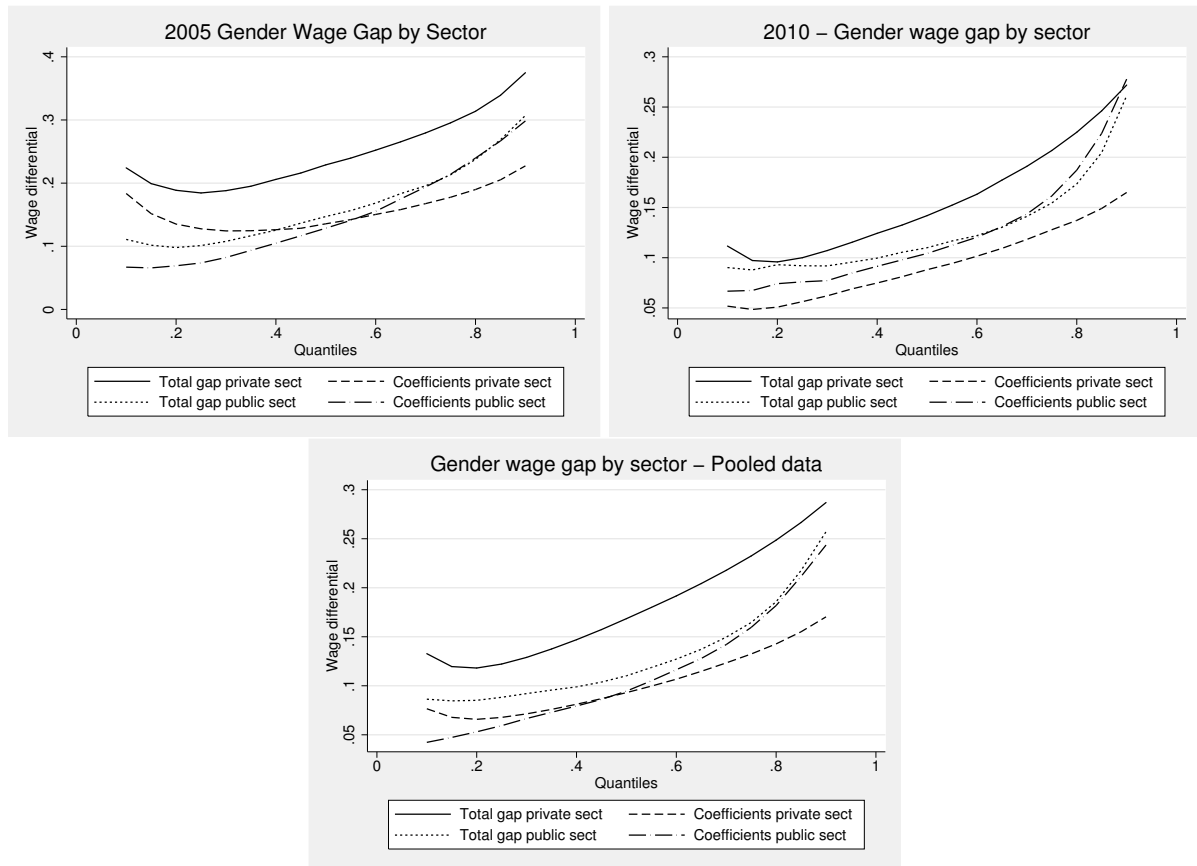[16]See also Blau and Kahn (1992) and Blau and Kahn (1996).

Figure 5: Gender wage gap by sector.

unexplained component of the conditional GWG decreases along the wage distribution. Therefore, it seems that employer discrimination is more prevalent among low-wage employees than among their high-wage counterparts. Thus, contrary to what found for the public-sector, the mechanism in action seems to be of sticky floors rather than glass ceilings.[17]

5 Longitudinal Analysis

5.0.1 Public Sector Wage Gap

The analysis in Figures 3 and 4 shows that using QR and Machado-Mata decomposition, both males and females appear to receive a premium in the public sector along all the wage distribution even if slightly decreasing. We have already stressed that this result has been confirmed by Lucifora and Moeurs (2006) and Campos and Centeno (2012), among others. To verify if these findings depends on a positive selection in the public sector we control for endogenous selection using fixed effect estimation provided by Canay (2011) and then we decompose the pay gap by Machado and Mata

---

[17]This result turned out to be robust to several specifications and different periods considered.

(2005)' method.

First, we estimate the following two fixed effects models for Private and Public sectors:

$$y_{it,Pub} = \alpha_i^{Pub} + x'_{it,Pub}\beta^{Pub} + \epsilon_{it,Pub} \tag{6}$$

$$y_{it,Priv} = \alpha_i^{Priv} + x'_{it,Priv}\beta^{Priv} + \epsilon_{it,Priv} \tag{7}$$

where Pub ($Priv$) stands for employees working in the public (private) sector.

Second, we estimate

$$Q_\theta(\hat{y}_{it,Pub}|x_{it}) = x'_{it,Pub}\beta_\theta^{Pub} \tag{8}$$

$$Q_\theta(\hat{y}_{it,Priv}|x_{it}) = x'_{it,Priv}\beta_\theta^{Priv} \tag{9}$$

where $\hat{y}_{it,k} = y_{it,k} - \hat{\alpha}_i^k$ for ($k = Priv, Pub$) is the log wage net out by the estimated individual heterogeneity. Last, we apply the Machado-Mata decomposition to compute the counterfactual distribution of $\hat{y}$.

In order to prevent any distortion due to the switching among sectors we compute the decomposition by restricting the sample to no sector switches observations with at least three waves. In this way we attenuate also the bias that could arise due to measurement errors due to misreported or miscoded value for the variable referring to working in the public sector.[18]

The comparison between Figure 5 and 6 shows that the relationship between the public and the private wage is reversed with respect to those obtained using cross section techniques. The estimated *net* wage differential ($\hat{y}_{Pub} - \hat{y}_{Priv}$) becomes negative along all the wage distribution; i.e. the public wage premium becomes negative and it increases along the wage distribution.

According to Bargain and Melly (2008) and Campos and Centeno (2012),[19] this evidence suggests a positive selection effect determining that better-endowed individuals choose to work in the public rather than in the private sector. This happens particularly for low wage workers. And this happens particularly for females than males. The Public premium sharply decreases for female sample with respect to male sample. The women hence are better off in the public sector and select themselves there. For the males the phenomenon is more attenuated but still important.

---

[18]We compute the decomposition also by restricting the sample to sectors switches in only one direction. The results are very closed to what presented in this Section.

[19]Both Bargain and Melly (2008) and Campos and Centeno (2012) find a weaker result; the public sector premium vanishes when the unobserved individual heterogeneity is taken into account.

The differential between the results obtained with cross section and panel technique suggests that the former may be hampered by an upward biased stemming from the omission of relevant factors contributing to the determination of wages (and sector of employment).
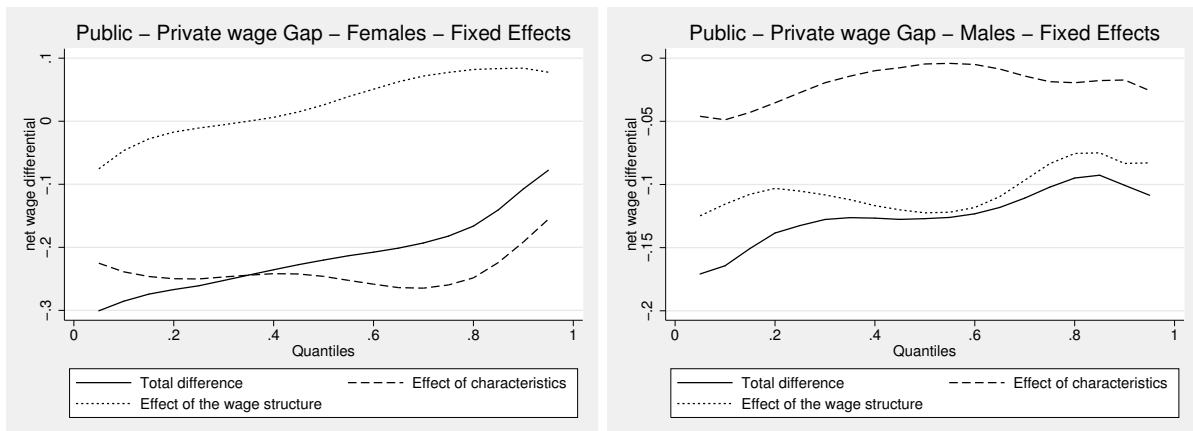
Figure 6: Public-Private sector wage gap decomposition, divided by gender. FE-quantile estimation.

5.0.2 Gender Wage Gap by Sector

Following the analysis presented in Section 5.0.1, we estimate the following two fixed effects models for female and male subsamples:

$$y_{it,f} = \alpha_i^f + x'_{it,f}\beta^f + \epsilon_{it,f} \tag{10}$$

$$y_{it,m} = \alpha_i^m + x'_{it,m}\beta^m + \epsilon_{it,m} \tag{11}$$

where f (m) stands for female (male) employee.

Second, we estimate

$$Q_\theta(\hat{y}_{it,f}|x_{it}) = x'_{it,f}\beta_\theta^f \tag{12}$$

$$Q_\theta(\hat{y}_{it,m}|x_{it}) = x'_{it,m}\beta_\theta^m \tag{13}$$

where $\hat{y}_{it,k} = y_{it,k} - \hat{\alpha}_i^k$ for $(k = f, m)$ is the log wage net out by the estimated individual heterogeneity. Last, we apply the Machado-Mata decomposition to compute the counterfactual distribution of $\hat{y}$. As in the previous section, we restrict the sample to no sector switches observations with at least three waves.

The longitudinal analysis of the GWG shown in Figure 7 confirms only partly the findings for the cross-section analysis presented in Section 4.2.

In terms of decomposition, the analysis shows that in both sectors the difference in pricing among genders is quite stable across the distribution; the unexplained component of the GWG strictly follows the increasing path of the wage gap in both sectors. In other words, the weight of the *discrimination component* is quite the same in both sectors. The difference relies only on the level of the GWG which is much higher (about two times) in the private sector with respect to the public one. The evidence found for the cross section analysis of a glass ceilings (sticky floors) in the public (private) sector vanishes when we control for unobserved heterogeneity, i.e. selection. However, within each sector there is a significant gender gap; this gap is larger in the private sector and increases along the wage distribution in both sectors. More importantly, the GWG is mostly unexplained in both sectors.
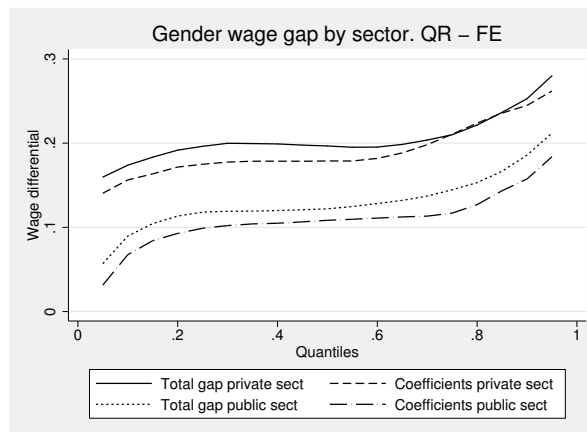
Figure 7: Gender wage gap by sector

6   Conclusion

In this paper, we investigate the public-private wage determination and the decomposition by gender of the wage in the Italian public sector. Using quantile regression methods we perform the analysis for both cross section and panel data. The main results are as follows. The cross section analysis shows the public wage gap is very similar for men and women. Both females and males are better off being in the public sector, particularly at the lowest quantiles.

However, when we control for unobserved heterogeneity, the public wage gap is very different among genders. We find a sharply decrease of the public premium which suggests that there is a positive selection in the public sector, especially for females. Second, when we look at the GWG

within sectors, the evidence of a glass ceilings in the public sector vanishes. However, in both sectors there is a significant unexplained GWG almost stable across the distribution.

References

ALTONJI, J., AND R. BLANK (1999): "Race and Gender in Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, chap. 48, pp. 3143–3259. North-Holland, Amsterdam.

ARULAMPLAM, W., A. BOOTH, AND M. BRYAN (2006): "Is There a Glass Ceiling in Europe? Exploring the Gender Pay Gap Across the Wage Distribution," *Industrial and Labor Relations Review*, 62, 163–186.

BARGAIN, O., AND B. MELLY (2008): "Public Sector Pay Gap in France: New Evidence Using Panel Data," Discussion paper, IZA DP No. 3427.

BARÓN, J. W., AND D. COBB-CLARK (2010): "Occupational Segregation and the Gender Wage Gap in Private- and Public-Sector Employment: A Distributional Analysis," *The Economic Record*, 86, 227–246.

BLAU, F., AND L. KAHN (2003a): "Understanding International Differences in the Gender Pay Gap," *Journal of Labor Economics*, 21(1), 106–144.

BLAU, F. D., AND L. KAHN (1992): "The Gender Earnings Gap: Learning from International Comparisons," *American Economic Review*, 82, 533–538.

——— (1996): "Wage Structure and Gender Earnings Differentials: An International Comparison," *Economica*, 63(250), S29–S62.

——— (2003b): "Understanding International Differences in the Gender Pay Gap," *Journal of Labor Economics*, 21(1), 106–144.

BUCHINSKY, M. (1995): "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo study," *Journal of Econometrics*, 68, 303–338.

——— (1998): "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *The Journal of Human Resources*, 33, 88–126.

CAMPOS, M. M., AND M. CENTENO (2012): "Public-private wage gaps in the period prior to the adoption of the euro: an application based on longitudinal data," Discussion paper, Banco de Portugal. Working Papers.

CANAY, I. (2011): "A Simple Approach to Quantile Regression for Panel Data," *The Econometrics Journal*, 14, 368–386.

CASTAGNETTI, C., F. CHELLI, AND L. ROSTI (2005): "Educational Performance as Signalling Device: Evidence from Italy," *Economics Bulletin*, 9, 1–12.

CASTAGNETTI, C., AND L. ROSTI (2009): "Effort Allocation in Tournaments: The Effect of Gender on Academic Performance in Italian Universities," *Economics of Education Review*, 28, 357–369.

CHERNOZHUKOV, V., I. FERNANDEZ-VAL, AND A. GALICHON (2010): "Quantile and probability curves without crossing," *Econometrica*, 78, 1093–1125.

CHERNOZHUKOV, V., I. FERNANDEZ-VAL, AND B. MELLY (2013): "Inference on Counterfactual Distributions," *Econometrica*, 81, 2205–2268.

DISNEY, R. (2007): "Public-private sector wage differentials around the world: methods and evidence," University of Nottingham.

DISNEY, R., AND A. GOSLING (1998): "Does it pay to work in the public sector?," *Fiscal Studies*, 19, 347–374.

FLORES, C., A. FLORES-LAGUNES, AND D. KAPETANAKIS (2014): "Lessons From Quantile Panel Estimation of the Environmental Kuznets Curve," *Econometric Reviews*, 8, 815–853.

GIORDANO, R., D. DEPALO, C. PEREIRA, B. EUGENE, E. PAPAPETROU, J. PEREZ, L. REISS, AND M. ROTER (2011): "The Public Sector Pay Gap in a Selection of Euro Area Countries," ECB Working Paper Series, No. 1406.

GRAHAM, B., J. HAHN, AND J. POWELL (2009): "The incidental parameter problem in a non-differentiable panel data model," *Economics Letters*, 105, 181–182.

GREGORY, R., AND J. BORLAND (1999): *Recent developments in public sector labor markets*vol. Handbook of Labor Economics, chap. 53, pp. 3573–3630. Ashenfelter, O. AND Card, D.

GUNDERSON, M. (1989): "Male-Female Wage Differentials and Policy Responses," *Journal of Economic Literature*, 27, 46–72.

HARDING, M., AND C. LAMARCHE (2014): "Estimating and Testing a Quantile regression model with Interactive Effects," *Journal of Econometrics*, 178, 101–113.

KATO, K., AND A. GALVAO (2011): "Smoothed quantile regression for panel data," Mimeo.

KEE, H. (2006): "Glass Ceiling or Sticky Floor? Exploring the Australian gender Pay Gap," *Economic Record*, 82, 408–427.

KOENKER, R. (2004): "Quantile Regression for Longitudinal Data," *Journal of Multivariate Analysis*, 91, 74–89.

——— (2005): *Quantile Regression.* Cambridge University Press, New York.

KOENKER, R., AND G. BASSETT (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

LUCIFORA, C., AND D. MOEURS (2006): "The Public Sector Pay Gap In France, Great Britain And Italy," *Review of Income and Wealth*, 52, 43–59.

MACHADO, J., AND J. MATA (2005): "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression," *Journal of Applied Econometrics*, 20(4), 445–465.

MELLY, B. (2005a): "Decomposition of Differences in Distribution Using Quantile Regression," *Labour Economics*, 12(4), 577–590.

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

NEYMAN, J., AND E. SCOTT (1948): "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16, 1–32.

Table 2: Descriptive Statistics

	Private				Public			
	2005		**2010**		**2005**		**2010**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Net hourly wage	1.97	0.41	1.92	0.49	2.17	0.37	2.18	0.38
Net hourly wage - women	1.84	0.36	1.82	0.45	2.08	0.31	2.10	0.32
Net hourly wage - men	2.10	0.42	1.99	0.51	2.25	0.39	2.25	0.41
Women	0.48	0.50	0.45	0.50	0.47	0.50	0.47	0.50
Age	36.92	11.88	37.22	13.19	45.90	10.86	45.38	13.10
Age - women	33.50	10.08	35.34	12.24	43.93	10.59	44.08	12.74
Age - men	40.07	12.52	38.73	13.74	47.66	10.80	46.53	13.31
Married	0.51	0.50	0.42	0.49	0.69	0.46	0.66	0.47
Number of Children	1.27	0.93	1.72	0.74	1.29	0.97	2.03	4.29
Children	0.77	0.42	0.37	0.48	0.75	0.43	0.66	0.47
Experience	16.13	12.14	15.25	12.70	23.42	11.29	23.77	12.48
Monthly hours worked	181.00	24.61	179.27	22.52	167.90	22.60	168.05	22.41
Permanent contracts	0.87	0.34	0.78	0.42	0.91	0.28	0.90	0.30
Tenure	10.46	10.34	10.36	10.83	18.47	11.16	19.08	12.18
Nord Region	0.55	0.50	0.55	0.50	0.38	0.49	0.40	0.49
Metropolitan area	0.30	0.46	0.31	0.46	0.34	0.47	0.33	0.47
Education								
University degree	0.21	0.41	0.26	0.44	0.35	0.48	0.39	0.49
University degree - women	0.21	0.41	0.29	0.45	0.39	0.49	0.45	0.50
University degree - men	0.21	0.41	0.24	0.43	0.30	0.46	0.34	0.48
University perf.§ - women	95.17	11.25	95.71	12.27	96.79	12.01	97.24	12.08
University perf.§ - men	90.86	12.68	92.26	13.48	94.10	13.18	94.95	12.63
High School	0.67	0.47	0.62	0.48	0.54	0.50	0.51	0.50
Primary Education	0.01	0.08	0.00	0.06	0.01	0.08	0.00	0.05
Secondary Education	0.11	0.32	0.11	0.31	0.11	0.31	0.09	0.29
Occupation								
Managers*	0.12	0.32	0.20	0.40	0.24	0.43	0.39	0.49
Intermediate professions**	0.13	0.34	0.32	0.47	0.18	0.39	0.28	0.45
White-collars workers***	0.75	0.43	0.48	0.50	0.58	0.49	0.33	0.47
Service Sector	0.70	0.46	0.79	0.41	0.96	0.20	0.94	0.25
Bigfirm	0.37	0.48	0.54	0.50	0.00	0.00	0.00	0.00
No of observations	3,591		3,968		2,890		2,413	

*Intellectual professions, scientific, and highly specialized occupations. **Intermediary positions in commercial, technical or administrative sectors, health services, technicians. ***Commercial, technical and administrative employees and clerks. §See footnote 8 for the definition of University Performance.

Table 3: 2005 Quantile Regression by Gender (Selected Covariates)

q	Gender	Experience	Experience squared	Years of Schooling	University Performance	Public Sector
.1	Females	0.026321*** (0.0049556)	−0.0004389*** (0.0001019)	0.0350356*** (0.007104)	0.0003165*** (0.0003786)	0.189488*** (0.0351926)
	Males	0.0271715*** (0.005234)	−0.0004642*** (0.0001038)	0.0214019*** (0.0055186)	0.0004144*** (0.0004613)	0.0686028*** (0.0252441)
.25	Females	0.0202326*** (0.0026141)	−0.0003639*** (0.0000552)	0.0250002*** (0.003535)	0.0001854*** (0.0002593)	0.146261*** (0.0218218)
	Males	0.020783*** (0.0025575)	−0.0003262*** (0.0000498)	0.0223116*** (0.0033917)	0.0001873*** (0.0003087)	0.0579894*** (0.0150811)
.5	Females	0.0150086*** (0.0020155)	−0.000255*** (0.0000435)	0.0230128*** (0.0030901)	0.000513*** (0.0001686)	0.1025329*** (0.0135995)
	Males	0.0209548*** (0.0024302)	−0.0003276*** (0.0000465)	0.0260577*** (0.0028282)	0.000614** (0.00031)	0.035471*** (0.0164718)
.75	Females	0.0161182*** (0.002265)	−0.0002542*** (0.000051)	0.022364*** (0.003013)	0.0009307*** (0.0002153)	0.0684795*** (0.0118829)
	Males	0.0267728*** (0.0027796)	−0.0004084*** (0.0000518)	0.0291174*** (0.0047843)	0.001129*** (0.0004252)	0.014133 (0.0203703)
.9	Females	0.0156806*** (0.0028149)	−0.0002495*** (0.0000702)	0.0269788*** (0.0045234)	0.001037*** (0.0003548)	0.0493987*** (0.0213018)
	Males	0.0221722*** (0.0043829)	−0.0002281*** (0.0000777)	0.0373113*** (0.004422)	0.0005047*** (0.0003813)	0.0063816 (0.0251095)

Bootstrap s.e. in parenthesis (800 replications). ***: significant at .99 level; **: significant at .95 level; *: significant at .90 level

Table 4: 2010 Quantile Regression by Gender (Selected Covariates)

q	Gender	Experience	Experience squared	Years of Schooling	University Performance	Public Sector
.1	Females	0.023244*** (0.012939)	-0.00043** (0.0002237)	0.0251557** (0.005723)	0.000573*** (0.000364)	0.277432*** (0.088483)
	Males	0.033224*** (0.010805)	-0.00051*** (0.000167)	0.014041*** (0.00443)	0.001198*** (0.000348)	0.120683 (0.039234)
0.25	Females	0.018098*** (0.005574)	-0.00029*** (0.000108)	0.018775*** (0.004882)	0.000768*** (0.000357)	0.150631*** (0.025984)
	Males	0.027458*** (0.006664)	-0.00042*** (0.00011)	0.012403*** (0.003705)	0.001301*** (0.000304)	0.09102*** (0.026591)
0.5	Females	0.01078*** (0.004411)	-0.00017*** (8.06E-05)	0.00933*** (0.003501)	0.001311*** (0.000259)	0.139693*** (0.018266)
	Males	0.029548*** (0.005119)	-0.00048*** (8.24E-05)	0.012061*** (0.003699)	0.001847*** (0.000383)	0.0539256* (0.032892)
0.75	Females	0.007959** (0.004592)	-0.00012* (0.0000876)	0.00995*** (0.005036)	0.00127*** (0.000275)	0.11068*** (0.025124)
	Males	0.030255*** (0.006124)	-0.00048*** (0.000108)	0.016693*** (0.005761)	0.002716*** (0.000476)	-0.02726 (0.038112)
0.9	Females	0.012375*** (0.005071)	-0.00019* (0.0001011)	0.013493* (0.007044)	0.001847*** (0.000579)	0.062548*** (0.043713)
	Males	0.0259107** (0.01167)	-0.00039** (0.0002012)	0.0151243** (0.006722)	0.004081*** (0.000645)	-0.09718* (0.055756)

Bootstrap s.e. in parenthesis (800 replications). ***: significant at .99 level; **: significant at .95 level; *: significant at .90 level