The Public Sector Wage Gap in Spain: Descriptive Evidence from Income Tax Data*

Laura Hospido
Banco de España and IZA
laura.hospido@bde.es

Enrique Moral-Benito
Banco de España
enrique.moral@bde.es

This draft: February 17, 2014

Preliminary and Incomplete. Please Do Not Quote.

Abstract

This paper studies the public sector wage gap by gender, skill level and type of contract in Spain using recent administrative data from tax records. We estimate wage distributions in the presence of covariates separately for men and women in the public sector and in the private sector. Then, we decompose the public sector wage gap along the wage distribution and isolate the part due to differences in the remunerations of observable characteristics. We find that the public sector raw wage gap is 0.35 for men and 0.40 for women on average, but 0.44 for both at the median. Once the contribution of differences in characteristics is net out, the conditional wage gap is 0.26 at the median, and only 0.06 at the top of the wage distribution. From 2008 to 2012, the public sector raw wage gap has decreased substantially. At the 90th percentile, in 2012 the conditional gap for men becomes negative and almost zero for women.

JEL Codes: C21, J31, J45.
Keywords: Public sector wage gap, Quantile regression, Wage distribution.

*We thank Samuel Bentolila, Sara de la Rica, Blaise Melly, and Ildefonso Méndez for useful comments. We also thank seminar participants at the Bank of Spain, the SAEe Meetings in Vigo, the ESPE Conference in Aarhus, the Jornadas de Economía Laboral in Madrid, and the EEA-ESEM meetings in Gothenburg. All remaining errors are our own. The opinions and analyses are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem. First draft: June 10, 2012.

1 Introduction

In 2012, more than 15% of the labor force received their wage from the public sector and compensation of employees represented around 30% of Spanish public consumption expenditures. Given that, in order to ensure fiscal sustainability under pressure from financial markets, the Spanish Government has undertaken huge fiscal consolidation efforts, the size of the public sector wage bill has been under scrutiny. Indeed, several measures aiming at reducing this public sector wage bill have been already implemented.[1] Under these circumstances, a deep understanding of the public-private wage gap and its distribution seems of paramount importance.[2]

Public and private sectors workers can be paid differently because of several reasons: (i) the monopolistic power of governments in the provision of public services results in non-competitive wage settlements (Reder, 1975); (ii) the public sector might have different objectives from those of the private sector, for instance, vote maximization rather than profit maximization; (iii) the wage setting environment substantially differs between both sectors, for example, union density is often higher in the public sector; (iv) productivity-enhancing characteristics of employees such as education or experience might be different between both sectors. In this paper we argue that the room for cutting public sector wages should be based on the public wage gap due to reasons (i)-(iii) so that we focus on the analysis of the public wage gap not explained by observable productivity-related characteristics of employees in the two sectors.

There exists and extensive literature analyzing the public - private wage gap based on average figures for different countries including Spain. However, the average public sector wage premium only provides an incomplete picture of the whole distribution. Therefore, there is also a more recent literature analyzing the whole distribution of the public-private wage gap based on quantile methods (see section 2 for an overview). We embed our paper into this strand of the literature. In particular, we analyze the distribution of the public-private wage gap in Spain using recently developed methods for estimating counterfactual distributions (i.e. Chernozhukov et al., 2013).

For that purpose we use a dataset based on tax records which allows us to overcome a potential drawback of previous empirical studies about the public-private wage gap based on survey data. To the best of our knowledge, all those studies are based on databases in which responses are provided by individual workers (e.g. the German Socio-Economic Panel, the European Community Household Panel, or the Wage Structure Survey in the Spanish case). Concerns about response errors in survey data and their implications for economic analysis date back to the fifties (e.g. Cohen and Lipstein, 1954; Miller and Paley, 1958). For instance, using two unique matched worker-employer data files, Mellow and Sider (1983) find that almost one-half of workers surveyed indicate a different detailed occupation than is reported by their

[1] In 2010 and 2012 the Spanish Government approved cuts in the nominal wages of public employees, and in 2011 and 2013 the decision was to freeze those wages.

[2] Furthermore, as a side-effect, cuts in public sector wages might induce reductions in private wages with the subsequent gains in terms of competitiveness (see Lamo et al., 2012).

employer. Zweimuller (1992) concludes that sample selectivity due to interviewees' refusal to answer to the survey-questionnaire is a significant problem, even of larger importance than the selectivity bias due to non-participation in the labor market.[3] Regarding the quality of survey measures of income, several studies (e.g. Herriot and Spiers, 1980; Gottschalk et al., 2008; Gottschalk and Huynh, 2010) use earnings reports from survey data (e.g. PSID or CPS) matched to tax records and find substantial evidence that measurement error in self-reported earnings is important and not classical. Moreover, an additional concern is that reporting biases may follow different patterns between public and private sector workers; while income sources for public sector employees are clearly determined and unambiguously-established, uncertainty surrounding income in the private sector is more important due to, for instance, bonuses or extra hours.

In this paper, we use recently released social security data for Spain. Social security records have several advantages compared to the survey-based datasets that have been previously used. These include large sample sizes, complete coverage of the part of the population that is affiliated to the social security administration, and accurate earnings measurements.. We focus on the period 2005-2012, for which the social security dataset has a proper longitudinal design (before 2005 the information is retrospective). In addition, in that period, annual income information from tax records are available for the same individuals as in the social security dataset. Contrary to the social security measure of labor earnings that is top- (and bottom-) coded, tax records are not subject to censoring, making them suitable to perform our study. On the other hand, the social security dataset do not record hours of work. To overcome this drawback, we match our dataset with information on hours from the Spanish Labor Force Survey.

We estimate wage distributions in the presence of covariates separately for men and women in the public sector and in the private sector. Then, we decompose the public sector wage gap along the wage distribution and isolate the part due to differences in the remunerations of observable characteristics. We find that the public sector gap in log hourly wages is 0.35 for men and 0.40 for women on average, but 0.44 for both at the median. Our preliminary results show that, once the contribution of differences in characteristics is net out, the conditional wage gap in favour of public employees is 0.26 at the median, and only 0.06 at the top of the wage distribution. From 2008 to 2012 we find that the public sector raw wage gap has decreased substantially both for men and women. At the 90th percentile, in 2012 the conditional gap for men becomes negative and almost zero for women.

The rest of the paper is organized as follows. We start by summarizing the relevant literature in Section 2. Section 3 explains our methodological approach. We describe the data in Section 4, and in section 5 we discuss our results. Lastly, Section 6 concludes.

[3] For more details on this issue see also Griliches et al. (1978), Atkinson and Micklewright (1983), or Groves (2006).

2 Related Literature

Several studies have already addressed the issue of the public - private wage gap in different countries. Some examples based on average gaps are Smith (1976) or Borjas (2002) for the United States, Dustmann and Van Soest (1997) for Germany, Panizza and Qiang (2005) for Latin American countries, Anghel et al. (2011) for OECD countries, De Castro et al. (2013) for the European Union countries, and Lassibille (1998), or García-Pérez and Jimeno (2007) for Spain. This strand of the literature has reached consensus in the following findings: (i) the public premium is positive for low-skilled male workers but negative for the high-skilled ones when observable characteristics are accounted for; (ii) the public premium remains positive for females even after controlling for individual characteristics; and (iii) the distribution of wages is more compressed in the public sector.[4]

Since the public sector apparently compresses the distribution of wages, the mean public sector wage premium only provides an incomplete picture of the whole distribution. In response to this concern, several authors, including ourselves, apply quantile regression methods to analyze the whole distribution of the public-private wage gap.

Mueller (1998) used quantile regressions to estimate the size of the public sector wage premium for Canada. He found that public sector pay differentials tend to be highest for federal government employees, females and individuals at the lower tail of the wage distribution. Similar results were reported by Cai and Liw (2008) for Australia. Utilizing quantile regression analysis, they show that the public sector pay premium declines at the higher spectrum of the wage distribution and becomes negative for male workers at the top half of the conditional wage distribution. Melly (2005) measures and decomposes the differences in earnings distributions between public and private sector employees in Germany for the years 1984-2001. Results suggest that conditional wages are higher in the public sector for women but lower for men; the "premium" is highest at the lower end of the distribution and then monotonically decreases by moving up the wage distribution. His findings are stable over the '80s and the '90s. Bargain and Melly (2008) estimate the public wage gap in France for the period 1990-2002 at the mean and at different quantiles of the wage distribution for both sexes. Controlling for unobserved heterogeneity by using fixed effects estimation on panel data they report that public sector premia or penalties are indeed much lower than commonly found. In particular, public wage premia for women and penalties for men are the result of the selection of the employees. Finally, only small pay differences between sectors remain over time, reflecting fluctuations due to specific public policies and the procyclical movement of private sector wages. Papapetrou (2006) using microdata from the European Community Household Panel Survey (ECHP) for Greece reports that average earnings are higher in the public sector than in the private sector and employees in the public sector at the lower end of the wage distribution earn a higher wage gap compared with their counterparts in the private sector, but this gap decreases at higher

[4] See Gregory and Borland (1999) for a survey of this literature.

quantiles. Furthermore, quantile regression estimation reveals that earnings differentials at the lower end of the wage distribution cannot be attributed to individual characteristics whereas at the highest quantiles pay differentials reflect differences in the employee's endowment. Boyle et al. (2004) report wage premia for public sector workers, greater for low-paid workers and smaller for public sector workers at the top of the earnings distribution using microdata from the European Community Household Panel Survey. Another study by Foley and O'Callaghan (2009), using micro data from the 2007 National Employment Survey, also find a sizable public sector wage premium, highest at the lower ends of the earnings distribution. Campos and Pereira (2009) for Portugal show that public sector employees earn higher wages than their private sector counterparts and this premium has risen over the 1996-2005 period from almost 10 per cent in 1996 to around 15 per cent in 2005. The premium is higher for female workers compared to male workers and decreases as one moves from the lower to the upper quantiles of the earnings distribution. Ramos et al. (2013) use data from the Spanish Wage Structure Survey in 2010 and also report that public sector employees earn higher wages than their private sector counterparts. However, once characteristics of both the worker and the firm are taken into account the premium is relatively small, specially for men workers under fixed-term contracts.[5] Giordano et al. (2011) use data from the European Union Statistics on Income and Living Conditions (EU-SILC) referring to the period 2004-2007. They evaluate the differential across countries, distinguishing by gender, educational level, sub-sectors and firm size. Other studies along these lines include Poterba and Rueben (1995), Nielsen and Rosholm (2001), and Jürges (2002).

3 Counterfactual Distributions

Blinder (1973) and Oaxaca (1973) proposed to decompose the difference in average earnings between public and private workers into a explained component given by differences in characteristics and an unexplained component given by differences in coefficients. This popular approach only provides information about average differences. However, statistical measures of the public-private wage gap based on average effects might mask important differences along the distribution of wages.

Since Koenker and Bassett (1978) the quantile regression approach has became relatively popular to study the effects of a covariate (X) on the whole conditional distribution of the dependent variable (Y). Quantile regression provides a more complete picture of the conditional distribution of Y given $X = x$ when both lower and upper quantiles are of interest. More concretely, we can specify the θth quantile of the conditional distribution of y_i given X_i as a linear function of the covariates,

$$Q_\theta(y_i|X_i) = X_i\beta_\theta, \qquad \theta \in (0,1). \tag{1}$$

[5]Similar results are obtained by Rahona et al. (2013), also using data from the Wage Structure Survey but applying different sample selection filters.

The quantile regression estimator of β_θ estimates the effect of the covariates on the θth quantile of the dependent variable and solves the following problem (Koenker and Bassett, 1978):[6]

$$\hat{\beta}_\theta = argmin_\beta \left[\sum_{i \in \{i: y_i \geq X_i\beta\}} \theta |y_i - X_i\beta| + \sum_{i \in \{i: y_i < X_i\beta\}} (1-\theta)|y_i - X_i\beta| \right]. \tag{2}$$

Given the quantile regression approach just discussed, we can now present the details on the generalization of the Blinder-Oaxaca decomposition to the whole distribution of wages based on Chernozukov et al. (2013). In particular, we can proceed in seven steps:

Step 1. Quantile regressions: We separately run two different sets of quantile regressions, one for the public sector (group 1) and one for the private sector (group 0) to obtain the two sequences of quantile coefficients $\hat{\beta}^1_{\theta_j}$ and $\hat{\beta}^0_{\theta_j}$ for $j = 1, ..., J$ with $\theta_j \in (0,1) \forall j$. Despite asymptotically one could estimate an infinite number of quantile regressions for each group (i.e. $J \to \infty$), following the suggestion in Portnoy (1991) we only estimate 150 different regressions to approximate the whole quantile function (i.e. $J = 150$).[7]

Step 2. Conditional quantile functions: Given the quantile regression coefficients obtained in the first step, it is straightforward to estimate the θ_j's conditional quantile of Y_g given X_i by computing $X_i'\hat{\beta}^g_{\theta_j}$ where $g = (0,1)$ represents the group (public or private workers). Hence we can construct the two conditional quantile functions as follows:

$$\begin{aligned} \hat{q}^1_{\theta_j} &= X_i'\hat{\beta}^1_{\theta_j} \quad \forall j = 1, ..., J \\ \hat{q}^0_{\theta_j} &= X_i'\hat{\beta}^0_{\theta_j} \quad \forall j = 1, ..., J. \end{aligned} \tag{3}$$

Step 3. Conditional distribution functions: We can also estimate the conditional distribution function by inverting the conditional quantile function obtained in step 2 so that:[8]

$$\begin{aligned} \hat{F}_{Y_1}(q|X_i) &= \int_0^1 (1(X_i'\hat{\beta}^1_{\theta_j} \leq q)d\theta) = \sum_{j=1}^{J} (\theta_j - \theta_{j-1}) 1(X_i'\hat{\beta}^1_{\theta_j} \leq q) \\ \hat{F}_{Y_0}(q|X_i) &= \int_0^1 (1(X_i'\hat{\beta}^0_{\theta_j} \leq q)d\theta) = \sum_{j=1}^{J} (\theta_j - \theta_{j-1}) 1(X_i'\hat{\beta}^0_{\theta_j} \leq q). \end{aligned} \tag{4}$$

where $F_Y(q)$ refers to the cumulative distribution function (CDF) of the random variable Y evaluated at q, $F_Y^{-1}(\theta)$ represents the inverse of the CDF, also known as quantile function evaluated at $0 < \theta < 1$, and $F_Y(q|X_i)$ refers to the conditional CDF of Y evaluated at q and given the realization $X = X_i$.

[6] Buchinsky (1998) provides an overview of the quantile regression estimator together with details on its asymptotic covariance matrix.

[7] In finite samples, Portnoy (1991) shows that given the set of points in which the vector of coefficients changes $(\theta_0 = 0, \theta_1, ..., \theta_J = 1)$, the coefficients estimate $\hat{\beta}_{\theta_j}$ prevails in the interval from θ_{j-1} to θ_j.

[8] Note that since the estimated quantile function might not be monotonic, we need to resort to the following property of the CDF: $F_{Y_g}(q|X_i) = \int_0^1 (1(F^{-1}_{Y_g}(\theta|X_i) \leq q)d\theta) = \int_0^1 (1(X_i'\beta^g_{\theta_j} \leq q)d\theta)$.

Step 4. Unconditional distribution functions: Therefore, we can now estimate the unconditional distribution function for public ($g = 1$) and private ($g = 0$) workers as follows:

$$\begin{aligned} \hat{F}_{Y_g}(q|g=1) &= \int \hat{F}_{Y_g}(q|x)dF_X(x|g=1) = \frac{1}{n_1}\sum_{i:g=1}\hat{F}_{Y_g}(q|X_i). \qquad (5)\\ \hat{F}_{Y_g}(q|g=0) &= \int \hat{F}_{Y_g}(q|x)dF_X(x|g=0) = \frac{1}{n_0}\sum_{i:g=0}\hat{F}_{Y_g}(q|X_i). \end{aligned}$$

where n_1 and n_0 are the number of public and private workers in the sample.

Step 5. Unconditional quantile functions: Given our interest in simulating counterfactual quantiles to decompose differences in the distribution of wages, we estimate the unconditional quantile function. For this purpose we take as an estimator of the θ^{th} quantile of the unconditional distribution from step 4 the minimum of the set as follows:

$$\begin{aligned} \hat{q}^1_\theta &= inf\left\{q : \frac{1}{n_1}\sum_{i:g=1}\hat{F}_{Y_1}(q|X_i) \geq \theta\right\} \qquad (6)\\ \hat{q}^0_\theta &= inf\left\{q : \frac{1}{n_0}\sum_{i:g=0}\hat{F}_{Y_0}(q|X_i) \geq \theta\right\}. \end{aligned}$$

Step 6. Counterfactual quantile functions: Armed with the previous function estimates, we are now able to estimate the counterfactual quantile function. That is, we estimate the θ^{th} quantile of the distribution that we would observe if public workers ($g = 1$) would be paid as private workers ($g = 0$):

$$\hat{q}^c_\theta = inf\left\{q : \frac{1}{n_1}\sum_{i:g=1}\hat{F}_{Y_0}(q|X_i) \geq \theta\right\}. \qquad (7)$$

where n_1 is the number of public workers in the sample. Note that for the construction of the conditional distribution $\hat{F}_{Y_0}(q|X_i)$ we used in step 3 the coefficients estimated for the private workers, i.e., $\hat{\beta}^0_\theta$; and we are computing the counterfactual quantile using the Xs among public workers, i.e., sum over individuals with $g = 1$. This counterfactual distribution is an interesting object per se that will deserve special attention in our empirical exercises.

Step 7. Decomposition: Analogously to the Blinder-Oaxaca approach for the mean, we can now compute a decomposition of the difference between the θ^{th} quantile of the unconditional distribution of public and private workers:

$$\hat{q}^1_\theta - \hat{q}^0_\theta = \underbrace{\left[\hat{q}^1_\theta - \hat{q}^c_\theta\right]}_{\text{Coefficients Effect}} + \underbrace{\left[\hat{q}^c_\theta - \hat{q}^0_\theta\right]}_{\text{Characteristics Effect}} \qquad (8)$$

4 Data

4.1 Data sources

Our main data source is the Continuous Sample of Working Histories (*Muestra Continua de Vidas Laborales*, MCVL, in Spanish). The MCVL is a micro-level dataset built upon Spanish administrative records with detailed information on labor earnings and days worked, in addition to other worker and firm characteristics. It is a representative sample of the population registered with the social security administration at any time in the reference year. The MCVL also has a longitudinal design. From 2005 to 2012, those individuals who are present in a wave and subsequently remain registered with the social security administration stay as sample members. In addition, the sample is refreshed with new sample members so it remains representative of the population in each wave. Finally, the MCVL tries to reconstruct the market labor histories of the individuals in the sample back to 1967. Besides the MCVL, we will use annual income information from tax files that have been matched to the social security sample. Contrary to the social security measure of labor earnings that is top- (and bottom-) coded, tax records are not subject to censoring. In addition, as mentioned before, the MCVL does not record hours of work. Hence, in order to compute a hourly wage measure, we combine the daily earnings from administrative records with information on hours of work from the Spanish Labor Force Survey (*Encuesta de Población Activa*, EPA, in Spanish).

4.2 Sample Selection

The population of reference of the MCVL consists of all individuals registered with the social security administration, including pension earners, recipients of unemployment benefits, employees and self-employed workers, but excluding those registered only as medical care recipients, or those with a different social assistance system. The raw data represents a 4 per cent non-stratified random sample of this reference population. It consists of nearly 1.1 million individuals each year.

We use data from working individuals in the 2005-2012 MCVL original samples with Tax Information.[9] We select prime-age employees enrolled in the General Regime of the Social Security Administration at any time in the sample period.[10] To ensure that we only consider income from wage sources, we exclude self-employees from our sample. We also exclude individuals younger than 25 and older than 54 years to avoid to get mixed with formal education enrollments issues and early retirement decisions, respectively.

In the empirical analysis, we use individual log hourly wages as our main dependent variable. To recover the information on hours of work from the EPA, we define cells given by year, age, gender, level of qualification, sector of activity, tenure in the firm, type of contract (fixed-term vs. open-ended), type of work schedule (full-time vs. part-time), and region. For each cell in the EPA, we compute the average number of usual weekly hours of work, and then we impute that number to those individuals belonging to an equally defined cell in the MCVL dataset. Then

[9] Basque Country and Navarra are excluded, because they enjoy a different system known as the Economic accord.

[10] In Spain, more than 95 per cent of employees are enrolled in the general scheme of the Social Security Administration. Separate schemes exist for domestic workers, some workers in fishing, mining and agricultural activities, and some government employees, such as the armed forces or the judicial power.

we divide those hours by 5 to obtain daily hours of work. With this procedure, we have been able to merge 88 per cent of the observations from our MCVL raw sample. Hourly wages are computed as the individual annual labor income from the tax record, divided by the individual annual days of work from the social security records and the average number of daily hours obtained from the EPA.

The merged final sample is a panel of 688,597 individuals and 3,232,598 annual observations for the period 2005-2012. We present descriptive statistics on sample composition in Tables A.1.

4.3 Definition of Public Employees

In our dataset public employees refer to those workers from either the central administration, the regional governments or the local corporations, as well as those working in public firms.[11] However, some public employees who belong to social assistance systems different to the General Regime of the Social Security Administration, such as the armed forces or the judicial power, are not generally included.

According to our dataset, in Spain 15 per cent of employees work in the public sector (see Table 1). In the case of women the incidence is higher (20 per cent), almost doubles the corresponding share for men (11 per cent). By skill groups,[12] we obtain that the share of public employees is higher among high-skilled relatively to less skilled workers. One particular feature of the Spanish case is the high proportion of public employees among workers with fixed-term contracts (more than 31 per cent for women).

The evolution of those shares over time, as shown in Figure 1, is clearly affected by the current crisis. We can see than before 2009 the public sector share was 14.6 per cent, then increases up to 16.2 from 2009 to 2011, before decreasing to 14.0 per cent in 2012. For men, the increase in the share was from 10.1 to 11.9, and then it decreases to 10.5 per cent, while in the case of women, the corresponding numbers are 20.3, 21.2, and 18.0, respectively.

4.4 Raw Wage Gaps

According to Table A.2 of the Appendix, *annual earnings* are on average 26 thousands euros in the public sector and 19 thousands euros in the private sector over the period. However, part of this raw gap of 33 percentage points is due to the different labor force composition of the two sectors. As reported in the Table A.2, public employees are on average older, more skilled, and

[11] The dataset includes two variables that allow us to distinguish workers in the public sector to those in the private sector: one from the point of view of the worker (so-called employee type), and another from the firm's perspective (type of legal entity). The results presented in the paper correspond to the first definition. We also use the second definition as a robustness check, and the results do not change.

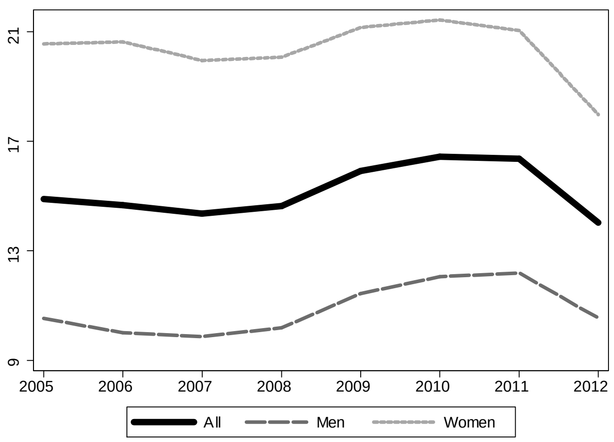
[12] In Spain, each worker affiliated to the social security is assigned to one of the ten contribution groups (for instance, Group 1 corresponds to workers with university degree). In particular, we label a worker as high-skilled (groups 1-3), medium-skilled (groups 4-7), or low-skilled (groups 8-10).

Table 1: Share of public employees (%)

	All	Men	Women
Overall	**15.16**	**10.82**	**20.36**
High-skilled	34.82	23.06	46.93
Medium-skilled	16.98	16.57	17.24
Low-skilled	7.16	5.29	10.95
Permanent	12.12	10.26	14.43
Temporary	21.17	11.99	31.51

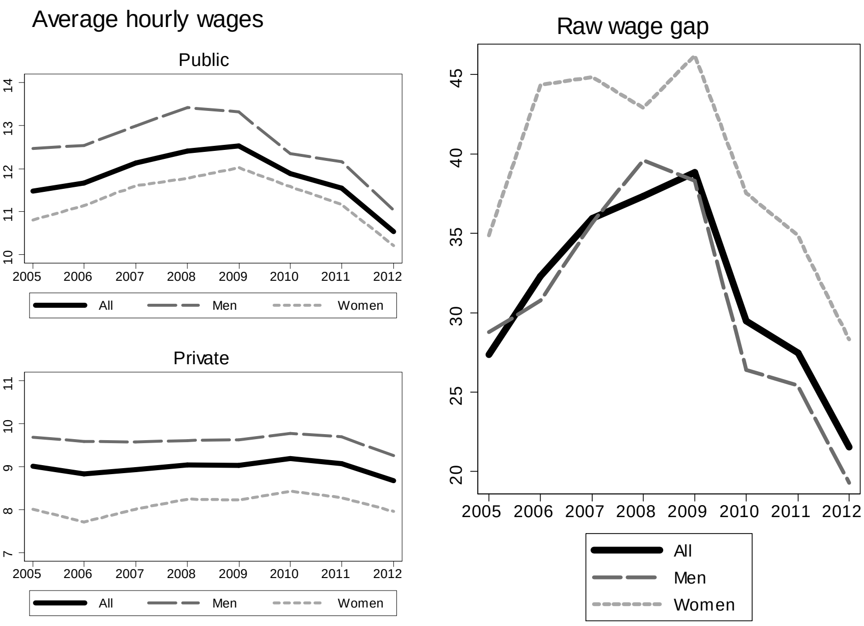
Notes: Whole sample (2005-2012). High-skilled (1-3), medium-skilled (4-7), low-skilled (8-10).

Figure 1: Share of public employees over time (%)



Notes: Whole sample (2005-2012).

Figure 2: Hourly wages and average gap (%) over time



work more on a full-time basis. On the other hand, they have temporary contracts in a higher proportion.

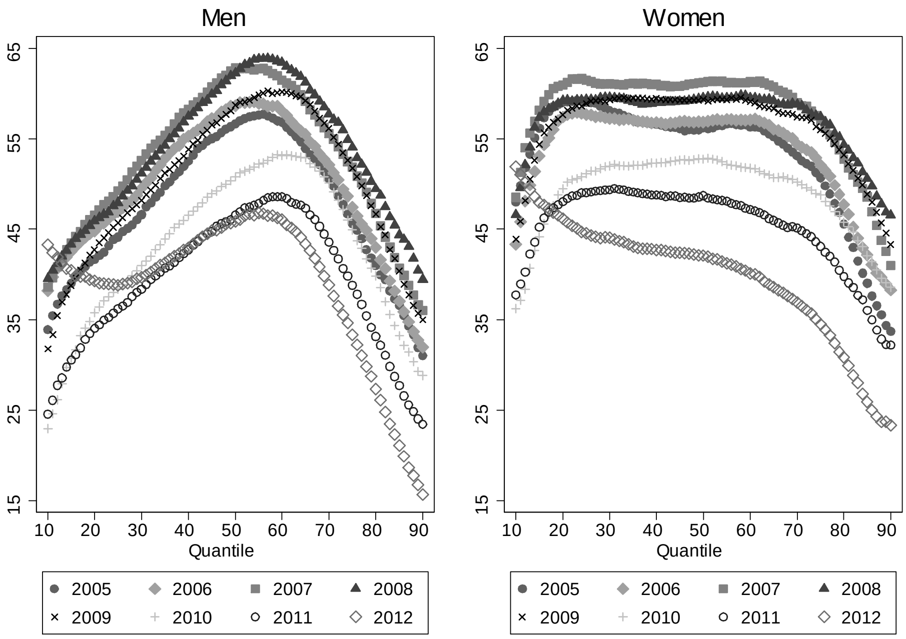
In addition, the gap in annual earnings includes differences in the total number of days worked in a year, and in the number of hours worked per day. On the one hand, the number of annual days of work is on average higher in the public sector. Given that, the raw public sector wage gap is lower in a *daily* basis than in annual terms (27% versus 33%). On the other, employees in the public sector work on average less hours than those in the private sector (7.3 and 7.6 hours per day, respectively), being then the *public sector hourly wage gap* on average equal to 31.7%. By gender, we obtain than the raw wage premium in the public sector is higher for females than for males (39.7% and 30.8%, respectively).

Figure 2 shows the evolution of the public sector wage gap over time. We can see that the average wage gap increased from 2005 to 2009 and then decreased (with the overall gap being the highest in 2009, 38.8 per cent, and the lowest in 2012, 21.5 per cent). This decrease in the public sector wage gap goes in line with the recent cuts in public wages.

Behind those differences by gender in the average public sector wage gap there are very different profiles along the wage distribution. As shown in Figure 3, for men we observe an inverse V-shaped pattern, whereas for women the profile is more compressed and similar to an inverse U. Over time, those profiles have changed in terms of the level and only recently also in their shapes.

Next, we consider the public sector wage gap in the presence of covariates - both in the mean and along the wage distribution - and we decompose this gap to isolate the part due to

Figure 3: Raw gaps (%) along the wage distribution over time



differences in the remunerations to those characteristics.

5 Results

5.1 The Public Sector Average Wage Gap

In this section, we present estimates of a standard Oaxaca-Blinder decomposition for the average wage gap. Let y_i be the individual i's log hourly wage in year t in real terms (in a given year t, or in the pooled data for the whole period). We denote Public $\equiv$ 1 and Private $\equiv$ 0, so that we consider the following regressions for each sector:

$$\begin{aligned} y_{i1} &= x_{i1}\beta_1 + u_{i1} \\ y_{i0} &= x_{i0}\beta_0 + u_{i0} \\ \bar{y}_1 - \bar{y}_0 &= \bar{x}_1\beta_1 - \bar{x}_0\beta_0 + \bar{x}_1\beta_0 - \bar{x}_1\beta_0 \\ \bar{y}_1 - \bar{y}_0 &= \underbrace{(\bar{x}_1 - \bar{x}_0)\,\beta_0}_{\text{Characteristics effect (Explained)}} + \underbrace{\bar{x}_1\,(\beta_1 - \beta_0)}_{\text{Coefficients effect (Unexplained)}} \end{aligned}$$

where x_i is the set of covariates, $\bar{z} = N^{-1}\sum_i zi$ a sample mean, and $\bar{x}_1\beta_0$ is a *counterfactual wage* that measures the average wage we would observe if public workers would be paid as private workers. This manipulation allow us to decompose the average difference between wages in the public and private sectors in two components: the *characteristics* effect (an explained component given by differences in characteristics), and the *coefficients* effect (an unexplained component given by differences in coefficients).

With respect to the vector of covariates (x_i), we consider three different specifications: first, we consider those variables often included in Mincerian models, namely, age, age squared, skill-groups, time and regional dummies; second, we add indicators for tenure in the firm (less than 1 year, between 1 and 2 years, between 2 and 4, between 4 and 7, between 7 and 15, and more than 15 years), the type of contract (fixed-term vs. open-ended), and the type of work schedule (full-time vs. part-time); and finally, we also include firm size as an additional categorical variable (less than 10 employees, 10-50, 50-200, more than 200).[13]

In Table 2 we present estimates of the *coefficients* effect, that is, the difference in average log hourly wages between public and private workers once the effects of differences in characteristics is net out. We show estimates for the whole period, in column 1 pooling men and women, and in columns 2 and 3 for each of them separately. We find that for an overall raw difference of 0.36 log points, between 0.121 and 0.160 log points (depending of the specification) are explained by differences in observed characteristics of public and private workers. However, there is still almost one half of the difference that remains unexplained. For men, the raw log difference is 0.35 and at least 53 per cent of the difference is due to the *coefficients* effect. For women, the raw difference is higher (0.40) but again the fraction unexplained is around one half of it.

Table 2: Average logwage difference

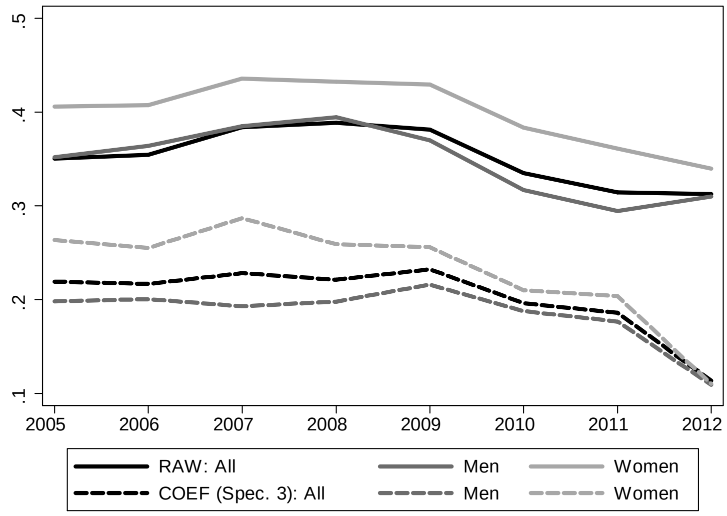
	All	Men	Women
Raw difference	0.355	0.349	0.401
	(0.001)	(0.001)	(0.001)
Coefficients Effect (Specification 1)	0.195	0.195	0.207
	(0.001)	(0.002)	(0.001)
Coefficients Effect (Specification 2)	0.234	0.213	0.267
	(0.001)	(0.002)	(0.002)
Coefficients Effect (Specification 3)	0.202	0.185	0.231
	(0.001)	(0.002)	(0.002)

Notes: Whole sample (2005-2012). SE in parentheses.

In Figure 4 we show the raw differences and the estimates of the*coefficients* effect from Specification 3 - overall and by gender - for each year. We find that the raw log difference increases from 0.35 in 2005 to 0.39 in 2008, and then diminishes to 0.31 in 2012. In addition, we estimate that in 2005, 63 per cent of the raw log difference was due to the *coefficients* effect, 57 per cent in 2008, and only a 36 per cent of the gap in 2012 remained unexplained. For males, the evolution of the raw gap is from 0.35 in 2005, to 0.39 in 2008 and 0.31 in 2012, whereas for

[13]These models correspond to *specification 1, 2* and *3*, respectively, in subsequent tables and figures. For regressions that pool men and women together we add a female indicator. Coefficient estimates of these regression are available upon request.

Figure 4: Average logwage difference over time



women the corresponding figures are 0.41, 0.43 and 0.34, respectively. With respect to the size of the *coefficients* effect, for men it moves from 56 per cent in 2005 to 50 per cent in 2008, and 35 per cent in 2012. For women, this effect is much lower and it moves from 26 per cent in 2005 and in 2008 to just 11 per cent in 2012.

5.2 The Public Sector Wage Gap along the Wage Distribution

Similarly to the comparison before at the mean, now we compare the estimated percentiles of the *total* public sector logwage gap, $\hat{q}^1_\theta - \hat{q}^0_\theta$, with the corresponding ones once the contribution of different characteristics has been net out (that is, the *coefficients* effect $\hat{q}^1_\theta - \hat{q}^c_\theta$).[14]

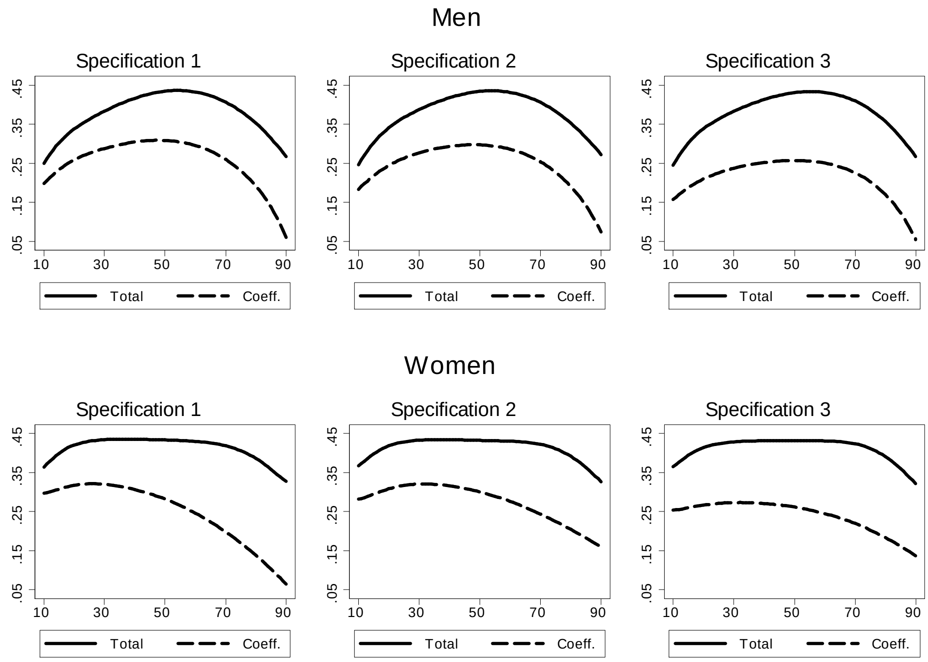
Figure 5 shows the percentiles of the two public sector wage gaps (total and coefficients) by gender for the three specifications considered. The solid lines stand for the estimated total wage gaps, while the dashed lines correspond to the estimated wage gaps once the contribution of the sample composition has been net out. Table 3 summarizes point estimates at selected quantiles.[15]

In the case of the conditional mean, as reported in Table 2, a simple Oaxaca-Blinder decomposition would tell us that around half of the public sector raw wage gap is explained by differences in observable characteristics. Similarly, we find that if workers in the private and in the public sectors had the same characteristics, the public sector wage gap along the wage distribution would be significantly lower, especially at the top. In fact, for men in the upper-

[14] These estimates are based on quantile regressions presented in Appendix B.

[15] Given the huge sample size we consider there is not need to include standard errors in the tables or figures. To illustrate this point Figure A.1 in the Appendix show how tight are the confidence intervals in the case of a 5 per cent random draw of the sample the we use. Standard errors are computed by bootstrap and the computational burden is very high.

Figure 5: Estimated gaps along the wage distribution



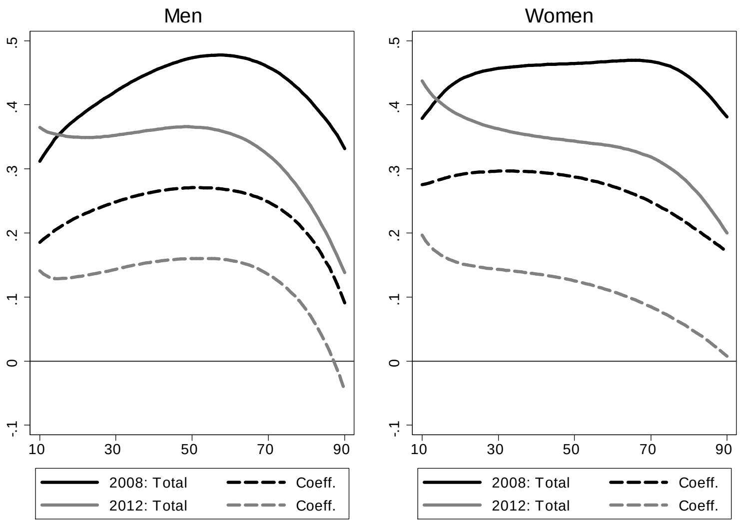
Notes: Whole sample (2005-2012).

Table 3: Estimated gaps along the wage distribution

			Spec. 1		Spec. 2		Spec. 3	
	Quantile	Sample	Total	Coef.	Total	Coef.	Total	Coef.
Men	10	0.286	0.250	0.198	0.247	0.184	0.245	0.158
	25	0.361	0.362	0.276	0.368	0.263	0.363	0.226
	50	0.440	0.435	0.309	0.434	0.297	0.431	0.258
	75	0.381	0.383	0.231	0.383	0.228	0.388	0.203
	90	0.266	0.268	0.060	0.273	0.074	0.268	0.055
Women	10	0.368	0.365	0.297	0.368	0.281	0.365	0.254
	25	0.438	0.430	0.321	0.429	0.317	0.424	0.270
	50	0.437	0.434	0.282	0.433	0.300	0.432	0.261
	75	0.403	0.406	0.169	0.411	0.224	0.412	0.202
	90	0.322	0.326	0.065	0.326	0.160	0.321	0.136

Notes: Whole sample (2005-2012). Sample reports the difference between the j quantile of log hourly wages in the public sector, q_j^1, and the one in the private, q_j^0. Total refers to $\hat{q}^1_{\theta_j} - \hat{q}^0_{\theta_j}$, and Coefficients to $\hat{q}^1_{\theta_j} - \hat{q}^c_{\theta_j}$.

Figure 6: Estimated gaps: 2008 vs 2012



part of the distribution, the positive wage gap practically disappears (the gap ranges 0.05-0.07 depending on the specification). This means that a substantial fraction of the public sector gap is due to the fact that public employees are in general better in terms of covariates than private sector employees. The table also shows that the three specifications perform similarly in terms of the fit (which is remarkably good), and that the three offer similar estimates of the unconditional quantiles. Results from here onwards are all obtained using specification 3.[16]

With respect to the evolution of the public sector wage gap over time, to easy the presentation and analyses of results, we focus in two particular years: 2008 and 2012. Figure 6 shows the percentiles of the two public sector wage gaps (total – solid lines; and coefficients – dashed lines) by gender for those two years. Table 4 summarizes point estimates at selected quantiles.

From 2008 to 2012 we see that the public sector raw wage gap has decreased substantially both for men and women, with the only exception of the 10th percentile. Once the contribution of observed characteristics is taken into account, we still observe significant decreases, with the conditional median wage gap for men moving from 0.27 in 2008 to 0.16 in 2012, and for women from 0.29 in 2008 to 0.12 in 2012. At the 90th percentile, in 2012 the gap for men becomes negative and almost zero for women.

5.3 The Public Sector Wage Gap by Subgroups of Workers

In order to better understand the evolution of the public sector wage gap, in this section we consider two different subgroups of workers. We first consider workers by skill groups, distinguishing between high, medium and low skilled individuals. Second, we separate workers by type of contract, that is, those workers with a permanent contract *versus* those with a

[16]Results from the two other specifications are available upon request.

Table 4: Estimated gaps along the wage distribution

		2008		2012	
	Quantile	Total	Coef.	Total	Coef.
Men	10	0.311	0.185	0.365	0.141
	25	0.401	0.238	0.349	0.137
	50	0.473	0.270	0.365	0.160
	75	0.439	0.229	0.292	0.113
	90	0.332	0.091	0.138	-0.046
Women	10	0.379	0.275	0.437	0.196
	25	0.451	0.294	0.371	0.146
	50	0.464	0.287	0.343	0.125
	75	0.460	0.232	0.301	0.070
	90	0.381	0.171	0.200	0.008

Notes: Sample refers to the difference between sample quantiles, Total to $\hat{q}^1_{\theta_j} - \hat{q}^0_{\theta_j}$, and Coefficients to $\hat{q}^1_{\theta_j} - \hat{q}^c_{\theta_j}$.

fixed-term or temporary position.

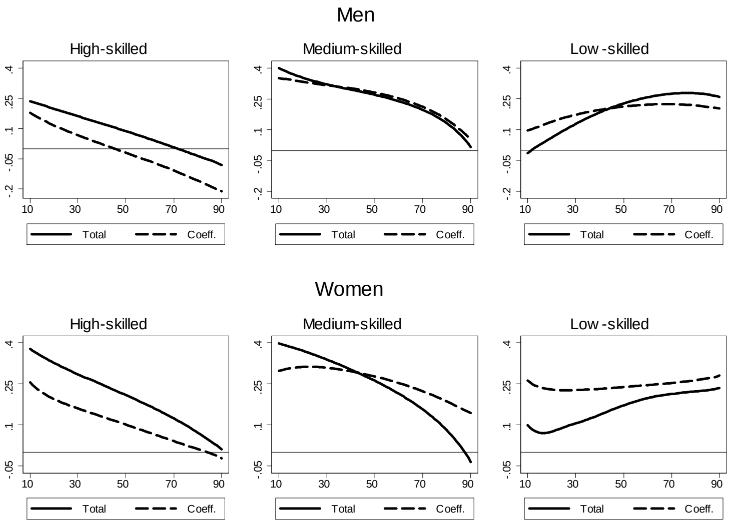
5.3.1 Skills

Figure 7 shows the percentiles of the public sector wage gaps by gender and skill level. As previously, the solid lines stand for the estimated total wage gaps, while the dashed lines correspond to the estimated wage gaps once the contribution of the sample composition has been net out. Table 5 summarizes point estimates at selected quantiles.

For high-skilled and medium-skilled workers the total gap is decreasing along the distribution of wages, whereas for low-skilled workers the slope is positive in the bottom half of the distribution and flat or slightly negative in the upper part. Once we condition on observables, we find that if high-skilled male workers in the private and in the public sectors had the same characteristics, the public sector wage gap would be negative already at the median. For high-skilled women is always positive, but substantially lower. For medium and low-skilled male workers the role of characteristics is rather limited. Finally, for medium and low-skilled female workers the conditional public sector wage premium is higher than the total gap for observationally comparable individuals.

As before, to see the evolution over time, we report in Figure 8 the percentiles of the public sector wage gaps in 2008 and 2012. From 2008 to 2012, we observe important decreases in all those gaps. The most salient facts are the following. For high-skilled male workers, the

Figure 7: Estimated gaps by skill level



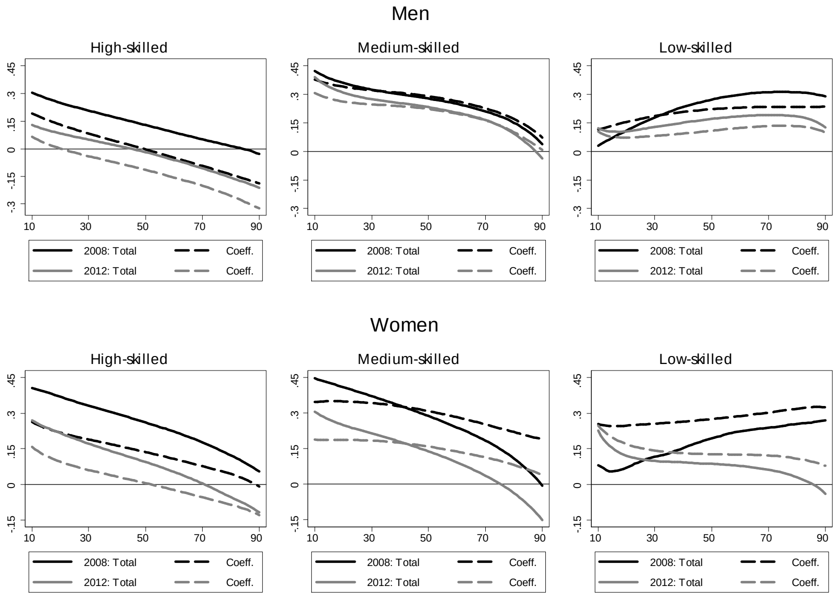
Notes: Whole sample (2005-2012).

Table 5: Estimated gaps by skill level

		High		Medium		Low	
	Quantile	Total	Coef.	Total	Coef.	Total	Coef.
Men	10	0.235	0.178	0.401	0.351	-0.013	0.096
	25	0.180	0.091	0.337	0.325	0.096	0.158
	50	0.087	-0.020	0.271	0.282	0.228	0.213
	75	-0.014	-0.135	0.170	0.185	0.278	0.222
	90	-0.083	-0.212	0.017	0.051	0.259	0.204
Women	10	0.378	0.255	0.398	0.298	0.098	0.261
	25	0.306	0.176	0.356	0.312	0.091	0.226
	50	0.208	0.101	0.261	0.277	0.171	0.238
	75	0.098	0.026	0.122	0.205	0.218	0.257
	90	0.011	-0.022	-0.035	0.144	0.234	0.281

Notes: Whole sample (2005-2012). Total refers to $\hat{q}^1_{\theta_j} - \hat{q}^0_{\theta_j}$, and Coefficients to $\hat{q}^1_{\theta_j} - \hat{q}^c_{\theta_j}$.

Figure 8: Estimated gaps by skill level: 2008 vs 2012



conditional public sector wage gap is negative already at the 21th percentile in 2012, and now also for women is negative from the 52th percentile onwards. For medium-skilled workers, the total gap in 2012 is negative at the very top of the distribution, but once composition is considered the gap is always positive. Finally, the uncommon increasing profile obtained for low-skilled workers in 2008 disappears in 2012.

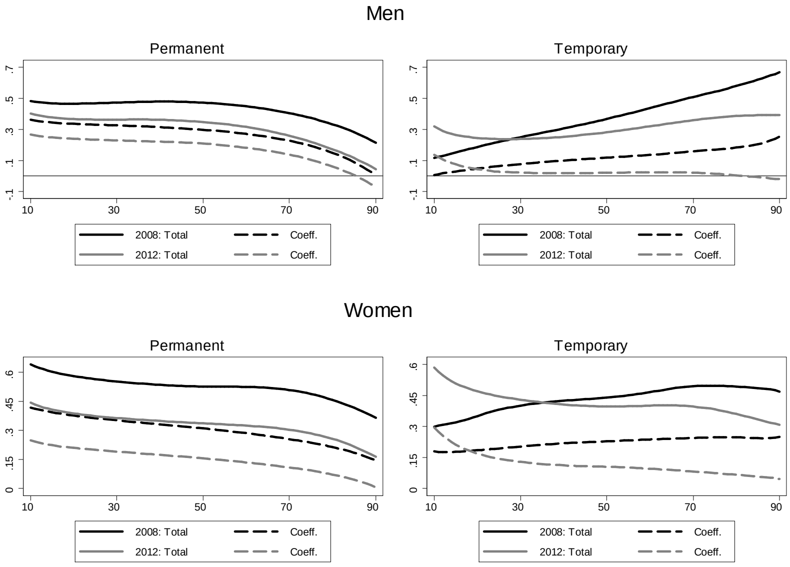
5.3.2 Type of contract

Figure 9 shows the percentiles of the public sector wage gaps by gender and type of contract. Again, the solid lines stand for the estimated total wage gaps, while the dashed lines correspond to the estimated wage gaps once the contribution of the sample composition has been net out. Table 6 summarizes point estimates at selected quantiles.

For workers with a permanent contract the public sector raw wage gap is in general decreasing, while - on the contrary - for temporary male workers the raw gap increases as wages also increase, and for temporary women it remains flat. Once composition is taken into consideration, the gap for indefinite positions falls in a parallel fashion, similarly to the case of women in temporary positions, whereas for men the gap adopts a concave shape.

Once again, to see the evolution over time, we depict in Figure **??** the percentiles of the public sector wage gaps in 2008 and 2012. For permanent workers the falls in the gaps, both total and in coefficients, are parallel. For temporary workers, however, we find that the gaps from 2008 to 2012 rotate downward, adopting a decreasing shape more in line with previous evidence.

Figure 9: Estimated gaps by type of contract



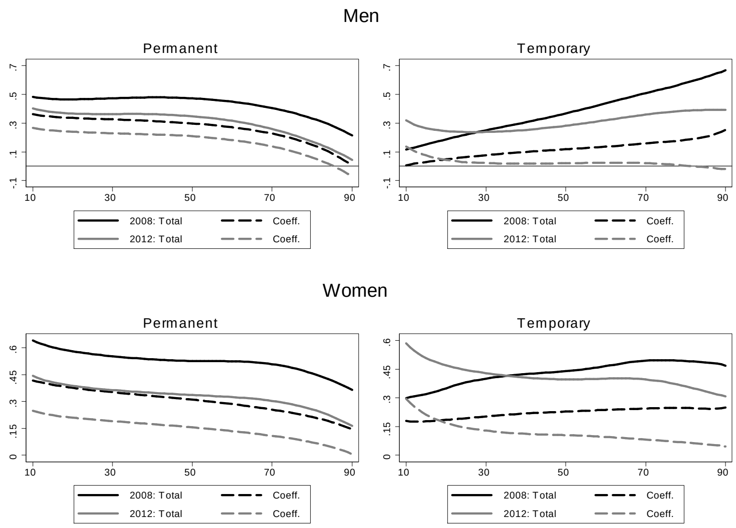
Notes: Whole sample (2005-2012).

Table 6: Estimated gaps by type of contract

Quantile		Permanent Total	Permanent Coef.	Temporary Total	Temporary Coef.
Men	10	0.465	0.349	0.080	0.029
	25	0.451	0.327	0.168	0.052
	50	0.446	0.292	0.306	0.094
	75	0.340	0.184	0.460	0.116
	90	0.190	0.006	0.555	0.140
Women	10	0.583	0.384	0.328	0.200
	25	0.522	0.348	0.377	0.182
	50	0.485	0.301	0.422	0.202
	75	0.438	0.226	0.456	0.204
	90	0.326	0.145	0.410	0.208

Notes: Whole sample (2005-2012). Total refers to $\hat{q}^1_{\theta_j} - \hat{q}^0_{\theta_j}$, and Coefficients to $\hat{q}^1_{\theta_j} - \hat{q}^c_{\theta_j}$.

Figure 10: Estimated gaps by type of contract: 2008 vs 2012



5.4 The Public Sector Wage Gap by Region

In this section we analyze the regional variation in the public wage gaps. In particular, we estimate the public-private wage gaps for all the fifteen regions with information available in our dataset (note that there are no data in the MCVL for the Basque Country and Navarra). Table 7 reports the estimated gaps (both total and due to returns) at the 25, 50, and 75 percentiles for each region. The highest gaps are observed in Murcia, Canary Islands, and Balearic Islands while Andalusia, Valencia, and Cantabria also present gaps above the national average for all the percentiles. Moreover, the group (and ranking) of high-gap regions remains virtually unaltered once we control for observable characteristics and focus in the part of the gap due to returns (coefficients). Only Cantabria (Galicia) switches from the above-average (below-average) group to the below-average (above-average) group once composition effects are accounted for. On the other hand, the lowest public-private gaps are observed in Extremadura, La Rioja, and Aragon.

Regarding the profiles of the estimated public wage gaps, we observe different patterns across the different regions. The nationwide gaps reported above present an inverted-U shape due mainly to the increasing profile of low-skilled workers combined with the decreasing profiles of medium- and high-skilled workers. Interestingly enough, this aggregate profile is only present in Valencia where the 25 and 75 percentiles are lower than the 50 percentile. However, for all the other regions the overall profile is either increasing or decreasing. Regions such as Madrid and Catalonia present a clearly decreasing (i.e. Q75<Q50<Q25) profile which is similar to the profile observed for medium- and high-skilled workers in the national aggregate. In contrast, other regions such as Extremadura present an increasing profile (i.e. Q75>Q50>Q25) similar to that of low-skilled workers. We tentatively argue that these marked differences represent an

Table 7: Estimated gaps by region

	Total			Coef.		
	Q25	Q50	Q75	Q25	Q50	Q75
Andalusia	0.38	0.45	0.48	0.28	0.30	0.29
Aragon	0.35	0.33	0.30	0.16	0.14	0.09
Asturias	0.41	0.40	0.37	0.28	0.26	0.18
Balearic Islands	0.46	0.53	0.56	0.26	0.31	0.32
Canary Islands	0.50	0.55	0.58	0.37	0.40	0.37
Cantabria	0.41	0.44	0.44	0.19	0.21	0.17
Castilla La Mancha	0.21	0.38	0.44	0.11	0.20	0.21
Castilla and Leon	0.34	0.38	0.39	0.18	0.19	0.14
Catalonia	0.41	0.41	0.34	0.24	0.23	0.17
Valencia	0.46	0.48	0.44	0.23	0.27	0.26
Extremadura	0.11	0.32	0.41	0.09	0.18	0.21
Galicia	0.37	0.42	0.46	0.25	0.26	0.22
La Rioja	0.36	0.37	0.36	0.04	0.12	0.12
Madrid	0.39	0.34	0.21	0.26	0.21	0.11
Murcia	0.55	0.59	0.59	0.36	0.40	0.38

Notes: This region-specific gaps are based on the 2008-2012 period.

Total refers to $\hat{q}^1_{\theta_j} - \hat{q}^0_{\theta_j}$, and Coef. to $\hat{q}^1_{\theta_j} - \hat{q}^c_{\theta_j}$.

indication of the heterogeneous composition of the workforce across regions.

We now turn to the relationship between the regional public wage gaps and economic activity. Theoretically, regions with higher unemployment and lower productivity should also present higher public wage gaps. A reduction in productivity or an increase in the unemployment rate should lead to lower wages (see e.g. García-Pérez and Jimeno, 2007); however, since private wages are more responsive to economic conditions than public wages, a deterioration of economic activity (increase in unemployment or reduction in productivity) should generate higher public wage gaps.

The upper panel of Figure 11 presents two scatter plots of unemployment and the logarithm of labor productivity against the public wage gaps for the year 2008. Both graphs support the hypothesis discussed above, regions with higher public wage gaps are those with higher unemployment rates and lower productivity levels. The slope coefficients for the estimated regression lines are highly significant in both cases even with only fifteen observations. However, in the bottom panel of Figure 11 we present the same scatter plots but considering 2008-2012 changes in the variables instead of their levels. Surprisingly enough, the significant relationship found for the levels completely vanishes for the changes; the slope coefficient is now indistinguishable from zero in both cases. Moreover, contrary to the theoretical arguments above, the public wage gaps have been reduced in all the regions over the 2008-2012 period, characterized by a severe economic recession in Spain.

We tentatively conclude from the findings above that, on the one hand, the evolution of the public wage gap in the Spanish regions over the last years has been dominated by the recent cuts in public sector wages rather than reductions in private wages in response to the economic downturn; on the other hand, causality claims from productivity/unemployment to public-private wages in Spanish regions are not granted.

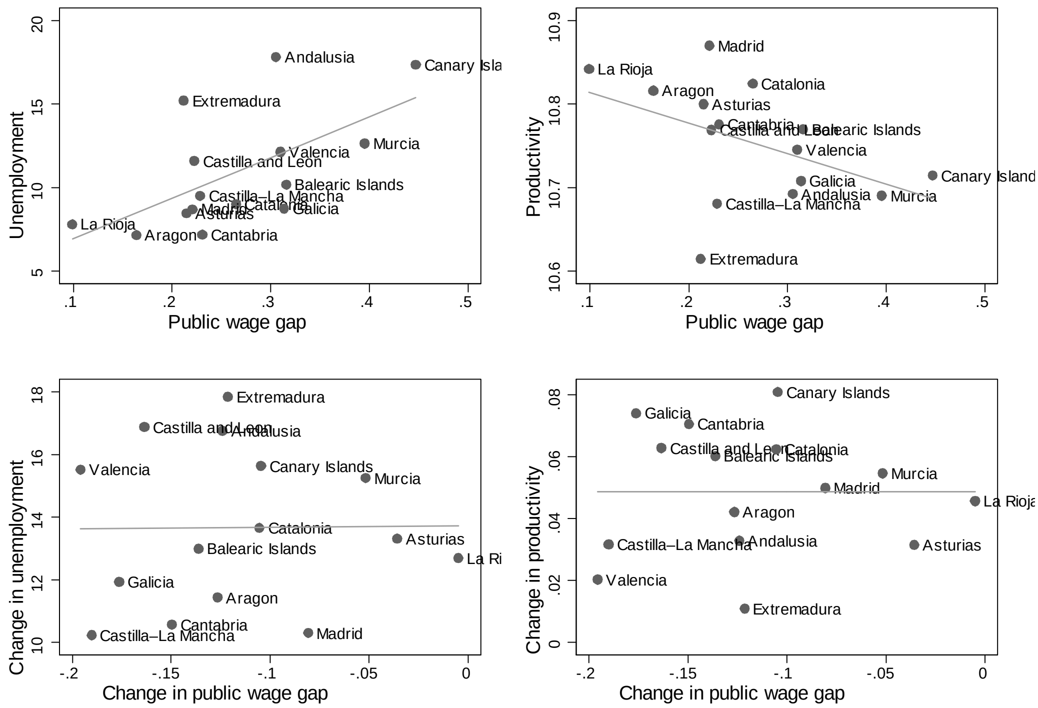
5.5 The Role of Unobservables

[TO BE COMPLETED]

6 Concluding Remarks

[TO BE COMPLETED]

Figure 11: Public wage gaps and economic activity across Spanish regions



The plots in the upper panel refer to levels for the year 2008 while the bottom panel plots refer to changes between 2008 and 2012. Productivity refers to log labor productivity from National Accounts.

References

[1] Anghel, B., S. de la Rica, and J. Dolado (2011) "The Effect of Public Sector Employment on Women's Labour Martket Outcomes", FEDEA Working Paper No. 2011-08.

[2] Atkinson, A. and J. Micklewright (1983) "On the reliability of income data in the family expenditure survey 1970-1977", Journal of the Royal Statistical Society, vol. 146, pp. 33-61.

[3] Bargain, O. and B. Melly (2008) "Public sector pay gap in France: New evidence using panel data", Discussion Paper Series, IZA DP, No. 3427.

[4] Blinder, A. (1973) "Wage Discrimination: Reduced Form and Structural Estimates", Journal of Human Resources, vol. 8, pp. 436-455.

[5] Borjas, G. (2002) "The Wage Structure and the Sorting of Workers into the Public Sector", NBER Working Paper No. 9313.

[6] Boyle, G., R. McElligott and J. O'Leary (2004) "Public-private wage differentials in Ireland, 1994-2001", ESRI Quarterly Economic Commentary, Special Article, Summer, Dublin: The Economic and Social Research Institute.

[7] Cai, L. and A.Y.C. Liu (2008) "Public-private wage gap in Australia: Variation along the distribution", Centre for Economic Policy Research Discussion Paper No. 581, (June), Canberra, Australia: The Australian National University.

[8] Campos, M. and C. Pereira (2009) "Wages and incentives in the Portuguese public sector", Economic Bulletin, Banco de Portugal, Summer, pp. 57-77.

[9] Chernozhukov, V., I. Fernandez-Val, and B. Melly (2013) "Inference on counterfactual distributions", Econometrica, 81 (6), 2205-2268.

[10] Cohen, S. and B. Lipstein (1954) "Response Errors in the Collection of Wage Statistics by Mail Questionnaire", Journal of the American Statistical Association, vol. 49, pp. 240-250.

[11] De Castro, F., M. Salto, and H. Steiner (2013) "The gap between public and private wages: new evidence for the EU", European Economy. Economic Papers 508.

[12] Dustmann, C. and A. Van Soest (1997) "Wage Structures in the Private and Public Sectors in West Germany", Fiscal Studies, vol. 18, pp. 225–247.

[13] Foley, P. and F. O'Callaghan (2009) "Investigating the public-private wage gap in Ireland using data from the National Employment Survey 2007", Statistical and Social Inquiry Society.

[14] García-Pérez, I. and J. Jimeno (2007) "Public Sector Wage Gaps in Spanish Regions", The Manchester School, vol. 75, pp 501-531.

[15] Giordano, R., D. Depalo, M. 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 WP Series No. 1406.

[16] Gottschalk, P. and M. Huynh (2010) "Are Earnings Inequality and Mobility Overstated? The Impact of Nonclassical Measurement Error", The Review of Economics and Statistics, vol. 92, pp. 302-315.

[17] Gottschalk, P., E. McEntarfer, and R. Moffitt (2008) "Trends in the Transitory Variance of Male Earnings in the U.S., 1991-2003: Preliminary Evidence from LEHD data", Boston College WP No. 696.

[18] Gregory, R. and J. Borland (1999) "Public Sector Labor Markets", In: Ashenfelter O, Card D (eds.) Handbook of Labor Economics, vol. 3c. Elsevier, Amsterdam, pp. 3573-3630.

[19] Griliches, Z., B. Hall, and J. Hausman (1978) "Missing data and self-selection in large panels", Annales de l'INSEE, vol. 30, pp. 138-176.

[20] Groves, R. (2006) "Nonresponse rates and nonresponse bias in household surveys", Public Opinion Quarterly, vol. 70, pp. 646-675.

[21] Herriot, R. and E. Spiers (1980) "Measuring the Impact on Income Statistics of Reporting Differences Between the Current Population Survey and Administrative Sources", Studies from Interagency Data Linkages, U.S. Department of Health, Education, and Welfare, Social Security Administration, pp. 39-49.

[22] Jürges, H. (2002) "The Distribution of the German Public–Private Wage Gap", LABOUR, vol. 16, pp. 347-381.

[23] Koenker, R. and G. Bassett (1978) "Regression Quantiles", Econometrica, vol. 46, pp. 33-50.

[24] Lassibille, R. (1998) "Wage Gaps Between the Public and Private Sectors in Spain", Economics of Education Review, vol. 17, pp. 83–92.

[25] Lamo, A., J. Perez, and L. Schuknecht (2012) "Public or private sector wage leadership? An international perspective", Scandinavian Journal of Economics, vol. 144, pp. 228-244.

[26] Mellow, W. and H. Sider (1983) "Accuracy of Response in Labor Market Surveys: Evidence and Implications", Journal of Labor Economics, vol. 1, pp. 331-344.

[27] Melly, B. (2005) "Public-private sector wage differentials in Germany: Evidence from quantile regression", Empirical Economics, Vol. 30, pp. 505-520.

[28] Miller, H. and L. Paley (1958) "Income Reported in the 1950 Census and on Income Tax Returns. An Appraisal of the 1950 Census Income Data", Princeton, NJ: Princeton University Press, pp. 179-201.

[29] Mueller, R. E. (1998) "Public-private sector wage differentials in Canada: Evidence from quantile regressions", Economics Letters, vol. 60(2), pp. 229-235.

[30] Nielsen, H. and M. Rosholm (2001) "The Public-Private Sector Wage Gap in Zambia in the 1990s: A Quantile Regression Approach", Empirical Economics, vol. 26, pp. 169-182.

[31] Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets", International Economic Review, vol. 14, pp. 693-709.

[32] Panizza, U. and C. Qiang (2005) "Public–private wage differential and gender gap in Latin America: Spoiled bureaucrats and exploited women?", The Journal of Socio-Economics, vol. 34, pp. 810-833.

[33] Papapetrou, E. (2006) "The public-private sector pay differential in Greece", Public Finance Review, vol. 35(4), pp. 450-473.

[34] Portnoy, S. (1991) "Asymptotic Behavior of the Number of Regression Quantile Breakpoints", SIAM Journal of Scientific and Statistical Computing, vol. 12, pp. 867-883.

[35] Poterba, J. and K. Rueben (1995) "The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods", NBER Working Paper No. 4734.

[36] Rahona, M., M. Salinas, and I. Murillo (2013) "Análisis de los diferenciales salariales por sector y género en España", mimeo.

[37] Ramos, R., E. Sanromá, and H. Simón (2013) "Diferencias salariales entre sector público y privado según tipo de contrato", mimeo.

[38] Reder, M. (1975) "The theory of employment and wages in the public sector", in D. Hamermesh (ed.) Labour in the Public and Nonprofit Sectors, Princeton University Press, Princeton, pp. 1-48.

[39] Smith, S. (1976) "Pay differentials between federal government and private sectors workers", Industrial and Labour Relations Review, vol. 29, pp. 233–257.

[40] Zweimuller, J. (1992) "Survey non-response and biases in wage regressions", Economics Letters, vol. 39, pp. 105-109.

A Additional information

Table A.1: Sample composition

	All		Men		Women	
	Ind.	Obs.	Ind.	Obs.	Ind.	Obs.
Prime-age individuals	716,838	4,104,291	381,467	2,190,919	335,371	1,913,372
Prime-age working individuals	716,706	3,662,603	381,380	1,978,177	335,326	1,684,426
Merged with hours data	688,597	3,232,598	366,027	1,762,890	322,570	1,469,708

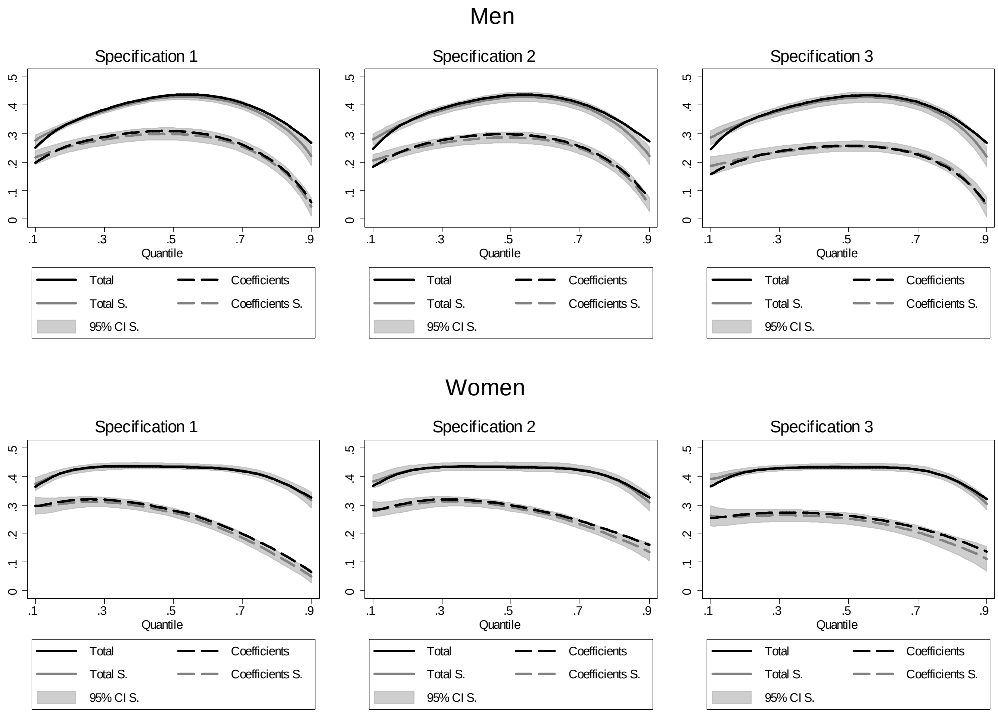
Notes: Whole sample (2005-2012). Ind. = Individuals; Obs. = Observations.

Table A.2: Summary statistics

	Public			Private		
	All	Men	Women	All	Men	Women
Age	40.45	40.76	40.26	37.20	37.63	36.63
High-skilled	38.21	32.96	41.56	12.78	13.35	12.01
Medium-skilled	38.75	37.81	39.35	33.86	23.11	48.31
Low-skilled	23.03	29.23	19.08	53.35	63.54	39.68
Tenure	4.75	5.60	4.22	4.01	4.33	3.58
Temporary	46.85	36.11	53.70	31.17	32.16	29.84
Part-time	5.47	2.70	7.24	15.48	4.66	30.02
Days of work	299.86	311.06	292.72	295.72	300.27	289.60
Hours worked	7.33	7.52	7.21	7.65	8.24	6.85
Annual earnings	25.90	29.17	23.82	19.38	22.41	15.32
Daily earnings	85.65	93.70	80.52	67.34	77.37	53.88
Hourly wages	11.81	12.56	11.33	8.97	9.60	8.11

Notes: Whole sample (2005-2012). Standard deviations of non-binary variables in parentheses. Annual earnings in thousands EUR 2012.

Figure A.1: Estimated gaps along the wage distribution



Notes: Whole sample (2005-2012).

B Quantile Regression Coefficients

In this Appendix, we present some estimates of the quantile regression coefficients for each group —public and private— based on both pooled and panel approaches. These pooled and fixed-effects quantile regressions represent the first step in the counterfactual decompositions reported in subsections 5.2 and 5.5, respectively.

B.1 Pooled Quantile Regressions

Our dependent variable (y_i) is individual i's log hourly wage in real terms. With respect to the vector of covariates (X_i), we consider three different specifications: first, we consider those variables often included in Mincerian models, namely, age, age squared, skill-groups, time and regional dummies; second, we add indicators for the type of contract (fixed-term vs. open-ended), and the type of work schedule (full-time vs. part-time); and finally, we also include firm size as an additional categorical variable (less than 10 employees, 10-50, 50-200, more than 200).[17]

We present estimates of quantile regression coefficients by gender for selected quantiles, a particular specification and for the pooled sample of the whole period 2005-2012. We have conducted separate quantile regressions for every year as well. Results in coefficient estimates do not change much when we consider different years, or alternative specifications.[18]

In particular, Table B.1 presents the estimation results for specification 3 and five different quantiles —10th, 25th, 50th, 75th, and 90th— of the wage distribution for private (columns 1-5) and public (columns 6-10) male workers. Similarly, Table B.2 presents the estimation results for females.

The age-earnings profiles are concave both in the public and the private sectors (only the top quantiles for women in the public sector do not present such a concave profile).

We now analyze the differences in "returns to schooling" across the wage distribution in both the private and the public sector. Our coefficient estimates, both for males and females, point to one striking difference between the public and the private sector; while the return to education in general increases with the quantile considered in the private sector, this is not the case in the public sector. This also implies that only at the top of the distribution returns to education are higher in the private sector (competitive) than in the public sector (non-competitive). In contrast, at the bottom of the distribution the return to education is always higher in the public sector. We also find that in the private sector, the profile of returns for low-skilled positions is flatter relative to high and medium-skilled jobs. In addition, for women in the private sectors the profile of returns is less steep than the one for men. In fact, for women in low skilled occupations the return to education also decreases with the quantile in the private sector.

The effect of working part-time on hourly wages is generally positive and, for women, slightly larger in the public sector.[19] On the other hand, temporary contracts in the private sector have a wage penalty for men, whereas for women the penalty is only present in the bottom-half of the distribution. For females with a temporary contract in the private sevtor, the wage premium

[17] These two correspond to *specification 1, 2* and *3*, respectively, in subsequent tables and figures. For regressions that pool men and women together we add a female indicator.

[18] They are available upon request.

[19] Only at the 10th quantile the part-time effect is negative in both sectors.

Table B.1: Quantile regression estimates for public and private sectors - Men

	Private Sector					Public Sector				
	Q10	Q25	Q50	Q75	Q90	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age	0.014	0.013	0.010	0.009	0.017	0.035	0.023	0.014	0.014	0.012
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Age squared	-0.013	-0.010	-0.004	-0.000	-0.007	-0.034	-0.021	-0.011	-0.012	-0.009
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Group 1	0.684	0.735	0.877	1.072	1.262	0.876	0.814	0.828	0.907	0.861
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)	(0.006)
Group 2	0.546	0.548	0.617	0.706	0.768	0.786	0.682	0.609	0.558	0.436
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)
Group 3	0.470	0.486	0.603	0.761	0.858	0.575	0.487	0.429	0.405	0.281
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.008)	(0.006)	(0.005)	(0.006)	(0.010)
Group 4	0.346	0.354	0.451	0.564	0.629	0.542	0.493	0.495	0.467	0.339
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.007)	(0.006)	(0.005)	(0.006)	(0.009)
Group 5	0.241	0.212	0.267	0.387	0.480	0.455	0.401	0.376	0.306	0.141
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)
Group 6	0.064	0.059	0.061	0.089	0.120	0.287	0.161	0.079	0.017	-0.123
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.008)
Group 7	0.069	0.029	0.055	0.140	0.234	0.399	0.295	0.261	0.264	0.120
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)
Group 8	0.211	0.149	0.122	0.129	0.122	0.391	0.290	0.242	0.195	0.042
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)
Group 9	0.111	0.065	0.051	0.067	0.069	0.303	0.184	0.117	0.068	-0.080
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.008)	(0.006)	(0.005)	(0.006)	(0.010)
Part-time	-0.232	0.012	0.226	0.427	0.618	-0.252	-0.010	0.184	0.365	0.538
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.006)	(0.005)	(0.004)	(0.005)	(0.008)
Temporary	-0.102	-0.049	-0.035	-0.034	0.020	-0.283	-0.222	-0.221	-0.190	-0.091
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)
Size 1	-0.195	-0.222	-0.262	-0.283	-0.260	-0.095	-0.098	-0.099	-0.079	-0.023
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
Size 2	-0.119	-0.166	-0.206	-0.223	-0.221	-0.091	-0.074	-0.053	-0.031	0.034
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
Size 3	-0.016	-0.052	-0.076	-0.089	-0.098	-0.039	-0.030	-0.028	-0.025	0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)
Constant	1.020	1.311	1.568	1.740	1.777	0.686	1.179	1.594	1.829	2.151
	(0.014)	(0.009)	(0.008)	(0.010)	(0.018)	(0.036)	(0.027)	(0.025)	(0.029)	(0.046)
Obs.			1,698,709					219,247		
Pseudo R-squared	0.112	0.138	0.206	0.250	0.239	0.316	0.281	0.274	0.254	0.239
Joint p-value	0.00	0.00	0.00	0.00	0.00					

Notes: Male sample (2005-2011). All regressions include regional and time dummies. Standard errors in parentheses.

Table B.2: Quantile regression estimates for public and private sectors - Women

	Private Sector					Public Sector				
	Q10	Q25	Q50	Q75	Q90	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age	0.014	0.017	0.018	0.019	0.025	0.027	0.013	0.009	0.002	-0.009
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Age squared	-0.013	-0.016	-0.017	-0.015	-0.020	-0.020	-0.007	-0.004	0.004	0.014
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Group 1	0.806	0.774	0.836	0.964	1.117	0.914	0.904	0.906	0.943	0.869
	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
Group 2	0.695	0.642	0.692	0.717	0.755	0.841	0.799	0.735	0.627	0.448
	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)
Group 3	0.644	0.601	0.704	0.871	0.976	0.538	0.531	0.474	0.388	0.240
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)	(0.007)	(0.006)	(0.004)	(0.005)	(0.008)
Group 4	0.390	0.314	0.388	0.556	0.693	0.463	0.412	0.362	0.312	0.191
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.008)	(0.006)	(0.005)	(0.006)	(0.009)
Group 5	0.439	0.331	0.340	0.389	0.478	0.467	0.444	0.387	0.287	0.112
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)	(0.005)	(0.004)	(0.003)	(0.003)	(0.005)
Group 6	0.245	0.144	0.136	0.153	0.190	0.368	0.304	0.227	0.109	-0.091
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)	(0.006)
Group 7	0.307	0.194	0.181	0.199	0.245	0.424	0.350	0.266	0.143	-0.061
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)
Group 8	0.273	0.159	0.130	0.120	0.129	0.330	0.286	0.267	0.194	0.019
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.009)	(0.007)	(0.005)	(0.006)	(0.010)
Group 9	0.209	0.101	0.064	0.051	0.049	0.327	0.262	0.175	0.058	-0.129
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.007)	(0.005)	(0.004)	(0.005)	(0.008)
Part-time	-0.122	0.034	0.169	0.273	0.320	-0.026	0.101	0.246	0.373	0.476
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)
Temporary	-0.094	-0.036	0.005	0.055	0.151	-0.152	-0.110	-0.095	-0.072	-0.034
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Size 1	-0.176	-0.186	-0.194	-0.182	-0.157	-0.154	-0.119	-0.087	-0.048	-0.010
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)
Size 2	-0.057	-0.090	-0.115	-0.126	-0.140	-0.206	-0.176	-0.113	-0.047	-0.002
	(0.003)	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)
Size 3	0.023	-0.013	-0.037	-0.047	-0.061	-0.110	-0.081	-0.039	-0.021	-0.003
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Constant	0.671	0.981	1.163	1.319	1.401	0.684	1.172	1.442	1.852	2.405
	(0.020)	(0.012)	(0.010)	(0.013)	(0.020)	(0.030)	(0.022)	(0.017)	(0.021)	(0.032)
Obs.			1,327,292					328,582		
Pseudo R-squared	0.105	0.115	0.159	0.183	0.168	0.330	0.306	0.315	0.304	0.272
Joint p-value	0.00	0.00	0.00	0.00	0.00					

Notes: Female sample (2005-2011). All regressions include regional and time dummies. Standard errors in parentheses.

in the upper-part increases along the wage distribution, reaching a maximum of 15.1% at the 90th percentile. In contrast, workers - both men and women - with a temporary contract earn significantly less than permanent workers in the public sector at all quantiles.

Regarding firm size, we find negative wage effects of working in smaller firms both in the private and in the public sector. However those penalties increase along the wage distribution in the private sector, whereas in the public sector the penalty is less as we move up in the distribution.

Finally, the last row in Tables B.1 and B.2 presents the p-values of a joint test of all public-private interactions, clearly pointing to the existence of a different wage determination process in the public sector.

B.2 Quantile Regressions with Fixed Effects

[TO BE COMPLETED]