ARE PUBLIC SECTOR WORKERS PAID TOO MUCH

DOMENICO DEPALO, RAFFAELA GIORDANO

Are Public Sector Workers Paid too Much?

Domenico Depalo & Raffaela Giordano

Bank of Italy *

June 1, 2010

Abstract

This paper investigates whether a public sector premium exists even after controlling for observable characteristics and for additional motivations, other than monetary, that may induce workers to prefer employment in the public sector. We do this by studying the entire conditional wage distribution on Italian micro data, covering the period 1998-2008. The evidence under random sampling shows the existence of a wage differential averaging at about 14% for women and less or equal to 7% for men, generally higher at the low tail of the wage distribution, for blue collars and in the southern regions. The premium significantly increases when possible sorting is considered; the correction is particularly large above the median of the wage distribution, therefore suggesting that extra-motives may play an important role above all at higher wage levels.

JEL classification:H50, J31, J45, J50.

Keywords: public employment, wage differentials, wage determination.

*Part of the analysis has been performed modifying the routines available in Chernozhukov and Hansen (2005). The views expressed in this paper are those of the author and do not imply any responsibility of the Bank of Italy. Corresponding author: Raffaela Giordano, Banca d'Italia, Research Department, Via Nazionale, 91 - 00184 Roma, Tel.: 39-06-4792 4124, Fax: 39-06-4792 2324, e-mail: `raffaela.giordano@bancaditalia.it`

"I am not worth £213,000. This wage bill is mad. ...
People join, and remain in, the public sector because of a sense
of vocation, to make a difference to society or to the quality of people's lives."
Sir Norman Bettison - Chief Constable of West
Yorkshire Police, in *The Times*, April 12, 2010

1 Introduction

Wage and employment characteristics in the public sector differ significantly from those in the private sector. Analyses comparing average earnings outcomes in public and private sectors generally find positive wage premia for public sector employees. The premium is larger for women and for low skilled public sector employees; it varies also across countries and with the econometric specification.

The theory of why such a wage differential exists is scant.[1] Most arguments invoked to justify the evidence of a wage differential are related to the peculiarity of the public employer's objective function, which encompasses lobbying (Gunderson, 1978), electoral motives (Fogel and Lewin, 1974) or budget-maximization (Niskanen, 1971; Reder, 1975). Other arguments focus on the relatively inelastic labor demand curve in the public sector that the unions would exploit, to the extent allowed by the government budget constraint, to gain higher wages for public sector workers.

Our contribution is to the empirical analysis. The existing literature either uses aggregate data on average wages, which are easily obtained from published sources, or uses individual data in an attempt to control for the quality and the composition of public employees relative to private employees. When controlling for employment characteristics the public sector wage premium typically decreases with the number of controls.

Several types of econometric techniques have been adopted to investigate the public sector wage

[1]The only theoretical paper on the sectoral wage differential is Homlund (1993). In his paper the public and private sector unions bargain over wages with an utilitarian government. The model predicts a wage premium only if the unions act uncooperatively. In this case, in fact, the unions do not internalize two sources of externalities: the first involves the increase in taxes needed to pay for the increase in public sector wages, that are spread across all workers; the second is due to the fact that increases in public sector wages will cause employment in the public sector to decrease, and therefore consumption of goods produces by the public sector to decrease, but again for all workers and not just for public sector workers.

premium using individual data. One approach consists of inferring the existence of a premium from lower quit rates in the public sector or the presence of queues for public sector jobs, rather than actually measuring and comparing the wages across sectors. This method has the advantage to capture the overall (and not just the monetary) compensation in the two sectors, but is not useful to estimate the size of the gap. A second technique foresees the estimation of a single earning equation augmented with a dummy variable indicating whether the worker is employed in the public sector or not. The coefficient of such a variable provides a measure of the wage differential after having controlled for other characteristics of the worker. In the same vein, another econometric specification allows the coefficients to vary across sectors. It consists of two wage equations, one for each sector, in order to capture different returns to worker characteristics. The wage gap is obtained by differencing the worker's actual wage and the estimate of what the worker would earn in the sector in which he is not employed. Through standard decomposition techniques it is then possible to disentangle the impact on the wage gap of differences in worker endowments and that associated to unexplained factors (that is usually interpreted as the 'rent to public sector'). More recently, quantile regression techniques have also been used to compare wages in the public and the private sectors along the entire wage distribution. These analyses usually find higher premia at the bottom tail of the distribution and lower or even negative premia at the top deciles.

The results obtained by all the above techniques should however be interpreted with caution. Indeed, there may be a sample selection bias due to the possibility that sorting of employees between sectors is not random but occurs on the basis of unobserved characteristics. This problem has been typically addressed by jointly estimating equations for the worker's sector of employment and for earnings (relying a set of exclusion restrictions), or by using longitudinal data. To choose appropriate instruments the literature has typically followed some studies that attempted to ascertain whether certain characteristics of individual employees increase the probability that they will seek employment in the public sector. For example, Bellante and Link (1981) find that workers employed in the public sector are more risk averse than private sector employees. Perry and Wise (1990) argue that public sector jobs are likely to provide a greater scope than in the private sector to engage in community or welfare-type activities and this would provide a motivation for public service that will affect individual workers' labor supply decisions; finally, imitation of parents and the presence of social networks, as summarized by variables related to family background, can be important channels to enter one sector or the other. Estimated wage gaps obtained by means of sample selection corrections are generally found to be larger than those not conditioned on these corrections.

Not only are the econometric techniques responsible for the results, but also the preferred point of view is fundamental. Indeed, the public-private sector pay gap can be evaluated from different perspectives, either cross-country or single country. Most of the early research focused on the US; only few studies based on macro data were carried on non-US countries. At the beginning of the '90s a number of studies started addressing wage differentials in Europe, Australia and some developing countries. Bender (1998) and Gregory and Borland (1999) provide extensive surveys of such analyses in a range of countries.

Among more recent studies focusing on the European context, Portugal and Centeno (2001) use the 1995 wave of the European Community Household Panel (ECHP) to compare wage differentials between the general government and the private sectors in the European Union member countries. Considering identical worker's characteristics, they find that the wage gap is wider in Portugal, Ireland, Luxembourg, Spain and Italy; it is narrower in Austria, Belgium, Germany and Denmark, where the differential turns out slightly negative.

Focusing on a cross-country comparison can be useful to assess the impact on wage differentials of different institutions, different wage setting schemes, or different culture. However, heterogeneity in data collection and, in particular, in the definition of the public sector may yield spurious results.

Within EU studies, a single country perspective is adopted by Melly (2005, 2006), who analyzes the public sector pay gap in Germany. In his 2005 paper, by measuring and decomposing the differences in wage distributions between public sector and private sector employees for the years 1984-2001, he obtains higher conditional wages in the public sector only for the women. Furthermore, wage premia are found only at the low tail of the wage distribution, whereas differences in worker's characteristics explain more that the raw wage gap at high wages. These features appear quite stable over both decades. Melly (2006) addresses the issue of possible sorting between the two sectors, along the entire conditional distribution. The qualitative results obtained in his previous paper are confirmed with the new technique, although the magnitude of the premium is substantially higher. Bargain and Melly (2008) estimate the public sector wage premium in France over the period 1990-2002, both at the mean and at different quantiles of the wage distribution, using data on Germany for the year 2003. They account for unobserved heterogeneity by using fixed effects estimations on panel data. Contrary to common results, they find that wages do not substantially differ across sectors after controlling for unobserved heterogeneity. They explain these results by arguing that public sector manages to attract better workers in the lower part of the distribution, in part due to non-monetary gains, but fails to retain the most productive ones at the top.

In this paper we take a single country perspective and investigate the public-private pay gap using Italian micro data covering the period 1998-2008.

Analyses on Italian data are provided, among others, by Bardasi (1996), Comi *et al.* (2002), Dell'Aringa *et al.* (2005). Brunello and Dustmann (1997) compare wage differentials in Italy and Germany; Lucifora and Meurs (2006) evaluate the wage premium in Italy, France and UK. The common finding is that, after controlling for relevant characteristics, in fact there exists a wage premium in Italy for those who work in the public sector that is generally lower for men, high-skilled workers, and in the Northern Italy. With the exception of Bardasi (1996) and Brunello and Dustmann (1997), these studies consider the public/private sector as an exogenous choice. As people can sort in one sector or the other according to unobserved earning-related characteristics, one should be careful while interpreting the results.

The goal of this paper is to investigate whether there exists a premium for public sector workers. In contrast to most of the existing literature, we analyze the entire distribution of (log) wage, by using conditional quantile regression techniques, while properly considering the possible endogeneity of the sector choice. To our knowledge, the only other attempt to jointly address these issues is provided by Melly (2006). As exclusion restrictions, in addition to variables related to parents' occupational status as in Melly (2006), we investigate the role of risk aversion, a measure of the inter-temporal preferences and pro-social attitude.

The paper is organized as follows. Section 2 presents the existing empirical evidence on Italy:

Section 3 briefly discusses our empirical strategy; Section 4 presents the data and some descriptive statistics; Section 5 discusses the empirical findings; Section 6 offers some concluding remarks.

2 The wage gap in Italy

In Italy the difference in pay between public and private sectors in the last decades has always been sizeable. If we look at aggregate data from national accounts the gap was about 20 percent in 1980, reached almost 40 percent in 1990, decreased to 22 percent in 1995; it started increasing again at the beginning of this decade, to reach 36 percent in 2008 (Figure 1).

This evolution seems to contrast with the introduction in 1993 of the current legislation, which aimed at a "privatization" of employment relations in the public sector, that is at making pay and employment condition determination mechanisms in the public sector closer to those in the private sector (by envisaging a greater role for negotiation, imposing tighter constraints to wage growth, replacing the automatic component for wage increases with schemes based on merit). The 1993 reform assigned a larger role to collective bargaining and created an independent agency (Agenzia per la Rappresentanza Negoziale nella Pubblica Amministrazione - ARAN) responsible for negotiating pay levels and working conditions for most public employees. Analyses on both the motivations and the effects of the 1993 reform are provided, among others, by Dell'Aringa (1997) and Lucifora (1999). In principle, higher cooperation between unions in public and private sectors during the wage determination process (which could have followed the greater role assigned to collective bargaining in the public sector) together with a more centralized bargaining process (due to the creation of a single agency representing a vast majority of public employees) should have fostered wage moderation in the public sector or, at least, determined wage dynamics in the public sector more in line with those in the private sector.[2] Fifteen years after the introduction of the reform, many analysts agree that the reform failed to achieve its main targets (see Dell'Aringa, 2006). The evidence reported in this paper and in the previous ones suggests that differences between public and private sectors, regarding not only pay levels but also labor conditions, are still present and maybe widened since then. Furthermore, wage differentials vary significantly both over

[2] Forni and Giordano (2003) present a theoretical model that delivers these results.

time and across different categories of workers and geographical areas.

Possible reasons for the gap are that public sector's workers are older, have higher education and take managerial positions more frequently than in the private sector. In fact, taking into account the significant heterogeneity in the composition (by age, gender, education and occupational level) of the labor force, the public sector premium turns out to be lower but still sizeable.

Most of the papers that analyze the public-private pay gap in Italy using micro data are based on the Bank of Italy Survey of Household Income and Wealth (SHIW), which contains information about personal and occupational characteristics, wages (net of income and payroll taxes) and type of economic activity.

Among them, Bardasi (1996) looks at 1991 incomes. She uses a two-stage approach to take into account the possibility that the distribution of people between sectors may not be random, but rather result from self-selection. She finds a wage premium in the public sector of 9 and 35 per cent for men and women, respectively. Although differences in observable human capital characteristics explain a small fraction of the overall gap, most of the gap is due to a residual component, which captures differences in compensation of given individual characteristics and the effect of geographical and occupational dummies. Furthermore, self-selection seems to negatively affect the wage differential, above all for men. If workers distributed randomly across sectors, the average wage in the public sector would be 25 and 17 per cent higher for men and women, respectively.

Brunello and Dustmann (1997) compare public-private wage differentials in Italy and in Germany. As for Italian data, they use the 1993 wave of the SHIW; German data are from the German Socio-Economic Panel and refer to 1989 incomes. They find a positive wage premium in both countries, higher in Italy than in Germany (21 and 7 per cent, respectively). Moreover, by decomposing the premium into two factors, one associated with individual characteristics and the other with price differentials, they find that the positive premium in Germany is entirely explained by the first factor, which more than offsets a negative price differential that is larger for workers with higher education. In contrast, in Italy the compensation of different characteristics adds to a positive price differential that is higher for less educated workers. Therefore, they conclude that, for given individual characteristics, working in the public sector is penalizing in Germany and rewarding in Italy. According to the authors, their preliminary findings do not reveal significant sample selection biases.

Comi *et al.* (2002) analyze the gap over the period 1977-1998 using data from the SHIW. They also find a positive wage premium for the public sector that is higher for women and for low-income workers. They observe that such premium, after having reached a peak in 1995 (20 per cent for women, after controlling for individual characteristics), started decreasing in 1998. They explain the wage moderation in the public sector observed in the last part of their sample period also by mentioning the reform in the pay determination mechanisms introduced in 1993, for its aim at a "privatization" of the employment relationship for most public sector employees.

Although gender is one key determinant of public sector pay gap, also the geographical heterogeneity turns out to be substantial. An analysis of pay differentials at regional level is provided by Dell'Aringa *et al.* (2005). They show that such differentials, which vary significantly across Italian regions, are partly explained by local labor market conditions that influence wages in the private sector and only marginally wages in the public sector. In the years 1991-2002 in most regions in the North the gap, after controlling for individual and occupational characteristics, turns out to be below 10 per cent, about 4-6 per cent in the largest and most industrialized ones (Piemonte, Lombardia and Emilia Romagna). In contrast, in the southern regions the differential almost always exceeds 15 per cent, reaching 20-25 per cent in the regions characterized by high unemployment levels and where public employment is widespread (Calabria and Sicilia).

Finally, Lucifora and Meurs (2006) investigate the public sector pay gap in 1998 in France, Great Britain and Italy. They use quintile regression (QR) method to analyze wage differentials along the wage distribution, and find that the pay gap significantly declines along the wage distribution. The estimated wage premium from the pooled model ranges between 10 and 12 per cent (lower when more control variables are included in the specification) in the lowest deciles of the distribution, and is about zero in the highest deciles (in the case of women the gap remains positive even at the top deciles, while for men it turns negative in the upper part of the distribution). Quite surprisingly, when conditioning on a larger set of variables they find the highest premium in Great Britain (about 6 per cent on average), where the private sector wage is systematically used as a reference for pay determination in the public sector. Estimates for France and Italy, where no application of this comparability principle is in place, are lower and do not differ considerably between each other (approximately 5 per cent).

To summarize, the presence of a public sector wage premium is largely documented. The existing analyses, most of which focus on a particular year, show a wage differential generally exceeding 20 percent for all employees, which decreases but remains quantitatively large and significant after controlling for individual characteristics of the workers. The premium varies a lot between men and women, workers with different educational and professional qualifications and across geographical areas. With the exception of Lucifora and Meurs (2006), these studies focus on averages rather than on the entire distribution of wages and, with the exception of Bardasi (1996) and Brunello and Dustmann (1997), they take the decision between private and public sector as random. In contrast with them, we relax this assumption and consider the entire wage distribution.

3 The empirical strategy

In our empirical analysis, in addition to Ordinary Least Square (OLS), we use quantile regression (QR) to assess the value of the pay gap along the entire wage distribution.

Indeed, there is in general no guarantee that the mean of a distribution successfully summarizes all its relevant features. This is particularly true in the presence of outliers, as the results obtained using standard approaches are very sensitive to them. Moreover, QR is useful when the impact of explicative variables is different across different parts of the distribution as, in contrast to OLS, it does not impose that covariates have the same effects over the entire distribution. In particular, we employ the quantile regression methodology developed by Koenker and Bassett (1978). It consists of a semi-parametric estimate that combines two parts: one is parametric and refers to the deterministic part, i.e. the covariates; the other is non parametric and concerns the random residuals about which no distributional assumption is imposed. In a simple model $y_i = x_i\beta + u_i$ with $i = 1, \ldots, N$, when we are interested in the θ-th conditional quantile of y_t, β is obtained from the minimization

$$\underbrace{\min}_{\beta\in R}\left[\sum_{i\in\{i:y_i\geq x_i\beta\}}\theta|y_i-x_i\beta|+\sum_{i\in\{i:y_i<x_i\beta\}}(1-\theta)|y_i-x_i\beta|\right].$$

Under some regularity conditions, the solution to the minimization problem gives a consistent and asymptotically normal estimator.[3] Moreover, it is shown that, whereas for a symmetric distribution function median and mean regressions are equivalent, for more complicated distribution function the median regression is uniformly more efficient than the mean regression. For the present study all these aspects are very useful and will be exploited throughout.

However, related to our specific goal, an additional concern regards the exogeneity of the decision of working in the public sector. Suppose for example that workers in the public sector are more risk adverse (Ballante and Link, 1981) and that managers in the private sector have more responsibility: if we don't control for the amount of responsibility in the OLS and QR, we would estimate as "private premium" what in fact is compensation for responsibility. Of course, it can be also vice versa: if, for example, private companies give fringe benefits to their workers as part of salary, we would conclude that there exists a "public premium" if we don't control for the amount of non monetary benefits. Hence, endogeneity of sector choice can bias our estimates and eventually can lead to a misunderstanding of the driving forces of the data generating process. It follows that the assumption of exogeneity must be checked and cannot be imposed.

To our knowledge, the first systematic treatment of endogeneity with median regression is due to Amemya (1982) who proposed a consistent estimator, whose asymptotic properties were further generalized in Powell (1983). The approach is a two stage least absolute deviations (2SLAD) and is analogous to the two stage least square (2SLS) for the mean.

More recently in a series of papers Chernozhukov and Hansen (2005) suggested a dual approach.

[3] The main conditions are that *i*) the distribution F and its density f are continuous and *ii*) the matrix $T^{-1}X'X$ is positive definite.

They consider a model

$$\begin{cases} y &= & D\alpha(U) + X\beta(U) & \text{U |X,Z} \sim \text{Uniform(0,1)} \\ D &= & \delta(X, Z, V) & \text{V is statistically dependent on U} \\ \tau &\mapsto & D'\alpha(\tau) + X\beta(U) & \text{is strictly increasing in } \tau \end{cases}$$

where D is statistically dependent on U (i.e., endogenous), X is a set of exogenous covariates and Z is a set of instruments related to D, but independent on U. Thus, the true structural parameters α cannot be estimated with usual quantile regression, but with an Instrumental Variable Quantile Regression (IVQR). The method is a GMM on

$$Q_n(\theta, \alpha, \beta, \gamma) = \frac{1}{n} \sum_{i=1}^{n} \rho_\theta(Y - D\alpha - X\beta + Z\gamma).$$

The spirit of the approach is that if we knew the true value of α we could obtain consistent estimates using ordinary quantile regression of $y - D\alpha$ on $X\beta$. Thus we can run a series of quantile regression for *given* values of α (grid search) to obtain

$$(\beta(\hat{\alpha, \tau}), \gamma(\hat{\alpha, \tau})) = \underbrace{\arg\min}_{\beta,\gamma} Q_n(\tau, \alpha, \beta, \gamma)$$

and select $\alpha(\theta)$ that makes γ as close to 0 as possible. More formally, it must be found a value of α that satisfies

$$\alpha(\hat{\tau}) = \underbrace{\arg\inf}_{\alpha \in A} [W_n(\alpha_n)], \qquad W_n(\alpha) = n[\gamma(\hat{\alpha, \tau})' A(\hat{\alpha}) \gamma(\hat{\alpha, \tau})],$$

with $A(\alpha)$ set equal to the asymptotic covariance matrix of $\sqrt{n}(\gamma(\hat{\alpha, \theta})' - \gamma(\alpha, \theta)')$, so that $W_n(\alpha)$ can be employed as Wald statistic to test $\gamma(\alpha, \theta) = 0$ (Chernozhukov, Hansen, Jansson, 2007). The estimator is shown to be consistent and asymptotically normal as

$$\sqrt{n}([\alpha(\hat{\theta}), \beta(\hat{\theta})]' - [\alpha(\theta), \beta(\theta)]') \rightarrow_{d.} N(0, \Omega_\theta),$$

for Ω_θ given in Chernozhukov and Hansen (2005). We conclude this section with a summary of the four steps required for the estimation of the coefficients:

1. define a set of (j=1,...,J) values for α;
2. save the covariance matrix to be used for $W_n(\alpha_j)$;
3. choose $\alpha(\tau)$ that minimizes $W_n(\alpha)$ and obtain an estimate for β;
4. obtain the inference on α using Ω_θ.

Finally, notice that this procedure works when the number of exogenous covariates is large, but the number of endogenous covariates is small (typically one, as in our case, or two), because a grid search is employed.

4 The data

Our data are taken from SHIW and refer to the period 1998–2008. Its target population is representative of the Italian population; the sample consists of a small part followed over time (i.e., panel) and a larger part refreshed at each wave. The data contain information about a wide range of personal and occupational characteristics (age, gender, marital status, educational level, region of residency, sector of economic activity and occupational level), about wages (net of income and payroll taxes) and type of activity (firm size, partime status, number of months worked in the year, average number of hours worked in a week). As for all survey data, the quality of SHIW is affected by unit and item non-response. Unit non-response is sizeable (it amounts to 50% and more of the selected households), whereas item non-response depends much on the variable under study. In fact the first issue is tackled by a post stratification weighting process, whereas the second issue is irrelevant for employees: as a consequence, the present study does not suffer from these drawbacks. Furthermore, the data quality is satisfactory when monetary aggregates are compared to Italian National Accounts (Banca d'Italia, 2008).

One critical aspect related to this study concerns the identification of the public sector, which can be defined using two different questions, one regarding the sector of economic activity and the other the firm size. In particular, a public worker can be identified when *i)* her sector of activity is, among others, "public administration, defence, education, health and other public services" or, up to 2006, when *ii)* firm size "is not applicable, because public employee". We checked the robustness of our results using both definitions, but the rest of the paper is based on sector of activity, as in Lucifora and Meurs (2006).

Another issue is related to the choice of the appropriate variable for wage comparison. Indeed, as the average number of hours worked in a week in the public sector is lower than that reported by workers in the private sector, on average 35 and 39 respectively, using monthly wage would under estimate the public sector wage premium. Thus we approximated the hourly wage as $(YW/M)/(4*\bar{H})$, where YW is the yearly wage, M the number of months worked in the year, and $\bar{H}$ the average number of hours worked in a week.

4.1 Descriptive statistics

In this section we illustrate some descriptive statistics for key wage related characteristics of the sample at hand. The statistics refer employees aged 15–65, as usual in this literature.

In Table 1 we tabulate the share of public sector workers in total employment by wave, gender and area. We consider five different areas: North West (NW), North East (NE), Center (C), South (S) and Islands (Is). Public sector employment in Italy is sizeable. It is more widespread among women than among men, in the South and in the Islands than in the rest of the country and declining from 1998 to 2008. In our sample the percentage of public sector employment was 43.5% for women and 28.5% for men in 1998, and went down to 35.4% and 20.9% in 2008, respectively.[4] The reduction occurred in all areas although in the North to a lesser extent, bringing about a "convergence" towards a less dispersed range across areas. The gender gap remained constant over time at about 14%, with again substantial differences across areas.

In Tables 2–3 we report some descriptive statistics for individual characteristics for the pooled 1998–2008 sample (the same evidence emerges in each single year). In Table 2 we present summary

[4] In National Account data the fraction of public sector employees in total employment was equal to 20% in 2008. The difference with our sample is likely due to the way in which the public sector is identified, in particular the sample may include employees working outside the general government, e.g. in education and health services.

statistics for the entire distribution of log of hourly wage and age. On average in Italy the hourly wage in the private sector is higher in the sample of men than in the sample of women by 12% (12% in the North East, 15% in North West, 16% in the Center, 21% in the South and 3% in the Islands). If we focus on selected quantiles, the picture is even more differentiated over the territory: for example, in the South the gender gap is 25% at 25th quantile, 21% at the median and 19% at 75th, i.e. decreasing over quantiles, whereas in the Center it does not vary substantially along the wage distribution (15% at 25th and 75th quantiles and 16% at the median). In the public sector the gender gap is smaller than in the private sector: 4% as national average, positive in the North and in the Islands, and negative or zero in Center and Southern Italy.

More important for the purposes of our analysis is the comparison of wages across sectors. On average the wage gap is in favour of the public sector by 34% for women and 25% for men;it is 30% and 36% at the 25th and the 75th quantile for women and 21% and 26% respectively for men. It is less than 30% in the North and much higher in the Center (36%), in the South (55%) and in the Islands (41%) in the sample of women. In the sample of men these differences are smaller: 20–25% in the North and in the Center, 31.5% in the South and 40% in the Islands. Thus not only the gender gap, but also the private/public pay gap differs significantly across regions.

A graphical summary of these characteristics is presented in Figure 2. In the upper panel we show the Kernel density estimation of the (log of) hourly distribution by gender and sector; the lower panel reports the cumulative distribution functions. In the bottom left figure we fix the gender to compare the (unconditional) gender wage differential: it is quite clear that the horizontal distance between the two curves in the sample of women is larger than in the sample of men, with the public sector cumulative distribution function lying to the right of the private sector cumulative distribution function (i.e. at each quantile the wage is higher in the public sector than in the private sector); moreover, such difference tends to be larger at higher quantiles than at others. In the bottom right figure we fix the sector: it is clear that the horizontal difference is larger in the private sector than in the public sector, with the curve for men lying to the right of the curve referring to women in the private sector (i.e., at each quantile men earn more than women), and virtually overlapping in the public sector.

Turning to workers' characteristics, the age of men and women is similar within each sector (with men being on average older by 1–2 years), but the difference across sectors is sizeable: on average public sector's employees are older by about 5–7 years. Other workers' characteristics are reported in Table 3. The share of married individuals is higher amongst public sector's employees than in the private sector. More relevant for the wage process are education and job position. Public employees are more educated.[5] The fraction of women with a graduate or post-graduate degree is 32% in the public sector as against 8% in the private sector; for men it is 23% and 7%, respectively. The educational gap is observed in all geographical areas, but is wider in Southern regions. As for job position, the great majority of both men and women are white collar in the public sector (up to 92% for women and 81% for men in the South), and blue collar in the private sector, with relative frequencies changing widely across areas. As expected, the fraction of managers is the lowest in both sectors, either for men and for women: it is three times as much in the public sector than in the private sector for men (7% as against 2%); it is 3% in the public sector and negligible in the private sector for women. In general, in the public sector occupational positions seem relatively more favourable than in the private sector compared to educational attainments, above all for men and in the Southern regions. In Italy the fraction of men with intermediate or high education is 71%, the fraction of white collar and managers is 85% in the public sector; the shares are 49% and 32%, respectively, in the private sector. For women the differences across sectors are less striking: the share of workers with intermediate and high education is 86%, that of white collars and managers is 87% in the public sector, as opposed to shares equal to 57% and 48% in the private sector. Finally the part time jobs are observed more frequently for women than for men, for whom it is virtually zero, and in the private sector more than in the public sector.

Given the high variability in job position across areas, gender and sectors, in Table 4 we enrich the analysis of hourly wage by also looking at the position. Table 4 shows that the greatest pay differential between men and women is for private sector managers, whereas for private sector blue and white collars the difference is about 13–15%. Such differences are much smaller in the public sector (for example, the gender difference for white collars is virtually zero). As for the public-

[5] In the table the figures for education do not some up to 1 because it does not include the share of individuals with minimal education.

private pay gap, it is highest for white collar women, in particular those living in the South and in the Islands. For men the premium is generally lower and negative for managers.

So far we have documented a public sector pay gap by gender, job position and geographical area; the interesting question is now whether it persists even after controlling for relevant characteristics determining the wage process. We do this in the rest of the paper.

5 Results

In this section we present the model specification (Section 5.1) and our empirical findings on the wage gap in Italy. In Section 5.2 we report the results for random sampling based on quantile and ordinary least square regressions; in Section 5.3 we remove the assumption of random sampling and control for possible endogeneity between the decision of working in the public sector and the earning process.

5.1 The model

To investigate the public sector pay gap we estimate a standard wage equation, extensively used in previous studies:

$$\ln W_{it} = \alpha X_{it} + \beta PUB_{it} + \epsilon_{it}, \tag{1}$$

where $\ln W_{it}$ is the log of hourly wages (net of income and payroll taxes) of individual i at time t, X_{it} is a vector of individual characteristics, PUB is a dummy variable, which takes value one if the individual works in the public sector, α and β are vectors of parameters, and ϵ_{it} is an error term.

We focus on the coefficient attached to PUB, which is the "return on public sector". All other things equal, if the coefficient is positive there exists a "wage premium" for working in the public sector.

Estimating equation 1 using OLS, or QR, yields consistent estimates if and only if the distribution of workers between the two sectors is random and the vector X consists of exogenous variables. In fact, the distribution of people between the two sectors may not be random, but rather depend on unobserved individual characteristics of the workers, such as different aptitude for risk or different preference for time flexibility; workers might be heterogeneous across sectors also with respect to the gratification that they derive from different types of jobs (e.g., they may desire to be a civil servant or work in the non-profit sector) and thus self-select themselves according to those features; finally, there may be other factors, both monetary and non-monetary, such as fringe benefits, occupational pension benefits, job tenure or wage variability, that can contribute to explain wage differentials. In these situations, interpreting the wage gap as a "premium" or a "penalty" may not be appropriate, unless a correction is made.

Hence we begin our analysis by making the assumption of random sampling, as most of the literature has done so far. We then relax the assumption and address the possibility of non-random sampling. From the previous section it should be clear that the wage differential may differ along the wage distribution; thus the following analysis is on selected quantiles as well as on the mean.

5.2 Random sampling

In Table 5 we present the estimates obtained with QR and OLS using data by year for all employees, and distinguishing by gender. We report only the coefficients attached to public sector, although the regressions are conditional on a set of characteristics, namely: age (a second degree polynomial); marital status; educational level (a set of dummies for basic education, i.e. primary school, lower secondary school, which corresponds to compulsory school according to Italian legislation, higher secondary school, and BA or post-graduate degree); job position (a set of dummies for blue collar, white collar and manager); whether part time or not; geographical area.[6] The reference individual is a woman with higher secondary school education, aged 43 years, working in the private sector as a full-time white collar in Center Italy.

The average premium ranges between 14–17% for women (10% in year 2000) and is much lower for men (7% in 1998, 6% in 2004, 3% in 2006, 7% in 2008, and not statistically significant in years 2000 and 2002). However, further insights can be gained from the investigation of the complete distribution. Indeed, the premium varies along the wage distribution as a formal test of equality of slopes across quantiles rejects the null hypothesis at standard confidence levels. As usual, much

[6] All complete estimates described in this and the following sections are available from the authors upon request.

of the heterogeneity is on the tails of the distribution. For women, over time, the premium ranges between 9.5–16% (28% in 2008) at the left tail, between 8–16% at the 90th quantile, and in general in a narrower interval (around 12–13%) from the 20th to the 80th quantile. Also for men the premium at the left tail varies substantially (from 7% to 17% at the 10th decile). But, unlike for women, the premium for men exhibits a clear decreasing pattern from lower to higher quantiles: the premium is around 5% at the median and even lower at higher quantiles; it is almost never significant from the 70th onward. As in Lucifora and Meurs (2006), we find penalties rather than premia for men at high quantiles, although they are not really significant. The gender difference is generally about 5% up to the 30th quantile and increases constantly to approximately 12% at quantiles 70th, 80th and 90th.

This set of results points towards a public sector premium, which is higher for women than for men, with the gender gap increasing from the lower to the upper tail of the wage distribution. This appears clearly in Figure 3, where the maximum (vertical) distance between the two lines is at the right-hand side of the graph. This feature appears in all years included in the study.

By splitting the sample between different geographical areas and occupational qualifications, other interesting features emerge from the analysis.

By geographical area, as for the sample of women (Table 6), there is a public sector premium in all areas that generally increases as going from Northern to Center and to Southern Italy. On average in the last decade the premium has been about 10% for women in the North, 15% in the Center, 28% in the South and 23% in the Islands. For men there is an impressive heterogeneity, above all in the left tail of the distribution. At high quantiles, working in the public sector sometimes entails a penalty rather than a premium, above all in the North (although not always significant). Moreover, while in Northern and Center Italy the public sector premium remains rather flat along a large part of the wage distribution, in the Islands and in the South it appears particularly high at low quantiles and decreases fast as we move to higher quantiles.

The other important conditioning is on position, whether “Blue collar”, “White Collar” or “Manager” (Table 7). Whereas for blue collar women there is a substantial public sector premium, although somewhat decreasing from lower to upper quantiles and not always significant, the earning premium for white collar women is always significant and generally much higher (the mean averaging above 16% over the sample period, as against 9% for blue collars). On the contrary, female managers are in general worse off in the public sector. Also for men the highest premium is observed for intermediate occupational positions, whereas for managers the disadvantage from working in the public sector is even stronger than for women (who are worse off from the median onward, while male managers are worse off along the entire wage distribution).

According to these results the most important determinant for the public premium is gender: although both men and women are generally better off in the public sector, the advantage for women is larger. Also geographical area and type of job are relevant. For men the premium is virtually zero in Northern Italy, whereas it is positive in other areas: for women it is positive in all areas and generally higher in the Center and in the South. At very high quantiles (80th and above) and for managers there is virtually no premium for public sector employees, and in some cases also emerges a penalty.

Although suggestive, these results must be evaluated with care because of a possible endogeneity between public sector decision and earning process. We try to solve this issue in the rest of the paper.

5.3 Non random sampling

In this section we specifically address the issue of non random sampling across sectors. We first focus on the mean, then we make use of the techniques discussed in Section 3 and investigate quantiles adopting a two step procedure. The first step purges from the endogeneity between public sector choice and earning process, and the second step uses the projection of the sector decision obtained in the first to estimate the premium, along with a grid search. The grid ranges between -1 and +1 (i.e., 100% penalty/premium) with a step of 0.01. Consistent estimates are obtained based on valid exclusion restrictions, i.e. *i)* restrictions that influence the probability of working in the public sector, while *ii)* being uncorrelated with the earning process.

Almost all the existing literature used only variables related to parents' sector of occupation as instruments for sector sorting, mainly due to data limitations. There are several motivations to control for the parents' sector of occupation, beside the fact that they are uncorrelated with the individual's wage level: *i)* the family background may shape the preferences of an individual to work in one sector or another; *ii)* the opportunity for some workers to leave their own job and be replaced by their sons may be relevant, especially in the Italian context; *iii)* the connections built up by parents during their working career may be important channels to enter one sector or the other. Among the studies focussing on Italian data, Brunello and Dustmann (1996) argue that the job of the fathers is key for the job of the sons. On the other hand, Bardasi (1996) argues that public sector workers are more risk averse and value more the stability, thus chooses as exclusion restrictions home ownership and the presence of children.

Undoubtedly family background, risk aversion and preference for stability arguments are all convincing exclusion restrictions. Thus, we exploit a detailed set of information available in SHIW regarding previous fathers' and mothers' job, home ownership and the presence of children (hereafter home related variables), as well as a direct measure of the attitude to take financial risks, which was never taken into account in previous studies. Unfortunately while the latter may be a better proxy of risk aversion than the former, due to questionnaire design it is available only for a smaller subset of respondents. As a consequence, should we use this information the sample size would shrink.

Furthermore, our dataset allows us to control for two other possible motivations not yet explored by the existing literature, although only for wave 2004. One is the rate of time preference, that may affect people's choice to join one sector or the other. We expect that lower future discounting increases the probability to be a public sector worker as deferred salaries, in the form of pension benefits, are generally more generous in the public sector, at least in Italy. The other instrument measures the propensity of the individual to be engaged in active pro social activity and is meant to capture the importance of a pro-social motivation to work in the public sector.[7] On the one

[7] The exact wording of the question for risk aversion is: "When managing your financial investments, would you describe yourself as someone who looks for:" and four categorical answer from "VERY HIGH returns, regardless of a HIGH risk of losing part of your capital" to "LOW returns, WITHOUT any RISK of losing your capital", which we recoded in high vs low risk aversion. The exact wording for the rate of time preference is "Imagine you were told you had won on the lottery the equivalent of your households net annual income. The sum will be paid to you in a year time. However, if you give up part of the sum you can have the rest immediately", and five categorical possible answers regarding the fraction the respondent is willing to give up are envisaged. The exact wording for pro social activity is "In the last year, have you taken an active part in gatherings of any of the following groups

hand including other possible motivations allows us to evaluate whether the existing practice to control for family background only is valid or not for the case of Italy; on the other hand we can sort the importance of motivations.

5.3.1 Sorting the sector of activity

In the first step we regress the public sector indicator on parents' sector of occupation, risk attitude measured by both home related variables and our direct measure of risk aversion, discount rate and pro-social vocation. In Table 8 we report the coefficients only for the exclusion restrictions: column F uses family background alone, column H home related variables alone, column FH both family background and home related variables; column D, column R and column S present the results for discount rate, risk aversion and pro-social motivation, respectively; column FHDRS reports the results obtained when all possible exclusion restrictions are accounted for, but since the inclusion of R implies a large loss of observations, we consider FHDS as the most inclusive.

In what follows, we focus our comments on the exogeneity and relevance of instruments, that are the two essential requisites for the IV technique.[8] Because in general we have more exclusion restrictions than endogenous variables, we can test the over-identifying restrictions for the validity of our instruments (Hansen, 1982). Except for FHDRS for women and FHDS for men, all our instruments are valid as they meet this requisite. The companion requisite is the relevance of instruments. To evaluate it we use the F-statistic as proposed by Stock and Yogo (2005): as a rule of thumb, with one endogenous variable the F-statistic is above 10 if instruments are relevant, below if they are weak. The F-statistic for F is always the highest: in particular having a father working in the public sector significantly increases the probability to work in the public sector for both men and women. The test statistic is higher than 10 for men but not for women. Certainly H alone is a weak indicator for sorting in the public sector, as the test does not reject the null hypothesis that the coefficients of home ownership and presence of children are jointly zero. However, when we focus on FH we reject, for both men and women, the null hypothesis that coefficients of exclusion restrictions

or associations: associations/groups involved in social, environmental, union policy, religious, cultural, sports or recreational, professional, or voluntary activities?", and a YES/NO answer.

[8] We performed these analyses on conditional mean, as this theory for quantiles is not yet established and extending it is beyond the scope of the paper. Notice however that *i)* this first step is common to both IV and IVQR techniques.

are jointly zero. For men the F-statistic is also close to 10, indicating that it is an appropriate set of instruments. For women this is not the case, but, in contrast to men, the presence of children significantly increases the probability of being in the public sector, may be capturing additional motivations related to preferences for time flexibility. The discount rate variables do not have much explanatory power: individually they are not significant not even at 10% confidence level, even though for men they are jointly significant at 5% confidence level. The risk attitude when investing in financial markets is important for men but not for women: notice however that in this case the sample size is dramatically reduced due to questionnaire design. Finally, pro-social motivation is significant for women at 5% confidence level, but largely insignificant for men. In all cases the F-statistics for joint significance of the coefficients of this additional instrument are well below 10.

From these preliminary statistics the only sets of instruments that proves fruitful are family background, whose F-statistic is about 20 and preserves the sample size, and family background joint with home related variables. In order to evaluate whether adding the other instruments would help improving the efficiency of the estimator, we also use the redundancy test (Breush et al, 1999) which basically consists in the comparison of the variances of the estimators based on different sets of instruments: a particular set of instruments is redundant if its inclusion does not reduce the variance of the estimator. The entries in Table 8 are calculated exclusively on column FHDS, because "the potential ambiguity [in ascertain redundant instruments] encountered in a start small approach does not occur in a start big approach" (Breush et al, 1999, p. 99). In column F, for example, the redundancy of family background instruments based on estimates from column FHDS is tested; in column H the redundancy of home ownership and children based on estimates from column FHDS is tested, etc. For women home related and pro-social instruments are not redundant at 10% confidence level (notice however that the size of the sample including S is low). For men, the null hypothesis of redundancy is strongly reject for F and FH; R is non redundant but in this case more than 70% of observations are dropped, whereas D or S are redundant other than weak.

Two conclusions can be drawn from this analysis. First, the instruments more successfully explain the sector choice for men than for women. This may be so because for women also the preceding decision to enter or not the work force should be taken into account, which requires a different model. Nevertheless, our results are valid conditional on individuals being employed. In the literature this problem is either removed by focussing on men, as in Dustmann and Van Soest (1998) who raised the issue explicitly, or completely ignored. Second, according to our tests, the relevant set of instruments is confirmed to be the family background and in particular the sector of the father is most important: leaving aside these restrictions brings to weak identification, or, even worse, all the asymptotic properties for a correct identification are not matched. Quite interestingly, the determinants of sector choice substantially differ across gender.

5.4 The pay gap

Based on this background we can now study the public sector pay gap. In Table 9 we report the results for wave 2004 using as instruments F, FH and FSHD. We investigate the averages and all the quantiles as we did for the exogenous case. For women the average premium is just below 80% when we consider either F or FH and not really significant when we also add the discount rate and the pro-social attitude (notice that in this case the sample consists of only 319 observation). For men the premium is 69% with F or FH and 46% when also accounting for the other instruments. As for quantiles, for women the premia are higher than under the assumption of random sampling although estimated imprecisely. For men the premia are positive and higher than those estimated under the exogeneity assumption, and strongly significant above all in the middle part of the earning distribution. In contrast to previous findings, the premium is relatively constant up to the 70th quantile, at around 40%. It is worth emphasizing that, whatever the set of instruments, the assumption of random sampling substantially under-estimates the true premium, both with averages and with quantiles.

As family background and home related variables are available for all the waves, while being the best instruments from a statistical viewpoint, we also exploit the time dimensionality. Although point estimates are not really stable over time, either when we focus on averages or on quantiles, the overall analysis clearly shows that there is a sorting mechanism at work in all the years, for both men and women. Indeed, based on the classical Hausman (1978) test, the coefficients estimated under the assumption of exogeneity are statistically different from those obtained under non random sampling (at standard confidence level). As a consequence the results obtained assuming exogeneity of sector choice must be taken with caution.

The average premia (Table 10) for women are positive and much higher with IV than with OLS in almost all years no matter which set of instruments is used; it is strongly significant in the period 2000–2004, whereas for 1998, 2006 and 2008 the point estimates are not really precise. For men there is a significant premium in 1998 and 2004 with both sets of instruments and a (almost never significant) penalty in 2000 and 2002. Waves 2006 and 2008 are relatively sensitive to the set of instruments and the point estimates are not really precise. Like for women, for men the size of the premium is in general larger under IV than under OLS. Turning to quantiles (Table 11), for women there exists an increasing pattern across quantiles in 2000–2004, a decreasing pattern in year 2006, whereas the estimation are rather imprecise in year 2008. However, while for year 2000-2004 the estimates obtained using F or FH instruments are quite similar, in year 2006 starting from the median the difference between the two estimates is non trivial, whereas in 2008 the two estimates are quite different across the entire distribution. Therefore for the period 2000-2004 the estimated premium is less sensible to the choice of instruments and in some sense is more reliable. For men the premium is relatively constant across quantiles, as in 1998 or 2004, or increasing, as in 2006 and 2008 (up to 80th quantile). Even when there is a penalty, at 90th quantile (or along a large part of the earning distribution, as in 2000 and 2002), it is not significantly different from zero, suggesting that, even at high wage level, public sector workers are not worse off than private sector workers. Apart from 2008, the estimates are very similar and the same qualitative conclusions hold across the two sets of the instruments.

With these results at hand, it is interesting to understand where, along the wage distribution, the endogeneity bias exerts its main power. Following Angrist et al. (2006), the bias of the QR estimates due to omitted variables, like family background or risk aversion, depends on the (weighted) correlation between the public sector choice and a remainder term, which incorporates family background or risk aversion. As can be seen from a comparison of the estimated coefficients, with and without correction for sector choice endogeneity, in general the bias increases as going from lower to higher quantiles, suggesting that, other things equal, the sorting process is more important at high quantiles than at low quantiles, i.e. precisely at those quantiles where previous studies found lower premia or even penalties. To provide an overall picture of the bias, in Figure 6 we plot the differences in absolute values averaged over time between the coefficients obtained controlling and without controlling for sorting. The bias is similar in size between men and women, and clearly increases along the quantiles, except at the very right end. As the raw difference between the coefficients averaged by year is always positive, under the assumption of exogeneity the premium tends to be under estimated, above all at the high wage levels.

In previous sections we argued that the premium differs across geographical areas and job positions. Thus in the rest of this section we investigate whether also the bias exhibits a systematic pattern. Unfortunately, for this purpose the IVQR would be inappropriate as it is reccomended only when the number of endogenous variables is one or two at most. The alternative solution of splitting the sample would, however, dramatically reduce the number of observations, so we perform this analysis only on the conditional means.

The finding that the penalty at 90th quantile under the assumption of random sampling becomes non significant when controlling for sorting suggests to explore this dimension by job position. Indeed, for both men and women the correction increases as going from blue collars to white collars and then to managers; within the same job position, it is larger for women than for men. As a consequence, even under endogeneity the highest premium is estimated for white collars, whereas for managers the penalty found under the assumption of random sampling becomes non significant, confirming the results obtained by quantiles. For blue collar the difference between OLS and IV is instead small. We propose the following interpretation for these findings. At low quantiles of the earning distribution what really matters is the monetary motivation, so that the bias for not controlling for the sorting mechanism is relatively small (even though still sizeable). In contrast, at high earning quantiles also other factors, mostly related to family background and home characteristics, matter and the bias from not controlling for them leads to quite imprecise conclusions. As for geographical area the correction is sizeable in all areas, and positive on average over time, with only one exception for women (in NE) and for men (in NW). Moreover the correction is particularly high in the Islands for women and in the South and Islands for men.

Our results are coherent with the only existing paper (to our knowledge) that employs instrumental variable on quantile regression to study the public wage pay gap (Melly, 2006). Using data on Germany in year 2003, he finds a positive premium at the low end of the conditional wage distribution and a significant negative premium at the upper tail of the distribution under the assumption of exogenous sector choice. When endogenous sorting is considered, the premium increases by roughly 40 percentage points (this is broadly in line with the correction that we find in 2004). While in his data the premium declines more or less monotonically from the low to the high end of the distribution (more or less as it happens in our 2004 wave), we do not find evidence of such a firm conclusion over time. Interestingly, the same difference between premia in Italy (the present paper) and in Germany (Melly, 2006) are found in Brunello and Dustmann (1997).

6 Conclusions and final remarks

In this paper we studied the public private sector pay gap in Italy in the period 1998–2008, using micro data from the Bank of Italy Survey of Household Income and Wealth. Compared to previous studies on Italy that control for relevant characteristics of employment, this paper provided an updated (up to 2008) analysis that traces the evolution of the wage premium over the last decade.

Unlike most of the existing literature that focuses on the average premium, we analyzed the entire distribution of (log) wage, by using conditional quantile regression techniques, and properly considered the possible endogeneity of the sector choice. Because the performance and the validity of IV estimators crucially depends on the exogeneity and relevance of instruments, we investigated several possible motivations invoked in the literature as key to enter the public sector. By exploiting a rich set of information provided in the 2004 wave of the Survey, we were able to discuss whether certain characteristics of employees, such as family background, risk aversion, degree of forward-lookingness and preferences for pro-social activities increase the probability that they will seek employment in the public sector. As a measure of risk aversion we considered the presence of sons and home ownership, like in the existing literature, as well as the answer to a specific question aimed at evaluating the attitude of the respondent in taking financial risks; to assess the degree of forward-lookingness or impatience we relied on direct measures of discounting; as an indicator of altruistic vocation we controlled for the engagement of the respondent into pro-social voluntary activities; finally, as for family background we accounted for parents' sector of occupation.

In fact, sample selection appears to play a key role. On the one hand, under the assumption of random sector choice quantile and mean regressions gave us a public sector premium averaging at about 14 per cent for women and less than or equal to 7 per cent for men. At bottom quantiles the premium is slightly higher for women than for men; then it decreases substantially along the conditional wage distribution for men, whereas it remains relatively constant for women, so that the gender gap increases towards higher quantiles. The premium is generally higher for white collars (while managers in the public sector seem to be at an earning disadvantage with respect to those in the private sector), and in Southern Italy. No clear trend emerges instead over time.

On the other hand, when we controlled for possible endogeneity of the sector choice we obtained substantially higher premia, suggesting that additional motives, other than monetary ones, may induce people to seek employment in the public sector and that ignoring them may result in an underestimation of the overall advantage. Such additional motivations appear to be particularly relevant above the median of the wage distribution, precisely where previous studies found lower premia or even penalties.

While documenting the existence of a public sector premium in Italy, this paper told little or nothing about its causes and consequences. Nonetheless, understanding whether the public-private pay gap simply depends on institutional features, or rather reflects the true preferences of the policy-makers, and what are its implications for the quality of public services as well as for educational choices and labor market developments is of crucial importance. Thus, we take the analysis in this paper as a necessary intermediate step and leave the other questions for future research.

References

Amemiya, T. (1982), "Two Stage Least Absolute Deviations Estimators", *Econometrica*, 50: 689–711.

Angrist J., Chernozhukov, V. and Fernndez-Val, I. (2006), "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure", *Econometrica*, vol. 74(2): 539–563.

Banca d'Italia, (2008) "Supplemento al bollettino statistico - I bilanci delle famiglie italiane nell' anno 2006", Banca d'Italia.

Bardasi, E. (1996), "I differenziali salariali tra i settori pubblico e privato: un'analisi microeconometrica", *Lavoro e relazioni industriali*,3, 3-51.

Bargain, O. and Melly, B. (2008), "Public Sector Pay Gap in France: New Evidence Using Panel Data", *IZA Discussion Papers* 3427, Institute for the Study of Labor (IZA).

Bellante, D. and Link, A. N. (1981), " Are public sector workers more risk averse than private sector workers?", *Industrial and Labor Relations Review*, 34: 408-412.

Bender, K. (2002) "The Central Government-Private Sector Wage Differential", *Journal of Economic Surveys*, 12: 177–220.

Breusch, T., Qian, H., Schmidt, P. and Wyhowski, D. (1999), "Redundancy of moment conditions", *Journal of Econometrics*, 91: 89–111.

Brunello, G. and C. Dustmann (1997), "Le retribuzioni nei settori pubblico e privato in Italia e in Germania: un paragone basato su dati microeconomici", in C. Dell'Aringa (edited by), *Rapporto ARAN sulle retribuzioni*, Collana ARAN, Franco Angeli.

Chernozhukov, V. and Hansen, C. (2005), "An IV Model of Quantile Treatment Effects", *Econometrica*, 73: 254–261.

Chernozhukov, V., Hansen, C. and Jansson M. (2007), "Inference approaches for instrumental variable quantile regression", *Economics Letters*, 95: 272–277.

Comi, S. P. Ghinetti and C. Lucifora (2002), "La distribuzione dei salari nei settori pubblico e privato: un'analisi disaggregata", in C. Dell'Aringa and C. Lucifora (edited by), *Dinamica occupazionale salariale*, Vita e Pensiero.

Dell'Aringa, C. (edited by), (1997), *Rapporto ARAN sulle retribuzioni - 1996*, Collana ARAN, FrancoAngeli.

Dell'Aringa, C. (2006), "Una regia nazionale per la PA", Lavoce.info.

Dell'Aringa, C. C. Lucifora and F. Origo (2005), "Public Sector Reforms and Wage Decentralisation: a First Look at Regional Public-Private Wage Differentials in Italy", mimeo.

Disney, R. A. Goodman, A. Gosling and C. Trinder (1998), "Public Pay in Britain in the 1990s", Commentary no. 72, London, Institute for Fiscal Studies.

Dustmann, C. and Van Soest A. (1998), "Public and private sector wages of male workers in Germany", *European Economic Review*, 42(8): 1417–1441.

Fogel, W. and Lewin, D. (1974), "Wage Determination in the Public Sector", *Industrial and Labor Relations Review*, 27: 410–431.

Forni, L. and R. Giordano (2003), "Employment in the Public Sector", CESifo, W. P. no. 1085.

Gregory, R. G. and J. Borland (1999), "Recent Developments in Public Sector Labor Markets", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics, Vol. III*, Amsterdam, The Netherlands, Elsevier Science Publishers.

Gunderson, M. (1978), "Public-Private Wage and Non-Wage Differentials in Canada: Some Calculations from Published Tabulations", in D. K. Foot (ed.) *Public Employment and Compensation in Canada: Myths and Realities*, 127–166). Ontario: Institute for Research on Public Policy.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50: 1029–1054.

Hausman, J. (1978), "Specification Tests in Econometrics", *Econometrica*, 46: 1251–1271.

Holmlund, B. (1993), "Wage Setting in Private and Public Sectors in a Model with Endogenous Government Behavior", *European Journal of Political Economy*, 9: 149–62.

Koenker, R. and Bassett, G. (1978), "Regression Quantiles", *Econometrica*, 46: 33–50.

Lucifora, C. (1999), "Rules vs. Bargaining: Pay Determination in the Italian Public Sector", in Elliott, R. C. Lucifora and D. Meurs (edited by), *Public Sector Pay Determination in the European Union*, Macmillan.

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

Melly, B. (2006), “Public and private sector wage distributions controlling for endogenous sector choice”, *Universität St. Gallen*, working paper.

Melly, B. (2005), “Public-private sector wage differentials in Germany: Evidence from quantile regression”, *Empirical Economics*, 30(2): 505–520.

Niskanen, W. A. (1975), “Bureaucrats and Politicians”, *Journal of Law and Economics*, 18: 617–643.

Perry, J. L. and Wise L. R. (1990), “The motivational bases of public service”, *Public Administration Review*, 50: 367-373

Portugal, P. and M. Centeno (2001), “Wages of Civil Servants”, in Banco de Portugal, *Economic Bullettin*, September.

Powell, J. (1983), “The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators”, *Econometrica*, 51:1569–1575.

Reder, M. (1975), “The theory of employment and wages in the public sector”, In D. Hamermesh (ed.), *Labor in the Public and Nonprofit Sectors*, 1–48). Princeton: Princeton University Press.

Stock J. and Yogo M. (2002), “Testing for Weak Instruments in Linear IV Regression”, *NBER Technical Working Papers*, 0284.

Table 1: The evolution of public sector over time. Relative frequencies

Area	1998	2000	2002	2004	2006	2008
			Women			
NW	0.370	0.354	0.365	0.385	0.349	0.328
NE	0.340	0.314	0.302	0.333	0.320	0.311
C	0.456	0.397	0.422	0.313	0.398	0.368
S	0.572	0.537	0.489	0.491	0.472	0.403
IS	0.674	0.547	0.601	0.515	0.599	0.460
ITALY	0.432	0.394	0.397	0.379	0.388	0.354
			Men			
NW	0.207	0.190	0.182	0.207	0.186	0.157
NE	0.178	0.184	0.180	0.159	0.127	0.136
C	0.357	0.272	0.318	0.239	0.250	0.231
S	0.397	0.359	0.331	0.324	0.352	0.271
IS	0.313	0.287	0.316	0.321	0.351	0.320
ITALY	0.283	0.251	0.255	0.239	0.237	0.209
			Total			
ITALY	0.344	0.309	0.314	0.297	0.302	0.273

Table 2: Descriptive statistics for continuous variables

Variable	Area	Mean	SD	p25	p50	p75	Mean	SD	p25	p50	p75
Women											
		Private					Public				
Ln(Wage)	NW	1.969	0.404	1.757	1.945	2.153	2.245	0.415	1.980	2.194	2.456
	NE	1.946	0.378	1.731	1.924	2.120	2.240	0.412	1.987	2.181	2.444
	C	1.912	0.427	1.683	1.889	2.120	2.269	0.405	2.011	2.227	2.504
	S	1.691	0.543	1.395	1.689	1.970	2.238	0.499	1.984	2.252	2.519
	IS	1.818	0.450	1.581	1.833	2.092	2.228	0.470	1.958	2.181	2.510
	ITALY	1.910	0.432	1.683	1.913	2.120	2.246	0.435	1.987	2.200	2.477
Age	NW	37.7	9.9	30.0	38.0	45.0	42.0	8.5	36.0	42.0	48.0
	NE	36.7	9.9	29.0	36.0	44.0	40.6	8.7	34.0	41.0	47.0
	C	37.9	9.8	30.0	38.0	44.0	43.4	9.4	37.0	42.0	51.0
	S	37.2	10.6	29.0	37.0	45.0	44.2	9.2	38.0	44.0	51.0
	IS	37.5	10.4	29.0	37.0	45.0	43.2	9.2	36.0	43.0	50.0
	ITALY	37.4	10.0	29.0	37.0	45.0	42.5	9.0	36.0	42.0	49.0
Men											
		Private					Public				
Ln(Wage)	NW	2.116	0.406	1.873	2.063	2.343	2.345	0.430	2.067	2.288	2.558
	NE	2.064	0.375	1.864	2.046	2.238	2.356	0.439	2.075	2.280	2.560
	C	2.071	0.436	1.833	2.056	2.270	2.265	0.438	2.005	2.200	2.506
	S	1.900	0.475	1.650	1.913	2.161	2.214	0.434	1.970	2.189	2.397
	IS	1.852	0.505	1.581	1.865	2.120	2.285	0.389	2.081	2.238	2.449
	ITALY	2.031	0.438	1.808	2.019	2.238	2.282	0.432	2.019	2.226	2.498
Age	NW	38.8	10.2	31.0	39.0	46.0	43.8	9.2	37.0	45.0	51.0
	NE	37.0	10.1	29.0	37.0	44.0	43.0	8.7	37.0	43.0	50.0
	C	39.6	10.8	31.0	39.0	48.0	43.9	10.0	37.0	44.0	51.0
	S	39.1	11.2	30.0	39.0	48.0	45.6	9.8	40.0	46.0	53.0
	IS	38.2	11.4	29.0	38.0	47.0	45.8	9.4	39.0	46.0	52.0
	ITALY	38.5	10.6	30.0	38.0	46.0	44.5	9.6	38.0	45.0	52.0

Table 3: Descriptive statistics for dichotomous variables

Area	Sector	Married	Education			Positions			Partime	Obs.
			Low	Int.	High	Blue	White	Man.		
					Women					
NW	Private	0.583	0.331	0.511	0.090	0.455	0.539	0.006	0.198	2515
	Public	0.714	0.139	0.529	0.310	0.170	0.790	0.041	0.085	1516
NE	Private	0.560	0.332	0.541	0.064	0.521	0.476	0.003	0.263	2549
	Public	0.669	0.118	0.580	0.285	0.143	0.833	0.024	0.161	1286
C	Private	0.544	0.374	0.488	0.080	0.538	0.456	0.006	0.235	1988
	Public	0.629	0.112	0.525	0.347	0.128	0.833	0.039	0.061	1221
S	Private	0.501	0.333	0.382	0.086	0.654	0.343	0.003	0.279	1112
	Public	0.668	0.089	0.533	0.338	0.062	0.918	0.020	0.061	1215
IS	Private	0.476	0.363	0.451	0.051	0.599	0.399	0.003	0.328	573
	Public	0.637	0.149	0.549	0.284	0.100	0.883	0.017	0.119	633
ITALY	Private	0.553	0.342	0.495	0.078	0.523	0.472	0.005	0.241	8737
	Public	0.670	0.120	0.542	0.315	0.128	0.842	0.031	0.095	5871
					Men					
NW	Private	0.599	0.377	0.442	0.100	0.608	0.365	0.027	0.016	3864
	Public	0.758	0.240	0.426	0.314	0.140	0.751	0.109	0.019	951
NE	Private	0.581	0.408	0.477	0.057	0.689	0.294	0.017	0.017	3644
	Public	0.707	0.194	0.514	0.274	0.142	0.771	0.087	0.036	778
C	Private	0.595	0.402	0.443	0.082	0.620	0.345	0.035	0.027	3034
	Public	0.740	0.282	0.473	0.219	0.172	0.764	0.064	0.019	1113
S	Private	0.664	0.443	0.351	0.032	0.775	0.217	0.008	0.059	3072
	Public	0.806	0.267	0.505	0.175	0.144	0.812	0.044	0.033	1539
IS	Private	0.653	0.503	0.257	0.027	0.770	0.220	0.010	0.076	1554
	Public	0.865	0.299	0.473	0.197	0.143	0.808	0.049	0.046	748
ITALY	Private	0.611	0.413	0.416	0.067	0.676	0.303	0.021	0.032	15168
	Public	0.776	0.259	0.478	0.231	0.149	0.782	0.069	0.029	5129

Table 4: Log of hourly wage by area, sector, position and gender

Area	Sector	Blue	White	Man.	Blue	White	Man.
		Women			Men		
NW	Private	1.838	2.039	2.391	1.987	2.190	2.647
	Public	1.957	2.277	2.560	2.053	2.305	2.675
NE	Private	1.836	2.045	2.298	1.969	2.195	2.549
	Public	2.002	2.267	2.405	2.116	2.331	2.629
C	Private	1.799	2.024	2.211	1.944	2.166	2.622
	Public	1.993	2.301	2.365	2.076	2.211	2.561
S	Private	1.598	1.821	2.178	1.831	2.058	2.503
	Public	1.805	2.265	2.298	1.976	2.222	2.444
IS	Private	1.770	1.872	2.148	1.772	2.001	2.591
	Public	1.939	2.253	2.381	2.061	2.277	2.616
ITALY	Private	1.790	2.011	2.305	1.920	2.154	2.605
	Public	1.960	2.275	2.419	2.047	2.260	2.580

Table 5: Quantile Regression for log of hourly wage. Coefficient of "Public" by Gender and Year

Women

	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
					1998					
Public	0.095**	0.130***	0.122***	0.135***	0.149***	0.148***	0.138***	0.151***	0.135***	0.144***
Obs.	2233	2233	2233	2233	2233	2233	2233	2233	2233	2233
					2000					
Public	0.117***	0.090***	0.104***	0.110***	0.112***	0.112***	0.104***	0.101***	0.080**	0.104***
Obs.	2420	2420	2420	2420	2420	2420	2420	2420	2420	2420
					2002					
Public	0.129***	0.128***	0.130***	0.117***	0.112***	0.107***	0.125***	0.111***	0.149**	0.152***
Obs.	2399	2399	2399	2399	2399	2399	2399	2399	2399	2399
					2004					
Public	0.159***	0.120***	0.096***	0.129***	0.130***	0.137***	0.122***	0.098***	0.076*	0.136***
Obs.	2417	2417	2417	2417	2417	2417	2417	2417	2417	2417
					2006					
Public	0.152***	0.131***	0.121***	0.110***	0.125***	0.116***	0.141***	0.131***	0.126***	0.143***
Obs.	2407	2407	2407	2407	2407	2407	2407	2407	2407	2407
					2008					
Public	0.282***	0.146***	0.135***	0.138***	0.155***	0.142***	0.163***	0.142***	0.159***	0.168***
Obs.	2502	2502	2502	2502	2502	2502	2502	2502	2502	2502

Men

	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
					1998					
Public	0.151***	0.103***	0.091***	0.061***	0.057***	0.044**	0.034***	0.051**	0.064**	0.070***
Obs.	3274	3274	3274	3274	3274	3274	3274	3274	3274	3274
					2000					
Public	0.077*	0.037*	0.024	0.006	-0.003	-0.004	-0.003	-0.028	-0.010	0.013
Obs.	3590	3590	3590	3590	3590	3590	3590	3590	3590	3590
					2002					
Public	0.074	0.047***	0.032	0.024	0.021	-0.002	-0.005	-0.008	0.006	0.007
Obs.	3328	3328	3328	3328	3328	3328	3328	3328	3328	3328
					2004					
Public	0.119*	0.083***	0.073***	0.068***	0.054***	0.022	0.019	0.010	-0.009	0.060***
Obs.	3366	3366	3366	3366	3366	3366	3366	3366	3366	3366
					2006					
Public	0.065*	0.062***	0.059***	0.055***	0.049**	0.038*	0.040	0.007	-0.013	0.030*
Obs.	3257	3257	3257	3257	3257	3257	3257	3257	3257	3257
					2008					
Public	0.166***	0.098***	0.095***	0.068***	0.058***	0.046**	0.021	0.013	-0.001	0.070***
Obs.	3225	3225	3225	3225	3225	3225	3225	3225	3225	3225

Note: Entries are the coefficients for dummy variable "Public".

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Table 6: Quantile Regression for log of hourly wage. Coefficient of “Public” by Gender and Area

Women

	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
1998										
NW	0.019	0.053	0.062 **	0.057 **	0.030	0.070	0.071	0.112 **	0.109 *	0.085 ***
NE	0.030	0.100 **	0.095 ***	0.092 ***	0.137 ***	0.108 **	0.074	0.113 **	0.160 ***	0.089 **
C	0.192 **	0.142 ***	0.147 ***	0.175 ***	0.176 ***	0.198 ***	0.212 ***	0.190 ***	0.140 ***	0.153 ***
S	0.587 ***	0.555 ***	0.472 ***	0.430 ***	0.415 ***	0.395 ***	0.379 ***	0.314 ***	0.313 ***	0.435 ***
Is	-0.264 ***	0.209 *	0.147 **	0.039	-0.002	0.085	0.000	-0.247 **	-0.284 ***	-0.056
Obs.	2233	2233	2233	2233	2233	2233	2233	2233	2233	2233
2000										
NW	0.069	0.041	0.082 **	0.085 ***	0.084 ***	0.056	0.053 **	0.050	-0.009	0.044
NE	0.115 **	0.007	0.048	0.070 ***	0.106 ***	0.131 ***	0.113 ***	0.113 ***	0.094 ***	0.065 *
C	0.207 ***	0.107 **	0.117 ***	0.110 ***	0.131 ***	0.135 ***	0.150 ***	0.177 ***	0.140 ***	0.163 ***
S	0.188 **	0.313 ***	0.315 ***	0.240 ***	0.242 ***	0.169 ***	0.217 ***	0.108 ***	0.065	0.177 ***
Is	0.429 ***	0.331 ***	0.300 ***	0.229 ***	0.123 **	0.155 **	0.098 ***	0.025	0.121 **	0.223 ***
Obs.	2420	2420	2420	2420	2420	2420	2420	2420	2420	2420
2002										
NW	0.015	0.069	0.101 **	0.093 **	0.074 **	0.069 **	0.062	0.082	0.099	0.122 ***
NE	0.118	0.085 *	0.075 *	0.044	0.069 **	0.076 ***	0.093 **	0.118 **	0.186 *	0.145 ***
C	0.099	0.100	0.108 **	0.068	0.111 ***	0.080 **	0.110 **	0.041	0.096	0.094 **
S	0.391 ***	0.364 ***	0.292 ***	0.304 ***	0.278 ***	0.288 ***	0.218 ***	0.242 ***	0.245 **	0.285 ***
Is	0.257 **	0.272 ***	0.271 ***	0.269 ***	0.177 ***	0.192 ***	0.198 ***	0.192 **	0.371	0.238 ***
Obs.	2399	2399	2399	2399	2399	2399	2399	2399	2399	2399
2004										
NW	0.313 ***	0.103 **	0.055 *	0.077 **	0.109 ***	0.105 ***	0.100 **	0.068	0.098	0.127 ***
NE	0.152 **	0.078 *	0.044	0.079 **	0.058 **	0.070 **	0.071	0.064	0.051	0.076 **
C	0.136 **	0.116 **	0.071 **	0.098 **	0.109 ***	0.139 ***	0.098 *	0.076	-0.047	0.095 **
S	0.511 ***	0.328 ***	0.338 ***	0.305 ***	0.293 ***	0.324 ***	0.275 ***	0.217 ***	0.284 ***	0.262 ***
Is	0.090	0.308 ***	0.208 ***	0.246 ***	0.301 ***	0.238 ***	0.204 **	0.294 ***	0.479 ***	0.270 ***
Obs.	2417	2417	2417	2417	2417	2417	2417	2417	2417	2417
2006										
NW	0.121 **	0.102 **	0.098 ***	0.105 ***	0.118 ***	0.077 ***	0.098 ***	0.077 **	0.055	0.112 ***
NE	0.133 **	0.065	0.054 ***	0.038	0.031	0.014	0.016	0.059	0.093 *	0.072 **
C	0.127 *	0.149 ***	0.175 ***	0.157 ***	0.193 ***	0.174 ***	0.224 ***	0.223 ***	0.193 ***	0.172 ***
S	0.573 ***	0.293 ***	0.248 ***	0.229 ***	0.262 ***	0.218 ***	0.273 ***	0.284 ***	0.097	0.257 ***
Is	0.435 ***	0.298 ***	0.210 ***	0.161 ***	0.211 **	0.238 ***	0.212 ***	0.127 *	0.180 **	0.259 ***
Obs.	2407	2407	2407	2407	2407	2407	2407	2407	2407	2407
2008										
NW	0.227 ***	0.069	0.096 **	0.062 *	0.080 **	0.079 **	0.084 ***	0.071 *	0.121 *	0.091 ***
NE	0.297 ***	0.135 ***	0.098 **	0.088 ***	0.103 ***	0.127 ***	0.109 ***	0.121 ***	0.219 ***	0.157 ***
C	0.496 ***	0.220 ***	0.195 ***	0.217 ***	0.217 ***	0.214 ***	0.199 ***	0.103 **	0.048	0.216 ***
S	0.517 ***	0.315 ***	0.356 ***	0.297 ***	0.280 ***	0.242 ***	0.259 ***	0.241 ***	0.267 ***	0.291 ***
Is	0.029	0.039	0.103	0.172 ***	0.167 ***	0.167 ***	0.151 ***	0.181 ***	0.192 **	0.146 ***
Obs.	2502	2502	2502	2502	2502	2502	2502	2502	2502	2502

Note: Entries are the coefficients for interactions between dummies of “Public” and of Areas, e.g. NW=Public × North West.

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Men

	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
					1998					
NW	0.161 *	0.017	0.031	0.018	-0.012	-0.003	-0.041	-0.057	-0.059	-0.004
NE	-0.068	0.021	0.007	0.051	0.036	0.013	-0.032	0.003	0.037	0.010
C	-0.002	0.044	0.043	0.022	0.007	0.022	0.065 **	0.102 ***	0.165 ***	0.018
S	0.375 ***	0.221 ***	0.164 ***	0.172 ***	0.118 ***	0.108 ***	0.065 **	0.044	0.066 *	0.148 ***
Is	0.573 ***	0.427 ***	0.280 ***	0.260 ***	0.211 ***	0.205 ***	0.139 ***	0.112 **	0.075	0.245 ***
Obs.	3274	3274	3274	3274	3274	3274	3274	3274	3274	3274
					2000					
NW	-0.022	-0.049	-0.019	-0.023	-0.027	-0.019	-0.021	-0.027	0.010	-0.045 *
NE	0.154 **	-0.014	-0.049	-0.026	0.004	0.015	0.017	0.077	0.086 **	0.018
C	0.082	0.032	0.013	-0.024	-0.028	-0.040	-0.028	-0.082 **	-0.053	-0.013
S	0.194 ***	0.112 ***	0.054	0.028	0.027	0.012	-0.006	-0.033	-0.052	0.017
Is	0.674 ***	0.304 ***	0.301 ***	0.143 ***	0.122 ***	0.116 ***	0.079	0.039	-0.027	0.207 ***
Obs.	3590	3590	3590	3590	3590	3590	3590	3590	3590	3590
					2002					
NW	0.090	-0.021	0.028	0.020	0.056 *	0.035	-0.003	-0.065 *	-0.143 ***	-0.028
NE	0.016	0.000	-0.025	-0.029	-0.023	-0.008	0.003	0.007	0.036	-0.000
C	-0.173 ***	0.061	0.036	0.031	0.030	0.052	0.086 *	0.100 ***	0.270 ***	0.068 **
S	-0.037	0.080 *	0.023	-0.009	-0.052 *	-0.100 ***	-0.073 *	-0.084 ***	-0.001	-0.034
Is	0.605 ***	0.142 **	0.173 ***	0.089 **	0.064	0.032	-0.138 **	-0.089 *	0.006	0.050
Obs.	3328	3328	3328	3328	3328	3328	3328	3328	3328	3328
					2004					
NW	0.083	0.050	0.032	0.059 **	0.041	0.021	0.008	-0.054 *	-0.050	0.038
NE	0.026	0.036	0.034	0.010	0.034	0.024	0.045	0.028	0.012	0.021
C	0.098	-0.035	-0.016	-0.005	0.012	0.058 *	0.040	0.012	-0.003	0.020
S	0.331 ***	0.283 ***	0.162 ***	0.126 ***	0.061	0.001	-0.025	0.010	-0.002	0.125 ***
Is	0.221 *	0.235 ***	0.141 ***	0.189 ***	0.135 **	0.052	0.081 **	0.053	-0.014	0.111 **
Obs.	3366	3366	3366	3366	3366	3366	3366	3366	3366	3366
					2006					
NW	0.100 *	0.042	0.027	0.032	0.055 *	0.020	0.004	-0.051	-0.050	-0.026
NE	-0.029	0.031	0.064 **	0.035	0.036	0.021	0.011	0.104 **	0.140 *	0.076 *
C	-0.089	-0.001	0.017	0.054 **	0.049	0.075 *	0.095	0.067	-0.042	0.003
S	0.204 ***	0.071 **	0.051 *	0.065 ***	0.052 *	0.058	0.049	0.007	0.015	0.060 **
Is	0.230 ***	0.259 ***	0.119 **	0.112 **	0.016	0.036	0.062	0.003	0.055	0.067 *
Obs.	3257	3257	3257	3257	3257	3257	3257	3257	3257	3257
					2008					
NW	0.147 **	0.045	0.062 **	0.043	0.035	-0.012	-0.062	-0.096 **	-0.033	-0.000
NE	0.094	0.034	0.084 ***	0.083 **	0.094 ***	0.075	0.129 ***	0.180 ***	0.342 ***	0.135 ***
C	-0.008	0.080 **	0.035	0.026	0.018	-0.021	0.008	-0.028	-0.016	0.030
S	0.212 ***	0.217 ***	0.152 ***	0.100 ***	0.062 **	0.039	0.028	-0.008	-0.094 *	0.087 ***
Is	0.442 ***	0.256 ***	0.163 ***	0.143 ***	0.121 ***	0.074	0.076 *	0.133 ***	0.064	0.141 ***
Obs.	3225	3225	3225	3225	3225	3225	3225	3225	3225	3225

Note: Entries are the coefficients for interactions between dummies of “Public” and of Areas, e.g. NW=Public × North West.

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Table 7: Quantile Regression for log of hourly wage. Coefficient of "Public" by Gender and Position

Women	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
1998										
Blue	0.081	0.085 *	0.091 **	0.055 *	0.083	0.078	0.072	0.045	-0.011	0.056
White	0.100 **	0.139 ***	0.138 ***	0.158 ***	0.178 ***	0.165 ***	0.176 ***	0.193 ***	0.169 ***	0.183 ***
Manager	-0.724 ***	-0.233	-0.022	-0.360 ***	-0.774 ***	-0.704 ***	-0.786 ***	-0.774 ***	-0.671 ***	-0.551 ***
Obs.	2233	2233	2233	2233	2233	2233	2233	2233	2233	2233
2000										
Blue	0.062	0.065	0.043	0.025	0.021	0.024	0.010	-0.026	0.038	0.010
White	0.145 ***	0.108 ***	0.128 ***	0.146 ***	0.132 ***	0.142 ***	0.135 ***	0.162 ***	0.078 **	0.129 ***
Manager	0.067	-0.069	0.061	0.054	-0.039	0.070	0.131	-0.108	-0.359 ***	0.002
Obs.	2420	2420	2420	2420	2420	2420	2420	2420	2420	2420
2002										
Blue	0.083	0.139 *	0.193 ***	0.150 ***	0.125 ***	0.113 ***	0.129 **	0.148 **	0.341 ***	0.184 ***
White	0.133 **	0.127 ***	0.122 ***	0.096 ***	0.100 ***	0.107 ***	0.123 ***	0.111 ***	0.136 *	0.145 ***
Manager	0.551 ***	0.615 ***	-0.169	0.157	0.084	-0.021	-0.345 **	0.011	0.088	0.091
Obs.	2399	2399	2399	2399	2399	2399	2399	2399	2399	2399
2004										
Blue	0.163 **	0.101 **	0.040	0.055 *	0.034	0.071 *	0.036	0.022	-0.043	0.066 *
White	0.160 **	0.148 ***	0.113 ***	0.155 ***	0.164 ***	0.167 ***	0.144 ***	0.134 ***	0.163 ***	0.157 ***
Manager	-0.128	-0.196	-0.278 *	0.005	0.088	0.040	0.188	0.347 ***	-0.086	-0.034
Obs.	2417	2417	2417	2417	2417	2417	2417	2417	2417	2417
2006										
Blue	0.182 ***	0.067 *	0.041	0.020	0.021	0.016	0.035	0.018	0.050	0.070 **
White	0.136 ***	0.177 ***	0.161 ***	0.152 ***	0.159 ***	0.168 ***	0.209 ***	0.177 ***	0.179 ***	0.170 ***
Manager	0.393 ***	0.340 ***	0.441 ***	-0.371 ***	-0.226 **	-0.155	-0.129 *	-0.067	0.003	-0.057
Obs.	2407	2407	2407	2407	2407	2407	2407	2407	2407	2407
2008										
Blue	0.306 ***	0.158 **	0.149 ***	0.125 ***	0.146 ***	0.126 ***	0.109 ***	0.103 **	0.067	0.144 ***
White	0.282 ***	0.146 ***	0.132 ***	0.148 ***	0.164 ***	0.158 ***	0.197 ***	0.161 ***	0.202 ***	0.193 ***
Manager	-0.472 ***	-0.550 ***	-0.453 ***	-0.417 ***	-0.377 ***	-1.054 ***	-0.977 ***	-0.903 ***	-0.908 ***	-0.661 ***
Obs.	2502	2502	2502	2502	2502	2502	2502	2502	2502	2502

Note: Entries are the coefficients for interactions between dummies of "Public" and of Position, e.g. Blue=Public × Blue Collar.

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Men

	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
					1998					
Blue	0.192 ***	0.137 ***	0.109 ***	0.091 ***	0.067	0.043	0.038	0.018	0.048	0.072 **
White	0.100 **	0.098 ***	0.080 ***	0.055 ***	0.068 ***	0.052 **	0.040 **	0.062 **	0.066 **	0.082 ***
Manager	0.378 ***	-0.108	-0.101 *	-0.160 ***	-0.206 ***	-0.211 ***	-0.303 ***	-0.241 ***	-0.262 ***	-0.119 *
Obs.	3274	3274	3274	3274	3274	3274	3274	3274	3274	3274
					2000					
Blue	0.177 ***	0.100 ***	0.090 **	0.054 *	0.061 *	0.051	0.044 *	0.020	0.035	0.090 ***
White	0.069	0.037	0.024	-0.001	-0.009	-0.015	-0.033 *	-0.067 ***	-0.020	0.005
Manager	-0.240 **	-0.184 **	-0.315 ***	-0.298 ***	-0.273 ***	-0.242 ***	-0.212 ***	-0.113	-0.013	-0.160 ***
Obs.	3590	3590	3590	3590	3590	3590	3590	3590	3590	3590
					2002					
Blue	0.213 ***	0.084 **	0.049	0.065 *	0.067 **	0.039	0.012	0.001	0.006	0.051
White	0.022	0.049 **	0.048 **	0.018	0.021	-0.004	0.001	0.010	0.036	0.014
Manager	-0.414 ***	-0.143 **	-0.166 **	-0.227 ***	-0.273 ***	-0.360 ***	-0.433 ***	-0.429 ***	-0.437 ***	-0.332 ***
Obs.	3328	3328	3328	3328	3328	3328	3328	3328	3328	3328
					2004					
Blue	0.048	0.139 **	0.133 ***	0.118 ***	0.131 ***	0.105 ***	0.101 **	0.062	0.148	0.136 ***
White	0.163 **	0.087 **	0.063 ***	0.055 **	0.053 **	0.013	-0.002	-0.014	-0.036	0.052 **
Manager	-0.089	-0.164	-0.243 ***	-0.226 ***	-0.141 *	-0.145 **	-0.069	-0.308 ***	-0.309 ***	-0.195 **
Obs.	3366	3366	3366	3366	3366	3366	3366	3366	3366	3366
					2006					
Blue	0.106 *	0.093 **	0.067 ***	0.059	0.038	0.046	0.023	-0.022	-0.028	0.047
White	0.036	0.043 *	0.054 ***	0.059 ***	0.055 **	0.047 *	0.068 **	0.027	0.004	0.029
Manager	0.170 **	0.190 ***	0.037	-0.088	-0.076	-0.012	-0.023	-0.122	-0.271 ***	-0.024
Obs.	3257	3257	3257	3257	3257	3257	3257	3257	3257	3257
					2008					
Blue	0.157 ***	0.068	0.069 *	0.041	0.018	0.049	0.012	0.035	-0.021	0.061 *
White	0.146 ***	0.105 ***	0.100 ***	0.084 ***	0.072 ***	0.051 **	0.022	0.003	0.019	0.081 ***
Manager	0.309 ***	0.089	0.124	0.079	0.009	-0.016	-0.062	-0.201 **	-0.468 ***	-0.009
Obs.	3225	3225	3225	3225	3225	3225	3225	3225	3225	3225

Note: Entries are the coefficients for interactions between dummies of “Public” and of Position, e.g. Blue=Public × Blue Collar.

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Table 8: Instrumental Relevance by Gender; Year 2004. Only coefficients attached to exclusion restrictions.

Instrument	F	H	FH	D	R	S	FHDS	FHDRS
				2004 – Women				
Father Pub.	0.111 ***		0.108 ***				0.058	0.003
Mother Pub.	-0.019		-0.019				-0.061	-0.169
No home		0.002	-0.015				-0.118 **	-0.005
Children		0.016	0.034 **				0.058	0.120 *
High Disc.				0.018			-0.093	-0.343 **
Lower Disc.				0.040			-0.104	-0.200 *
Risk Av.					-0.002			0.074
Pro Social						0.137 **	0.121	0.178
Obs.	1396	2417	1396	808	291	380	319	134
RMSE	0.422	0.420	0.422	0.426	0.438	0.437	0.448	0.441
R2	0.229	0.252	0.230	0.245	0.232	0.210	0.187	0.226
J test	0.079	0.054	0.932	0.000	.	.	4.382	14.402 **
Joint sig.	5.287 ***	0.784	3.677 ***	0.590	0.001	3.906 **	1.456	2.000 *
Redundancy	0.748	5.859 *	6.266	2.943	0.835	2.839 *		
				2004 – Men				
Father Pub.	0.132 ***		0.132 ***				0.149 ***	0.253 ***
Mother Pub.	0.049		0.049				0.146 **	0.178 *
No home		-0.001	-0.003				-0.014	-0.015
Children		0.014	-0.003				-0.014	0.047
High Disc.				0.038			-0.017	0.132 *
Lower Disc.				-0.025			-0.036	-0.047
Risk Av.					0.069 **			0.099 *
Pro Social						-0.016	-0.047	-0.114 *
Obs.	2355	3366	2355	1856	653	851	712	260
RMSE	0.365	0.366	0.365	0.392	0.396	0.373	0.376	0.378
R2	0.280	0.265	0.280	0.266	0.269	0.341	0.341	0.373
J test	1.272	1.322	2.153	0.725	.	.	14.157 ***	5.875
Joint sig.	19.682 ***	1.187	9.857 ***	3.815 **	4.408 **	0.178	3.296 ***	3.950 ***
Redundancy	20.131 ***	0.707	21.555 ***	1.081	3.987 **	1.372		

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

F: Family, H: House ownership&Children, D: Discount rate, R: Risk adv., S: Pro-Social.

Table 9: Instrumental Variable Quantile Regression for log of hourly wage. Coefficient of "Public" by Gender. Year 2004

2004										
	10th	20th	30th	40th	50th	60th	70th	80th	90th	Mean
Women										
F	0.18	0.04	0.22	0.14	0.13	0.27	0.23	0.30	0.83	0.78 **
F,H	0.90 *	0.18	0.34	0.24	0.18	0.30	0.44	0.38	0.43	0.79 **
FHDS	-0.36	0.12	0.07	0.11	0.29	0.34	0.18	0.23	0.36 **	0.38
Men										
F	0.44	0.46	0.48 ***	0.42 ***	0.58 **	0.34 *	0.40	0.97	-0.03	0.69 ***
F,H	0.44	0.46 *	0.48 ***	0.41 ***	0.59 **	0.38 *	0.40	0.87	-0.18	0.69 ***
FHDS	0.24	0.22	0.14	0.36 **	0.09	0.15	0.15	0.80 * **	0.08	0.46 **

Note: Entries are the coefficients for dummy variable "Public".

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Table 10: Instrumental Variable Mean Regression for log of hourly wage. Coefficient of "Public" by Gender and Year. Heading of rows is the set of instruments.

	1998	2000	2002	2004	2006	2008
			Women			
Fam.	0.270	0.475 **	1.067 **	0.775 **	0.037	0.491
Obs.	1245	1351	1361	1396	1328	1218
F,H	0.297	0.523 ***	0.524 ***	0.788 ***	0.229	-0.257
Obs.	1245	1351	1361	1396	1328	1218
			Men			
Fam.	0.712 ***	-0.196	-0.158	0.690 ***	0.010	0.077
Obs.	2892	2680	2365	2355	2335	2295
F,H	0.671 ***	-0.264 **	-0.127	0.686 ***	-0.028	0.641 ***
Obs.	2892	2680	2365	2355	2335	2295

Note: Entries are the coefficients for dummy variable "Public".

Note: *** is 1% sig. level; ** is 5% sig. level; * is 10% sig. level.

Table 11: Instrumental Variable Quantile Regression for log of hourly wage. Quantile Coefficient of "Public" by Gender and Year. Heading of rows is the set of instruments.

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Women									
1998									
F	0.57	0.24	0.22	0.51	0.40	0.48 **	0.35 **	0.21	0.02
F,H	0.57	0.12	0.22	0.51 *	0.39	0.49 ***	0.35 **	0.21	0.04
2000									
F	0.13	0.37	0.44 **	0.50 **	0.57 **	0.77 ***	0.78 **	0.63	0.79
F,H	0.14	0.53 ***	0.44 ***	0.45 **	0.72 ***	0.77 ***	0.83 ***	1.00 **	0.95 **
2002									
F	0.28	0.40 *	0.36 *	0.24	0.29	0.44 *	0.69 *	0.63 *	-0.14
F,H	0.92	0.12	0.19	0.31 *	0.37 **	0.40 **	0.69 **	0.82 **	0.85
2004									
F	0.18	0.04	0.22	0.14	0.13	0.27	0.23	0.30	0.83
F,H	0.90 *	0.18	0.34	0.24	0.18	0.30	0.44	0.38	0.43
2006									
F	0.28	0.15	0.21	-0.03	-0.14	-0.23	-0.22	-0.50	0.08
F,H	0.37	0.25	0.25	0.07	0.04	0.40 **	0.26 **	0.47 *	0.52
2008									
F	-0.03	-0.10	0.20	0.16	-0.23	0.33	1.00	0.31	-0.41
F,H	-1.00	-0.77	-0.33	-0.39	-0.25	0.25	0.23	0.13	0.37
Men									
1998									
F	0.50	0.37 ***	0.49 ***	0.52 **	0.56 **	0.64 ***	0.56 **	0.40 *	0.24 *
F,H	0.41 ***	0.31 ***	0.47 ***	0.49 ***	0.55 ***	0.62 ***	0.95	0.55	0.24 *
2000									
F	-0.05	-0.03	0.02	-0.05	-0.17	-0.22 *	-0.09	0.08	0.39
F,H	-0.95	-0.96	-0.17	-0.13	-0.18 *	-0.19	-0.10	-0.03	0.13
2002									
F	-0.20	-0.17	-0.23	-0.17	-0.17	-0.07	-0.01	0.13	-0.03
F,H	-0.48 ***	-0.34 *	-0.38 *	-0.38 *	-0.27	-0.15	0.10	0.23	0.25
2004									
F	0.44	0.46	0.48 ***	0.42 ***	0.58 **	0.34 *	0.40	0.97	-0.03
F,H	0.44	0.46 *	0.48 ***	0.41 ***	0.59 **	0.38 *	0.40	0.87	-0.18
2006									
F	-0.40	0.18	0.16	0.33	0.44	0.50 **	0.97	0.97	0.25
F,H	-0.07	0.18	0.16	0.38	0.42 **	0.48 **	0.70	0.70	0.23
2008									
F	0.13	0.15	0.16	0.20	0.22	0.28	0.40	0.35	-0.14
F,H	0.39 **	0.66 ***	0.63 ***	0.64 ***	1.00 *	1.00	0.89	0.95 *	0.68 **

Note: Entries are the coefficients for dummy variable "Public".

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.

Figure 1: The wage gap in National Account data

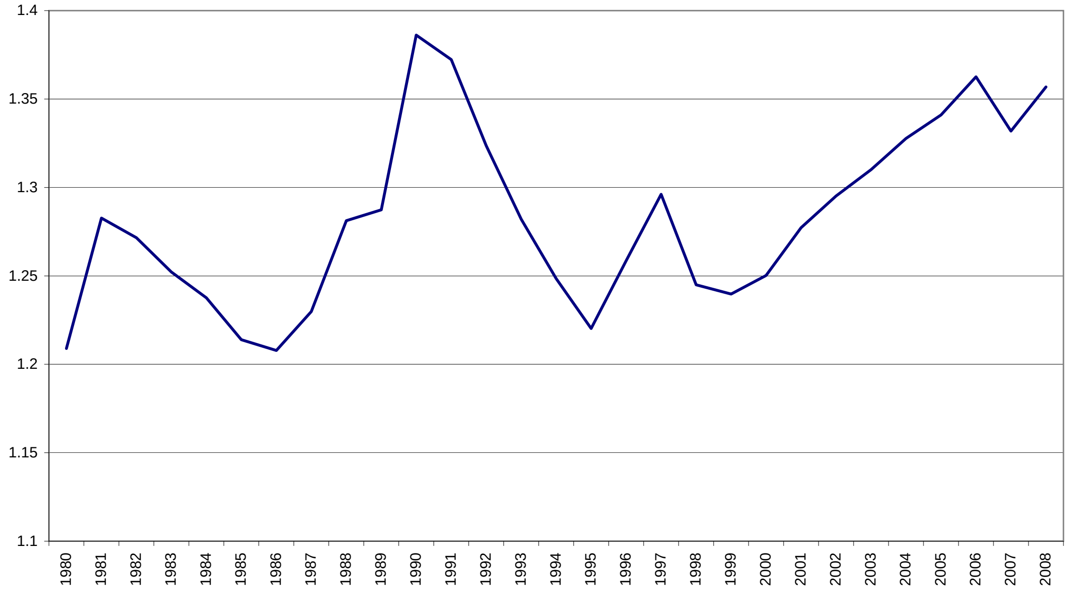

Figure 2: Distribution function for (log of) hourly wage, by Gender and Sector

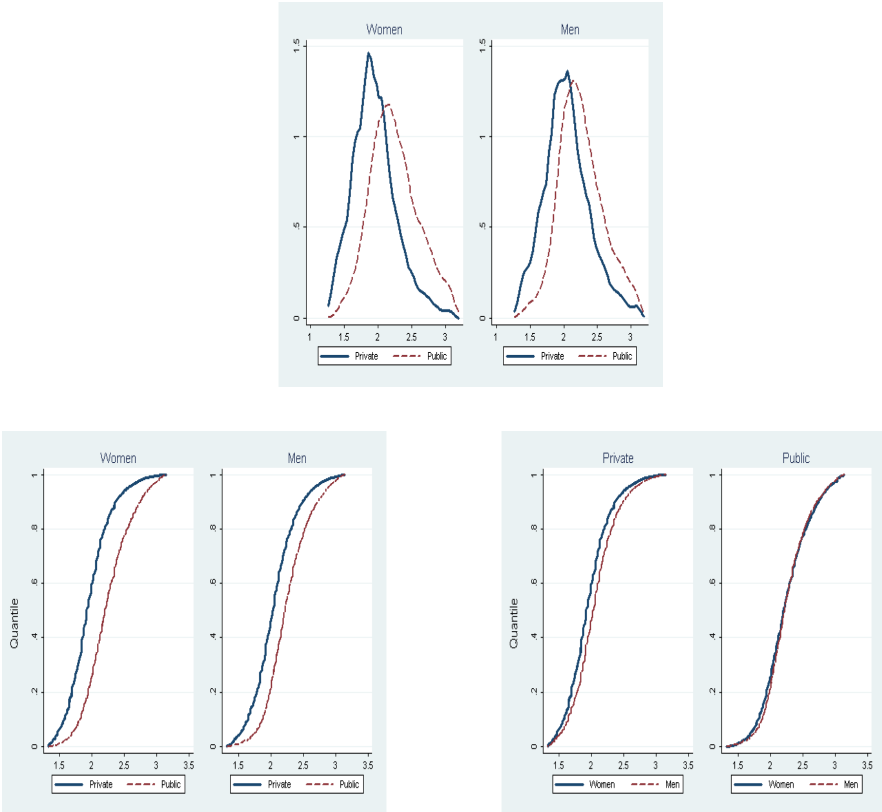

Figure 3: Coefficient of PUBLIC across quantiles

Figure 4: Mean coefficient of PUBLIC across quantiles, by area. Women and Men

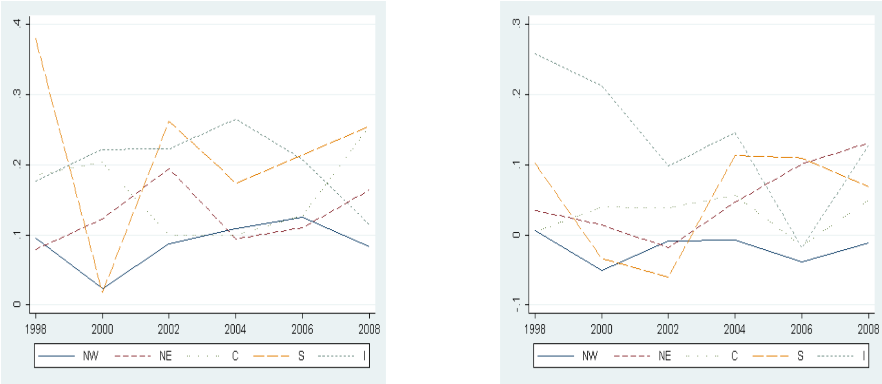

Figure 5: Mean coefficient of PUBLIC across quantiles, by position. Women and Men

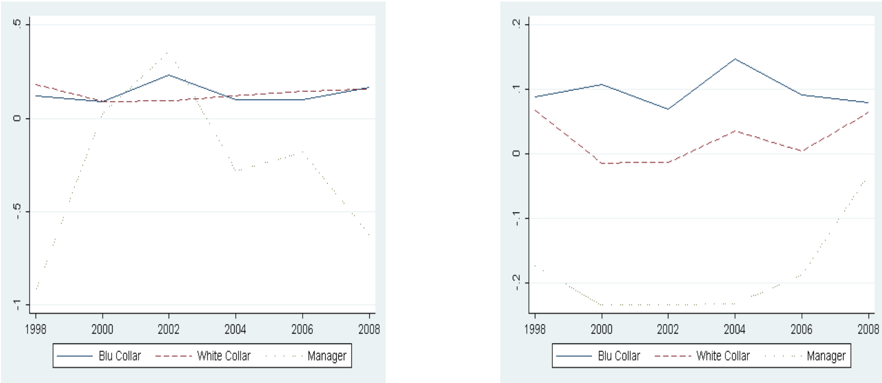

Figure 6: The bias averaged by year, by quantiles and gender.