UNIVERSIDAD TORCUATO DI TELLA

MASTER IN ECONOMETRICS

Evaluating the Public - Private Wage Differential in Chile

Raul Jimenez Mori

May 2012

Evaluating the Public - Private Wage Differential in Chile

Raul Jimenez Mori

Abstract

This paper estimates and analyzes the wage gap between the public and private sector employees in Chile for the years 2000 and 2009. We first concentrate on conditional mean estimation and on a control function method. Then, the Machado and Mata (2005) decomposition method is used to discriminate the differences in earnings distribution between both sectors in each year. In general our results tend to suggest an heterogeneous gap along the wage distribution and by group of educational attainment. They also tend to indicate an alignment of the public sector with the market in the sense of reducing the wage premium for the low skilled worker and increasing the wage to those with higher qualifications.

May 2012

Key words: Public-private wage differential, quantile regression, wage gap decomposition, endogenous selection, counterfactual decomposition, Chile.

1 Introduction

As an efficient and competitive public sector has been an area of increasing concern, the public compensation system has been a particular subject of political attention and empirical research. Over the last 20 years, Latin America has undertake several public reforms in order to improve the efficiency and transparency of the public compensation system. Chile is one of the most progressive experiences, which has implemented reforms in public administration since 1990 and is recognized as one of the most professional and capable public administration in the region.[1]

In this context, the objectives of this paper are the following: First, we estimate the average wage gap between public and private workers in Chile. Then, we are interested in estimate and decompose such gap along the entire wage distribution. Finally, for each sector (public and private), we decompose the change in distribution over a period of time in several factors contributing to those changes. Since the professionalization of the public administration has been an area of particular interest, we also perform the previous analysis by educational attainments.

The estimations are performed for 2000 and 2009. For the mean conditional estimations, we relay on ordinary least square (OLS) and selection bias model (SBM). The estimations and decomposition by quantiles are based on the methodology proposed by Machado and Mata (2005). They combine quantile regression and a bootstrap approach to stochastically simulate the counterfactual wage distribution, which is then used to estimate the wage differential along the distribution.

On the last two decades, the literature on this topic for Latin America has received the following main contributions. Panizza (2001) analyzes the differences between the public and private wages in 17 Latin American countries (LA) over the 80s and 90s. He found not large differences in male workers but a significant public sector premium for females. His results also suggest that the premium are often higher for those with low education. Panizza & Qiang (2005), by using a cross section data set for 13 LA countries in the mid-nineties, focus on the gender wage gap by sector. They found a public sector premium in five countries, a public sector penalty in three countries, and not statistically significant premium in the other five countries. By gender, in general,

[1] Echevarria and Cortazar (2007) offers a detail description of the public sector reforms in Brazil, Chile and Peru. For the case of Argentina, Iacoviello and Tommasi (2002), discuss the civil service dysfunctions and the main factors limiting the development of the civil service administration.

they found that women are paid less than men in both sector, but the gender gap is usually larger in the private sector.

More recently, Coppola and Calvo-Gonzalez (2011) analyze a cross section 2009 survey in Peru finding not significant pay differential between public and private workers. However, the authors found that those with the highest level of education receive larger wages in the formal private sector. They also found that even when female workers earn a lower salary, this gender penalty is less significant in the public sector. Mizala, Romaguera & Gallegos (2011) analyze seven LA countries in 1999 and 2007 finding that, on average, public sector workers earn more than their private counterparts, and this differential increased over that period. Besides, by analyzing the wage distribution, they observe that public wage gap tend to decrease as it move to the higher percentiles. They also observe that over time, the wage gap tend to favor the higher percentiles. However, the most qualified public sector workers still face a wage penalty.[2]

The above contributions usually perform OLS estimations and control by the selection bias through the Heckman methodology (1979). These estimations are also performed separately by sector, gender and education level. The Oaxaca-Blinder decomposition is used to split the wage gap between the explained variables and unexplained factors. Panizza (2001) and Panizza & Quiang (2005) also perform simulations that assume different levels of correlation between wage and the selection equation. They found that OLS estimations are robust to a large range of values of selection bias. Taking advantage of matching techniques, Mizala et al (2011) estimate the public-private wage gap arguing that these methods do not require any estimation of earnings equations and hence no validity-out-of-the-support assumptions are needed. Furthermore, this approach allows them to estimate not only the average wage gap but also its distribution.

Other contributions that specifically take distributional approaches are; Hyder and Reilly (2005) examine the magnitude of public premium in Pakistan using quantile regression analysis, finding that the premium declines monotonically as it move to the right of the conditional wage distribution. Cai and Liu (2010) also employ quantile regressions to examine whether the sectoral wage effect varies along the wage distribution in Australia. They find that the wage effect varies mainly for males, showing a premium for the bottom half and a penalty for the top half of the distribution. For females, public sector wage premiums are relatively stable for almost the entire distri-

[2] Mizala et al (2011) offers a concise and complete description of the literature on (public-private) wage differential.

bution. For the case of Germany, Melly (2005a) employs the counterfactual quantile decomposition, proposed by Machado and Mata (2005), to measure and decompose the differences public-private wage gap. As in the previous cases, he found that public sector wage premium is highest at the lower end of the distribution and then decreases monotonically as it move up the wage distribution.

In this context, the main contribution of the paper is to implement the counterfactual quantile distribution to evaluate the wage gap along the entire distribution as well as its evolution over time. At our concern, this is the first attend to apply this method to the public-private wage differential in a Latin American country.

The paper is organized as follows: Section 2 describes the data set used in the paper. Section 3 presents a general empirical model and the approaches to address the estimations. Section 4 explores the public sector wage gap at the mean and along the wage distribution. Section 5 present the final remarks.

2 Data sets:

We employ the National Socioeconomic Survey; CASEN[3] for 2009 and 2000. This survey was designed to make inference at country and regional level, both in rural and urban areas, at 95% of statistical confidence. For 2009, the cross section data set offers a sample of 246,924 individuals and 71,460 households. For 2000, the survey contains 252,784 individuals and 48,632 household. The design of both surveys relies on stratified, poli-phase, sampling by conglomerates. The stratification was carry at geographical level to include urban and rural cities. It is important to note that these surveys exclude those cities with less than 40 thousand habitants. After that, there were three phases of random selection: 1) selection of cities , 2) selection of conglomerates, and 3) selection of households.

We include urban and rural workers, between 15 and 65 years old. We considered important to include workers living in rural areas since around 1/5 of public sector employees belong to such areas (for 2009). The analysis is focused on monthly labor earnings from the main job. We work with the broadest definition of public sector, which includes all individuals that declare to work in the public sector. However, we do not include those that work in public firms. In this sense, our characterization

[3]Encuesta de Caracterizacion Socioeconomica Nacional

is similar to the one used by Panizza (2005) and Mizala (2010). Our final sample include only wage earners and exclude the informal sector and self-employees, that is, the estimation are performed in order to compare the public sector with the formal private sector. Even when self-employment and informal worker are both important activities, including those dimensions add another sources of selection biases. For this reason, our results must be interpreted conditional on the selected sample.

Our final sample has around 14% of employees that belong to the public sector. They are mainly concentrate in the area of social services which consists of six sub-sectors: i) public administration and defense; ii) Education; iii) Social and Health Services; iv) Other Community Services; v) Board of Directors Buildings and Condominiums; and vi) International organizations. By profession, the public sector employees are mainly concentrate in highly qualify occupations, while private sector employees are concentrate in non-skill workers[4](see descriptive statistics in tables A).

Tables A in appendix describes the variables we use by sector (public/private). Since the dependent variable, the wage (w), is very right-skewed, we take the natural logarithmic of monthly labor income from the main work, in current American dollars. This helps fitting the variable into a model, making the dependent variable more close to a normal distribution. In the table A-I, we can see that the mean get closer to the median after the transformation. The graph A presents the kernel density after the transformation; showing the wage to behave more as a normal distribution. It is observed that the wage dispersion is higher in the public sector. Over time, such dispersion tend to reduce similarly in both sectors. The descriptive statistics also shows that the workers in the public sector earn more than the private ones at different percentiles.

Our control variables (X) include: age, education, experience, gender, ethnicity, marital status, area of living, economic sector of employment and profession. On average, public worker have 2 year more of schooling, around 5 year more of tenure and they are 4 years older. More than the 50% of worker in the public sector are females, almost the double than in the private sector.

[4]Both; occupation sector and profession are according to the CASEN classification.

3 Methodology

In order to estimate the wage gap we employ three strategies: first we estimate the conditional mean by OLS. Then we correct for selection bias by applying a control function method. Finally, to analyze the drivers contributing to changes in the wage distribution, we apply quantile regression counterfactual decomposition.

For the first two approaches, the setup of our problem can be specified as a sample selection model. By assuming linearity we can express the general model as follows:

$S = Z\delta + \varepsilon^s$ .... (a)

$w^{pu} = X^{pu}\beta^{pu} + \varepsilon^{pu}$.... (b)

$w^{pr} = X^{pr}\beta^{pr} + \varepsilon^{pr}$.... (c)

Where, $S = 1$ for public employees and $S = 0$ for private workers. Z represents those variables influencing the employment sector decision. X are the covariates as explained above. ε represent the error terms. In our case the values of the wage (W) is always observed.

$$S = \begin{cases} 1, S* > 0 \\ 0, S* < 0 \end{cases} \quad ...(a.1)$$

$$W = \begin{cases} w^{pu}, S = 1 \\ w^{pr}, S = 0 \end{cases}$$

If we would not be interested in particular specifications for each sector, the previous system can be reduced to:

$S = Z\delta + \varepsilon^s$

$W = X\beta + \alpha S + \varepsilon$ ... (d)

As explain Cameron et al (2010), under this specification only α varies across the two possible outcomes providing a measure of the gap for belong to the public sector.

i) Conditional Average

OLS estimations would return $E(\alpha|X, ED, S = 1)$. This estimation would be unbiased in absence of selection problem, that is assuming that S is orthogonal to ε. In other terms, assuming that on average there is not selection on unobservables. Support for this strategy is found in Panizza et al (2001, 2005), who perform a sensitivity analysis with different degrees of selection bias, showing that the conditional mean returns unbiased estimators of the wage gap. Another potential problem under this approach is the potential lost of efficiency due to non-constant variance, however, we sort this by estimating the heroskedasticity-robust estimator of the variance-covariance matrix of the estimator.

Since we are interested in estimate the gap by educational attainment, we take the approach of perform the estimation in the sub-sample of each education categories. In this way, we avoid to deal with the problem of endogeneity of the education variable.

ii) Selection Bias Correction Method

The control function approach allow us to take into account the possibility of self-selection. This refers to the probability that unobserved individual characteristics could affect both: the choice of the employment sector and the individual earns in the chosen sector. Although we are considering only the formal private and public dependents[5], it could be the case, for example, that a person choose to work in a sector where his/her abilities allowing to earn more. If this happens, the error terms (ε^S,ε) will be correlated, and the OLS estimation of α would be bias[6]. The control function approach add structure to the problem by modeling the nature of the selection rule and taken it into consideration in the estimation of the main equation. This approach follows Heckman (1979) by treating the endogeneity in S as an omitted variable problem.

It is a binary endogenous regressor model that specify the "decision" of work for the public sector as the outcome of an unobserved latent variable, $S*$ $(a.1)$. It is assumed that $S*$ is a linear function of the exogenous covariates Z and a random component ε^s (see $eq.\ a$). Then, this approach is based on the assumption that all relevant information is contained in ε^S; that is $(\varepsilon, \alpha)\_|\_(S, Z)|\varepsilon^S$. That is, once we were able to control for ε^S, S would be exogenous in the main equation. This is the role of the control function.

From equation (a) we can estimate the control functions for the each sector ($w^{pu} = X^{pu}\beta^{pu} + \varepsilon^{pu}$,$w^{pr} = X^{pr}\beta^{pr} + \varepsilon^{pr}$). This control function refer to the conditional ex-

[5] Since we are not considering entrepreneurs, self employees and informal sector; we do not deal with the biases arising of all this elections.

[6] Our first stage estimation indicates that both error terms are correlated suggesting a selection bias problem

pectation of ε given S and Z, and are equivalent to the assignment propensity. These control function take the form of:

$P(d=1|Z) = E[\varepsilon|s=1,Z] = \rho\lambda^{pu} = \rho\frac{\phi(Z^{pu}\delta^{pu})}{1-\Phi(Z^{pu}\delta^{pu})}$

$P(d=0|Z) = E[\varepsilon|s=0,Z] = \rho\lambda^{pr} = \rho\frac{\phi(Z^{pr}\delta^{pr})}{1-\Phi(Z^{pr}\delta^{pr})}$

Where $\rho = \sigma_{\varepsilon} corr(\varepsilon, \varepsilon^{s})$. σ_{ε} is the standard error of ε, while $\sigma_{\varepsilon^{s}}$ is standardized to 1. ϕ refers to the probability density function and Φ to the cumulative density function. That is, we are assuming joint normality of the errors terms $(\varepsilon^{S}, \varepsilon)$.

The other assumption refers that Z is independent of (ε, α) given ε^{S}. However, this is a strong assumption; often it is only required a weaker conditional mean restriction. That is, the assumption that Z is mean independent of (ε, α) given ε^{S}. Given that we are not considering the endogeneity of education this last assumption make more plausible the consistency of our estimates[7].

Then, we add these terms to the wage equation in order to control for selection bias on observables and unobservables:

$W = X\beta + \alpha S + \lambda(Z\delta) + \varepsilon \; ...(d')$

Applying OLS to this last equation allow us to estimate $E(\alpha|X, Z)$. The main problem with this approach relate to the greater chance of mis-specifications (Cameron et al, 2010 and 2009; Minsky, 1989). This is due to the difficulty for finding a (set) of variable(s) that affect the probability of obtaining a public sector job but do not affect the wage. Without this exclusion restriction, the above system would only be identified by the non-linearity of the selection equation. At respect, Manski (1989) highlight that since there are no clear theoretical reasons on why one should use a particular functional form, such assumption could lead to greater bias in the estimated coefficients. As Coppola et al (2011), we use the variable population size of the living area to satisfy the exclusion restriction[8]. Our first stage estimations return a statistical and significant coefficient for this variable, suggesting it is appropriate to use such variable in the selection equation. However, our main concern is that, as stress by Cameron et al (2009), mis-specifications could persists; presenting heteroskedastic errors that lead to inconsistent estimations[9].

[7]It could be argued for example that ability is in ε but not in ε^{S} . As education is one of the variables in Z , it would violate the strong exogeneity assumption. It would be a little more credible the mean independent assumption.

[8]We use a dummy indicating living area with more than 100 Townsend population.

[9]In fact, a graphical inspection of the main equation error suggest that we are in presence of

iii) Counterfactual decomposition

This section follows the method proposed by Machado & Mata (M-M, 2005) to decompose the change in the wage distribution over a period of time in several factors contributing to those changes[10]. As it was noted previously, by using mean estimations, we have been restricting the effects of the covariates to operate in the form of a simple location shift. The approach suggested by M-M not only allow to study the effects of the covariates on the whole conditional distribution of the dependent variable; but also simulate a counterfactual wage distribution to perform the comparisons. This allows to study the heterogeneous effects of belong to the public sector. Even more, this method is based in the estimation of the marginal density function of wages in a given year, or sector, implied by counterfactual distribution of some or all the observed attributed.

We implement this decomposition in order to estimate and decompose the public-private wage gap in 2000 ($t = 0$) and 2009 ($t = 1$). Since we are also interested in the evolution of this gap over time, we also estimate, for each sector, the wage density in 2009, corresponding to the 2000 distribution of one or all covariates.

This method starts by establishing the conditional quantile of the wage in a given year. By assuming linearity of the quantile regression model, we have the following expression:

$Q_\theta(w_i|X_i) = X_i\beta_\theta,\ u\epsilon(0,1)\ ...(e)$

Where $\beta(\theta)$ is the vector of coefficients by quantile. We refer to a specific worker by i. θ specifies of the quantile distribution. The estimation procedure is the following[11]:

1. Generate a random sample of size m from a uniform distribution: $U[0,1]$: $[u_1, ..., u_n]$.
2. For each sector and year we estimate m different quantile regression coefficients: $\hat{\beta}^{s,t}_{u_i}$, $\hat{\beta}^{s,t}_{u_i}$, $i = 1, ...m$, $s = (pu, pr)$, $t = (0, 1)$
3. Generate, for each sector, a random sample of size m with replacement from the covariates of X, denoted by $\{\tilde{X_i}^{s,t}\}_{i=1}^m$ and $\{\tilde{X_i}^{s,t}\}_{i=1}^m$.
4. $\{\tilde{w_i}^{s,t} = \tilde{x_i}^{s,t}\hat{\beta}^{s,t}_{u_i}\}_{i=1}^m$are random sample of size m from the marginal wage distribution of w consistent with the linear model defined by (e).

non-constant variance.

[10] An excellent summary is also found in Melly (2005a)

[11] Here we follow closely to Melly (2005a)

5. Estimate the counter-factual density based on a random sample of the distribution. For the case of the decomposition between sectors in t, the counterfactual density take the form of $\{\tilde{w}_i^{cf,t} = \tilde{x}_i^{pu,t}\hat{\beta}_{u_i}^{pr,t}\}_{i=1}^m$. This is a random sample from the wage distribution that would have prevailed in the private sector if all covariates had been distributed as in the public sector. For the case the decomposition between years in a given sector, the counterfactual density take the form of $\{\tilde{w}_i^{s,cf} = \tilde{x}^{s,00}\hat{\beta}_{u_i}^{s,09}\}_{i=1}^m$. This represent the distribution that would have had the public (private) sector if their covariates had been distributed as in 2000.

Note that the counterfactual nature of the exercise requires the estimation of the wage distribution conditional on the variables of interest in a given year/sector. The changes of interest can be decompose into the contribution of the coefficients and the contribution of the covariates. For a given year, such decomposition could be expressed as:

$$Q_\theta(w^{pu,t}) - Q_\theta(w^{pr,t}) = [Q_\theta(\tilde{w}^{pu,t}) - Q_\theta(\tilde{w}^{cf,t})] + [Q_\theta(\tilde{w}^{cf,t}) - Q_\theta(\tilde{w}^{pr,t})] + residual$$

The above expression represent the total change in the wage in quantile θ (left side), which is decomposed into the contribution of coefficients (first component of right side) and the contribution of characteristics (second component of right side). That is, the first component of the right side represents the rent of belong to a specific sector, in this case public.

Equivalently, for a given sector (public, private), the contribution of the coefficient and covariates over time can be expressed as:

$$Q_\theta(w^{s,09}) - Q_\theta(w^{s,00}) = [Q_\theta(\tilde{w}^{s,cf}) - Q_\theta(w^{s,00})] + [Q_\theta(\tilde{w}^{s,09}) - Q_\theta(\tilde{w}^{s,cf})] + residual$$

In this case the first component of the right side represents the change in the rent over time to those characteristics.

In the same way, we may measure the contribution of an individual covariate by looking at indicators such as:

$$[Q_\theta(\tilde{w}^{s,09}) - Q_\theta(\tilde{w}^{s,cf})]$$

This procedure described relies on the probability integral transformation theorem. Thus, if $\mu_1, \mu_2, ...\mu_m$ are drawn from a uniform $(0,1)$ distribution, the corresponding m estimates of the conditional distribution quantiles of wages at X, $\{X'\hat{\beta}(\mu_i)\}_{i=1}^m$, constitute a random sample from the estimated conditional distribution of wages given X. That is, the estimated density function has the same distribution as the empirical density function. Then, the consistency of the estimator is based in that, under regularity condition, the estimated conditional quantile function is a strongly consistent estimator of the population quantile function, uniformly in μ on a compact interval $(0, 1)$. That is $sup_{\mu\epsilon[\xi,1-\xi]}|X'\hat{\beta}(\mu) - X'\beta(\mu)| \to 0$ for some $\xi > 0$ (M-M 2005; Chernozhukov et al, 2009; Melly 2005a,b) .

It is important to emphasize that M-M estimate the asymptotic distribution by bootstrapping the generated sample. In this sense, they do not present a formal proof of the consistency of this estimator. However, the validity of the bootstrap is shown by Chernozhukov, Fernandez-Val and Melly (2009). They derive a functional limit theory and offer inference procedures for a range of estimators, including the proposed by M-M. Even more, they show that if the estimators of the conditional and marginal distributions satisfy a functional central limit theorem, then the estimators of the counterfactual distributions and quantiles also obey a functional central limit theorem.

However, it is important to note that this approach does not account for self selection bias. That is, here we have the probability that workers choose the employer based on characteristic not explicit in our specification.

4 Empirical Results

Wage gap estimation at the mean:

For the year 2009, starting from a simple average estimation, it seems to exist a premium of about US$165 in favor of public workers. As we noted in tables A, this premium is present along different percentiles and it is excepted since the public servants have on average more years of schooling, age and tenure. At the mean, figure A2 and table A-III (profession) support the intuition of a premia for public sector as it has a higher concentration of employees with more educational attainment.

However, after controlling for the set of covariates described in section 2, it reduces the gap to a slightly positively, not statistically significant, public sector wage premium of US$6. As mentioned in the previous section this estimation could be bias for self-selection of the employees into the public sector. This is why we adopt a control function approach, which turns the gap into a significant penalty of US$43 for the public workers. The contrast between the OLS and the control function estimates suggest that the omitted variable could be positively correlated with S. [12]

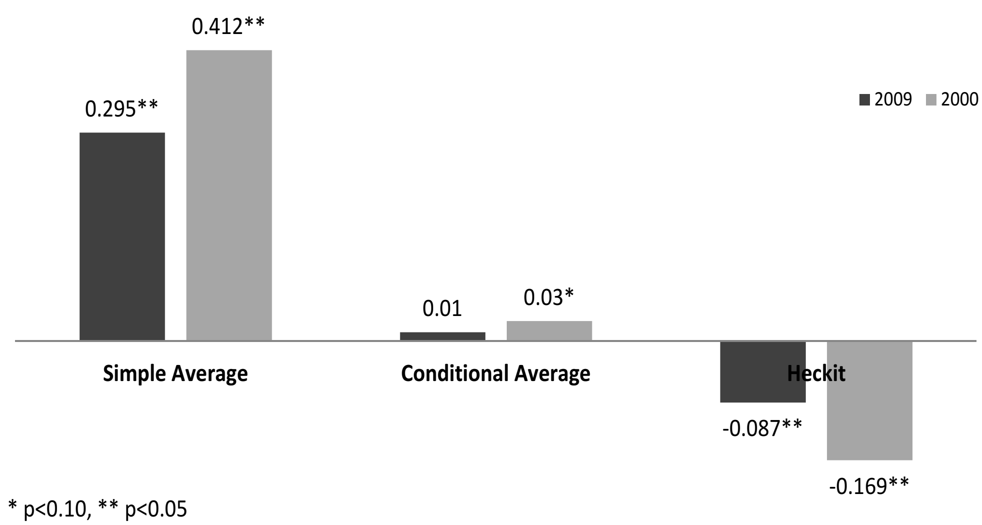
[12]However, given that the error of the 2SLS estimations seems to behave heteroskedastically, the

That is, as a percentage of the average wage, our conditional mean models return a gap between 1% and -8% for employees in the public sector. Other coefficients for the covariates seems to be according with the literature. For example; higher levels of education returns higher salaries. Gender and ethnicity are negative correlated with wage. Those who live in urban areas tend to earn more. These estimations behave robustly across different specification and samples[13]. See figure 1 and table B.

Figure 1 also compares the estimated gap for 2000 and 2009; showing a reduction in the wage differential between both sector. In general, the gap estimation tends to reduce both in magnitude and in statistical significance. For example, the OLS estimation was significant at 5% in 2000, but not significant in 2009. The control function approach was significant at 1% in 2000, but significant at 5% in 2009.

Figure 1: Wage Gap Estimates

(Coefficient from the regression; gap in Ln of montly income)



Gap by education attainment:

Now we estimate the wage gap by educational attainment. This allow us to evaluate if the gap behave heterogeneously between employees with different qualifications. As

control function estimations should be taken carefully because of the inconsistency of the result, as mentioned in the section 3.

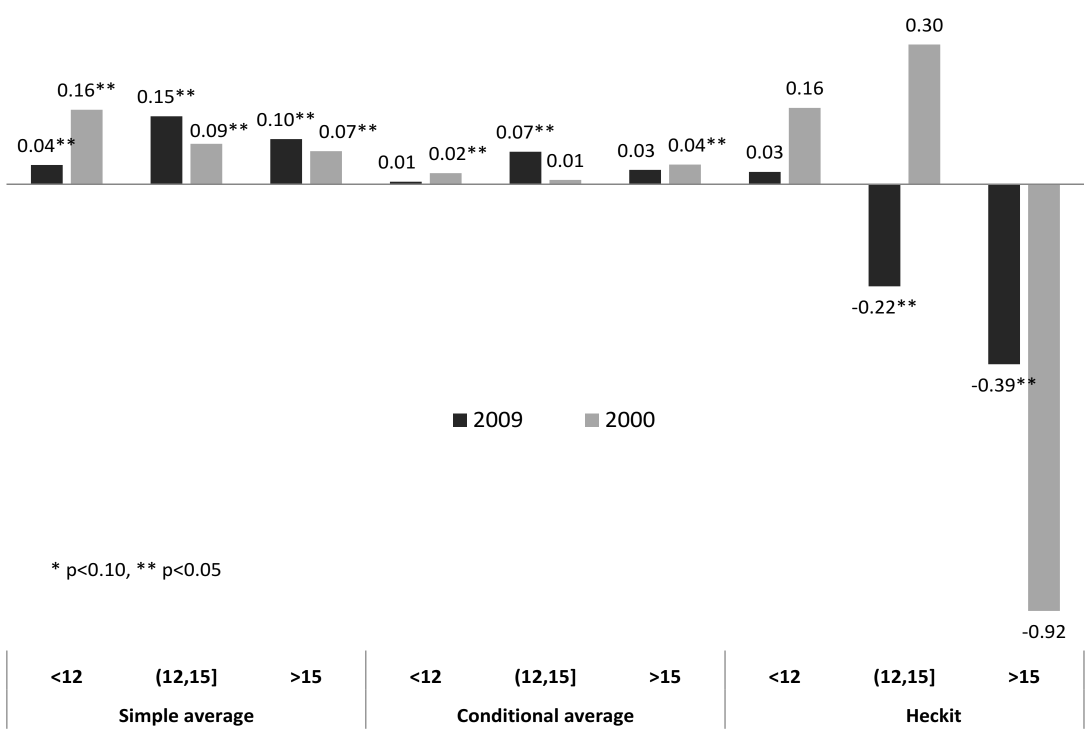
[13]We found similar estimations even after including the 1% superior and inferior of the log-wage distribution. We also found similar estimations after excluding the three provinces where the participation of public employees was greater than the 40%.

this dimension add another source of bias, that of education endogeneity, we sort such difficult by performing the estimations for each level of schooling. Figure 2 and table C show the estimations for years 2000 and 2009.

For 2009, the OLS estimations shows statistically significant estimations only for those with $(12, 15]$ years of schooling. This segment has an average differential of US$40 more for the public employees. However, after correcting by selection bias, we obtain statistical significant penalties for employees with higher educational attainment in the public sector. That is, those between $(12, 15]$ years of schooling earn on average US$120 less in the public sector. For those with university degree (> 15) the penalty reach US$330.

It is important to note that those gap are less pronounced that in 2000. Figure 2 presents the estimation for each education level in each year, showing that the wage gap between public and sector employees tend to reduce in the last 10 years.

Figure 2: Wage Gap by Educational Attainment
(Coefficient from the regression; gap in Ln of montly income)



In general, these findings are in line with the literature; Panizza (2002) and Mizala et al (2011) show that workers with low educational levels tend to have some wage premium in the public sector, while at the highest educational level it became into a wage penalty. However, the fact that the wage penalty tend to reduce (in the last 10 years) for those with higher education could support the intuition that the Chilean public administration has made an effort to modify the gap in order to attract more qualify employees.

Counterfactual Analysis

We perform the following counterfactual exercises: (i) for each year we decompose the difference in wage distribution (between public and private employees) into rent payment and characteristics differential; (ii) for each sector (public and private) we estimate the density function of wage in 2009, corresponding to the 2000 distribution of covariates; and (iii) in order to evaluate the contribution of each covariates to the change in distribution, we also estimate the density of wages in 2009 if only one covariate was distributed as in 2000. We perform these exercises by educational level. We set 50 and 100 replications (m) for the last and the former exercises respectively[14].

(i) Wage differential decomposition between public and private employees

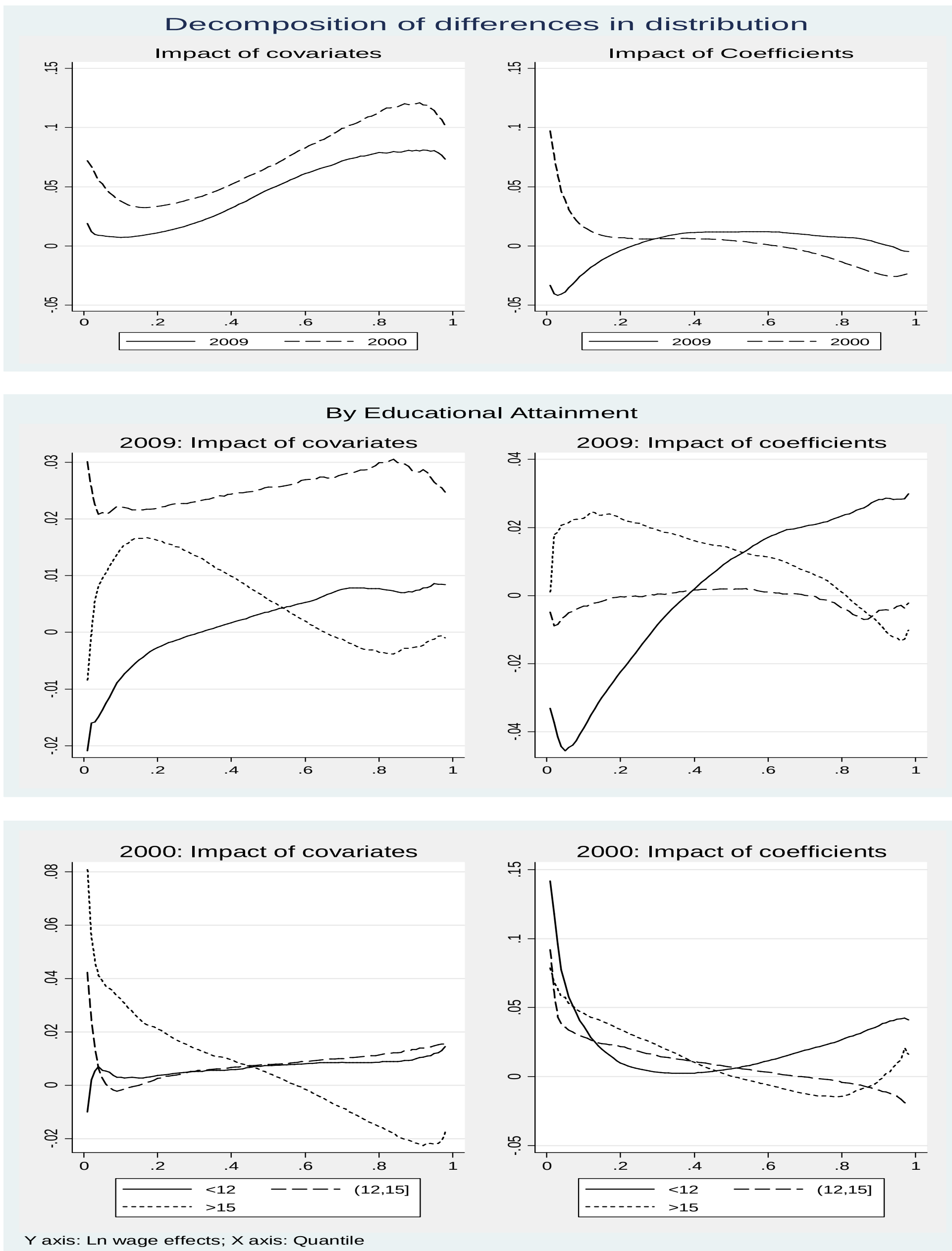
The first step is to decompose the wage gap into rent payment and characteristics differential. Figure 3, plots such decomposition for 2009 and 2000. It is observed that the public sector wage premium tends to increase until the eighth quantile and then decrease. In this context an important issue is that those under the eighth quantile have 13 years of educational attainment or less. This means that the premia stars to becoming a penalty for those with technical and college. In term of the average wage by quantile, the wage penalty for 2009 is about -2% at $\mu = 0.1$(US$3) and a public sector rent premium of 1% at $\mu = 0.8$ (US$12).

It is important to note the change in the distribution of the coefficients. While for 2000, it has a premium for the lower and a penalty for the higher for the higher quantiles; for 2009, such distribution shows a penalty for the lower quantiles and it shift to the right increasing the premia (or reducing the penalty) for the higher quantiles.

[14]For performing the estimations we use the stata program rqdeco.do, provided by Melly (version 2010). He shows that the M-M estimator is numerically identical to his approach if the number of simulations goes to infinity. M-M (2005) and Melly (2005a,b) in their empirical applications set the number of replication in 4500; however because of computational limitations we had to reduce such number to 50 and 100.

With respect to the contribution of the covariates it is highlighted that its distribution shifts downwards between 2000 and 2009.

Figure 3



Public sector wage premium by education attainment

Since the wage differential may vary across educational levels, the public wage premium is now estimated separately for three ranges of schooling: a) less than 12 years, refers to those with mandatory education; b) between 12 and 15 years, which includes technicians and incomplete university; and, c) above 15 years, generally including university graduates. The other regressor are the same as above. Figure 3 shows different patterns by educational categories.

The contributions of characteristics and coefficients tend to behave similarly for 2009. For those with less years of schooling, the impact of coefficients and characteristics tend to increase as we move up the income distribution. Those with "medium" education present not significant gap, while those with college studies start with a gap of around the 2% and decline along the income distribution. Since the average wage increases with the number of years of schooling, we can infer different effects of the gap behavior in each group. For example, in the group with 12 years of schooling or less, by increasing the wage gap, the public sector tends to increase also the inequality. For those with medium education, the gap is mostly horizontal; maintaining the wage structure along the wage distribution. At the university level, as the wage premium reduce, it tends to reduce the inequality.

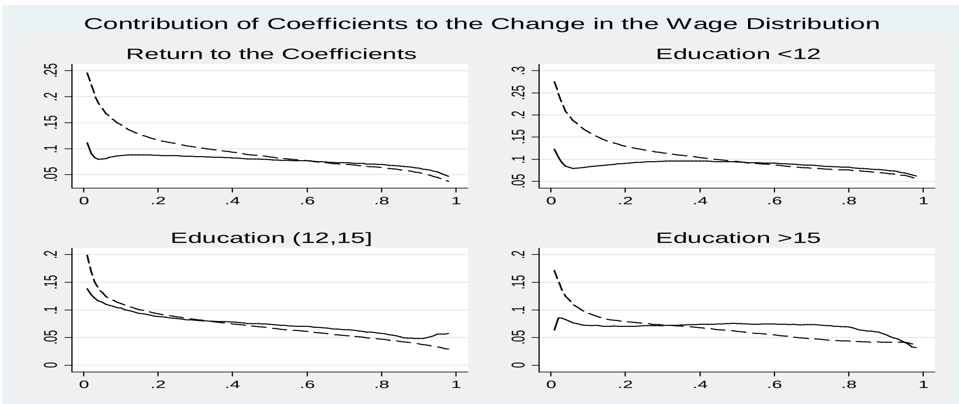
(ii) Wage differential decomposition over time

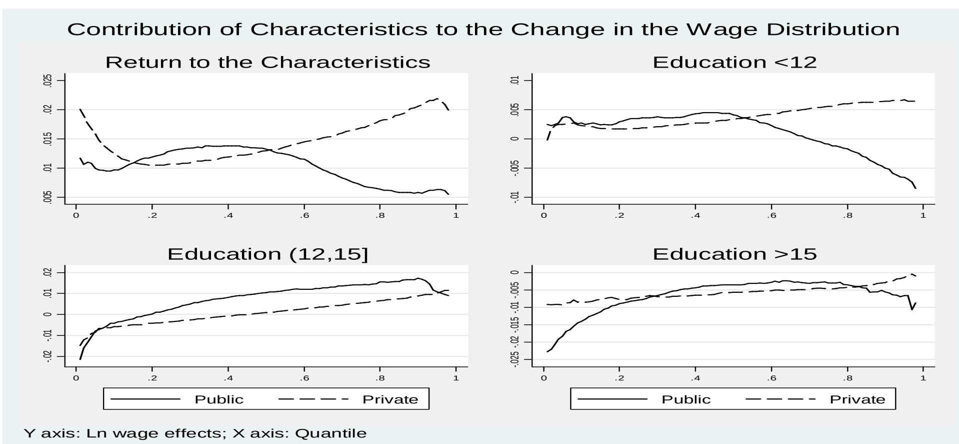
In order to assess the impact of the public administration reforms that Chile started in 1990, this section intends to decompose the changes in the wage distribution of the public sector between 2000 and 2009. Again we rely on the M-M decomposition to discriminate, over time, between changes in the characteristics of the workers and changes in the returns to these characteristics. This means that in difference with the previous sub-sections, here the wage counterfactual conditional distribution (cf) to be used is built, for each sector (public/private), with the covariates distributed as in year 2000. Then, the change of the rent to the covariates (the change in coefficients) is estimated by comparing the cf with the wage conditional distribution of 2000. Similarly, the change in covariates over this period of time is estimated by comparing the cf with wage conditional distribution of 2009.

Figure 4, shows the change (between 2000 and 2009) in coefficient and covariates along the wage distribution for the public and private sector. It can be appreciate that, in the last 10 year, the market increased much more the returns for those in the lowest quantiles, while the public sector increases comparative more the returns for higher quantiles. This would reinforce previous observations suggesting an effort of the public sector to attract more qualify employees.

The contribution of the covariates seems to be important only for the lower quantiles. Even more, when we estimate this contribution by education attainment, it it seems to be positive mainly for those in the range $(12, 15]$.

Figure 4





(iii) Covariate contribution over time

Table D decomposes the overall effect of the covariates into its constituents. To that end we estimated the density of wages in 2009 if only one covariate were distributed as in 2000. For the public sector, the main characteristic is the education variable, it contributes for most of the change (in covariates) in the wage distribution. At the quantile 0.25, 0.5 and 0.9, it explains the 10%, 22% and 9% of the total change. The effect of age also increase the wage but in less magnitude than education. The gender covariate has not effect on the wage distribution of the public sector. The effect of tenure shift the distribution towards the left.

Table D also presents the estimates for the private sector. In particular it call the attention that the effect of the education is greater in the private sector, especially for the upper quantiles (from $\theta = 0.5$ onwards). This could suggest that the moving up in the schooling contribution is lead by a labor market trend and that it is absorbed in the higher quantiles by the private sector. This is reasonable since, as we have seem before, the private sector tend to pay more to those with higher years of schooling. This could be an issue of deeper and specific research.

Two more results call our attention. In the public sector, the tenure variable had a negative effect on the change in wage distribution for those quantiles below $\theta = 0.9$. However, for the private sector it had a positive effect for all quantiles above $\theta = 0.25$. This could be evidence of an effort of the public administration to renovate its staff in the less qualify occupations.

The other result refers to the gender effect on the wage change along the distribution. It maintained stable over the last decade, while in the private sector it decays both: in contribution to the wage change, as well as, in its participation on the private labor force. This is aligned with the literature, in the sense that, the public sector tends to pay equally across gender or at least it tends to present a lower gender penalty.

5 Conclusions:

This paper has concentrated in estimate the wage gap between public and private employees in one of the most progressive examples of public administration reforms in Latin America. Since 1999, Chile has started a process of public sector reforms in different areas and now is recognized as one of the most competent public sectors in the Region. Taking this into consideration, we were particularly interested in estimate, not only the wage gap at the mean, but also analyze how such policy affected the wage distribution over time.

At the mean, for 2009, our results tend to suggest a penalty for the public servants. Even more, this penalty tend to concentrate in the group with higher educational attainments, while those with less years of schooling tend to present some wage premia with respect to the private sector. These penalties and premiums tend to decrease over the last 10 years, maybe, implying an effort of the public sector to get aligned with the market.

The counterfactual quantile regressions, in a given year, show that the distribution of the public sector rent to the covariates (the effects of coefficients) changed between 2000 and 2009 in favor of the upper quantiles. For 2000, such rent tend to be positive for the lower quantiles, but it became negative by 2009. What is more, the impact of coefficients moved upwards over this period favoring to the higher quantiles. This suggest an effort of the public sector to attract more qualify employees

The wage differential decomposition between 2000 and 2009 add important insides. It suggests that the public sector tend to increase more the rent to the covariates for the higher quantiles (while the private tend to increase the rent for the lower quantiles). That is, it suggest that the public sector tend to increase the between-group inequality . Since the average wage is highly correlated with the years of schooling, it could be interpreted as a signal of the public sector to attract more qualify employees. These estimations also shows that the tenure covariate had a negative effect on the wage, suggesting a trend of the public sector to renew its staff.

Our estimations are aligned with previous literature and with the view of Chile as a remarkable example of public sector improvement. However, our distributional approach does not take into account the possibility of self-selection bias and the mean conditional estimations, corrected by selection bias, could be inconsistent because of mis-specifications. These caveats could be issues of further empirical research.

References

Bender, K. (2003), “Examining Equality between Public and Private Sector Wage Distributions”, Economic Inquiry, Vol. 41 No 1 pp. 62-79;

Cai, L.and Liu, A. (2011). "Public–Private Sector Wage Gap in Australia: Variation along the Distribution". British Journal of Industrial Relations Volume 49, Issue 2, pp. 362–390;

Cameron, C. and Trivedi, P. (2009). Microeconometrics using Stata. Stata press;

Cameron, C. and Trivedi, P. (2010). Microeconometrics Methods and Application. Cambridge University Press;

Chernozhukov, V., Fernandez-Val, I. and Melly, B. (2009), “Inference on counterfactual distributions”, CeMMAP Working paper;

Coppola, A and Calvo-Gonzalez, O. (2011), “Higher Wages, Lower Pay Public vs. Private Sector Compensation in Peru”, World Bank Policy Research Working Paper Series No 5858;

Echevarria and Cortazar (2007), “Public Administration and Public Employment Reform in Latin America”. Chapter 4. IADB;

Fortin, N., Lemieux, T. and Firpo, S. (2010), “Decomposition Methods in Economics”, NBER Working Paper Series;

Glinskaya, E. and Lokshin, M. (2005), “Wage differentials between the public and private sector in India”, World Bank Policy Research Working Paper Series No 3574;

Heckman, J. (1979), “Sample Selection Bias as a Specification Error”, Econometrica, Vol. 47, No. 1, pp. 153-161;

Hyder, A. and Reilly, B. (2005). "The Public and Private Sector Pay Gap in Pakistan: A Quantile Regression Analysis". The Pakistan Development Review. Vol. 44, No. 3, pp. 271-306;

Iacoviello, M. and Tommasi, M. (2002). “Diagnostico Institucional de Sistemas de Servicio Civil”. IADB.

Melly, B. (2005a). “Public-private sector wage differentials in Germany: evidence from quantile regression”, Empirical Economics, 30(2), pp. 505-520;

Melly, B. (2005b). "Decomposition of differences in distribution using quantile regression," Labour Economics, Elsevier, vol. 12(4), pp. 577-590;

Katz, L. and Krueger, A. (1991), “Changes in the Structure of Wages in the Public and Private Sectors", NBER Working Paper Series, No 3667;

Lokshin, M. and Sajaia, Z. (2004), “Maximum likelihood estimation of endogenous switching regression models", Stata Journal, Vol. 4, pp. 282-289;

Machado, J. and Mata, J. (2005). "Counterfactual decomposition of changes in wage distributions using quantile regression," Journal of Applied Econometrics, vol. 20(4), pp. 445-465;

Manski, C. (1989), “Anatomy of the Selection Problem”, The Journal of Human Resources, Vol. 24, No. 3, pp. 343-360;

Mizala, A.,Romaguera, P. and Gallegos, S. (2011), “Public–private wage gap in Latin America (1992–2007): A matching approach”, Labour Economics, Vol.18, No. 1, pp. S115-S131;

Panizza, U. (2001), “The Public Sector Premium and the Gender Gap in Latin America: Evidence from the 1980s and 1990s”, RES Working Papers 4229, Inter-American Development Bank, Research Department;

Panizza, U., Di Tella, R. and Van Rijckeghem, C (2001), “Public Sector Wages and Bureaucratic Quality: Evidence from Latin America”, Economía, Vol. 2, No. 1, pp. 97-151;

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

Payne, J. and Carlson, I. (2002), "Cross Country Comparisons of Public Employment and Pay of 26 Latin American and Caribbean Countries" IDB Publications No 5438;

Poterba, J. and Rueben, K. (1994), "The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods", NBER Working Paper Series No 4734;

Robins, J., and Rotnitzky, A. (1995), "Semiparametric Efficiency in Multivariate Regression Models with Missing Data", American Statistical Association, Vol. 90, No. 429, pp. 122-129;

Wooldridge (2010). Econometric Analysis of Cross Section and Panel Data. Second ed. The MIT Press.

ANNEXES

Table A-I: Descriptive statistics

		2000						2009					
Sector	**Variable**	**p10**	**p25**	**p50**	**mean**	**p75**	**p90**	**p10**	**p25**	**p50**	**mean**	**p75**	**p90**
Private	Monthly Wage	111	160	195	279	297	482	261	316	355	487	513	790
	Ln(monthly wage)	4.7	5.1	5.3	5.4	5.7	6.2	5.6	5.8	5.9	6.0	6.2	6.7
	Years of schooling	4.0	6.0	10.0	9.3	12.0	14.0	6.0	8.0	12.0	10.5	12.0	15.0
	Age	22.0	27.0	35.0	35.8	43.0	52.0	23.0	28.0	38.0	38.2	47.0	55.0
	Tenure	0.0	0.0	2.0	4.7	6.0	14.0	0.0	0.0	2.0	4.9	7.0	14.0
Public	Monthly Wage	148	186	297	456	557	872	257	326	513	728	888	1382
	Ln(monthly wage)	5.0	5.2	5.7	5.8	6.3	6.8	5.5	5.8	6.2	6.3	6.8	7.2
	Years of schooling	6.0	10.0	12.0	12.2	16.0	17.0	8.0	12.0	12.0	13.0	16.0	17.0
	Age	25.0	31.0	40.0	39.9	48.0	55.0	25.0	31.0	41.0	41.1	50.0	57.0
	Tenure	0.0	1.0	5.0	9.2	15.0	25.0	0.0	1.0	5.0	8.9	14.0	25.0
All	Monthly Wage	111	167	204	304	334	557	258	316	365	517	553	888
	Ln(monthly wage)	4.7	5.1	5.3	5.4	5.8	6.3	5.6	5.8	5.9	6.1	6.3	6.8
	Years of schooling	4.0	7.0	10.0	9.7	12.0	15.0	6.0	8.0	12.0	10.8	12.0	16.0
	Age	22.0	27.0	35.0	36.4	44.0	53.0	23.0	28.0	38.0	38.6	48.0	55.0
	Tenure	0.0	0.0	2.0	5.3	7.0	15.0	0.0	0.0	2.0	5.4	8.0	16.0

Table A-II: Descriptive statistics

	Variable	2000	2009
Private	% of female employees	24%	28%
	% of married employees	50%	43%
	% Workers living in urban areas	63%	68%
	% Workers of indigenous ethnicity	0%	7%
Public	% of female employees	53%	52%
	% of married employees	55%	48%
	% Workers living in urban areas	73%	79%
	% Workers of indigenous ethnicity	0%	8%
All	% of female employees	28%	31%
	% of married employees	51%	44%
	% Workers living in urban areas	64%	70%
	% Workers of indigenous ethnicity	-*	7%

* The survey for 2000 does not consider ethnicity questions.

Table A-III: Descriptive statistics

	Private		Public	
Economic sector:	**2000**	**2009**	**2000**	**2009**
Not specified	0%	1%	0%	0%
Agriculture, etc.	34%	27%	2%	1%
Mining	3%	4%	0%	0%
Industry	15%	12%	1%	0%
Electricity, water, gas	1%	1%	1%	0%
Construction / Building	9%	10%	5%	0%
Commerce	15%	17%	1%	0%
Transport and communications	7%	8%	1%	1%
Financial services	5%	7%	1%	0%
Social services	11%	13%	88%	97%
Profession:				
Managers, Professionals and intellectuals	5%	7%	31%	30%
Technical and medium level professional	5%	7%	13%	21%
Blue collars	9%	8%	14%	12%
Services workers	12%	14%	14%	10%
Agriculture	12%	6%	2%	2%
Artisans and skilled workers	13%	13%	5%	2%
Operators	13%	13%	4%	4%
Non skill workers	32%	31%	18%	18%
Others (includes non-specified)	0%	0%	0%	1%

Figure A1: Income distribution by sectors and years

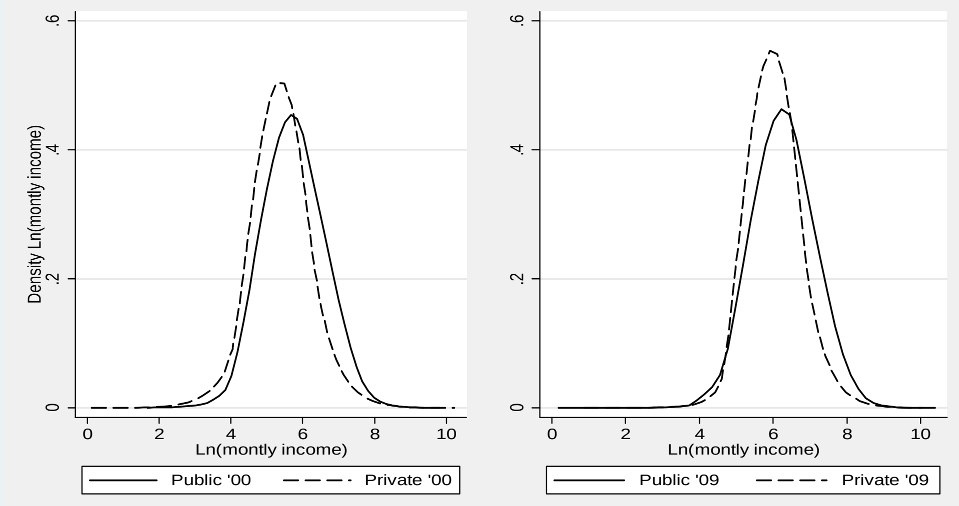


Figure A2: Schooling distribution by sectors and years

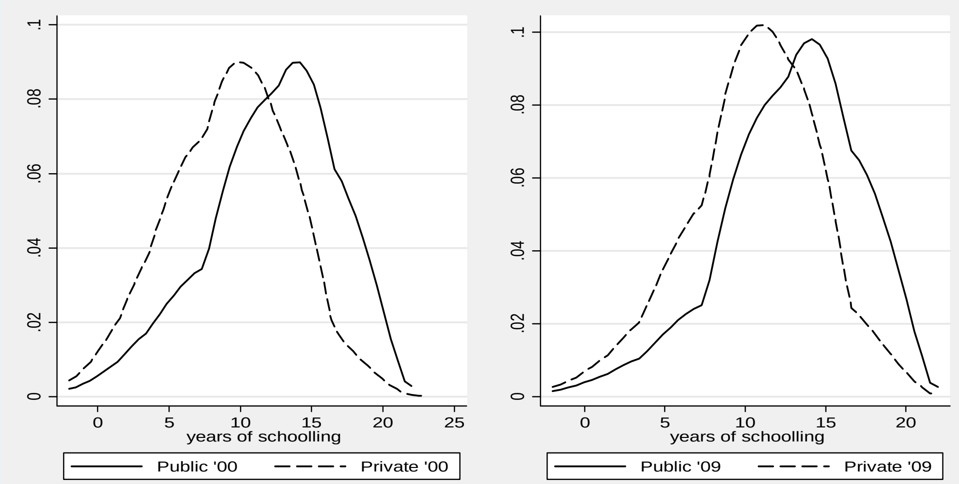


Table A-IV: Dispersion statistics of the monthly income

	Public		Private		All	
	2000	2009	2000	2009	2000	2009
Mean	456	728	279	487	304	517
Median	297	513	195	355	204	365
kurtosis	194	124	356	98	331	137
Coefficient of variation	1.25	1.00	1.13	0.88	1.21	0.93
Gini Index	0.44	0.40	0.38	0.31	0.40	0.34

Table B: Wage gap estimation at the mean

	Uncond. Aver. gap		OLS		Heckit	
	2009	2000	2009	2000	2009	2000
Public Sector Dummy	0.295**	0.412**	0.01	0.0283**	-0.0870**	-0.169**
	-0.01	-0.01	-0.01	-0.01	-0.03	-0.04
Education level 1			0.01	0.110**	0.01	0.111**
			-0.02	-0.02	-0.02	-0.02
Education level 2			0.0825**	0.148**	0.0831**	0.149**
			-0.01	-0.01	-0.01	-0.01
Education level 3			0.136**	0.166**	0.137**	0.169**
			-0.01	-0.01	-0.01	-0.01
Education level 4			0.190**	0.248**	0.186**	0.243**
			-0.01	-0.01	-0.01	-0.01
Education level 5			0.346**	0.295**	0.341**	0.278**
			-0.02	-0.02	-0.01	-0.02
Age			0.0242**	0.0316**	0.0243**	0.0321**
			0.00	0.00	0.00	0.00
Age square			-0.000260**	-0.000347**	-0.000260**	-0.000350**
			0.00	0.00	0.00	0.00
Gender			-0.192**	-0.230**	-0.194**	-0.228**
			-0.01	-0.01	-0.01	-0.01
Tenure			0.00752**	0.0110**	0.00777**	0.0116**
			0.00	0.00	0.00	0.00
Urban dummy			0.0460**	0.0962**	0.0446**	0.0892**
			0.00	-0.01	0.00	-0.01
Ethnicity			-0.0348**		-0.0325**	
			-0.01		-0.01	
Civil status dummy			0.0620**	0.0907**	0.0616**	0.0907**
			0.00	-0.01	0.00	-0.01
Controls for economic sector	Y	Y	Y	Y	Y	Y
Controls for Profession	Y	Y	Y	Y	Y	Y
Constant	6.018**	5.379**	5.315**	5.032**	6.051**	5.090**
	0.00	0.00	-0.09	-0.04	-0.07	-0.04

Note: First entry in each cell is the point estimated. The second entry is the standard error. * p<0.10, ** p<0.05

Table C: Wage gap estimation at the mean by educational attainment

	Year 2009									Year 2000								
	Unconditional gap			OLS			Heckit			Unconditional gap			OLS			Heckit		
	< 12	(12,15]	>15	< 12	(12,15]	>15	< 12	(12,15]	>15	< 12	(12,15]	>15	< 12	(12,15]	>15	< 12	(12,15]	>15
Public Sector Dummy	0.0414**	0.146**	0.0970**	0.00553	0.0699**	0.0308	0.0267	-0.219**	-0.386**	0.160**	0.0871**	0.0713**	0.0238**	0.00931	0.0422*	0.164	0.3	-0.915
	-0.0109	-0.0251	-0.0192	-0.0114	-0.0276	-0.0199	-0.0315	-0.0986	-0.142	-0.0101	-0.0228	-0.0219	-0.0106	-0.0263	-0.0234	-0.169	-0.286	-0.591
Education level 1				0.00766			0.00779						0.107**			0.00959		
				-0.0159			-0.0156						-0.018			-0.0703		
Education level 2				0.0780**			0.0780**						0.141**			0.111**		
				-0.00597			-0.00618						-0.0066			-0.0273		
Education level 3				0.121**			0.121**						0.140**			0.127**		
				-0.00517			-0.00505						-0.00665			-0.0309		
Education level 4					0.0511**			0.0476**						0.109**			0.287**	
					-0.0167			-0.0168						-0.0194			-0.117	
Education level 5					0	0.264**			0.264**						0.221**			0.191
					0	-0.018			-0.0184						-0.0193			-0.127
Age				0.0198**	0.0495**	0.0425**	0.0197**	0.0496**	0.0457**				0.0273**	0.0571**	0.0576**	0.0241**	0.120**	-0.0613
				-0.001	-0.006	-0.006	-0.001	-0.005	-0.006				-0.002	-0.007	-0.008	-0.006	-0.033	-0.053
Age square				-0.000216**	-0.000570**	-0.000393**	-0.000216**	-0.000569**	-0.000428**				-0.000302**	-0.000620**	-0.000579**	-0.000318**	-0.00160**	0.000586
				0.000	0.000	0.000	0.000	0.000	0.000				0.000	0.000	0.000	0.000	0.000	-0.001
Gender				-0.182**	-0.229**	-0.209**	-0.182**	-0.239**	-0.222**				-0.224**	-0.212**	-0.235**	-0.224**	0.0562	-0.184
				-0.005	-0.019	-0.018	-0.005	-0.018	-0.018				-0.007	-0.020	-0.020	-0.032	-0.106	-0.112
Tenure				0.00671**	0.0141**	0.00414**	0.00667**	0.0158**	0.00650**				0.0115**	0.0172**	0.00311**	0.0110**	0.0243**	0.0380**
				0.000	-0.002	-0.001	0.000	-0.002	-0.001				0.000	-0.002	-0.001	-0.002	-0.009	-0.013
Urban dummy				0.0480**	-0.00376	0.0683**	0.0482**	-0.0168	0.0392				0.104**	0.0807**	-0.00384	0.0889**	0.142	0.0513
				-0.005	-0.023	-0.022	-0.005	-0.024	-0.025				-0.006	-0.025	-0.024	-0.027	-0.101	-0.155
Civil status dummy				0.0473**	0.0751**	0.137**	0.0474**	0.0765**	0.130**				0.0742**	0.141**	0.164**	0.0613**	0.204**	0.042
				-0.005	-0.019	-0.018	-0.005	-0.019	-0.019				-0.006	-0.018	-0.020	-0.025	-0.094	-0.132
Controls for economic sector	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls for Profession	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	5.921**	6.204**	6.748**	6.085**	5.791**	5.788**	5.296**	5.833**	5.741**	5.270**	5.767**	6.401**	4.522**	4.770**	4.622**	4.253**	2.292**	11.08**
	-0.00235	-0.00998	-0.0124	-0.107	-0.143	-0.247	-0.0551	-0.181	-0.323	-0.00307	-0.0114	-0.0164	-0.0668	-0.239	-0.216	-0.243	-0.809	-1.596
N	34953	4059	5274	34362	3975	5183	34362	3975	5183	39869	4552	4451	39729	4546	4437	2302	151	119

Note: First entry in each cell is the point estimated. The second entry is the standard error. * p<0.10, ** p<0.05

Table D: Decomposition of the changes in wage distribution over time

Public Sector

Quantile	Marginals			Aggregate contributions		Individual Covariates			
	2009	2000	Δ	Covar.	Coeffic.	Age	Educ.	Gender	Tenure
0.1	5.470	4.933	0.537	0.054	0.483	0.001	0.000	0.000	-0.006
			0.015	0.010	0.011	0.006	0.000	0.001	0.008
				10.1%	89.9%	0.2%	0.0%	0.0%	-1.1%
0.25	5.829	5.257	0.572	0.075	0.497	0.005	0.058	0.000	-0.004
			0.009	0.009	0.009	0.007	0.021	0.000	0.006
				13.1%	86.9%	1.0%	10.2%	0.0%	-0.6%
Median	6.271	5.726	0.544	0.078	0.467	0.018	0.118	0.000	-0.010
			0.011	0.010	0.011	0.011	0.025	0.021	0.011
				14.3%	85.7%	3.2%	21.6%	0.0%	-1.8%
0.75	6.776	6.280	0.496	0.045	0.451	0.026	0.017	0.000	-0.009
			0.012	0.012	0.012	0.011	0.025	0.008	0.014
				9.1%	90.9%	5.2%	3.3%	0.0%	-1.9%
0.9	7.234	6.772	0.462	0.039	0.423	0.028	0.041	0.000	0.000
			0.015	0.017	0.016	0.017	0.004	0.028	0.015
				8.5%	91.5%	6.0%	8.8%	0.0%	0.0%

Private Sector

Quantile	Marginals			Aggregate contributions		Individual Covariates			
	2009	2000	Δ	Covar.	Coeffic.	Age	Educ.	Gender	Tenure
0.1	5.558	4.679	0.879	0.069	0.810	0.000	0.000	0.000	0.000
			0.003	0.007	0.007	0.004	0.036	0.000	0.005
				7.9%	92.1%	0.0%	0.0%	0.0%	0.0%
0.25	5.733	5.042	0.691	0.061	0.630	0.004	0.034	0.000	0.003
			0.001	0.003	0.003	0.007	0.000	0.006	0.004
				8.9%	91.1%	0.6%	4.9%	0.0%	0.5%
Median	5.920	5.346	0.573	0.076	0.497	0.007	0.134	-0.009	0.003
			0.002	0.003	0.003	0.006	0.000	0.004	0.004
				13.3%	86.7%	1.2%	23.3%	-1.7%	0.6%
0.75	6.246	5.716	0.530	0.107	0.423	0.014	0.118	0.000	0.013
			0.004	0.005	0.004	0.006	0.000	0.004	0.006
				20.2%	79.8%	2.6%	22.2%	0.0%	2.5%
0.9	6.668	6.162	0.507	0.140	0.367	0.039	0.143	0.000	0.013
			0.006	0.007	0.006	0.009	0.000	0.022	0.012
				27.6%	72.4%	7.7%	28.2%	0.0%	2.6%

Note: First entry in each cell is the point estimated. The second entry is the standard error. The third entry is the proportion of the total change explained by the indicated factor.