Decomposing wage inequality: Public and private sectors in Vietnam 1993-2006

Clément Imbert

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2011 - 05

Decomposing wage inequality: Public and private sectors in Vietnam 1993-2006

Clément Imbert

JEL Codes: J45, J31, P31
Keywords: Transition, inequality decomposition, public sector

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Decomposing Wage Inequality: Public and Private Sectors in Vietnam 1993-2006. *

Clément Imbert
PSE-Paris School of Economics

January 25, 2011

Abstract

This paper studies the labor market in Vietnam during the transition towards market economy (1993-2006): we show that the public-private sector wage gap markedly increased, but that wage inequality decreased overall. Our aim is to assess how much of this evolution can be explained by workers' productive skills and their allocation between sectors. We use a simple, yet innovative, method that allows us to take into account workers' unobservable characteristics and their remuneration in each sector. Throughout the period we consider, workers are more skilled than private sector workers. However, rising returns to workers' skills in the public sector play a major role in the increase of the public-private sector gap. Against all expectations, the public sector grew richer as Vietnam moved towards market economy. Finally, a greater homogeneity among labor market participants seems to explain the overall decline in wage inequality.
JEL classification codes: J45, J31, P31

1 Introduction

Does the transition towards a market economy lead to higher earnings inequality on the labor market? The answer crucially depends on the context in which the transition occurs and on the policies that shape the way it unfolds. In Poland for example Keane and Prasad (2006) showed that the transition increased earnings inequality because the public and the private sectors both adopted a "competitive" wage setting, where workers' pay is more closely related to their productivity. The difference in earnings between the public and the private sector increased because the less productive employees moved from the public to the private sector. A quick look at Table 2 shows that Vietnam seem to have followed a different path: between-sector inequality indeed rose but within-sector inequality declined.

Of course Vietnam differs in many respects from the well-studied Eastern European countries. First, the private wage sector is much less developed, because the work

*I wish to thank Abhijit Banerjee, François Bourguignon, Gary Fields, Thomas Lemieux, Martin Rama, Martin Ravallion and Dominique Van de Walle for helpful comments. I am grateful to Bryan Mac Caig for sharing his codes on the VHLSS panels, and to Marc Gurgand for his patient supervision.

force is massively engaged in self-employed activities, both in urban and rural areas. Second, the Communist party Congress of 2001 decided against the drastic privatization agenda proposed by the World Bank and the International Monetary Fund in the aftermath of the 1998 financial crisis (Painter 2005). Ideological considerations were important in this decision. But the dramatic reorganization of public enterprises in "General Companies" indicates that the public sector was entrusted with a leading role in Vietnam's development strategy (Eglinger 2005). In this sense Vietnam's transition towards market economy closer to China's (Iyigun & Rodrik 2004).

This paper focuses on public-private sectors differences, and identifies the effect of public sector reform on labor earnings inequality. From the literature on public sector labor markets, one could support two very different views. On the one hand, if there is a fixed "public sector premium" given to workers independently of their productive characteristics, public-private differences are inequality-increasing. A reform that would make public sector pay similar to a competitive wage setting would hence reduce inequality (Bales and Rama 2001). On the other hand if the public sector wage setting dampens inequality between skilled and unskilled workers (Disney Gosling 1998), and given that public sector workers are on average more skilled, then a public sector reform could increase inequality (Keane and Prasad 2006)[1].

However, in order to identify the effect of changes in the wage setting, one has to control for compositional effects, because workers have different skills in the two sectors[2]. These skills could be observed (diploma, experience) or unobserved (cognitive skills, communication skills). This motivates the use of a new model with unobserved workers' heterogeneity and different returns to skills in two employment sectors. The model was initially developed by Lemieux (1998) to investigate the causes of the union-non union sector wage gap[3]. This model allows us to simulate counterfactual wage distributions that show separately the effect of selection (differences in workers' characteristics), and the effect of wage setting policies (differences in returns to these characteristics).

Using two panels from the Vietnam Health and Living Standard Survey (1993-2006), we show that the public-private hourly earnings gap increased dramatically between the 1990s and the following decade, while within sector inequality decreased in both sectors. In line with the rest of the literature, we find that selection is a major component of inter-sector differences: public employees are better paid because they are more skilled than their private counterparts. But we show that changes in between-sectors wage inequality are mostly explained by rising returns to workers' skills in the public as compared to the private sector. Initially underpaid as compared to private sector workers - other things being equal -, public employees now enjoy much higher returns to their productive characteristics. These institutional changes were offset by a decrease in workers' heterogeneity on the labor market so that wage inequality decreased overall.

Our paper is structured as follows. In the section 2, we delimit our sample and document the fact that the private-public wage gap increased between the 1990s and the 2000s, while overall wage inequality decreased. In section 3, we present our model. In section 4, we present our results, and compare them to those of other methods which

[1] Liu (2004) has the same line of argument regarding the gender wage gap in Vietnam, which is lower in the public sector. The author is concerned that a shrinking public sector would lead to higher gender inequality.

[2] The literature on developing countries typically finds large public private pay gaps that are mostly explained by selection (see for example Van der Gaag and Vijverberg 1988, Glinskaya and Lokshin 2005).

[3] Suri (Forthcoming) uses a similar model to study the effect of selection and comparative advantage in the adoption of fertilizers in rural Kenya.

are commonly used to decompose wage inequality. In section 5, we draw counterfactual distributions to decompose the public-private earnings gap at each period, and the overall change in wage inequality across periods.

2 Rising public/private earnings differences

2.1 Data

The data we use come from the national representative surveys Vietnam Living Standard Surveys VLSS (1993 and 1998) and Vietnam Health and Living Standard Surveys VHLSS (2002, 2004 and 2006). Except for 2002, they have been carried out by the General Statistical Office (GSO) in partnership with the World Bank. The strength of this dataset is its panel dimension. We can rely on three panels of two observations each: 1993-1998, 2002-2004 and 2004-2006. Changes in survey methodology and sample do not allow households to be tracked throughout the 1993-2006 period. Even if the 2002, 2004 and 2006 are included in the same panel, it rotated in 2004, so that a very small fraction of the households on which we have data in 2002 and 2004 is still present in 2006. For purposes of analysis, we will construct a 2002-2006 panel by simply pooling the two 2002-2004 and 2004-2006 panels, ignoring the fact that some households have been interviewed in 2002, 2004 and 2006[4].
The VLSS and the VHLSS are representative of the whole country on a given year, provided we use the appropriate sample weights. In every second year of the two years panels, an additional sample is drawn to compensate selective attrition from the panel and adapt to changing demographics of the total population. Since our analysis focuses on the panel sample, selective attrition can threaten its internal validity. In particular, private-public sectors comparison could be biased if private sector workers are less stable geographically than public employees[5]. We will discuss this problem whenever there is a possibility that it may bias our results. However, we do not find any significant difference in observable workers' characteristics nor in public-private sector shares between the cross-sectional and the panel sample. More importantly, the wage distributions are perfectly similar.

2.2 Public and Private employment

While the focus of our study is on wage and wage related benefits, only a small share of the labor force in Vietnam is actually working for wage. In the early 1990s out of five workers only one was employed for her main occupation outside of her household, and this share is no higher than one third in 2006 (see Table 1)[6]. As small as it is at the beginning of the period, wage employment is growing rapidly and much to the benefit of the private sector. The private sector share increased from 10% to 23% of the labor force (rural and urban taken together) between 1993 and 2002. Public employment also increased, but at a lower rate, from 8 to 11% of the labor force. In the following years, public employment share did not decline[7].
What we consider here as public employees is a heterogeneous pool of workers, with a

[4] Some 2004 individual data are hence *de facto* duplicated. Standard errors account for individual clustering.
[5] The direction of the bias is ambiguous, since we could be missing both the most fragile and the most successful of private sector employees.
[6] Our definition of labor market participation is restrictive because we focus on the main occupation over the last 12 months. Gallup (2002) finds 25% in 1998 by including secondary occupations
[7] This finding is confirmed in the official reports based on 2006 Labor force survey.

large variation in terms of status and working conditions. A growing majority of them work in government, education and health services: 51% in 1993, 62% in 2006. The remaining 49% and 38% are employed by public enterprises. Throughout the 1993-2006 period, public ownership is present in almost all industries: Electricity and Water Production, Mining, Food and Beverages, Textile etc. Even when they operate in the same sector, public firms are larger than private companies: in 1998 public workers in Paper Metal and Plastic Industries declared having 200 co-workers on average while private employees declared having only 35.
Over the period, gradual but considerable efforts have been done to reduce over- employment, increase profitability and harden budget constraints in the public sector. During the 1990s, the number of public enterprises shrank as losses-making firms were closed or integrated into bigger ensembles, named "General Companies" (see Eglinger 2005). According to 2006 enterprises' surveys, public firms are more capitalistic and generate more profits than private domestic firms (Vietnam Socio-Economic Development 2008). This may be seen as an achievement of the reforms, but it may also reflect unequal access to credit on the one hand (see J. Mac Millan and C. Woodruff 1999) and quasi-monopolistic positions of public firms on the other.

2.3 Public and Private sector earnings distributions

The evolution of public and private sectors' earnings distributions can be seen on figure 1: The two distributions followed each other closely in 1993, whereas in 2006, the public sector earnings distribution clearly dominated the private sector one. In 1993, private and public employees had the same hourly earnings of 2000 Dongs on average. In 2006, public employees earned on average 10,000 Dongs, against 6,000 for private employees. Across time, the private wage distribution became less dispersed around its mean, whereas the public wage distribution kept the same shape. In a nutshell, within-sector inequality fell in the private sector, leading to a fall in overall inequality despite rising between-sectors inequality and constant within-sector inequality in the public sector (see the Theil decomposition in table 2).
Two statistical artifacts may explain these findings. The sample may be unable to cope with an increasing mobility of private sector workers (missing the working poor in urban areas for example). The survey may also be unable to prevent under-reporting from high-wage private workers. Unfortunately, the lack of any other statistical source on the period we consider (except for Labour Force Surveys in 2006) prevents us to check the representativeness of our sample. This limitations will be important while interpreting changes in workers' selection.
Our measure of earnings is hourly compensation as declared by the workers themselves, and includes wage (in cash and in kind), together with benefits received from the employer, (in cash and in kind) Being able to include benefits is crucial for our purposes: if we take only wages into account, we do not find any public sector premium in 1998 because all the difference in compensation we observe in Figure 1 is due to more generous benefits in public enterprises. However the post 1998 evolution is mostly due to wages, with the difference in benefit provision remaining constant.
Since we consider hourly earnings, the rise in the public-private earnings gap can be decomposed between two factors First, the public-private difference in the number of hours worked per year shrank. Private employees used to work 1900 hours per year in 1993 and public employees 2100 hours; in 2002 both public and private employees work 2100 hours on average. Second, the difference in yearly earnings increased steadily over the period. It follows a continuous trend and doubles from 25% in 1993

to 58% in 2006. Summing these two effects gives us the specific pattern of a moderate increase of the public/private hourly earnings gap from 1993 to 1998, and a much more substantial increase between 1998 and 2006, because private workers from 2002 onwards work as many hours on average as their public counterparts, but for a much lower level of yearly compensation.

3 Model

3.1 Motivation

The stylized fact that we have to explain is a decreasing wage inequality combined with an increase in the public-private gap between the 1990s and the following decade. In the spirit of the Roy model, a possible explanation would be that workers were allowed to re-allocate in the sector they wanted, or that employers were able to decide more freely whom they hired. Heckman Honoré (1990) show that for a log normal distribution, if workers used to be assigned "randomly" and are now specialized in the sector in which they have a comparative advantage, inequality should decrease within each sector and increase between them. The panel structure of our data gives us an interesting opportunity to take into account the sorting of workers. Our model estimates different returns to the (unobserved) individual productivity term in each sector. This structure allows us to gauge at each period the extent to which workers are sorted, i.e. allocated to the sector where their comparative advantage lies . Since we have two panels, we can test for a change in workers' sorting between the two phases of the transition towards a market economy.

3.2 Model

The structure of our model is the following[8]. Individual *i*'s earnings at time *t*, noted Y_{it}, depend on four factors: a time effect τ_t, a time-varying sector effect η_t and θ_i, which is an unobserved and time invariant characteristic. Returns to workers' unobserved characteristic are normalized to one in the private sector, and are equal to α in the public sector. Let P_{it} be a dummy equal to 1 if the individual *i* is working in the public sector at time *t* and ϵ_{it} an idiosyncratic shock with zero mean.

$$\begin{cases} Y_{it} = \tau_t + \theta_i + \epsilon_{it} \ \ if \ \ P_{it} = 0 \\ Y_{it} = \tau_t + \eta_t + \alpha * \theta_i + \epsilon_{it} \ \ if \ \ P_{it} = 1 \end{cases} \tag{1}$$

In order to explain wage differentials between the public and the private sector, this simple setting allows us to capture both the effects of selection (through individual characteristics θ_i) and of compensation patterns (captured by the η_t and α). Note that for identification purposes we have to assume α constant within a panel whereas η_t varies for each date. But since we have two panels we will estimate different α for each of them.

Our model also helps uncovering the determinants of between and within-sector wage inequality for a given distribution of θ. If, on average, public sector workers are more skilled than private sector employees, a positive η_t will increase between-sector inequality (it has no effect on within-sector inequality). This corresponds to the public sector "wage premium" found in the literature. An α smaller than one implies that

[8] It is similar to Lemieux(1998) except for the time-varying constant sector premium.

intra-sector inequality is smaller in the public than in the private sector, because wage differentials between workers of different skills are compressed in the public sector. If public sector workers are also more skilled, then an α smaller than one also implies that the public sector wage setting reduces inter-sector inequality. An α greater than one would have opposite effects.

3.3 Identification

Within each panel we observe workers at two different dates. We define four groups of employment history H: some workers stayed in the public sector at both dates (H=11), some switched to the private sector (H=10), some came from the private sector (H=01),some stayed in the private sector (H=00). For identification purposes, we assume that each sector treats movers and stayers the same way at a given date. This allows us to use information on workers who move between sectors to identify parameters that are common to all workers in a given sector (like α for public workers). We do not assume that the distribution of θ_i is the same for movers and stayers i.e. that they are similar in terms of productive characteristics. But conditionally on their individual, time invariant characteristic, the two groups' earnings are the same. Formally our assumptions writes:

$$Y_{it}|_H \perp \epsilon_{it}$$

This assumption is similar to the strict exogeneity assumption in other panel models. But it rules out some plausible mechanisms, for example if the idiosyncratic component of the wage at time *t*,ϵ_{it}, is systematically lower for workers who will leave one sector at the period *t+1* as compared to those who will stay in the same sector.
Let m_{00}, m_{01}, m_{10} and m_{11} denote the mean of θ_i in the different groups. Since it is unobserved, we normalize the mean of θ to 0 over the whole population. Identification relies on eight equations, since we observe the first moments of the four groups at two different periods:

$$\begin{array}{ll}
\bar{Y}_{t+1}|_{H=11} = \tau_1 + \eta_1 + \alpha * m_{11} & \bar{Y}_t|_{H=11} = \tau_0 + \eta_0 + \alpha * m_{11} \\
\bar{Y}_{t+1}|_{H=00} = \tau_1 + m_{00} & \bar{Y}_t|_{H=00} = \tau_0 + m_{00} \\
\bar{Y}_{t+1}|_{H=01} = \tau_1 + \eta_1 + \alpha * m_{01} & \bar{Y}_t|_{H=01} = \tau_0 + m_{01} \\
\bar{Y}_{t+1}|_{H=10} = \tau_1 + m_{10} & \bar{Y}_t|_{H=10} = \tau_0 + \eta_0 + \alpha * m_{10}
\end{array}$$

We have eight equations which allow to identify eight unknown parameters: τ_0, τ_1, α, η_0,η_1, m_{00}, m_{01} and m_{10}. If π_{ij} denotes the weight of each group $H = ij$, m_{11} is then given by:

$$m_{11} = -(\pi_{00} * m_{00} + \pi_{01} * m_{01} + \pi_{10} * m_{10}) * \frac{1}{\pi_{11}}$$

It is interesting to note as in Lemieux (1993) that:

$$\alpha = -\frac{(\bar{Y}_{t+1}|_{H=01} - \bar{Y}_t|_{H=10}) - (\bar{Y}_{t+1}|_{H=11} - \bar{Y}_t|_{H=11})}{(\bar{Y}_{t+1}|_{H=10} - \bar{Y}_t|_{H=01}) - (\bar{Y}_{t+1}|_{H=00} - \bar{Y}_t|_{H=00})} \quad (2)$$

Which is the ratio of the difference in earnings between "movers to public" and "movers to private" at the time they were in the public sector to the difference in earnings between "movers to private" and "movers to public" at the time they were in the private sector. The change in "private stayers"'s earnings is used to deflate private workers wages in the denominator. The change in "public stayers"'s earnings is used to deflate public workers wages in the numerator.

3.4 Specification

While discussing identification, we did not discuss the role of observed characteristics, such as gender, education or experience in our model. As in every standard panel data models, the effect of time invariant observable characteristics cannot be identified separately, unless we have reasons to believe that they are uncorrelated with our index of unobserved heterogeneity θ. We will not make any orthogonality assumption of that kind because we are precisely interested in the selection that is correlated with observable characteristics and which would explain why coefficients on this variables in OLS may be biased.
As noted by Lemieux (1993), a non parametric solution would consist in estimating the model for different groups of education and experience. However, given the small number of observations, we cannot split the sample in as many cells. In the basic specification, we do not include any observable characteristics:

$$ln\ w_{it} = \tau_t + \eta_t * P_{it} + (1 + (1-\alpha)P_{it}) * \theta_i + \epsilon_{it} \quad for\ t = 0, 1$$

In that case, the factor θ sums up all of workers' heterogeneity between sectors, be it observable or unobservable and α captures differences in returns to education as well as returns to experience between the public and the private sector.
Given the well documented fact that male and female earnings are more similar in the public than in the private sector, it seems crucial to consider gender separately. Since we cannot assume that gender is orthogonal to the productivity component, it has to be interacted with the unobserved term as in the following specification:

$$\begin{aligned} ln\ w_{it} = {} & \tau_t + \eta_t * P_{it} + (1 + (1-\alpha)P_{it}) * \theta_i \\ & + \delta_0 * G_i + \delta_1 * G_i * P_{it} + \gamma_0 * \theta_i * G_i + \gamma_1 * \theta_i * G_i * P_{it} + \epsilon_{it} \end{aligned}$$

Where G_i is a dummy for being a male worker. In this specification, relative returns to skills for women are 1 and α in the private and in the public sector respectively. For men returns to skills write $1 + \gamma_0$ and $\alpha + \gamma_1$ in the private and in the public sector respectively. The parameters δ_0 and δ_1 are sector-specific constant shifters between female and male workers.

3.5 Estimation framework

In order to estimate our model, we restrict ourselves to the sample of workers observed as working for wage for two consecutive years: 1993 and 1998, 2002 and 2004 or 2004 and 2006. As discussed earlier, VLSS-VHLSS samples are representative of the Vietnamese population, but contain few wage earners. Our final sample is hence relatively small, with 1440 individuals in 1993-1998, and 4018 individuals after pooling 2002-2004 and 2004-2006 surveys (See Table 6). Our model's identification further relies on workers who switched between the public and the private sector between the two surveys. There are 261 of workers who switch between 1993 and 1998 and there are 351 in total for the 2002-2006 surveys. The lack of observations exerts severe constraints on our estimation strategy, because it prevents us from estimating our model separately for different sub-samples defined by gender, education or experience. On the other hand it forced us to adopt a simple estimation method, non-linear least squares, which has better small-sample properties than the GMM used in Lemieux (1998).
Using a dummy variable for each of the four different groups of workers defined earlier as H_{00} ("private stayers"), H_{11} ("public stayers"), H_{01} ("movers to public") and

H_{10} ("movers to private"), the following equation is estimated through non linear least squares[9]:

$$Y_{it} = \tau_t + \eta_t * P_{it} + (1 + (1-\alpha)P_{it}) * (m_{00} * H_{00} + m_{01} * H_{01} + m_{10} * H_{10} + m_{11} * H_{11}) + \epsilon_t$$

Under the normalization constraint:

$$\pi_{11} * m_{11} + \pi_{00} * m_{00} + \pi_{01} * m_{01} + \pi_{10} * m_{10} = 0$$

4 Results

The literature on earnings inequality with a model with two sectors (public and private, formal and informal) is extremely rich and active, with some widely-used techniques, such as quantile regressions with a sector dummy, or the well-known Oaxaca-Blinder decomposition (Oaxaca (1973)). Before going to the estimation of our model, in this part we first present and discuss the results that can be obtained through the models that are more common. This presentation will help us to establish the robustness of some stylized facts, and to make the case for using a different approach.

4.1 Controlling for workers' characteristics: "public sector dummy approach"

Public and private sector workers have very different characteristics, reflecting differences in the requirement of the jobs offered in the two sectors (see Table 3). As one would expect, public sector workers have spent more time in school (5 years more in 1993), and are more likely to have received a vocational training (28% more in 1993). There are also proportionally more female workers in the public sector (12% more) suggesting that public employment is more "women friendly" than private employment, in Vietnam (Liu 2005) as in many other countries. Regarding the evolution over time, it appears that private and public sector workers moved towards convergence in terms of observable characteristics: the years of education-gap decreased to 3.5 years in 2006, the difference in the proportion of workers who attained vocational training went down to 25%. A major exception to this common trend is experience, with public workers becoming more experienced on average, suggesting that fewer hires happened after the reform.

In order to compute the public-private wage gap once differences in workers' characteristics are taken into account, one can simply regress earnings on workers observable characteristics, with a dummy variable for being in the public sector. This method is not very informative of the origin of public-private sector differences: it only allows for a specific intercept in the equation of public sector workers. If one estimates wage equations at different quantiles however, one can study how coefficients on the public sector dummy vary across the distribution of earnings. Our quantile regression estimates of the public sector premium are shown in Figure 2. The public sector premium was essentially zero in 1993 and 1998, and rose steeply afterwards. Most interestingly, in the years 2000 the public premium became much higher for the top half of the distribution (more than 30% at all quantiles above the median). Once observable characteristics are taken into account, it seems that compensation increased in the public sector for the

[9] The extended model we described before has exactly the same structure, but each variable is interacted with a dummy for being a male.

best workers, but less so for workers at the bottom of the distribution.
These findings are coherent with the rise in returns to skills in the public as compared to the private sector that we document in the next session. We will not pursue the "public sector dummy approach" however, because it constraints returns on workers' characteristics to be the same across sectors[10].

4.2 Selection on and returns to observable characteristics: Oaxaca-Blinder decomposition

A second approach, which is by far the most common in the literature, consists in estimating separate wage equations for the private and the public sector: returns to workers' observable characteristics are allowed to differ between the two sectors. Results from the OLS estimation of wage equations in the public and in the private sectors are shown below (see Table 4). In 1993, the public sector had higher returns to education and lower returns to experience than the private sector. Between 1993 and 2006, returns to education rose and returns to experience declined in both sectors, which is a common finding in other transition economies. At the end of the period, returns to education are still higher in the public sector, while returns to experience are almost comparable in both sectors. This helps to understand why the public-private wage gap widened, since public workers, who are better educated and more experienced than their private counterparts saw the returns to their skills improve as compared to the private sector.
Once returns to skill are estimated separately for each sector, one can use a decomposition similar to Oaxaca (1973) and decompose the public-private sector gap into two components: the effect of differences in skills on the one hand (selection effect) and the effect of differences in the returns to these skills on the other (price effect). We build a counterfactual wage for public employees by applying to their observed characteristics the coefficient of the wage equation estimated for the private sector. The difference between the wage observed for public workers and this counterfactual reflects the price component of the public-private wage gap, because difference in workers characteristics have been neutralized. We find a substantial price effect for the 1990s, which is consistent with Bales and Rama (2001). Extending their analysis to the following years however, we find that this premium rose steeply after 1998, from 20% to 37% (results of this decomposition are presented in the first rows of Table 6 [11]. These results suggest that changes in returns to workers' characteristics played a role in the rise of public-private pay differentials.
How insightful as it can be, the Oaxaca-Blinder decomposition method is biased if there are unobservable characteristics which impact workers' earnings, and if public and private sector employees are systematically different with respect to these characteristics[12]. This motivates the use of our model, because it allows for differences in workers' unobserved characteristics, and for differences in returns to these characteristics between the two sectors.

[10] Note that if one is willing to make this very strict assumption, the method can be made robust to unobserved heterogeneity using panel models (see Bargain & Melly 2008).

[11] In order to test for path-dependance, we present under "decomposition 2" the results one obtains from comparing private employees earning with the counterfactual wage that they would be paid in the public sector, given their observable characteristics.

[12] Albeit more sophisticated, the technique developed by Machado & Mata (2005), which uses sector-specific quantile regressions to draw counterfactuals suffers from the same bias (for an application to the public-private sector gap, see Azam & Prakash (2010)).

4.3 Estimation results

We now turn to the estimation of our model. An intuition of our results can be read from Table 7, which draws changes over time in average log earnings for the four groups of workers. Note that changes between 1993 and 1998 happened in five years, whereas changes in 2002-2006 are all measured on two-year intervals. The latter are hence smaller. In the 1990s, the group of workers which experienced the highest earnings gain is the group that joined the private sector (73% increase). In comparison, workers who joined the public sector only gained 58% between 1993 and 1998. Over the two panels of the 2002-2006 period, workers joining the public sector gained more (29%) than those who joined the private sector (19%). Using equation 2, Table 7 suggests that returns to workers' skills increased in the public as compared to the private sector between the periods 1993-1998 and 2002-2006 .

Model estimates obtained by non linear least squares (minimum distance estimation) are shown in Table 8. Very different patterns emerge for the 1993-1998 and for the 2002-2006 period. In the 1993-1998 panel, the factor loading α on workers' fixed effect in the public as compared to the private sector is significantly inferior to 1. In the 2002-2006 panel, α is not significantly different from 1 and the point estimate is greater than 1. This suggests that the public sector, which used to offer lower returns to workers' skills than the private one, has now adopted compensation patterns that reward them at least as much as the private sector. Interestingly the point estimate of η_t are not significantly different from zero and negative for most dates t. The higher pay in the public sector is not adequately described as a single "wage premium" which would be unrelated to productive characteristics.

Finally, even if our estimation lacks precision for the first panel, there is clear evidence of positive selection in the public sector. Indeed, the first moments of the distribution of θ are negative for employees who stay in or transit through the private sector: our normalization to zero implies that public sector stayers have on average higher (positive) values of θ. Our estimation results also indicate a negative selection into the public sector during the 1990s, with the average of θ among public sector joiners being significantly less skilled than the average worker (which by normalization has a θ equal to zero). This negative selection stops in the latter decade, but workers quitting the public for the private sector have still higher skill levels on average. These findings could be explained by an improved screening of workers hired in the public sector, and by a constant flow of skilled workers into the business sector.

As compared to the models presented at the beginning of this part, our model accounts for a broader range of individual characteristics, both observed and unobserved, which could explain differences in compensation across sectors. In this sense it allows for a greater effect of workers' selection into one sector or another. The fact that we still find significant differences in relative returns to workers' comparative advantage between the public and the private and a significant evolution of these differences over time makes a strong case for the role of compensation patterns in explaining the rising public-private wage gap, along with the selection of better workers in the public sector. In order to assess the respective effect of each factor, in the next section we use our model estimates to perform inequality decompositions for each period and between them.

5 Counterfactual analysis

In this section, we first compute for each worker the value of her unobserved comparative advantage θ_i. This comparative advantage is constant within each panel (1993-1998 or 2002-2006). Using the parameters we estimated in the last section, we then decompose the public-private average wage gap at each period in two components: a selection and a price effect. Finally, we consider the evolution of wage inequality from the 1993-1998 panel to the next 2002-2006. This evolution can be explained by changes in the distribution of workers' comparative advantage θ_i and by changes in relative returns to workers' comparative advantage α.

5.1 Recovering the distribution of workers' comparative advantage

If we note $\widehat{Y_{it}}$ the earnings predicted by our model for individual *i* at date *t*, θ_i is estimated using the following expression:

$$\widehat{\theta_i} = \frac{1}{2}\left(\frac{Y_{it} - \widehat{Y_{it}}}{1 + (\widehat{\alpha} - 1)P_{it}} + \frac{Y_{it+1} - \widehat{Y_{it+1}}}{1 + (\widehat{\alpha} - 1)P_{it+1}}\right)$$

The distribution of θ in the different employment groups is displayed in the two graphs 3 and 4. Positive selection is evident from the fact that the distribution of unobserved heterogeneity of "public stayers" stands to the right of the distribution of "private stayers". The distributions for the two groups of "sector switchers" lie in-between, because they have better skills than the average of private employees, and worse than the average of public employees.

5.2 Explaining the average public-private earnings gap

Our model allows us to decompose the public-private wage gap at each date into two components. The first is due to differences in workers' skills between the two sectors (selection effect), the second is due to the constant earnings difference between public and private employees, and to difference in returns to productive skills in the two sectors (price effect)[13]. Using our estimates, we build the predicted distribution of wages for each panel:

$$\begin{aligned} \widehat{y_{it}} &= \widehat{\tau_t} + \widehat{\eta_t} + \widehat{\alpha} * \widehat{\theta_i} \quad \text{if } P_{it} = 1 \\ \widehat{y_{it}} &= \widehat{\tau_t} + \widehat{\theta_i} \qquad\qquad \text{if } P_{it} = 0 \end{aligned}$$

From comparing each predicted distribution to the real one we obtain residuals $\widehat{\epsilon_{it}}$ and then compute two counterfactual distributions. The first counterfactual simulates public employees' wage without the public premium η_t and with returns to characteristics similar to the private sector. The second simulates private employees' pay if they were to receive the public premium and the same returns to their productive skills α as public employees.

$$\begin{aligned} Counterfactual\ 1 &= \widehat{\tau_t} + \widehat{\theta_i} + \widehat{\epsilon_{it}} \qquad \text{for } P_{it} = 1 \\ Counterfactual\ 2 &= \widehat{\tau_t} + \widehat{\eta_t} + \widehat{\alpha} * \widehat{\theta_i} + \widehat{\epsilon_{it}} \qquad \text{for } P_{it} = 0 \end{aligned}$$

[13] Although the price effect could be decomposed into two effects: the effect of the constant premium on the one hand, the effect of relative returns to workers' comparative advantage on the other, we keep the two categories to ease the comparison with the results previously obtained with the Oaxaca-Blinder method.

We first compare public employees' observed wage to counterfactual 1: this gives us the pure effect of public-private differences in returns to workers' skills, because the workers' characteristics are held constant (Decomposition 1 in Table 9). Conversely, Decomposition 2 compares private employees' observed wage to counterfactual 2. In line with the rest of the literature, we find that selection on workers' characteristics, both observable and unobservable, is a major component of the public-private earnings gap in Vietnam throughout the period. According to the first decomposition, in 1993-1998, positive selection into the public sector accounted for a 50% difference in public-private earnings, its effect decreased to 35% in 2002-2006.
Along with the effect of selection, our results emphasize the impact of returns to observable and unobservable characteristics (α) and of the constant premium from being in the public sector (η) on the public-private sector earnings gap. Both elements had a negative contribution to the public-private gap in 1993-1998: it means that inequality would have increased in the 1990s, would the public sector pay its workers the same as the private sector. In 2002-2006, their contribution has turned into a positive one. The subsequent reform in public sector's compensation pattern unambiguously widened the public-private wage gap.

5.3 Explaining changes in overall wage inequality

We now turn to the broader question of how public-private differences and their change over time impacted the overall distribution of earnings. To simplify, we will focus only on the first and last dates for which we have data, 1993 and 2006. We will note respectively τ^{93}, η^{93}, α^{93} and τ^{06}, η^{06}, α^{06} the constant, the public premium and the price factor on individual characteristics previously estimated for each date in their respective panel. P_i^{93} and P_i^{06} are dummies for being in the public sector at each date. Predicted distributions and residuals are obtained as before, with residuals now noted ϵ_i^{93} and ϵ_{i1}^{06}. Finally we simulate changes in the distribution of θ through a rank-preserving transformation, that is we assign to each 1993 (respectively 2006) individual the value of θ that has the same rank in 2006 (respectively 1993) distribution as the value of θ that we estimated for her in 1993 (respectively 2006) distribution. We note these values respectively θ_i^{93} and θ_i^{06}.
We now have all ingredients to build the three following counterfactuals[14]:

$$\begin{aligned}
Log\ wage\ in\ 1993 &= \widehat{\tau^{93}} + (1 + (\widehat{\alpha^{93}} - 1) * P_i^{93}) * \widehat{\theta_i} + \widehat{\eta^{93}} * P_i^{93} + \widehat{\epsilon_i^{93}} \\
Counterfactual\ 1 &= \widehat{\tau^{06}} + (1 + (\widehat{\alpha^{93}} - 1) * P_i^{93}) * \widehat{\theta_i} + \widehat{\eta^{93}} * P_i^{93} + \widehat{\epsilon_i^{93}} \\
Counterfactual\ 2 &= \widehat{\tau^{06}} + (1 + (\widehat{\alpha^{93}} - 1) * P_i^{93}) * \widehat{\theta_i^{06}} + \widehat{\eta^{93}} * P_i^{93} + \widehat{\epsilon_i^{93}} \\
Counterfactual\ 3 &= \widehat{\tau^{06}} + (1 + (\widehat{\alpha^{06}} - 1) * P_i^{93}) * \widehat{\theta_i^{06}} + \widehat{\eta^{06}} * P_i^{93} + \widehat{\epsilon_i^{93}} \\
Log\ wage\ in\ 2006 &= \widehat{\tau^{06}} + (1 + (\widehat{\alpha^{06}} - 1) * P_i^{06}) * \widehat{\theta_i} + \widehat{\eta^{06}} * P_i^{06} + \widehat{\epsilon_i^{06}}
\end{aligned}$$

The first counterfactual adds to wages in 1993 the average increase in wages between 1993 and 2006. Under the fundamental assumption that the rank in the distribution of θ adequately describes each individual characteristics, the second counterfactual estimates how the 1993 employees, with 2006 employees characteristics, would have been paid if the allocation process and the difference in compensation patterns had not changed between the public and the private sectors. The third one further changes both

[14] We also built the 3 symmetric counterfactuals, using the distribution of wages in 2006 as a starting point. With the order of decomposition, our results do change quantitatively but not qualitatively.

the constant premium and the price of individual characteristics in the public sector as compared to the private sector. All three counterfactuals, with the real distributions for 1993 and 2006 are drawn in Figure 5 and 6.
Starting from the real 1993 earnings distribution, the change in the constant τ shifts the curve to the right (Counterfactual 1). Our decomposition show that changes in compensation patterns in the public sector as compared to the public sector had an inequality increasing effect (Counterfactual 2). But this effect is offset by a shrinking dispersion of individual characteristics, which explains why the 2006 distribution looks so similar to the 1993 distribution (Counterfactual 3). If we compare this last counterfactual with the real distribution for 2006 individuals, we see at least two important reasons for which they could differ. First, if the allocation process between the public and the private sector conditional on individual characteristics has changed over time. Second, if the distribution of exogenous shocks is not the same for the two periods. Interestingly, these two factors taken together do not seem to explain a large public-private gap.

5.4 Discussion and extension

Our model confirms and elaborate on the results drawn with the Oaxaca-Blinder method described in Section 4.2: selection on workers' characteristics is important but returns to skills are the driving factor behind the widening public-private earnings gap. The value added by our model is that it offers a much more comprehensive view of selection on workers' characteristics, and hence strengthens the case against it playing a major role in the evolution observed throughout the 1993-2006 period. The fact that α is significantly inferior to 1 in the first panel suggests that neither a simple wage equation model nor a more elaborated random effect model (which de facto assumes constant returns to unobserved characteristics across sectors) could adequately describe differences in returns to workers' skills between the public and the private sector.
A potential threat to our conclusions is selective attrition. If the most productive workers in the private sector dropped out of the panel in the 2002-2006 period, we would be underestimating the average and the dispersion of private sector earnings. In order to affect the results we draw by comparing the 1993-1998 period to the 2002-2006 period, however, the attrition would have to be present in the latter but not in the former panel. Since the panel and the cross-sectional sample have similar earnings distribution at each date, the data we dispose of do not support this argument. It is still possible that the cross-sectional sample itself is biased. Unfortunately, we do not have any means to gauge how biased the 1993-1998 sample is, because there is no other data source on wages during this period that we could use for comparison purposes.
Another shortcoming of our method lies in the fact that the effect of observed characteristics such as education, experience and gender is subsumed under a single index. Table 10 shows correlation coefficients between the estimated θ_i and individual characteristics. The highest coefficient is for years of schooling, followed by residence (urban or rural), and experience. This reinforces the intuition that our method is capturing the observable characteristics one usually includes in earnings models together with some unobservable characteristics that are strongly correlated with them. These unobserved factors contributing to workers's comparative advantage could include cognitive abilities, social capital or political connections.
Out of all observable characteristics, gender may play a particular role in the evolution of wage inequality during the transition. As noted earlier, the public sector discriminates less against women than the private sector. This phenomenon may interfere with the increase of returns to workers' skill in the public sector relatively to the private

sector that we established before. Following the specification described in section 3.4, we estimate our model with different coefficients for female and male workers. As expected, the estimated increase in returns to skills in the public sector is higher than in the first specification: relative returns to female workers' skills rose dramatically from .38 to 2.1 (see Table 11). The increase for men has been more moderate (as the negative coefficient on γ_1 shows). Even though the lack of precision is more acute, these results seem to confirm the mechanisms explained above.

6 Conclusion

In order to explain differences in earnings between the public and the private sectors, the two usual candidates are differences in workers' skills and differences in returns to workers' skills in the two sectors. The two factors have very different interpretations, and a considerable amount of research has been devoted to disentangle the former, which suggests sorting by abilities, from the latter, which gives evidence of labor market segmentation. A plausible explanation of the rising public-private wage gap between 1993 and 2006 was that selection into this public sector had increased, with a greater flexibility in hiring and firing workers in the public sector. However, using a model of comparative advantage and differential returns between the two sectors, we found that sorting seemed to decrease between the first and the second period. Workers became more homogeneous across sectors. In order to explain the increased between-sector inequality, we showed that, in the first period, workers with a comparative advantage in the public sector were paid less (with respect to their skills) in the public than in the private sector. This is coherent with the existence of an institutional preference for low wage inequality in the public sector, similar to the unionized sector studied by Lemieux (1998). It could also be linked to the existence of (to the econometrician) unobserved benefits from being in the public sector, such as job stability or social influence, or unobserved costs of joining that sector, such as passing the examination or having personal connections with the administration. In the second period, on the opposite, workers with a comparative advantage to be in the public sector were indeed earning higher returns in the public than in the private sector. It seems a common feature of the transition to market economies that workers' compensation became more closely related to their productivity in each sector (see Keane and Prasad 2006). Unlike the "sorting scenario", our explanation does not imply that a massive reallocation of workers happened between the 1990s and the 2000s, and anecdotal evidence suggest that it did not. It involves a policy change in the public sector, with more productive workers receiving higher benefits. This change could be explained by pure economic considerations, such as the need for the public sector to retain its workers, or it could be due to the political economy of the reform process, which needed a strong support from State employees. Finally, further research is needed to understand how the distribution of workers' productive skills became more homogeneous, leading to an overall decrease in wage inequality. It could be simply due to the lack of representativeness of the sample we use: unregistered migrants are not covered by the VHLSS-VLSS survey. More interestingly, it may be explained by increased selection of workers into the labor market, with a better sorting on ability between self-employment and wage work. Alternatively, it could be a positive feature of Vietnam's transition towards market economy that the increasing valuation of skills by employers was accompanied by an improvement and homogenization of workers skills.

References

[1] M. Azam, N. Prakash (2010) "A Distributional Analysis of the Public-Private Wage Differential in India", *IZA Discussion Paper*, No. 5132.

[2] S. Bales and M. Rama (2001), "Are Public Sector Workers Underpaid? Appropriate Comparators in a Developing Country" *World Bank Policy Research Working Paper*, No. 2747.

[3] O. Bargain and B. Melly (2008), "Public Sector Pay Gap in France: New Evidence using Panel Data" *IZA Working Paper*, No. 3427.

[4] P. Belser and M. Rama (2001), "State Ownership and Labor Redundancy? Estimates based on Enterprise-Level Data from Vietnam" *World Bank Policy Research Working Paper*, No. 2599.

[5] R. Disney and A. Gosling (1998), "Does It Pay to Work in the Public Sector?" *Fiscal Studies*, Vol.19, No. 4, pp. 347-374.

[6] J.-P. Eglinger (2005), *Genèse, situation actuelle et perspectives d'avenir des compagnies générales (Töng cöng Ty) en République socialiste du Vietnam à travers les sources vietnamiennes*, Thèse d'histoire sous la direction de Michel Fournié, Institut National des Langues et Civilisations Orientales.

[7] R. Frydman, C. Gray, M. Hessel and A. Rapaczynski (1999), "When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies" *The Quarterly Journal of Economics*, Vol. 114, No. 4, pp. 1153-1191.

[8] J. L. Gallup (2002), "The Wage Labor Market and Inequality in Vietnam in the 1990s" *World Bank Policy Research Working Paper*, No. 2896.

[9] Glinskaya and Lokshin (2007), "Wage Differentials Between the Public and the Private Sectors in India" *Journal of International Development*, Vol. 19, Issue 3, pp. 333-355.

[10] A. Gosling and T. Lemieux (2001), "Labor Market Reforms and Changes in Wage Inequality in the United Kingdom and the United States" *NBER Working Papers*, No. 8413.

[11] M. Iyigun and D. Rodrik (2004), "On the efficacy of reforms: Policy Tinkering, Institutional Change, and Entrepreneurship", *NBER Working Papers*, No. 10455

[12] M. P. Keane, E. S. Prasad (2006), "Changes in the structure of earnings during the Polish transition" *Journal of Development Economics*, Vol. 80, pp 389Ű 427.

[13] T. Lemieux (1993), "Estimating the Effects of Unions in a Two-Sector Model with Comparative Advantage and Non-random Selection." *Working Paper: Centre de Recherche et Développement en Economique,Montreal*, No. 0493.

Journal of Labor Economics, Vol. 16, No. 2, pp. 261-291.

[14] T. Lemieux (1998), "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Nonrandom Selection" *Journal of Labor Economics*, Vol. 16, No. 2, pp. 261-291.

[15] T. Lemieux (2006), "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *The American Economic Review*, Vol. 96, No. 3, pp. 461-498.

[16] A. Y. C. Liu. (2004), "Sectoral gender wage gap in Vietnam" *Oxford Development Studies*, Vol. 32, Issue 2, pp. 225-240.

[17] J. Mac Millan and C. Woodruff (1999), "Interfirm Relationships and Informal Credit in Vietnam" *Quarterly Journal of Economics*, Vol. 114, No. 4, pp. 1285-1320.

[18] J. Machado and J. Mata (2005), "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression", *Journal of Applied Econometrics*, Vol.20, 445-465.

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

[20] M. Painter (2005), "The Politics of state sector reforms in Vietnam: Contested Agendas and uncertain trajectories" *Journal of Development Studies*, Vol. 41, No. 2, pp. 261-283.

[21] A. Shleifer and R. W. Vishny (1994), "Politicians and Firms" *The Quarterly Journal of Economics*, Vol. 109, No. 4, pp. 995-1025.

[22] P. Siminski (2008), "What Would the Average Public Sector Employee be Paid in the Private Sector?", *University of Wollongong, Economics Working Paper Series*, WP 08-05.

[23] T. Suri (Forthcoming), "Selection and Comparative Advantage in Technology Adoption", *Econometrica*.

[24] Vietnam's Socio-Economic Development (2008), *Vietnam's Socio-Economic Development: A Social Science Review No.55 September 2008*, Vietnam Institute of Economics, Hanoi.

Tables and Figures

Table 1: **Employment sector of the main occupation for working adults**

	Urban		Rural	
	1993	2006	1993	2006
Public employment	20.71%	23.23%	3.54%	6.28%
Private employment	8%	20.45%	20.06%	29.94%
Self-employed	88.24%	73.27%	59.22%	46.83%

Source: Author's calculations on VLSS data 1993 and VHLSS 2006.

Table 2: **Theil 1 Entropy index**

	1993	**1998**	**2004**	**2006**
Within public	0.25	0.25	0.26	0.20
Within private	0.31	0.24	0.22	0.20
Within sectors	0.28	0.24	0.24	0.20
Between sectors	0.00	0.00	0.04	0.03
Total	0.29	0.24	0.28	0.23

Source: Author's calculations on VLSS data 1993 and VHLSS 2006.

Table 3: **Descriptive statistics on public and private sector workers**

	1993			2006		
	Public	Private	Diff	Public	Private	Diff
Years of schooling	10.77	5.87	4.91	11.02	7.54	3.48
	(3.42)	(3.62)	***	(2.01)	(3.48)	***
Years of experience	14.76	15.29	-0.53	19.84	17.62	2.23
	(9.02)	(12.31)		(12.31)	(13.04)	***
Male worker	0.50	0.61	-0.11	0.56	0.63	-0.07
	(0.50)	(0.50)	***	(0.50)	(0.48)	***
Vocational training	0.32	0.04	0.28	0.37	0.12	0.25
	(0.47)	(0.20)	***	(0.48)	(0.32)	***
Number of workers	881	1401		2387	5227	

*Source: OLS estimates on VLSS 1993, 1998, VHLSS 2002, 2004 and 2006 surveys data. Standard deviation in parentheses. Stars denote significance levels of differences in means: *** for 1% significance level, ** for 5% and * for 10%.*

Table 4: **Wage equations in the public and private sector in 1993 and 2006**

Private sector	1993			2006		
	Coefficient	s.d.	Signif.	Coefficient	s.d	Signif.
Years of Schooling	0.021	0.010	**	0.046	0.005	***
Years of Experience	0.046	0.009	***	0.024	0.004	***
Experience square	-0.103	0.019	***	-0.001	0.000	***
Male	0.274	0.067	***	0.213	0.030	***
Vocational training	-0.200	0.135		0.190	0.051	***
South	0.440	0.082	***	0.117	0.031	***
Urban area	0.099	0.064		0.011	0.030	
Observations		438			1213	
Public sector	1993			2006		
	Coefficient	s.d.	Signif.	Coefficient	s.d.	Signif.
Years of Schooling	0.040	0.009	***	0.067	0.010	***
Years of Experience	0.024	0.011	**	0.022	0.006	***
Experience square	-0.032	0.030		0.000	0.000	
Male	0.091	0.059		-0.018	0.044	
Vocational training	0.007	0.062		-0.061	0.070	
South	0.289	0.059	***	0.130	0.044	***
Urban area	0.044	0.059		0.099	0.044	**
Observations		397			806	

*Source: Author's calculations on VLSS 1993, 1998, VHLSS 2002, 2004 and 2006 surveys data. Stars denote significance levels for t-test of mean equality between the public and the private sector: *** for 1% significance level, ** for 5% and * for 10%.*

Table 5: **Oaxaca Blinder decomposition of the public-private wage gap**

	1993	**1998**	**2002**	**2006**	**1993-2006**
Real Gap	0.16	0.23	0.44	0.51	0.35
Decomposition 1					
Price effect	0.20	0.19	0.27	0.37	0.17
Selection	-0.06	0.03	0.20	0.14	0.20
Decomposition 2					
Price effect	0.01	-0.15	0.16	0.25	0.25
Selection	0.14	0.37	0.31	0.25	0.12

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006.

Table 6: **Panel sample broken down by employment history**

	1993-1998		2002-2006	
	Number	Percent	Number	Percent
Private stayers	778	54.03	2,478	61.67
Movers to Public	210	14.58	236	5.87
Movers to Private	51	3.54	115	2.86
Public stayers	401	27.85	1,189	29.59
Total	1,440	100	4,018	100

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006.

Table 7: **Employment history and earnings gains**

	1993-1998			2002-2006		
	t	t+1	Change	t	t+1	Change
Private stayers	0,41	0,86	**0,45**	1,14	1,39	**0,25**
	(0,66)	(0,55)		(0,60)	(0,54)	
Movers to Public	0,13	0,71	**0,58**	1,18	1,47	**0,29**
	(0,65)	(0,56)		(0,80)	(0,82)	
Movers to Private	0,37	1,10	**0,73**	1,43	1,62	**0,19**
	(0,81)	(0,59)		(0,73)	(0,67)	
Public stayers	0,55	1,13	**0,58**	1,62	1,96	**0,34**
	(0,60)	(0,60)		(0,66)	(0,65)	
Total	0,46	1,00	**0,54**	1,34	1,62	**0,28**
	(0,65)	(0,60)		(0,68)	(0,66)	

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006. Average Log hourly earnings (Standard errors in parentheses). In the 2002-2006 part of the table, the t column refers to 2002 for 2002-2004 panel individuals, and to 2004 for 2004-2006 panel individuals.

Table 8: **Non linear least square estimation results**

	1993-1998			2002-2006		
	Estimate	s. d.	Sign.	Estimate	s. d.	Sign.
Factor α	0.454	0.272	*	1.584	0.902	*
Movers to public m_{01}	-0.462	0.163	***	-0.075	0.087	
Movers to private m_{10}	0.058	0.121		0.112	0.036	***
Private stayers m_{00}	-0.177	0.127		-0.115	0.039	***
Time 0 dummy τ_0	0.592	0.133	***	1.253	0.043	***
Time 1 dummy τ_1	1.040	0.126	***	1.503	0.039	***
Public premium at time 0 η_0	-0.118	0.158		0.083	0.087	
Public premium at time 1 η_1	-0.251	0.153		-0.005	0.087	
Sample size	1700			6156		

*Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006. Stars are used to denote significance levels: *** for 1% significance level, ** for 5% and * for 10%.*

Table 9: **Decomposition of the public private wage gap based on our simulations**

	1993	**1998**	**2002**	**2006**	**1993-2006**
Real Gap	0.16	0.19	0.47	0.53	0.37
Decomposition 1					
Price effect	-0.48	-0.32	0.10	0.18	0.66
Selection	0.64	0.51	0.37	0.35	-0.29
Decomposition 2					
Price effect	-0.13	-0.04	-0.06	0.03	0.16
Selection	0.29	0.23	0.53	0.50	0.21

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006.

Table 10: **Correlation of the estimated individual fixed effect θ_i with observable characteristics**

	1993 -1998	2002-2006
Male worker	0.0897***	0.0958***
Years of schooling	0.3464***	0.3993***
Years of experience	0.1227***	0.0769***
Urban	0.1814***	0.2501***
Vocational training	0.076***	0.1611***
South	0.0626***	-0.0168

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006.

Table 11: **Extended model with gender as an explanatory variable**

	1993			2006		
	Coef.	Std. Err.	Sign.	Coef.	Std. Err.	Sign.
Factor α	0.389	0.328		2.113	0.911	**
Movers to public m_{01}	-0.663	0.385	*	-0.103	0.083	
Movers to private m_{10}	-0.118	0.322		0.102	0.042	**
Private stayers m_{00}	-0.379	0.320		-0.193	0.053	***
Time 0 dummy τ_0	0.589	0.328	*	1.169	0.055	***
Time 1 dummy τ_1	1.037	0.324	***	1.411	0.052	***
Public premium at time 0 η_0	-0.128	0.341		0.001	0.121	
Public premium at time 1 η_1	-0.260	0.335		-0.079	0.121	
Male premium (private) δ_0	0.337	0.121	***	0.220	0.056	***
Male premium (private) δ_1	0.142	0.093		0.306	0.105	***
Male difference in returns (private) γ_0	0.114	0.298		-0.258	0.264	
Male difference in returns (public) γ_1	-0.054	0.209		-1.161	0.608	*
Sample size	1700			6156		

Source: Author's calculations on VLSS 1993, 1998 VHLSS 2002 2004 2006.

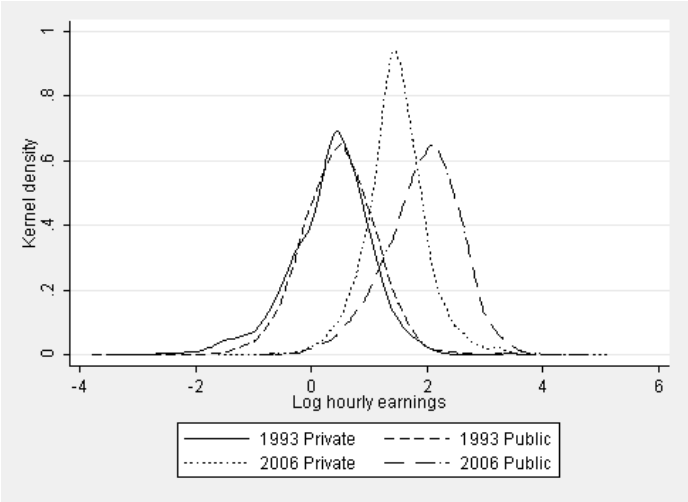

Figure 1: Changes in the distribution of log hourly earnings in the public and private sectors between 1993 and 2006 (Kernel estimates)

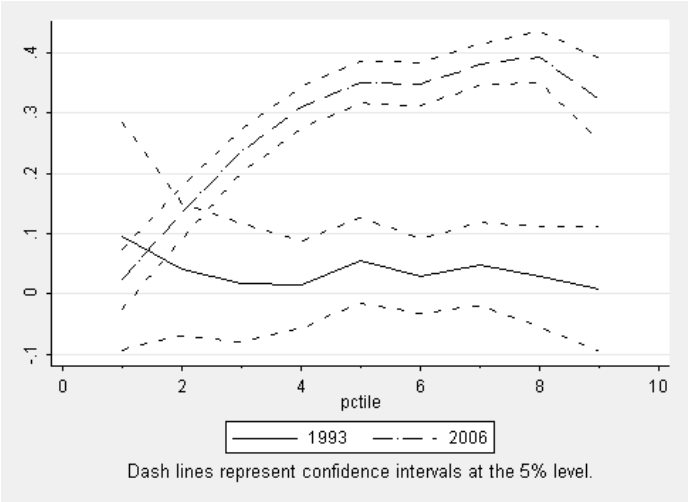

Figure 2: Public premium estimated in a wage equation for different earnings quantiles

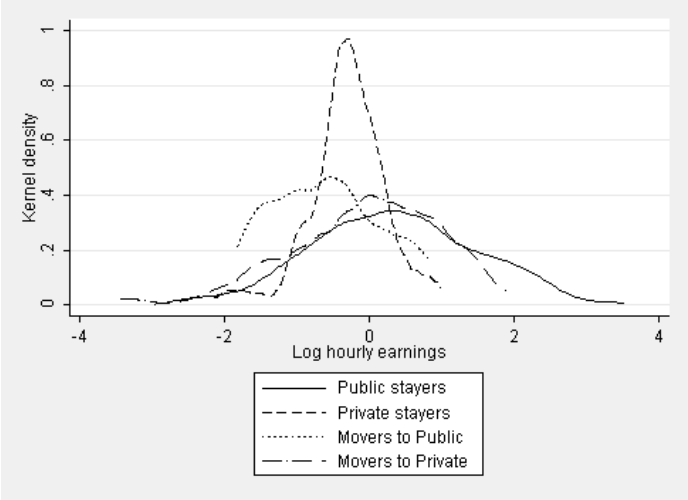

Figure 3: Distribution of unobserved heterogeneity for the 4 employment groups in 1993-1998

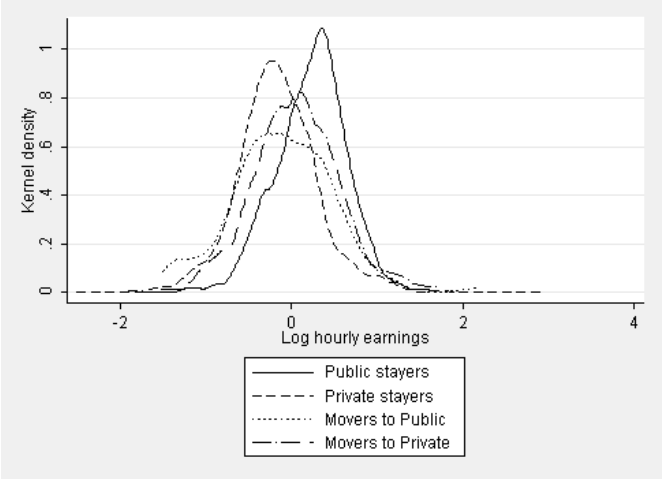

Figure 4: Distribution of unobserved heterogeneity for the 4 employment groups in 2002-2006

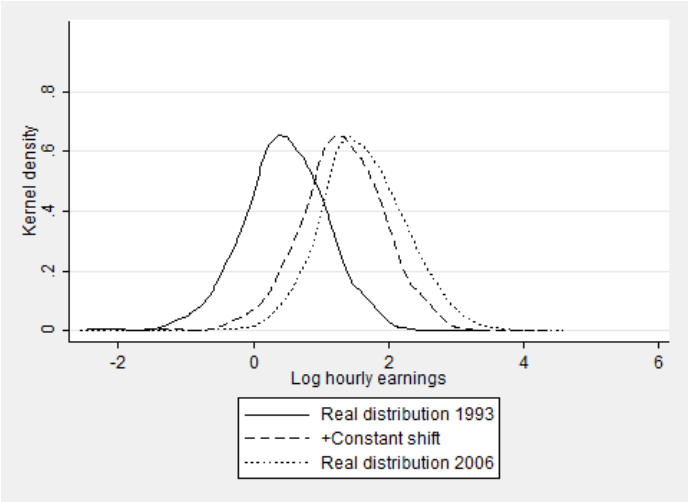

Figure 5: Whole distribution of wages in 1993 and 2006, with two counterfactuals

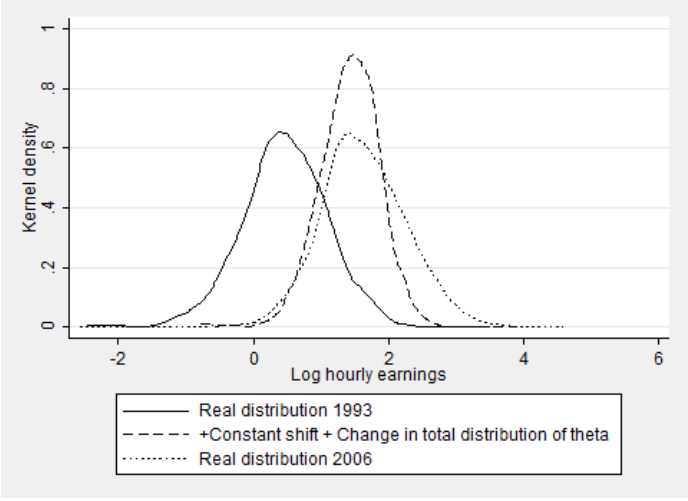

Figure 6: Whole distribution of wages in 1993 and 2006, with two counterfactuals