

# Decomposing the Labor Market Earnings Inequality

The Public and Private Sectors in Vietnam, 1993–2006

*Clément Imbert*

The World Bank  
Development Economics Vice Presidency  
Partnerships, Capacity Building Unit  
January 2013



## Abstract

: In contrast with the typical transition to a market economy, earnings inequality in Vietnam between 1993 and 2006 appears to have decreased, and the earnings gap in favor of public employees appears to have widened. The paper uses a comparative advantage model to disentangle the effect of sorting workers across sectors from the effect of the differences in returns to workers' skills. The selection of the best workers into the public sector is clearly an important component of the explanation for the public-private sector earnings gap,

but the widening of this gap over time is primarily due to changes in the compensation patterns. The paper finds that, in the 1990s, public employees were underpaid compared with their earning potential in the private sector whereas, in the early 2000s, public employees earned similar returns to their comparative advantage in the public and private sectors. The increasing homogeneity in returns to skills in the Vietnamese labor market appears to explain both the increase in the public-private pay gap and the decrease in overall inequality.

---

This paper is a product of the Partnerships, Capacity Building Unit, Development Economics Vice Presidency. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at [imberty@pse.ens.fr](mailto:imberty@pse.ens.fr).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Decomposing the Labor Market Earnings Inequality: The Public and Private Sectors in Vietnam, 1993–2006

*Clément Imbert\**

Keywords: Labor market inequality, Public sector, Transition.

JEL classification codes: J45, J31, P31

Sector Board: Public Sector Governance (PSM)

\* Clément Imbert is a PhD student at the Paris School of Economics, 48 bd Jourdan 75014 Paris. His email address is [imbert@pse.ens.fr](mailto:imbert@pse.ens.fr). The author wishes to thank Elisabeth Sadoulet, Abhijit Banerjee, François Bourguignon, Gary Fields, Thomas Lemieux, Martin Rama, Martin Ravallion, and Dominique Van de Walle as well as three anonymous referees for their helpful comments. This paper also benefited from the questions of the participants at the Northeast Universities Development Consortium Conference, the European Economic Association Conference, and the Paris School of Economics Development Lunch. Thanks also to Bryan McCaig for sharing his codes on the VHLSS panels and to Thomas Breda, Barbara Coello, Lucie Gadenne, and Luke Haywood for many fruitful discussions. Finally, the author is extremely grateful to Marc Gurgand for his patient supervision.

What happens to labor market earnings inequality during the transition from a communist to a communist capitalist regime? When we attempt to understand Vietnamese history over the last two decades, the lessons from previous experiences in Eastern Europe of transitions to a market economy might be misleading. Vietnam did not experience massive layoffs from public firms or ownership transfers to private stakeholders (Frydman et al. 1999). The Communist Party Congress of 2001 decided against the drastic privatization agenda that was proposed by the World Bank and the International Monetary Fund in the aftermath of the 1998 financial crisis (see Painter 2005). Instead, the Communist Party entrusted public enterprises, rebaptized as “general companies,” with a leading role in Vietnam's development strategy (Eglinger 2005; Fforde 2007). Private enterprises, in contrast, were only given legal existence gradually during the 1990s and were slow to develop in a very unfavorable environment (Mac Millan and Woodruff 1999). In this sense, Vietnam's transition toward a market economy is similar to China's transition.

Another important difference between Vietnam and the typical transition case is the small size of the labor market: in the 1990s, Vietnam was a predominantly agrarian economy composed of self-employed farmers. Major developments in inequality and poverty were brought about by land distribution (Ravallion and van de Walle 2006) and trade liberalization, which boosted rice exports (McCaig 2009). In turn, these groundbreaking changes in (mostly self-employed) agricultural production fueled the growth of the wage employment sector, as shown by Edmonds and Pavcnik (2006). In this sense, “scrutinizing the evolution of the labor market in the 1990s gives us clues about how economic development in Vietnam will continue to affect households and

society in the future” (Gallup 2002). The present paper focuses on one prominent feature of the Vietnamese labor market: the public-private pay gap.

Our aim is to identify the effect of the public sector reform on the labor market earnings inequality. From the literature on public labor markets, two opposing views can be supported. On the one hand, if there is a fixed “public sector premium” that is given to workers independently of their productive characteristics, the public-private differences increase inequality. Hence, a reform that would make public sector pay similar to competitive wage setting would reduce inequality (which was the expected benefit of the public sector reform in Vietnam, as detailed in Bales and Rama (2001)). On the other hand, if the public sector wage setting compresses the wage differentials between the skilled and unskilled workers, given that the public sector workers' skills are higher, on average, then public sector reform could increase inequality (see Gosling and Lemieux 2001). Similarly, Liu (2004) argues that the wage gap between male and female workers, which is typically lower among public employees, could rise in Vietnam after the public sector reforms.

Identifying the effect of any changes in wage setting on inequality is an empirical challenge. In particular, compositional effects must be controlled for because workers have different productive characteristics across sectors. Not all of these characteristics are observed, and it is difficult to account for selection on the unobservable characteristics. Fortunately, panel data provide an opportunity to credibly address this issue by identifying individual, specific, and time-invariant components of workers' pay. However, fixed effects models, which are most commonly used with panel data, implicitly assume that unobserved skills have the same returns in both sectors. Our model

goes one step further by introducing a latent variable of comparative advantage, which makes workers more or less productive in the public as compared with the private sector and allows for different returns in the public and the private sectors.<sup>1</sup> This latent variable summarizes the effect of all of the characteristics, observed or unobserved, that are important in workers' selection between sectors. Once we have estimated the model, we are able to simulate the counterfactual wage distributions that separately show the effect of selection (the differences in workers' characteristics) and the effect of wage-setting policies (the differences in the returns to these characteristics).

We show that the public-private gap in hourly earnings increased dramatically between the 1990s and the following decade. Consistent with the literature, we find that selection is a major component of the between-sector differences: public employees have a higher comparative advantage when working in the public sector than private employees do. However, the widening of the public-private sector gap is mostly driven by changes in compensation patterns in the public sector as compared with the private sector. Given their comparative advantage, public sector workers were underpaid in comparison with their private counterparts in the 1990s; however, in the 2000s, they were paid at least as well as private sector employees.

Our paper is structured as follows. Section 1 describes our data and documents the increase in the private-public wage gap between the 1990s and the 2000s and the decrease in overall wage inequality. Section 2 explains our empirical strategy. Section 3 presents our results and compares them to the results from other methods that are commonly used to study between-sector wage inequality. Section 4 draws counterfactual distributions to decompose the public-private earnings gap. Section 5 suggests an

interpretation for the decline in overall wage inequality. In section 6, we discuss the robustness of our results, and section 7 concludes the paper.

## **STYLIZED FACTS**

In this section, we describe the data that we use and present a few stylized facts on labor markets in Vietnam. Wage work represents a small but growing fraction of the total labor force. Among wage earners, public sector employees are the best paid, and this advantage has increased over time. Despite the increase in inequality between the public and private sectors, overall wage inequality has decreased because within-sector inequality has decreased.

### *Data*

Our analysis depends on the availability of reliable data with a panel dimension. We use two separate datasets, the Vietnam Living Standard Surveys (VLSS) for 1993 and 1998 and the Vietnam Health and Living Standard Surveys (VHLSS) for 2002, 2004, and 2006. These high-quality household surveys were implemented by the General Statistical Office (Hanoi) in collaboration with the World Bank. For the VLSS, the households surveyed in 1993 were interviewed for a second time in 1998. The VHLSS is a separate panel with a rotative design so that most households were interviewed twice, either in 2002 and 2004 or in 2004 and 2006. Very few households were surveyed three times. For the purposes of our analysis, we add the 2002–2004 and the 2004–2006 observations into a single dataset, which, by abuse of language, we will call the “2002–2006” panel and compare it to the 1993–1998 panel. This manipulation is necessary to ensure that we observe a sufficient number of transitions between employment sectors.

The VLSS and VHLSS are representative of Vietnam for a given year provided that we use the appropriate sample weights. A multiple-stage sampling design was used, and within each geographic sample unit, households were randomly drawn from the official population registers. For every second round, additional households were sampled to compensate for attrition from the panel. Selection into the survey sample was a concern; workers who had migrated to urban areas without a residency permit were not sampled. These workers were presumably employed in the private sector. Thus, the sample that we used may provide a biased view of the earnings distribution in that sector. To test for the representativeness of our sample, we compared the repartition of the workforce in the VLSS 1998 with the statistics published by the General Statistical Office based on the Labor and Job Survey of 1997. The repartition across economic sectors was exactly the same: 65 percent in agriculture, 13 percent in industry, and 22 percent in services. The proportion of public workers was smaller in the VLSS, which conflicts with the hypothesis of a sampling bias for private sector jobs.<sup>2</sup>

The issue of selective attrition from the panel sample is also important. Workers who were surveyed on two successive dates and who were working for wages on both dates were likely to differ from those who were not. If we compare the individual characteristics of the workers in the panel sample to those of the workers in the cross-section, the only significant difference is the years of experience, as expected. In contrast, the repartition of the workforce by sector in the panel sample is clearly biased toward the more stable forms of employment, and public employees and white-collar workers are overrepresented. This overrepresentation implies that the attrition bias attenuates the differences between the public and the private sectors and, in general, works against our

findings. If selection decreases over time, however, it may explain the rising public-private gap that we document. When comparing public and private earnings distributions in the panel sample and in the entire sample, we do not find a stronger positive selection bias for the private workers than for the public workers. Hence, the selection bias should not affect the public-private sector comparisons.

### *Labor markets in Vietnam*

This paper focuses on wage work, which comprises a small but growing share of total employment. If we assign each worker to the sector of his or her main occupation in the last 12 months, we obtain the allocation described in table 1. In 1993, 10.7 percent of the total workforce in rural areas worked for wages, and in 2006, 23.9 percent worked for wages. Work for wages is also increasing in urban areas, from 39.6 to 48.3 percent.<sup>3</sup> Within wage employment, we select the public sector or the private sector according to the workers' declarations about their employers: households, private companies, and cooperatives are considered private employers. The legal framework of private enterprises evolved during the 1990s, which translates into subtle changes in the survey nomenclature that must be accounted for.<sup>4</sup> The proportion of workers employed in the public sector increased from 3 to 6 percent in rural areas and from 20 to 21 percent in urban areas. The number of private employees increased as well, so the public share of the labor market remained stable: about one-third in rural areas and one-half in urban areas.

TABLE 1. Workforce by Employment Sector (%)				
	1993		1998	
	Rural	Urban	Rural	Urban
Public employees	3.12	20.01	5.90	21.51
Private employees	7.60	19.58	18.03	26.82
Self-employed	89.26	60.37	76.07	51.67

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

Public sector and private sector workers are engaged in very different economic activities. A growing majority of public employees work in government, education, and health services (51 percent in 1993, 62 percent in 2006) whereas only 8 percent of private employees participated in these sectors in 1993. The remainder of public employees work in public firms, which operate in many industries: electricity and water production, mining, food and beverages, textiles, and other fields. Even when they operate in the same industry, however, public and private firms have very different characteristics. In 1998, the average public worker in the paper, metal, and plastics industries had 200 coworkers whereas the average private sector worker in this industry had only 35 coworkers. According to the 2006 Enterprises Survey, public firms are more capitalistic and generate more profits compared with private domestic firms (Vietnam Socio-Economic Development 2008). Within the private sector, foreign-owned enterprises are an exception, with a much larger size and more capital compared with domestic firms, but they represent only a small minority of the employed workers in the period that we consider (up to 5 percent in 2006).

#### *Earnings in the public and private sectors*

On average, public employees earn more compared with private sector workers. In nominal terms, public employees earned 1460 dong per hour in 1993, and private sector workers earned 1320 dong. In 2006, public employees earned 9320 dong per hour, and

private sector workers earned 5750 dong per hour (see table 2). Our measure of earnings is the hourly compensation declared by the workers and includes the wages and benefits received from employers in cash and in kind.<sup>5</sup> The ability to account for workers' benefits is crucial for the purpose of this study. Without benefits, public employees' compensation would be much lower than the compensation of private sector employees (see table 2). Interestingly, wages, rather than benefits, drive the increase in the public-private pay gap. Finally, the last row of table 2 shows that the number of hours worked increased slightly more in the private sector, contributing to lower growth in hourly earnings compared with the public sector.<sup>6</sup>

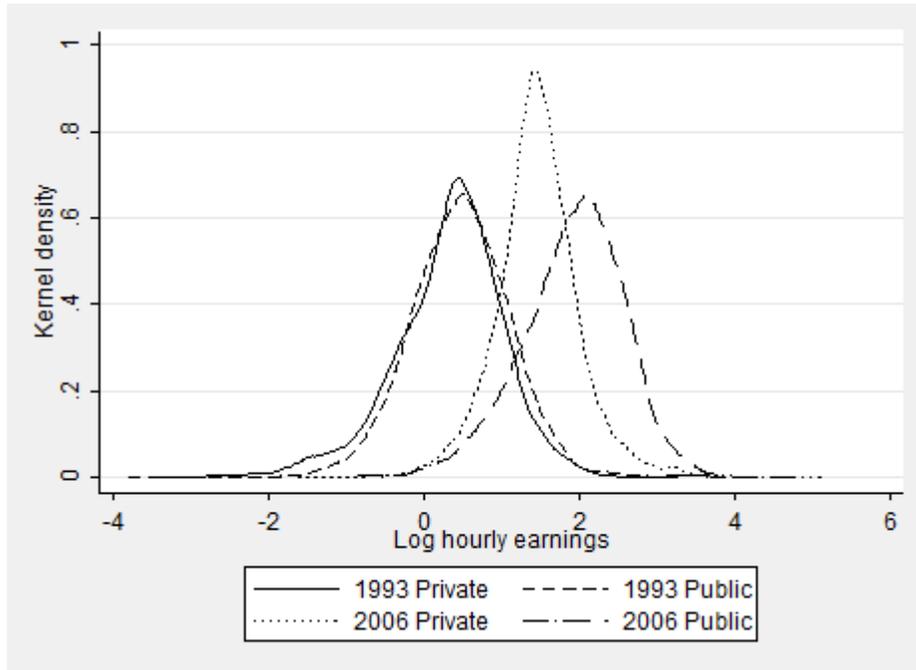
TABLE 2. Public and Private Sector Earnings in 1993 and 2006 (Dongs, in Nominal Terms)						
	1993		1998		Change	
	Public	Private	Public	Private	Public	Private
Hourly earnings	1,460	1,321	9,321	5,747	538%	335%
Hourly benefits	315	81	1,203	414	282%	414%
Hours worked	1850	1621	2047	1892	11%	17%

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

The distribution of earnings for public and private employees and their changes over time is displayed in figure 1. In 1993, the public and private sectors were difficult to distinguish whereas, in 2006, the public sector wage distribution clearly dominated the private sector distribution. Another striking change is the shape of the two distributions: the dispersion of the earnings distribution decreased between 1993 and 2006, strongly in the private sector and mildly in the public sector. The decrease in within-sector inequality dominated, so the overall earnings inequality was, in fact, falling. The main fact that we attempt to explain in this paper is the widening of the public-private pay gap combined with an increase in the relative dispersion of earnings in the public sector as compared

with the private sector. We will suggest explanations for the decrease in overall inequality in section 5.

FIGURE 1. The distribution of earnings in the public and private sectors (Kernel estimates)



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

## EMPIRICAL STRATEGY

### *Model*

In the spirit of the Roy model, we assume that workers differ in their productive characteristics, or skills, which are rewarded differently across the sectors of employment. In this study, we consider only two sectors of employment: public and private. In the simplest framework, each worker is characterized by a single index of comparative advantage from the public sector, denoted by  $\theta_i$ . This index is composed of all of the productive characteristics that influence a worker's pay in one sector relative to the other. Because  $\theta_i$  is a relative term, we can normalize its price to one in the private sector. We use  $\alpha$  to denote the relative returns to workers' comparative advantage in the

public sector compared with the private sector. At a given date  $t$ , in addition to the returns to the worker's productive skills, each public sector worker earns a constant premium  $\eta_t$ . The log labor market earnings of individual  $i$  at time  $t$ , which we denote as  $y_{it}$ , depend on the idiosyncratic term  $u_{it}$ , which has mean zero and is independent of the worker's skills (i.e.,  $u_{it} \perp \theta_i$ ).

Let  $P_{it}$  be a dummy that takes the value of one if worker  $i$  is employed in the public sector at date  $t$ , and let  $\tau_t$  be a time-fixed effect. The log labor market earnings are equal to the following:

$$y_{it} = \tau_t + \eta_t * P_{it} + \theta_i + (\alpha - 1) * P_{it} * \theta_i + u_{it}. \quad (1)$$

We impose no restriction on the correlation between  $\theta_i$  and  $P_{it}$ . In other words, the model allows the workers to sort into one sector or the other on the basis of their comparative advantage. Each worker's earnings are determined by the sector in which she works and her level of comparative advantage. Our objective is to determine whether the observed differences in the distribution of earnings in the two sectors are due to the wage-setting function (parameters  $\eta$  and  $\alpha$ ) or to the sorting of the workers ( $\theta_i$ ). The parameter  $\eta$  corresponds to a premium paid to all public employees independently of their productive characteristics; it only affects the between-sector inequality. In contrast,  $\alpha$  affects both the within- and between-sector inequality. An  $\alpha$  that is smaller than one implies that the wage scale is more compressed in the public sector. If, as we will show, public sector workers have a higher comparative advantage, then an  $\alpha$  smaller than one also implies that there is lower inequality between sectors. As  $\alpha$  rises toward one, the public sector earnings distribution becomes as unequal as the private sector distribution, and the public-private earnings gap increases.

### Identification

Similar to most econometric models using panel data, our model uses information on workers who switch between the public and the private sectors to identify the differences between the sectors. Our assumption is that once the comparative advantage and the employment sector at any date are taken into account, the residual determinants of workers' earnings are exogenous. Our assumption is similar to the strict exogeneity assumption in the fixed effect models, and it is written as follows:

$$E(u_{it} | \theta_i, P_{is}) = 0 \text{ for all dates } s \text{ and } t.$$

Each worker in our sample is observed at two dates:  $t$  and  $t+1$ . Let  $H_i$  denote each worker's employment history:  $H_i = (P_{it}, P_{it+1})$ . The workers belong to one of the four categories  $H_i = \{11,00,01,10\}$ : “stayers in public,” “stayers in private,” “switchers to public,” and “switchers to private.” Let us use  $m_H$  to denote the population average of the comparative advantage  $\theta_i$  in group  $H$ . For each worker in group  $H$ , we have the following:

$$\theta_i = m_H + \zeta_i, \text{ with } (E\zeta_i | H_i) = 0.$$

For identification purposes, we use the first moments of the distribution of earnings, measured in each group at each date:

$$\begin{aligned} E(y_{it+1} | H_i=11) &= \tau_{t+1} + \eta_{t+1} + \alpha * m_{11} & E(y_{it} | H_i=11) &= \tau_t + \eta_t + \alpha * m_{11} \\ E(y_{it+1} | H_i=00) &= \tau_{t+1} + m_{00} & E(y_{it} | H_i=00) &= \tau_t + m_{00} \\ E(y_{it+1} | H_i=01) &= \tau_{t+1} + \eta_{t+1} + \alpha * m_{01} & E(y_{it} | H_i=01) &= \tau_t + m_{01} \\ E(y_{it+1} | H_i=10) &= \tau_{t+1} + m_{10} & E(y_{it} | H_i=10) &= \tau_t + \eta_t + \alpha * m_{10}. \end{aligned}$$

We normalize the mean of  $\theta_i$  for the whole population to zero. Let  $\pi_H$  be the proportion of workers with employment history  $H$ . We impose the following:

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

The model is just identified; we have nine equations in total and nine parameters to estimate. These parameters are the five model parameters  $\alpha$ ,  $\eta_t$ ,  $\eta_{t+1}$ ,  $\tau_t$  and  $\tau_{t+1}$  and the average of  $\theta_i$  in the four employment history groups  $m_{11}$ ,  $m_{00}$ ,  $m_{01}$ , and  $m_{10}$ . Although the constant shifters ( $\tau_t$ ,  $\eta_t$ ) enter linearly in the equations above,  $\alpha$  makes the problem clearly nonlinear. The parameters are estimated through a regression method, but some intuition on how  $\alpha$  is identified can be gained from the following equation:

$$\alpha = - \frac{(E(y_{it+1}|H_i=01) - E(y_{it}|H_i=10)) - (E(y_{it+1}|H_i=11) - E(y_{it}|H_i=11))}{(E(y_{it+1}|H_i=10) - E(y_{it}|H_i=01)) - (E(y_{it+1}|H_i=00) - E(y_{it}|H_i=00))}$$

To see how the identification works, one can rewrite the first two terms of the numerator:

$$(E(y_{it+1}|H_i=01) - E(y_{it}|H_i=10)) = \alpha(m_{01} - m_{10}) + [\tau_{t+1} + \eta_{t+1} - (\tau_t + \eta_t)].$$

The difference between the public sector earnings of the switchers to the public (at time t+1) and the switchers to the private (at time t) depends on two components. The first component expresses how different their comparative advantage is, on average, and the second expresses how much the public sector earnings have changed between t and t+1. The latter is equal to the change in earnings for those who stayed in the public sector, which is used to “deflate” the numerator:

$$(E(y_{it+1}|H_i=01) - E(y_{it}|H_i=10)) - (E(y_{it+1}|H_i=11) - E(y_{it}|H_i=11)) = \alpha(m_{01} - m_{10}).$$

The denominator is the difference between the private sector earnings of the switchers to the private sector (observed at time t+1) and the switchers to the public sector (observed at time t) deflated by the change in earnings of the “stayers in private.” The denominator

is equal to the difference between the average comparative advantage of the switchers to private and the switchers to public:

$$\left(E(y_{it+1}|H_i=10) - E(y_{it}|H_i=01)\right) - \left(E(y_{it+1}|H_i=00) - E(y_{it}|H_i=00)\right) = m_{10} - m_{01}.$$

Finally, it is important to note that our model allows for the free sorting of workers into employment sectors as long as the sorting is driven by their comparative advantage. For example, the probability of working in the public sector could be an increasing function of  $\theta_i$ : workers with a low (high) level of comparative advantage would always work in the private (public) sector whereas the sector switchers would be found at the intermediate levels of  $\theta_i$ . However, we do not allow the past idiosyncratic shocks (i.e., the lagged values of  $u_{it}$ ) to influence the sectoral choice. The reason that we need this assumption is straightforward: we want to use data on the workers who switch sectors between the two surveys to disentangle the effect of the workers' fixed characteristics and the effect of the sector-specific wage settings. However, the sector switchers' earnings are informative about the entire workforce if and only if, at any given date, they are paid the same as the nonswitchers who are working in the same sector with the same level of comparative advantage (the same  $\theta_i$ ).

### *Estimation*

Drawing on the identification results presented in the previous section, we can now estimate the model described in equation (1) even though the index of the workers' productive characteristics  $\theta_i$  is not observed. We regress the log earnings on a set of dummies for the public and private workers and for the four groups of workers with different employment histories. The estimation yields the model parameters and the four

averages of  $\theta_i$ :  $m_{00}$ ,  $m_{11}$ ,  $m_{01}$ , and  $m_{10}$ . We use the nonlinear least squares method to estimate the following equation:

$$y_{it} = \tau_t + \eta_t * P_{it} + m_{00} * \mathbb{1}\{H_i = 00\} + \alpha * m_{11} * \mathbb{1}\{H_i = 11\} + \quad (2)$$

$$(1 + (\alpha - 1) * P_{it}) * (m_{01} * \mathbb{1}\{H_i = 00\} + m_{10} * \mathbb{1}\{H_i = 10\}) + \epsilon_{it}$$

$$\text{s.t.} \quad m_{11}\widehat{\pi}_{11} + m_{00}\widehat{\pi}_{00} + m_{10}\widehat{\pi}_{10} + m_{01}\widehat{\pi}_{01} = 0.$$

The residual  $\epsilon_{it}$  is composed not only of the idiosyncratic shock  $u_{it}$  but also the individual term  $\zeta_i$ . By construction,  $\zeta_i$  is not correlated with the employment history dummies and, hence, with the public sector dummies; our estimates remain unbiased. However,  $\zeta_i$  induces serial correlation, which we account for by clustering the standard errors at the individual level.

Thus far, we have not discussed the role of the specific observable characteristics in our model because we do not want to make an a priori assumption about the correlation between the observable characteristics and the comparative advantage. For example, if  $X_{it}$  is a vector of individual characteristics, it cannot enter linearly in equation (2) unless one assumes that it is not correlated with  $\theta_i$ . An entirely nonparametric solution would consist of estimating the model separately for all of the values taken by  $X_{it}$ , but the lack of data prevents us from implementing this solution. An intermediate solution consists of interacting  $X_{it}$  with  $\theta_i$ . For example, let us assume that one individual characteristic, described by the time-invariant binary variable  $D_i$ , is an important dimension of the public-private sector differences in the wage-setting policies. We rewrite equation (1) as follows:

$$y_{it} = \tau_t + \eta_t * P_{it} + \theta_i + (\alpha - 1) * P_{it} * \theta_i \quad (3)$$

$$+ \delta_t * P_{it} * D_i + (\gamma - 1) * P_{it} * \theta_i * D_i + u_{it}.$$

The new parameters express the differences between the earnings equations of workers with  $D_i = 1$  and  $D_i = 0$ . Parameter  $\delta_t$  denotes the difference in the public premium, and  $\gamma$  denotes the difference in the relative returns to the workers' comparative advantage.

The estimated equation is a simple extension of equation (2), with the interaction terms between the employment history dummies and the dummy  $D_i$ .

## RESULTS

The literature on earnings inequality between 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 wage equations with different coefficients for each sector. In this section, before estimating our model, we present and discuss the results that can be obtained with these models. These results will help us to identify some robust stylized facts and to make the case for using a different approach.

### *Sample and descriptive statistics*

We restrict ourselves to the workers who worked for wages for two consecutive surveys: 1993 and 1998 for the first panel and 2002 and 2004 or 2004 and 2006 for the second panel. As discussed earlier, the VLSS and VHLSS samples are representative of the Vietnamese population but contain few wage earners. Hence, our final sample is relatively small, with 875 individuals in 1993–1998 and 3893 individuals after pooling the 2002–2004 and the 2004–2006 surveys. As we saw, our model's identification further relies upon the workers who switched to the public or to the private sector between the

two surveys. There are 105 such switchers between 1993 and 1998 and 266 in the 2002–2006 surveys.

The public and the private sector workers have very different characteristics, which reflect the differences in the job requirements in the two sectors (see table 3). As expected, the public sector workers spent more time in school (5 years more in 1993) and are more likely to have received vocational training (28 percent more in 1993). There are also proportionally more female workers in the public sector (12 percent more), suggesting that public employment is more “women friendly” than private employment (see Liu 2004). Considering the evolution between 1993 and 2006, it appears that the private and the public sector workers have become more similar in terms of their observed characteristics, with the exception of experience; the public sector workforce is aging.

TABLE 3. Descriptive Statistics on Public and Private Sector Workers						
	1993			2006		
	Public	Private	Diff.	Public	Private	Diff.
Years of schooling	10.96 (3.43)	6.03 (3.65)	4.93 ***	10.75 (2.26)	7.15 (3.46)	3.60 ***
Years of experience	15.50 (8.72)	16.39 (11.33)	-0.89	22.14 (11.33)	19.28 (12.19)	2.85 ***
Proportion of men	0.51 (0.50)	0.67 (0.49)	-0.16 ***	0.58 (0.49)	0.66 (0.47)	-0.09 ***
Proportion with vocational training	0.36 (0.48)	0.05 (0.22)	0.31 ***	0.34 (0.47)	0.11 (0.32)	0.23 ***
Number of workers	564	824		1429	2664	

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

### *<<B>>The “public sector dummy approach”*

We first consider the simplest method to compute the public-private wage gap, controlling for the differences in the workers' characteristics between the two sectors.

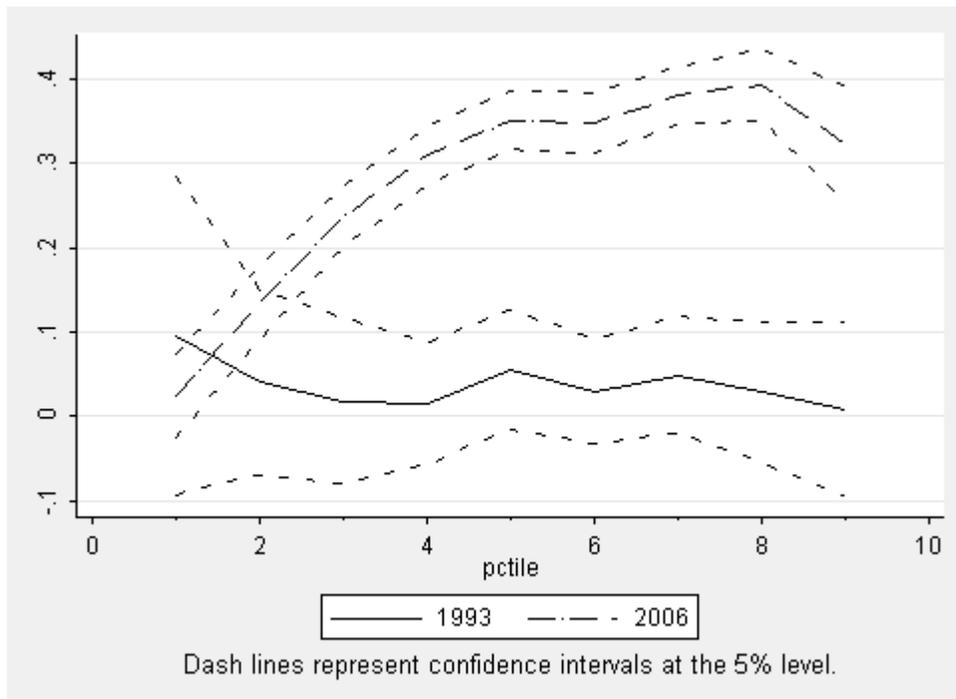
This method implicitly assumes that the returns to the workers' characteristics are the

same across the sectors. This method consists of a regression of the log earnings on the workers' observable characteristics, with a dummy variable that represents working in the public sector. As before,  $X_{it}$  denotes the vector of workers' characteristics. We estimate the following equation by quantile regression:

$$y_{it} = \tau_t + \eta_t * P_{it} + \beta * X_{it} + \epsilon_{it}.$$

The quantile regression estimates of the public sector premium are shown in figure 2. The public sector premium is essentially zero in 1993 and 1998 and rises steeply afterward. Interestingly, in the 2000s, the public premium becomes much higher for the top half of the distribution (over 30 percent at all quantiles above the median). These findings are coherent with the rise in the returns to skills in the public sector as compared with the private sector, which we document in the next section.

FIGURE 2. The public premium estimated through quantile regression



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

### *Difference in returns to workers' characteristics*

A second approach, which is by far the most common in the literature, consists of allowing the returns to the workers' observable characteristics  $X_{it}$  to differ between the two sectors. Let  $\beta_1$  represent the returns in the public sector and  $\beta_0$  represent the returns in the private sector. Using the same notations as above, the model is as follows:

$$y_{it} = \tau_t + \eta_t * P_{it} + \beta_0 * X_{it} + \beta_1 * P_{it} * X_{it} + \epsilon_{it}.$$

We include the workers' observable characteristics that have the most significant effect on earnings.<sup>7</sup> The OLS estimation results are shown in table 4. In 1993, the public sector has higher returns to education and lower returns to experience than the private sector. Between 1993 and 2006, the returns to education increase, and the returns to experience decline in both sectors, which is a common finding in other transition economies. At the end of the period, the returns to education are even higher in the public sector whereas the returns to experience are comparable in both sectors. This shift suggests an explanation for why the public-private wage gap increases. Public workers, who are better educated and more experienced than their private counterparts, benefit from a rise in the returns to skills. Note that once the differences in the returns to the workers' characteristics are taken into account, the public sector dummy becomes insignificant; if anything, it decreases over time. We will use these results to perform the Oaxaca-Blinder decomposition. However, an important limitation of this approach is that any unobserved characteristics that are correlated with  $P_{it}$  and  $X_{it}$  (e.g., cognitive abilities or political connections) bias the estimate of  $\beta_1 - \beta_0$ .

TABLE 4. OLS Estimation of Wage Equations in the Public and Private Sector in 1993 and 2006						
	1993			2006		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Public	0.193	0.181		-0.052	0.124	
Years of schooling	0.012	0.010		0.038	0.003	***
Years of experience	0.056	0.010	***	0.022	0.003	***
Square of experience	-0.118	0.024	***	0.000	0.000	***
Male	0.276	0.074	***	0.230	0.024	***
Vocational training	-0.332	0.195	*	0.135	0.040	***
Public * Years of schooling	0.024	0.013	*	0.047	0.009	***
Public * Years of experience	-0.035	0.016	**	0.006	0.007	
Public * Square of experience	0.092	0.045	**	0.000	0.000	
Public * Male	-0.184	0.095	*	-0.229	0.039	***
Public * Vocational training	0.260	0.203		-0.226	0.052	***
Constant	-0.321	0.107	***	0.771	0.045	***
Observations	856			3893		
R-squared	0.129			0.264		
Mean squared error	0.62			0.55		

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

### *Our model*

We now turn to the estimation of our model. The lack of observations leads us to adopt a simple estimation method, the nonlinear least squares (minimum distance estimation) method, which has better small-sample properties than the Generalized Method of Moments used by Lemieux (1998). The model estimates are shown in table 5.

Very different patterns emerge for the 1993–1998 and 2002–2006 periods. In the 1993–1998 panel, the returns to the workers' comparative advantage in the public as compared with the private sector ( $\alpha$ ) are significantly inferior to one. In the 2002–2006 panel, they are not significantly different from one, and the point estimate is greater than one. This change suggests that the public sector, which initially offered lower returns on workers' comparative advantage than the private sector, has now adopted compensation patterns that reward workers at least as much. Interestingly, the point estimate of  $\eta_t$  is also rising but is never significantly different from zero. Therefore, the higher pay in the

public sector cannot be adequately described as a simple “wage premium” that is unrelated to productive characteristics.

Finally, we find clear evidence that workers sort according to their comparative advantage  $\theta$ . As shown in table 5, the means of the distribution of  $\theta$  are negative for the employees who stay in or who switch to the private sector ( $m_{00}, m_{10}$ ). The public sector stayers have, on average, higher (positive) values of  $\theta$ . Our estimation results also suggest that the sorting of workers across sectors responds to the changes in the returns to their comparative advantages. During the 1990s, when the returns to workers' comparative advantage are lower in the public than in the private sector, the average of  $\theta$  among the “switchers to public” ( $m_{01}$ ) is much lower than it is for any other group. However, in the second panel, the relative returns are higher ( $\alpha \geq 1$ ).

TABLE 5. Non-Linear Least Squares Estimation Results						
	1993			2006		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Relative returns $\alpha$	0.505	0.267	*	1.448	0.725	**
Public premium $\eta_t$	-0.222	0.187		0.010	0.055	
Public premium $\eta_{t+1}$	-0.125	0.188		0.089	0.055	
Time dummy $\tau_t$	0.620	0.168	***	1.315	0.047	***
Time dummy $\tau_{t+1}$	1.103	0.164	***	1.559	0.046	***
Public stayers $m_{11}$	0.309	0.197		0.231	0.079	***
Private stayers $m_{00}$	-0.256	0.164		-0.158	0.046	***
Switchers to public $m_{01}$	-0.524	0.199	***	-0.108	0.103	
Switchers to private $m_{10}$	-0.021	0.186		0.061	0.046	
Sample size	1750			7786		

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

Compared with the other models presented above, our model accounts for any characteristic that contributes to workers' comparative advantage from being in the public sector rather than in the private sector. In this sense, our model allows for a greater effect from workers' selection into one sector or the other. However, we find significant differences in the relative returns to workers' comparative advantage between the public

and the private sectors (at least in the 1990s) and a significant change in these differences over time. These differences strengthen the argument that compensation patterns play a role in the rise of the public-private wage gap. To assess the respective effect of each factor, we now use our model estimates to perform decompositions of the labor market earnings inequality.

### *Explaining the average public-private earnings gap*

In this section, we aim to assess the relative importance of workers' selection and wage setting in explaining the rise in the average public-private earnings gap. We show that an important part of the gap is explained by differences in workers' characteristics because of positive selection into the public sector. The differences in the returns to these characteristics between the two sectors are the driving factors behind the widening of the public-private gap. For the sake of exposition, we first review the Oaxaca-Blinder method before presenting our proposed extension.

### *Oaxaca-Blinder method*

Let us write the difference in the mean log earnings between the public and the private sector:  $\Delta_t = E(y_{it}|P_{it}=1) - E(y_{it}|P_{it}=0)$ . Once the returns to skills are estimated separately for each sector, Oaxaca (1973) proposes decomposing  $\Delta_t$  into two components: the effect of the differences in the workers' skills, on the one hand (selection effect), and the effect of the differences in the returns on these skills, on the other (price effect). There are two possible ways of implementing this decomposition, which correspond to the two equations below. One way is to compute the counterfactual earnings for public employees by applying the coefficient of the wage equation estimated

for the private sector to their observed characteristics. The difference between the wage observed for the public workers and this counterfactual reflects the price component of the public-private wage gap, and the remaining part of the gap is due to selection (Equation 4). The other path compares private employees' earnings with the counterfactual earnings that they would be paid in the public sector, given their observable characteristics (Equation 5).

$$\Delta_t = \underbrace{\widehat{\eta}_t + (\widehat{\beta}_1 - \widehat{\beta}_0) * E(X_{it} | P_{it}=1)}_{Price} + \underbrace{\widehat{\beta}_0 * (E(X_{it} | P_{it}=1) - E(X_{it} | P_{it}=0))}_{Selection} \quad (4)$$

$$\Delta_t = \underbrace{\widehat{\eta}_t + (\widehat{\beta}_1 - \widehat{\beta}_0) * E(X_{it} | P_{it}=0)}_{Price} + \underbrace{\widehat{\beta}_1 * (E(X_{it} | P_{it}=1) - E(X_{it} | P_{it}=0))}_{Selection} \quad (5)$$

However insightful it may be, the Oaxaca-Blinder decomposition method is biased if there are unobservable characteristics that affect workers' earnings and if the public and private sector employees are systematically different with respect to these characteristics.<sup>8</sup> This bias motivates the use of our model because it allows for selection on all of the workers' characteristics.

#### *Our decomposition*

We use our model to decompose the public-private wage gap at each date into two components. The first component corresponds to sorting of workers into the public sector according to their comparative advantage  $\theta$  (selection effect), and the second component corresponds to differences in returns to this comparative advantage  $\alpha$  and the constant earnings difference between public and private employees  $\eta$  (price effect). As in the Oaxaca-Blinder method, there are two methods for implementing our decomposition. We can compute public employees' wages without the public premium  $\widehat{\eta}_t$  and with the

returns to the comparative advantage set to one, which simulates what they would be paid in the private sector (Equation 6). Alternatively, we can simulate private employees' pay if they were to receive the public premium and the same returns to their productive skills  $\hat{\alpha}$  as the public employees (Equation 7).

$$\Delta_t = \underbrace{\hat{\eta}_t + (\hat{\alpha} - 1) * E(\theta_{it} | P_{it}=1)}_{Price} + \underbrace{(E(\theta_{it} | P_{it}=1) - E(\theta_{it} | P_{it}=0))}_{Selection} \quad (6)$$

$$\Delta_t = \underbrace{\hat{\eta}_t + (\hat{\alpha} - 1) * E(\theta_{it} | P_{it}=0)}_{Price} + \underbrace{\hat{\alpha}(E(\theta_{it} | P_{it}=1) - E(\theta_{it} | P_{it}=0))}_{Selection} \quad (7)$$

The results of our decomposition are shown in table 6. We find that the selection on workers' comparative advantage is a major component of the public-private earnings gap in Vietnam. If there were no differences in wage setting between sectors, the earnings in the public sector would be 30 to 40 percent higher than in the private sector throughout this period. This effect is much larger than the effect one would obtain with a simple Oaxaca-Blinder method (results not shown here) because our model allows for the selection of workers on the basis of any of their characteristics. Our results also emphasize the impact of the returns to the comparative advantage ( $\alpha$ ) and the impact of the constant premium from being in the public sector ( $\eta$ ). Most strikingly, our decomposition displays a negative contribution for these parameters to the public-private gap in 1993–1998: public sector employees are underpaid compared with private workers. In 2002–2006, their contribution becomes positive, and the reform in the public sector's compensation policy causes the earnings gap to widen.

TABLE 6. Decomposition of the Public-Private Wage Gap Based on our Simulations					
	1993	1998	2002	2006	Change 1900–2000s
<b>Real Gap</b>	0.21	0.21	0.48	0.53	0.30
<b>Simulation 1</b>					
Price effect	−0.36	−0.23	0.11	0.18	0.44
Selection	0.56	0.44	0.37	0.35	−0.14
<b>Simulation 2</b>					
Price effect	−0.08	−0.01	−0.06	0.03	0.03
Selection	0.28	0.22	0.54	0.51	0.27
<b>Average 1–2</b>					
Price effect	−0.22	−0.12	0.02	0.10	0.23
Selection	0.42	0.33	0.45	0.43	0.07

*Source:* Author's calculations based on VLSS and VHLSS data.

Compared with the Oaxaca-Blinder method, our model offers a much more comprehensive view of selection on workers' characteristics. Hence, our model strengthens the case against explaining the widening of the public-private sector gap only by selection. The fact that  $\alpha$  is significantly inferior to one in the first panel suggests that neither a simple wage equation model nor a fixed effect model (which, de facto, assumes constant returns to the unobserved characteristics across the sectors) could adequately describe the differences in the returns on workers' skills between the public and private sectors.

#### **EXPLAINING THE CHANGES IN THE OVERALL LABOR MARKET EARNINGS INEQUALITY**

In this section, we turn to the entire distribution of earnings and its evolution over time. As we saw in section 1, between 1993 and 2006, the overall inequality falls despite the rising public-private gap because the within-sector inequality decreases sharply in both the public and the private sectors. This stylized fact stands in stark contrast with the other experiences of transition into a market economy, such as the Polish case studied in Keane and Prasad (2006).

We expect the labor markets after the transition to be more competitive than before and to align workers' earnings to their productivity. This expectation does not necessarily imply that the distribution of the workers' earnings should become more unequal. Examining the wage equation estimates presented in table 4, we see that the R squared is higher in the 2000s than in the 1990s and that the residual variance has declined. On the one hand, these findings suggest that workers' compensation is more closely determined by their individual characteristics and their public or private employee status after the transition than it was before. On the other hand, these findings suggest that the decline in earning inequality is not due to a decline in the variance of observable characteristics nor in their price. Within the simple OLS model, however, we cannot say whether this decline is due to a greater homogeneity in the compensation of workers with similar unobserved characteristics or to a falling variance of idiosyncratic shocks.

*Estimation of each worker's comparative advantage*

We can use our model to disentangle the effect of the changes in the relative compensation patterns, the changes in the distribution of comparative advantage in the population and the changes in the distribution of idiosyncratic shocks. This exercise requires an estimation of each worker's comparative advantage. If we denote  $z_{it} = \tau_t + \eta_t * P_{it}$ , then  $\theta_i$  is estimated using the following expression:

$$\theta_i = \frac{1}{2} \left( \frac{y_{it} - \widehat{z}_{it}}{1 + (\widehat{\alpha} - 1)P_{it}} + \frac{y_{it+1} - \widehat{z}_{it+1}}{1 + (\widehat{\alpha} - 1)P_{it+1}} \right).$$

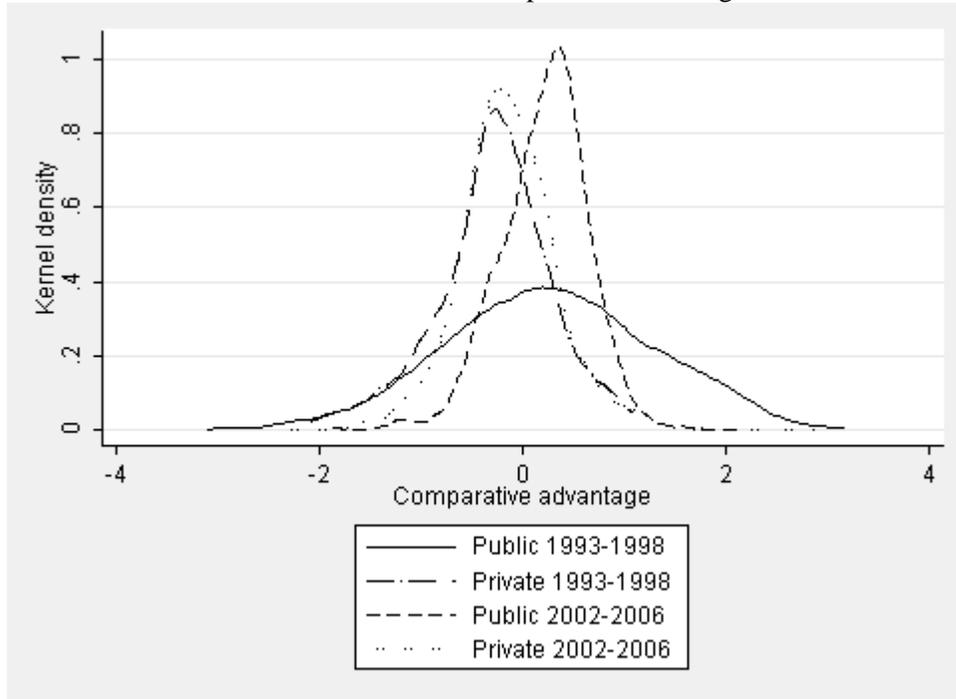
It is easy to show that this estimator is unbiased.<sup>9</sup> Indeed, using the orthogonality assumption between residuals  $u_{it}$ , the factor  $\theta_i$  and the public sector dummy  $P_{it}$ , the

expectation of the difference between the observed earnings and  $z_{it}$  can be written as follows:

$$E[y_{it} - z_{it}] = E[\theta_i + (\alpha - 1) * P_{it} * \theta_i + u_{it}] = (1 + (\alpha - 1) * P_{it}) * \theta_i.$$

The distributions of  $\theta_i$  for the public and private workers in 1993 and 2006 are displayed in graph 3. If we consider only the first moment of each distribution, this graph illustrates the findings presented in section 3. The public workers have a higher comparative advantage from being in the public sector, and the difference between the public and private employees decreases slightly between the first and the second panel. What we learn from estimating  $\theta_i$  is that its variance shrinks between the first and the second panel in both the public and the private sectors. There may be several factors in this evolution, such as a greater homogeneity of workers or of workers' pay. Our model cannot disentangle these factors because they affect both of the sectors in the same way. However, we can isolate their impact from the effect of the relative changes in workers' pay and the changes in the workers' allocation across sectors.

FIGURE 3. The distribution of workers' comparative advantage



Source: Author's calculations based on VLSS 1993–1998 and VHLSS 2002–2006 data

### *Decomposition of overall inequality*

A few steps are needed before implementing our decomposition. We first compute the residuals for each individual observation as follows:

$$\widehat{u}_{it} = y_{it} - \widehat{y}_{it} = y_{it} - (\widehat{\tau}_t + \widehat{\theta}_i + \widehat{\eta}_t * P_{it} + (\widehat{\alpha} - 1) * P_{it} * \widehat{\theta}_i).$$

We then simulate the changes in the distribution of  $\theta$  through a rank-preserving transformation; that is, we assign to each 1993 individual the value of  $\theta$  that has the same rank in the 2006 distribution as the value of  $\theta$  that we estimated for the individual in the 1993 distribution. We denote this value  $\widehat{\theta}_i^{06}$ . We next perform a similar rank-preserving transformation to assign a residual error term  $\widehat{u}_i^{06}$  to each worker observed in 1993.

Finally, we use  $\alpha_{93}$  and  $\alpha_{06}$  to denote relative returns to the workers' comparative advantage estimated for the 1993–1998 and 2002–2006 panels, respectively.

We are now able to build the three following counterfactuals:<sup>10</sup>

$$\text{Counterfactual 1} = \widehat{\tau}_{06} + \widehat{\theta}_i + \widehat{\eta}_{06} * P_{i93} + (\widehat{\alpha}_{06} - 1) * P_{i93} * \widehat{\theta}_i + \widehat{u}_{i93}$$

$$\text{Counterfactual 2} = \widehat{\tau}_{06} + \widehat{\theta}_i^{06} + \widehat{\eta}_{06} * P_{i93} + (\widehat{\alpha}_{06} - 1) * P_{i93} * \widehat{\theta}_i^{06} + \widehat{u}_{i93}$$

$$\text{Counterfactual 3} = \widehat{\tau}_{06} + \widehat{\theta}_i^{06} + \widehat{\eta}_{06} * P_{i93} + (\widehat{\alpha}_{06} - 1) * P_{i93} * \widehat{\theta}_i^{06} + \widetilde{u}_i^{06}.$$

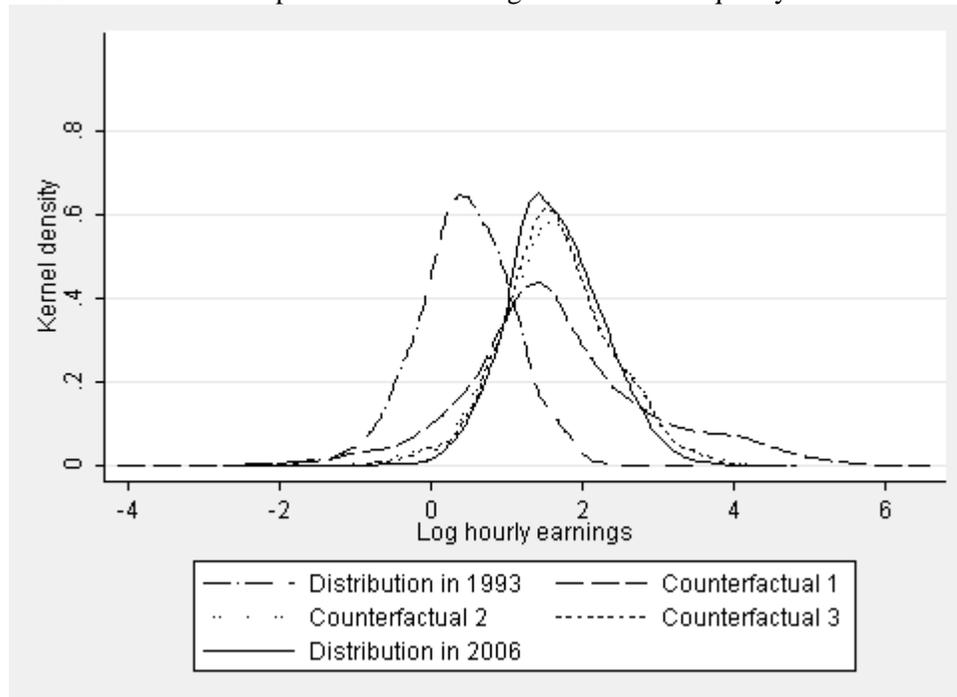
All three counterfactuals are shown in figure 4, together with observed earnings distributions in 1993 and 2006.

Let us first compare the observed 1993 earnings distribution to Counterfactual 1. The first counterfactual takes into account the new constant and changes both the public sector premium and the price of individual characteristics in the public sector as compared with the private sector. The rise in the constant  $\tau$  shifts the curve to the right whereas changes in the compensation patterns in the public sector compared with the private sector have an inequality-increasing effect.

Next, we compare Counterfactual 2 to Counterfactual 1. The second counterfactual simulates the 1993 workers' pay as though they had 2006 workers' characteristics, keeping the allocation process between the public and the private sectors constant. We see that the compression of the distribution of the workers' comparative advantage offsets the inequality-increasing effect of the rise in the relative returns to the workers' comparative advantage. If we compare Counterfactuals 2 and 3, the changes in the distribution of the idiosyncratic shocks  $u_{it}$  do not appear to make a difference.

Finally, Counterfactual 3 differs only slightly from the real earnings distribution in 2006.

FIGURE 4. The decomposition of the changes in overall inequality



Source: Author's calculations based on VLSS 1993–1998 and VHLSS 2002–2006 data

Our decomposition shows that two factors exert a significant influence on the evolution of overall inequality: the relative returns on the workers' comparative advantage and the distribution of this comparative advantage. We can rule out any major influence by the two other components, the allocation of the workers across the sectors and the residual variance in the earnings.

## DISCUSSION AND EXTENSION

As argued throughout this paper, including an index of the comparative advantage  $\theta_i$  allows our model to be more comprehensive with respect to the workers' selection into the public sector. One 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. In this section, we correlate the estimated comparative advantage of each

worker with her observable characteristics. We also test whether our results hold for different subpopulations of workers, defined by gender and experience.

#### *Correlation with observable characteristics*

In the previous section, we estimated the comparative advantage of each worker. It is easy to compute the correlation coefficients (not shown here) between  $\hat{\theta}_i$  and the characteristics that affect workers' productivity. We find high coefficients for years of schooling, years of experience, residence (urban or rural), and vocational training. This finding confirms that our method captures the effect of the observable characteristics that are usually included in wage equations as well as unobservable characteristics that are strongly correlated with them. Such unobservable characteristics could be cognitive abilities, social capital, or political connections.

#### *Returns to skill and experience*

The model extension presented in section 2 allows us to test whether the increase in workers' relative returns to their comparative advantage in the public sector holds for different subpopulation of workers. Interestingly, we can test for potential segmentation between new entrants on the labor market and more experienced workers. Let us define the dummy variable  $E_i$ , which takes the value zero for the workers with less than 10 years of experience and one otherwise. We follow the model described by equation (3), replacing  $D_i$  with  $E_i$ . In examining the results of the estimation in table 7, our previous finding, that public sector workers have very low relative returns to their comparative advantage in the 1990s, does not seem to hold for both the younger and the older workers. However, the more experienced workers benefit from a significant public premium in the 1990s ( $\delta$  positive) and from a larger rise in returns to their comparative

advantage in the public sector in the 2000s. In the 2000s, the labor market for young workers appears competitive, with no significant difference in the public and private sectors' pay ( $\alpha = 1$  and  $\eta = 0$ ). This finding suggests that public sector rents only benefit older workers.

TABLE 7. Extended Model: Interaction with Experience						
	1993			2006		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Relative returns $\alpha$	0.381	0.271		1.013	0.591	*
Public premium $\eta_t$	-0.288	0.220		-0.029	0.078	
Public premium $\eta_{t+1}$	-0.292	0.227		0.107	0.078	
Old workers * Returns $\gamma$	0.080	0.206		0.976	0.756	
Old workers * Public premium $\delta_t$	0.118	0.083		0.053	0.117	
Old workers * Public premium $\delta_{t+1}$	0.202	0.076	***	-0.026	0.119	
Time dummy $\tau_t$	0.118	0.083		1.294	0.037	***
Time dummy $\tau_{t+1}$	0.202	0.076	***	1.538	0.035	***
Sample size	1750			7786		

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

#### Returns to skill and gender

Another interesting method for extending our model is to consider the differences in the labor market earnings between the male and female workers. Confirming the previous findings by Liu (2004), we have seen that in the 1990s, women were paid more favorably in the public than in the private sector. After controlling for the workers' characteristics, the gender gap was not significant in the public sector in the 2000s (table 4). We use the model described by equation (3), replacing  $D_i$  with the gender dummy  $G_i = 0$  for men and  $G_i = 1$  for women. As before, the estimated returns to the comparative advantage in the public sector are about half of those in the private sector in the 1990s and are not significantly different from each other in the 2000s (see table 8). We find that in the 2000s, the returns to the comparative advantage are much higher for women, with  $\gamma$  being a positive and significant coefficient, whereas the public premium

is significantly lower ( $\delta$  is negative). This result suggests that women with a high comparative advantage are discriminated against in the private sector and that they have a larger incentive to enter the public sector whereas male employees in the public sector mostly benefit from a constant premium, independent of their comparative advantage.

TABLE 8. Extended Model: Interaction with Gender						
	1993			2006		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Relative returns $\alpha$	0.499	0.277	*	1.053	0.559	*
Public premium $\eta_t$	-0.153	0.200		0.124	0.061	**
Public premium $\eta_{t+1}$	-0.098	0.198		0.195	0.061	***
Women * Returns $\gamma$	0.009	0.190		1.474	0.879	*
Women * Public premium $\delta_t$	-0.162	0.076	**	-0.318	0.122	***
Women * Public premium $\delta_{t+1}$	-0.079	0.066		-0.300	0.122	**
Time dummy $\tau_t$	-0.162	0.076	**	1.307	0.043	***
Time dummy $\tau_{t+1}$	-0.079	0.066		1.551	0.041	***
Sample size	1750			7786		

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

## CONCLUSION

To explain the differences in earnings between the public and private sectors, the two usual candidates are the differences in the workers' skills and the differences in the returns to the workers' skills. The two factors have very different interpretations, and a considerable amount of research has been devoted to disentangling the former, which implies sorting by abilities, from the latter, which is an indicator of labor market segmentation. A plausible explanation for the rising public-private wage gap between 1993 and 2006 is that the selection into the public sector increased, with a greater flexibility in the hiring and the firing of workers. Using a model of comparative advantage and the differential returns between the two sectors, we find evidence that workers in the public sector have a comparative advantage in working there. Even if

selection contributes to explaining the average public-private pay gap, it does not explain the increase in this gap between the 1990s and the 2000s.

Our method emphasizes the role of the differences in returns to the workers' comparative advantage in the public and the private sectors. We find that, in the first period, workers with a comparative advantage in working in the public sector are paid less (with respect to their skills) in the public sector than in the private sector. This result 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). In the second period, in contrast, workers with a comparative advantage in working in the public sector earn slightly higher returns in the public sector than in the private sector. Because of the imprecision of our estimation, we cannot exclude the hypothesis that the returns are the same in the two sectors.

Unlike the “sorting scenario,” our explanation does not imply a massive reallocation of workers between the sectors in the 1990s and the 2000s, which would not match historical accounts. Our explanation involves a change in the public sector wage-setting policy, with workers who are more productive receiving higher benefits. This change could be explained by increased competition between the public and private sectors to attract the best workers in the market. This change may occur for younger workers (with less than 10 years of experience), who have the same returns to skills in the public and private sector. However, we show that the relative rise in returns to workers' skills in the public sector is stronger for older workers. A political economy interpretation would suggest that in the early 2000s, at the most critical stage of the reform process, larger rents were paid to state employees to secure their political support.

We also extend our model to explore the differences between male and female earnings in the public and private sectors. We show that the public sector pays almost equal wages to men and women in the 2000s. The returns to female workers' comparative advantage are much higher in the public sector than in the private sector, suggesting that there is little discrimination against women with good qualifications in the public sector. Unfortunately, because of severe data constraints, we cannot fully exploit the benefits of our methodology to examine the effect of public-private sector differences for many other subcategories of workers. More important, the lack of precision does not allow us to estimate our model separately for urban and rural workers, who are obviously employed in very different labor markets. When we do estimate our model separately for urban and rural workers, we find similar point estimates, and thus, we are confident that our findings hold for both urban and rural areas, but the estimates are too imprecise to prove this statistically.

Finally, we document the surprising fact that despite the increase in the public-private gap, the overall labor market earnings inequality fell between the 1990s and the 2000s. This finding conflicts with the intuition that more competition in the labor market leads to more inequality (see Keane and Prasad 2006). However, a major finding of the literature on the transition to a market economy is that workers' compensation becomes more closely related to their productivity. Indeed, the results from the wage equations and a decomposition based on our model suggest that the decline in overall inequality is due to a greater homogeneity in workers' compensation within each sector and within groups with similar education or experience or those of the same gender. This change offsets the increase in the public-private gap. A complete analysis of this phenomenon goes beyond

the framework of our model, which only takes into account the unobserved workers' heterogeneity in relative terms. More research is needed to fully understand the causes of this phenomenon.

## REFERENCES

- Azam, M., and N. Prakash. 2010. "A Distributional Analysis of the Public-Private Wage Differential in India." IZA Discussion Paper, No. 5132.
- Bales S., and M. Rama. 2001. "Are Public Sector Workers Underpaid? Appropriate Comparators in a Developing Country." World Bank Policy Research Working Paper, No. 2747.
- Edmonds, E., and N. Pavcnik. 2006. "Trade Liberalization and the Allocation of Labor between Households and Markets in a Poor Country." *Journal of International Economics* 69 (2): 272-295.
- Eglinger, J.-P. 2005. Genèse, situation actuelle et perspectives d'avenir des compagnies générales (Tông công 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.
- Fforde, A. 2007. *Vietnamese State Industry and the Political Economy of Commercial Renaissance: Dragon's tooth or curate's egg?* Oxford: Chandos.
- Frydman, R., C. Gray, M. Hessel, and A. Rapaczynski. 1999. "When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies." *Quarterly Journal of Economics* 114 (4): 1153-1191.
- Gallup, J. L. 2002. "The Wage Labor Market and Inequality in Vietnam in the 1990s." World Bank Policy Research Working Papers, No. 2896.

- Gosling, A., 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.
- M. P. Keane and E. S. Prasad. 2006. "Changes in the Structure of Earnings During the Polish Transition." *Journal of Development Economics* 80: 389– 427.
- Lemieux, T. 1998. "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Nonrandom Selection." *Journal of Labor Economics* 16 (2): 261-291.
- Liu, A. Y. C. 2004. "Sectoral Gender Wage Gap in Vietnam." *Oxford Development Studies* 32 (2): 225-240.
- Machado, J., and J. Mata. 2005. "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression." *Journal of Applied Econometrics* 20: 445-465.
- McCaig, B. 2009. "Exporting out of Poverty: Provincial Poverty in Vietnam and US market access." Australian National University Working Papers, No. 502.
- Mac Millan, J., and C. Woodruff. 1999. "Interfirm Relationships and Informal Credit in Vietnam." *Quarterly Journal of Economics* 114 (4): 1285-1320.
- Oaxaca, R. 1973. "Male-Female Differentials in Urban Labor Markets." *International Economic Review* 14: 693-709.
- Painter, M. 2005. "The Politics of State Sector Reforms in Vietnam: Contested Agendas and Uncertain Trajectories." *Journal of Development Studies* 41 (2): 261-283.

Ravallion, M., and D. van de Walle. 2006. "Land Reallocation in an Agrarian Transition." *Economic Journal* 116 (514): 924-942.

Suri, T. 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79 (1): 159-209.

Vietnam's Socio-Economic Development. 2008. *Vietnam's Socio-Economic Development: A Social Science Review*, No.55, Vietnam Institute of Economics, Hanoi.

---

<sup>1</sup> The model was initially developed by Lemieux (1998) to compare the unionized and nonunionized sectors in the US. Suri (2011) uses a similar model to study the effect of selection and comparative advantage in the adoption of fertilizers in rural Kenya.

<sup>2</sup> Unfortunately, we could not analyze the Labor and Job Survey data ourselves, and we do not know of any other source of data that could be used to perform the same test.

<sup>3</sup> Gallup (2002) counts workers employed for wages in both primary and secondary occupations and finds that 25 percent of the work force worked for wages in 1998.

<sup>4</sup> For example, “mixed government /private” (1993) and “joint ventures with the government” (1998) were primitive forms of private enterprises at a time when complete private ownership was not allowed. We classify these enterprises as private sector.

<sup>5</sup> The VLSS and VHLSS questionnaires are comprehensive in this respect. Even the least detailed questionnaire, the VHLSS 2002, uses four categories: “new year and holidays,” “social subsidy (for sickness, maternity or working accident),” “business trip allowance,” and “other.”

<sup>6</sup> Public workers may have worked more hours, on average, than private workers because of the higher propensity of private workers to engage in more than one activity.

<sup>7</sup> We do not include occupation in the wage equations, however, because a worker's occupation is endogenously determined by the sector in which she works.

<sup>8</sup> Albeit more sophisticated, the technique developed by Machado and 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 and Prakash (2010)).

<sup>9</sup> We only have two observations for each worker; hence, measurement error is an issue in the estimation of each individual  $\theta_i$ . However, we have a sufficient number of individuals for our estimates to provide a reasonable approximation of the moments of the distribution of  $\theta_i$ .

<sup>10</sup> We developed the three symmetrical counterfactuals using the distribution of wages in 2006 as a starting point. Our results do not change with the order of decomposition.