Is civil-servant human capital sector-specific?*

Magali Beffy† Thierry Kamionka‡

January 28, 2009

Abstract

Using the French European Household panel, the wage differentials between public and private sector workers are investigated by controlling for both employment and self-selection into the public sector. The model we estimate also allows for unobserved heterogeneity in the propensity to be employed, in either job sector, and in the specific ability to be rewarded. Evidence based on quantile analysis is found on a large public-private pay premia for women and for low public wages whereas male civil servants would earn more in the private sector.

Key words: wage differentials, public and private sector, unobserved heterogeneity

Journal of Economic Literature classification: J31, J45, J64, C33, C35 .

*We would like to thank E. Coudin and A. Maurel for their helpful suggestions. The normal disclaimers applies.
†CREST-INSEE-DEPP-IZA (magali.poncon@polytechnique.org).
‡CNRS and CREST (kamionka@ensae.fr).

1. Introduction

There has been an increasing interest in studying the public sector employment for more than two decades. This process was initiated by Smith (1977) and this is justified by the large share of public employment in total employment. Over that period of time, very few countries have seen this share decrease but for UK. For instance, in France, in 2006, this proportion amounts to one quarter of total employment, and it has been relatively stable since the beginning of the nineties (source OECD data). But this large part of the public sector employment may have adverse effects on private sector efficiency. Moreover a shortage of public sector workers can lead to difficulties, such as an excess can lead to fiscal outcomes.

In France, this interest has been renewed recently for budgetary reasons. On the one hand the French debt burden does not diminish, and on the other hand, the wage bill constitutes the largest item in the public sector spending. Furthermore the retirement of many public servants raises the following issue: should everyone departing be replaced? Is the cost of every public-servant job justified? The French government looks for efficiently spent money, hence every euro spent must be necessary and efficiently allocated. Hence public sector employment and wages come under close scrutiny. Therefore it is important to compare both public and private earnings.

However, few studies deal with the French case (Fougère and Pouget (2003), and Bargain and Melly (2008)). Fougère and Pouget (2003) concentrate on the main determinants of the entry into the public sector. Bargain and Melly (2008) focus on the public sector pay gap using quantile regressions on a short panel data set. This paper aims to contribute to the classical analysis of the public wage gap, especially in France. Do relationships between wages and wage determining factors differ by sector of work? Do public sector employees earn a premium? However usual the issue, the methods we develop extend some previous approaches.

Many recent studies rely on cross sectional switching regression, endogenous or not (see Disney and Gosling (1998) and Gyourko and Tracy (1988) for UK, Dustman and Van Soest (1998) for Germany, Hartog and Oosterbeek (1993) and Van Ophem (1993) for the Netherlands, Fougère and Pouget (2003) for France and Heitmueller (2006) for Scotland). Hartog and Oosterbeek (1993) stress that neglecting selectivity effects are likely to give a false picture of the relative earnings position of public-sector workers (see Goddeeris (1988)). Moreover, Dustman and Van Soest (1998) underlines that, even when the sector choice is controlled for, instruments must be chosen with particular care and exogeneity assumptions can lead to different results. In a different way, Heitmueller

(2006) controls for participation and sector selections, but in cross-sectional analysis.

In order to overcome these potential biases, Disney and Gosling (2003) uses the natural experiment that happened in the UK in the nineties with the privatization programme. And they show that their results are robust to self-selection. Bargain and Melly (2008) use panel data to control for both sector choice and individual fixed effects, and compare the quantiles of both distributions. Raising close but different issues, Bell, Elliott, and Scott (2005) exploit the mobility between both sectors, and study the wage incentives to change sectors. They identify the wage premium after a job change. Other studies focus on the link between the wage distribution and mobility. Postel-Vinay and Turon (2007) and Cappellari (2002) focus on earnings dynamics and lifetime values of employment in both sectors. They argue that public and private sectors differ not only in their log wage distribution but also in their income mobility. They conclude, for UK and resp. for Italy, that life cycle of earnings matters in the private sector whereas it does not in the public sector.[1]

In this paper, the way we proceed is more in line with Dustman and Van Soest (1998) such as Heitmueller (2006). We extend their approach by considering a panel framework, controlling both for self-selection -people do choose to work in the public sector - and employment -people choose to participate. We question the earnings differences by modeling the double selection i.e. employment and sector choice, and we account for unobserved heterogeneity by using the method of Heckman and Singer (1984). Unobserved heterogeneity allows us to control for individual tastes, and individual abilities. Moreover, we observe each individual for 8 years, ensuring convergence properties that can not be ensured with the method used by Bargain and Melly (2008), based on quantile regressions, as the LFS surveyed identical people at most three times.

We answer the following issue: do people who work in the public sector have a different propensity to get high wages or low wages given their educational level and other characteristics? Moreover we analyze the sector selection.[2] We find that mimicking the parents is determinant and as expected, that the public sector attracts more people when the local labor market is depressed.

[1] Another trend of the literature on public/private differences concerns queue models (Venti (1987), Heywood and Mohanty (1995) and Fougère and Pouget (2003)) and analysis the individual propensity to seek employment in public sectors. But these models require very detailed information about the offer and the demand for public jobs. Furthermore, they are very sensitive to the instruments chosen and it turns out difficult to share and identify the characteristics determining the search for a public job from those determining the access to a public job (more risk adverse workers in the public sector -see Bellante and Link (1981), but maybe a *taste* for public services -see Goddeeris (1988)).

[2] We use parental background information to ensure identification of the sector choice equation following Dustman and Van Soest (1998) advice.

We also find that there exists a wage public premium at the bottom of the public wage distribution, whereas it is not true at the upper tail. These results are in line with empirical observations: low-wage civil-servants are weakly mobile, whereas high-wage civil servants move more frequently from the public to the private sector.

Further, the female public-sector workers have a comparative advantage in the public sector. Their counterfactuals wages in the private sector are lower than their current wages. Unlike women, men would have higher wages in the private sector excepted at the bottom of the distribution. For men, this wage gap worsens with the education level; in the case of graduated women, the public wage premium is closer to zero. These results may reflect that motivations underlying sector choice differ between men and women.

Finally we find that wage differences between the public and the private sectors result from three factors: first the raw mean is greater in the public sector, second the returns to different observable characteristics differ between both sectors (for instance, for most of them, degrees are better rewarded in the private sector), finally the unobserved productivity civil-servants would have in the private sector is a bit inferior to the one they have in their current occupation. Unlike them, a part of the workers employed in the private sector seem to acquire specific human capital, that they could not transfer to the public sector.

The paper is structured as follows: in the next section, the structural model is presented. Section 3 reports a descriptive analysis of the data. Section 4 describes the econometric model and the estimation methods. Sections 5 and 6 discuss results and simulations. Finally section 8 concludes.

2. The Search Model

We derive a continuous time search model in a stationary labor market environment. There are three possible states: employment in the public sector, employment in the private sector and nonemployment. The individuals employed in the private sector can be laid off, and they may search for a private job even when they are employed. Those employed in the public sector do not search for a job and can not be laid off. This is justified by the fact that in France, a public servant can not be laid off. A public employee can be fired only in very few cases, the rules of which greatly differ from private-sector ones (disciplinary sanctions in the public sector). Hence firings in the public sector are neglected. Finally the nonemployed may search either for a public or for a private job.

Agents are infinitely alive. At each point in time, they can be either nonemployed (denoted by n, i.e. unemployed or out of the labor force), or employed either in the private sector (denoted by Pr) or in the public sector (denoted by Pu). Unemployment and nonparticipation are assumed to be non distinct labor force states. This assumption is not restrictive since we focus on the choice between the public and the private sector and the data set is restricted to people under 60. Nonemployed individuals enjoy a real return b and receive job offers at a Poisson rate depending on the sector of research $\lambda_{n,Pr}$ and $\lambda_{n,Pu}$. b may represent unemployment benefit or private rents. Nonemployed individuals support research costs depending on the type of job they search for: c_{Pr} (resp. c_{Pu}) for private jobs (resp. for public jobs). They can decide to restrict their research to private jobs. Agents who decide to search for a job in the public sector can fail to enter public services although they get a public offer since they have to succeed the entrance exam to become a civil servant. p_S denotes the probability to succeed conditional on searching for a job in this sector, and it implicitly depends on individual covariates.

When employed in the private sector, individuals receive a real wage w, and they continue to receive private job offers at a Poisson rate $\lambda_{Pr,Pr}$. They are assumed to restrict their job research to the private sector and they face search costs c.[3] Existing private jobs are hit by idiosyncratic (productive) shocks that occur at a Poisson rate δ. The instantaneous discount rate is ρ and the horizon is infinite. This assumption implies that an individual can not transit directly from the private to the public sector, he has to go through an unemployment period.

Finally, when employed in the public sector, agents can not search for a private job. We detail later what the consequences of this assumption are.

F_{Pr} denotes the wage distribution in the private sector on $[0, \overline{w}]$, F_{Pu} in the public sector on $[0, \overline{w}]$. We assume that these distributions are different from one another but that they rely on the same finite support for sake of simplicity. But the tails of the distributions can be far different (see the empirical part for some illustration, figure 1).

In the sequel V denotes the value function.

A worker currently employed in the public sector with starting wage w receives net income w and

[3] To search for a public job requires a lot of time since entrance exams need preparation.

cannot loose his job. He receives no external nor internal offer. His function value is

$$V_{Pu}(w) = \frac{w\Delta t}{1+\rho\Delta t} + \frac{V_{Pu}(w)}{1+\rho\Delta t}, \tag{2.1}$$

which yields

$$V_{Pu}(w) = \frac{w}{\rho}. \tag{2.2}$$

A worker currently employed in the private sector receives a net income $w-c$ and may be forcibly separated from his job with probability δ. He also receives private job offers at rate $\lambda_{Pr,Pr}$ which are accepted when their value exceeds the expected discounted lifetime utility stream in the current job:

$$\begin{aligned} V_{Pr}(w) &= \frac{(w-c)\Delta t}{1+\rho\Delta t} + \delta\frac{V_n\Delta t}{1+\rho\Delta t} + (1-\delta\Delta t-\lambda_{Pr,Pr}\Delta t)\frac{V_{Pr}(w)}{1+\rho\Delta t} \\ &+ \frac{\Delta t\lambda_{Pr,Pr}}{1+\rho\Delta t}\int_0^{\overline{w}} \max\left(V_{Pr}(x), V_{Pr}(w)\right) dF_{Pr}(x), \end{aligned}$$

which can be simplified in

$$\begin{aligned} V_{Pr}(w)\left(\rho+\delta\right) &= w - c + \delta V_n \\ &+ \lambda_{Pr,Pr}\int_w^{\overline{w}} \left(V_{Pr}(x) - V_{Pr}(w)\right) dF_{Pr}(x). \end{aligned} \tag{2.3}$$

A currently nonemployed person enjoys a net flow of income $b-c-d_{Pu}c_{Pu}$ depending on her choice to search for a public job ($d_{Pu}=1$) or not ($d_{Pu}=0$).

$$\begin{aligned} V_n^{d_{Pu}} &= \frac{(b-c-d_{Pu}c_{Pu})}{1+\rho\Delta t}\Delta t + \frac{\Delta t}{1+\rho\Delta t}\lambda_{n,Pr}\mathbb{E}_{F_{Pr}}\left(\max\left(V_n, V_{Pr}(x)\right)\right) \\ &+ \frac{\Delta t}{1+\rho\Delta t}\lambda_{n,Pu}d_{Pu}\mathbb{E}_{F_{Pu}}\left\{\max\left(V_n, p_S V_{Pu}(x) + (1-p_S)V_n\right)\right\} \\ &+ \frac{1}{1+\rho\Delta t}\left(1-\lambda_{n,Pr}\Delta t - \lambda_{n,Pu}d_{Pu}\Delta t\right)V_n. \end{aligned} \tag{2.4}$$

For the nonemployed, the optimal acceptance rule consists in accepting the first job, whatever the sector, that pays more than a reservation wage w^* that is specific to the sector: w^*_{Pu} for the public sector and $w^{*,d_{Pu}}_{Pr}$ for the private sector ($w^{*,d_{Pu}=1}_{Pr}$ denotes the reservation wage

when the unemployed search for both public and private jobs, $w_{Pr}^{*,d_{Pu}=0}$ when the unemployed search only for private jobs).

$$\begin{aligned} d_{Pu} \in \{0,1\} \quad V_{Pr}(w_{Pr}^{*,d_{Pu}}) &= V_n^{d_{Pu}} \\ V_{Pu}(w_{Pu}^*) &= V_n^{d_{Pu}=1}. \end{aligned}$$

These reservation wages exist and are unique because the private and public value functions are continuous and increasing functions. Therefore the equation (2.4) can be rewritten:[4]

$$\begin{aligned} \rho V_n^{d_{Pu}} &= (b - c - d_{Pu}c_{Pu}) + \lambda_{n,Pr}\int_{w_{Pr}^*}^{\overline{w}} (V_{Pr}(x) - V_n)\, dF_{Pr}(x) \\ &\quad + \lambda_{n,Pu} d_{Pu} p \int_{w_{Pu}^*}^{\overline{w}} (V_{Pu}(x) - V_n)\, dF_{Pu}(x). \end{aligned} \tag{2.5}$$

When does a nonemployed worker decide to search for a public job? Searching for a public job is actually costly and risky since individuals are not sure to succeed to enter public services. This cost of searching for a public job is induced by the fact that individuals have to pick up information, to prepare entrance exams, and by the fact that they anticipate private sector opportunities. Hence nonemployed will search for a public job when their expected gains, which may depend on individual characteristics and unobserved ability, exceed the cost of searching (see figure (6)). So $d_{pu} = 1$ when:[5]

$$\begin{aligned} c_{Pu} \quad &\leq \quad p_S \frac{\lambda_{n,Pu}}{\rho} \int_{h(w_{d=0}^*)}^{\overline{w}} \overline{F}_{Pu}(x)dx, \\ &\text{where} \quad p_S \frac{\lambda_{n,Pu}}{\rho} \int_{h(w_{d=0}^*)}^{\overline{w}} \overline{F}_{Pu}(x)dx = \mathcal{U}_{pu} \\ &\text{and} \quad h(w_{d=0}^*) \text{is a function of } \left(b, \lambda_{n,Pr}, \lambda_{Pr,Pr}, \overline{w}, \rho, \delta, \overline{F}_{Pr}\right). \end{aligned} \tag{2.6}$$

$h(w_{d=0}^*)$ can also be written as

$$h(w_{d=0}^*) = \frac{\lambda_{n,Pr}}{\lambda_{n,Pr} - \lambda_{Pr,Pr}} w_{d=0}^* - c - \frac{\lambda_{Pr,Pr}}{\lambda_{n,Pr} - \lambda_{Pr,Pr}} b \tag{2.7}$$

[4]Note that $\max(x, zp + (1-p)x) = p\max(x, z) + (1-p)x$ with $p \in]0, 1[$

[5]Details are in appendix A

The probability to search for a public job is $\mathbb{P}\left(c_{Pu} \leq \mathcal{U}_{pu}\right)$. When the cost of searching for a public job is greater than $\mathcal{U}_{pu}$, the nonemployed had rather search for a private job only. $\mathcal{U}_{pu}$ depends on individual characteristics, on the probability of passing the exam and finally, on the reservation wage via $h(w^*_{d=0})$.

Equation (2.6) entails that the value of the threshold, determining the search for a public job, is lower when the probability to succeed the entrance exams decreases or when the public offer is lower. $\mathcal{U}_{pu}$ depends on $\lambda_{n,Pu}$ and p_S in a multiplicative way, since $h(w^*_{d=0})$ does not depend on these variables.

The model teachings

First, the higher the probability to pass the exam, the higher the probability to search for a public job. This feature is empirically illustrated by the fact that individuals with higher degree are more likely to pass the exam. For instance, among individuals who take the "*CAPES*" exam (entrance exam for high-school and secondary school) in 2006, 53.2% have a "*licence*" degree and 44.0% a "*maîtrise*" (which requires an additional year of studies). And when considering those who pass the exam, the proportions switch: 57.8% who pass the exam have a maîtrise and 40.7% a licence. Similar features hold for other public-sector exams.

An unemployment growth is captured by the parameter δ. And $h(w^*_{d=0})$ depends on δ only by the way of $w^*_{d=0}$, and $w^*_{d=0}$ diminishes when δ grows (see equation (2.7)). Hence the probability to search for a public job increases with δ.

The possible limits of the model

This model presents several limits: unemployed and nonparticipants are not distinguished. As the sample is composed of people aged between 17 and 59, largely economically active, this assumption is not very restrictive. Furthermore, the model assumes that public workers can not search for a private job. This assumption is made for sake of simplicity but if relaxed, our core results would not be modified. The function value associated to the public sector would be larger, since more flexible as the French civil servants would be able to work a few years in the private sector and get back to public services. Henceforth they could not suffer from their private-job experience: even if laid off, they could recover a job in public services. In such a case, their wage would be the wage they had when departing for a private experience.

Finally the model does not take into account possible parental or sabbatic leaves for civil servants. Moreover we do not enable direct transitions from the public to the private sector, and vice versa. These limits would modify both the public and the private function values but the core results would not change.

3. The data

3.1. A brief description

The data used are taken from the French European Household Survey which was set up by the European Union via Eurostat. This survey analyzes and follows the wage and employment dynamics. Individuals were interviewed annually over 8 years, from 1994 to 2001.

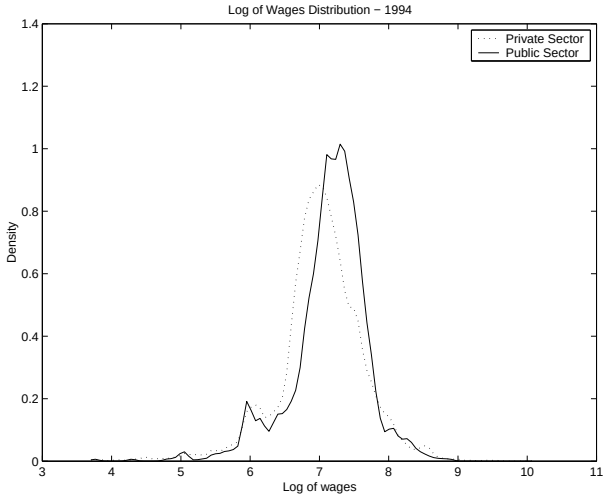
In the following, we consider a balanced panel. The individuals who work, but for whom the wage is not declared, are excluded. Self-employed workers and unpaid workers in a family business are also excluded. Finally the data consist in $5,092$ individuals, between 16 and 59, who are followed from 1994 to 2001.

Further, we consider the highest education degree instead of the educational level. In France, civil-servant exams are actually conditional on degrees and not on a certain number of years of studying. The activity sector is declared by individuals. We consider that they work in the public sector when they answered that they work for state or local governments. Otherwise, they are said to work in the private sector, namely that they are employed by national or private firms, or by their own firm.[6]

The French European Household survey provides annual wage earnings and a monthly description of the occupation but it does not provide the specific wage of the job occupied during the month of the interview. We divide the annual wage earnings by the total number of months employed during the year. In order to get an accurate estimate for this monthly wage, cautiousness is required when people had two jobs or more in a given year, but this actually concerns less than 5% of the employed people in a given year. Hence, for these people, we divide the annual wage earnings by the number of months they are employed whatever the number of jobs they got throughout the year. When there is a single employment spell in a year, no problem arises, since the data set distinguishes benefits from wages. Finally, monthly wages are assessed in euros at 1994 prices.

[6]There is no cross-validation of their employment status as in Card (1996) and the different public services can not be distinguished.

Figure 1 Density of the log of wages in 1994.



3.2. Some descriptive statistics on balanced data

The initial sample (without age restriction) contains half nonemployed people and half employed people. Logically the sample restricted to people aged less than 60 over-represents employed people (table (1)). In the data set, the public sector is over represented compared to usual national statistics. The public sector amounts to one third of employed people, versus one fourth for national statistics. This can be due to the fact that short-term jobs in public services may be occupied by individuals who declare themselves as civil servants, and who are not. And this over representation of civil-servants is not due to the balanced panel, as this proportion is also higher when assessed on the non balanced data set.

The proportion of female is traditionally higher in the public than in the private sector, since it is easier to conciliate professional and family lives while employed in public sectors (table (2)). Few foreigners work in the public sector because entrance exams often require French citizenship. Entrance exams also require a minimum level of degree and a certain age, so degree distribution differs within sectors. There are less graduates in the private sector compared to the public one, and there are less high-school drop-outs. The same remarks hold for the age distribution.

The log monthly wage variance is lower in the public sector than in the private one whatever the degree considered. The log monthly wage mean is also higher in the public sector when degree is not controlled for, see (figure (1), appendix B). When degree is controlled for, this still holds for high-school dropouts and vocational technical degrees, but this mean wage is roughly equivalent in

both sectors for high-school graduates, and university degrees seem to be less rewarded in the public sector. The public sector would play an insurance role for lower degrees, it would protect them from too low wages and it would guarantee them a lower variance. As descriptive statistics do not control for selectivity, the model we estimate goes one step further. It controls for participation selection and sector choice, such as unobserved and observed heterogeneity: age, experience,[7] region, number of children less than 3 years old, between 3 and 6...

Finally consider some figures about transitions. From 1994 to 2001, 816 individuals are non-employed, and $2,384$ are employed. Over these $2,384$ individuals, 775 are employed in public services, and $1,544$ are employed in the private sector. Moreover, $4,276$ individuals are employed at least once throughout this period. Few direct transitions from the public to the private sectors and from the private to the public sectors are observed: 93 transitions from the public to the private sector over the whole period (experienced by 92 individuals), against 67 from the private to the public sector (experienced by 67 individuals). 217 individuals are employed at least one period in the public sector and at least one period in the private sector. 334 individuals previously employed in the public sector become non-employed, which confirms that very few individuals transit from public employment to nonparticipation.

Table 1 Nonemployment, employment in public and private sectors.

Activity status	1994	1995	1996	1997	1998	1999	2000	2001
Non Employed	1,692	1,653	1,663	1,671	1,655	1,637	1,692	1,740
	(33%)	(32%)	(33%)	(33%)	(32%)	(32%)	(33%)	(34%)
Employed	3,400	3,439	3,429	3,421	3,437	3,455	3,400	3,352
	(67%)	(68%)	(67%)	(67%)	(68%)	(68%)	(67%)	(66%)
Sector								
Public	1,061	1,084	1,086	1,057	1,052	1,042	1,014	993
	(31%)	(32%)	(32%)	(31%)	(31%)	(30%)	(30%)	(30%)
Private	2,339	2,355	2,343	2,364	2,385	2,413	2,386	2,359
	(69%)	(68%)	(68%)	(69%)	(69%)	(70%)	(70%)	(70%)

Source: French European Household.

[7] Experience can not be precisely measured in the data set. Hence it is assessed by age minus age at the end of studies. It is a proxy for general experience and not for sector-specific experience.

Table 2 General descriptive statistics

	Whole Sample		**Private Sector**		**Public Sector**	
	Number	Percent	Number	Percent	Number	Percent
Gender						
Women	2,750	54.01	955	40.83	628	59.19
Men	2,342	45.99	1,384	59.17	433	40.81
Nationality						
French	4,887	95.97	2,220	94.91	1,053	99.25
Not French	205	4.03	119	5.09	8	0.75
Region						
Paris	698	13.71	351	15.01	170	16.02
Out of Paris	4,394	86.29	170	84.99	891	83.98
Age						
16-29	1,442	28.32	555	23.73	170	16.02
30-39	1,395	27.40	742	31.72	340	32.05
40-49	1,378	27.06	741	31.68	372	35.06
50-59	877	17.22	301	12.87	179	16.87
Highest diploma						
No secondary degree	1,838	36.10	742	31.72	276	26.01
Vocational technical school (Basic)	1,449	28.46	851	36.38	242	22.81
High school degree (general or vocational)	745	14.63	308	13.17	143	13.48
Technical College, undergraduate university, or *Licence, Maitrise*	765	15.02	299	12.78	270	25.45
Graduates	295	5.79	139	5.94	130	12.25
Part-time	541	15.9	360	15.4	181	17.1

The statistics given above are assessed on the first year 1994.

Source: French European Household.

4. Econometric model and estimation principles

4.1. Model

The structural model described above confirms that the sector choice equation can be modeled by a binary variable: nonemployed search for both public and private jobs when $c_{Pu} \leq \mathcal{U}_{pu}$. And this threshold depends on explanatory variables such as individuals covariates and local unemployment rate.

In addition to the sector choice equation, we consider an employment equation, and a switching wage equation. Hence the reduced form model is composed of four equations. The first one describes employment, namely the fact that the individual works or not ($y_{it} = 1$ if the individual

works). The second one describes the sector choice: public versus private ($z_{it} = 1$ when the individual works in the private sector, $z_{it} = 0$ otherwise). Finally, the third (resp. fourth) equation is the log monthly wage in the private sector w_{it}^{Pr}(resp. public sector w_{it}^{Pu}).

$$y_{it} = \mathbb{I}\left(X_{it}^{Y}\beta^{Y} + \theta_{i}^{Y} + u_{it}^{Y} > 0\right) \tag{4.8}$$

And if the individual works,

$$z_{it} = \mathbb{I}\left(X_{it}^{Z}\beta^{Z} + \theta_{i}^{Z} + u_{it}^{Z} > 0\right) \tag{4.9}$$

$$w_{it}^{Pr} = z_{it}\left(X_{it}^{Pr}\beta^{Pr} + \theta_{i}^{Pr} + u_{it}^{Pr}\right) \tag{4.10}$$

$$w_{it}^{Pu} = (1 - z_{it})\left(X_{it}^{Pu}\beta^{Pu} + \theta_{i}^{Pu} + u_{it}^{Pu}\right) \tag{4.11}$$

where

$\theta = \left(\theta^{Y}, \theta^{Z}, \theta^{Pr}, \theta^{Pu}\right)$ denotes the vector of the unobserved heterogeneity components,

$X = \left(X^{Y}, X^{Z}, X^{Pr}, X^{Pu}\right)$ denotes the vector of the observable characteristics,

$\beta = \left(\beta^{Y}, \beta^{Z}, \beta^{Pr}, \beta^{Pu}\right)$ denotes the vector of parameters,

$u = \left(u^{Y}, u^{Z}, u^{Pr}, u^{Pu}\right)$ denotes the vector of residuals, which are assumed to be independent and normally distributed across time and individuals, with variance σ_{Pr}^{2} (resp. σ_{Pu}^{2}).[8]

As unobserved heterogeneity is crucial to understand different economic behaviors and unobserved productivity, we integrate unobserved terms in each equation. And following the method of Heckman and Singer (1984), we model unobserved heterogeneity θ via a discrete random variable whose distribution has a given number of support points and has to be estimated. The model is estimated using an EM algorithm and standard errors are obtained by a parametric bootstrap. The effects of unobservables on wages vary between public and private sectors. It means that an individual may have a sector specific ability and his unobserved ability will be differently rewarded.

The model does not include lagged dependent variables. It could be included to capture state dependence. In such a case, the introduction of initial conditions could be solved using the method

[8]The assumption of independence across residuals is not restrictive since we follow Heckman and Singer (1984) to model the unobserved heterogeneity, see also Cameron and Heckman (1998) and Arcidiacono (2005).

proposed by Wooldridge (2005) or the one proposed by Heckman (1981). This would entail to compute the likelihood recursively, conditional on the initial conditions. Finally, the model is estimated on a balanced panel, it could have been estimated easily on an unbalanced one.

4.2. Likelihood and estimation principles

We do not use *Simulated Maximum Likelihood* to estimate the model, since it is too time consuming and presents convergence failures. Even with precise and accurate initial conditions, the program fails to converge quickly and it seems to get trapped in some regions.

Instead of that, as previously mentioned, we follow the method proposed by Heckman and Singer (1984) and we consider a discrete distribution for the heterogeneity terms. The heterogeneity is modeled by K distinct types $(\theta_1, \theta_2, ..., \theta_K)$ with $\theta_k = \left(\theta_k^Y, \theta_k^Z, \theta_k^{Pr}, \theta_k^{Pu}\right)'$. Let π_k denote the unconditional probability that an individual belongs to the type k. As it is usual for finite mixture distributions, we rely on the Expectation-Maximization (EM) algorithm (Dempster, Laird, and Rubin (1977)) to estimate the model.

This algorithm iterates the two following steps until the stability of the log-likelihood (for detailed explanations, see appendix C). At each iteration n of the algorithm, we use the values π_k^n and the values $(\theta_1^n, \theta_2^n, ..., \theta_K^n)$ of the mixture distribution, and the values $\left(\beta^n, \sigma_{Pu}^{2,n}, \sigma_{Pr}^{2,n}\right)$ of the parameters of interest.

E-step

For each type $k = 1, ..., K$ and each individual i, the posterior probability of type k is:

$$\mathbb{P}\left(T_i = k | \underline{y}_i, \underline{z}_i, \underline{w}_i\right) = \frac{\pi_k \mathbb{P}(\underline{y}_i, \underline{z}_i, \underline{w}_i | T_i = k)}{\sum_{l=1}^{K} \pi_l \mathbb{P}(\underline{y}_i, \underline{z}_i, \underline{w}_i | T_i = l)}. \tag{4.12}$$

where T_i is the random variable representing the type of the individual i. $\pi_{ik}^{(n)}$ denotes these posterior probabilities.

M-step

The expected completed log likelihood is maximized:

$$\max_{\substack{\beta,\sigma^2_{Pu},\sigma^2_{Pr},\\(\pi_k)_{k=1,\dots,K},\\(\theta_k)_{k=1,\dots,K}}} \sum_{i=1}^{N}\sum_{k=1}^{K} \pi_{ik}^{(n)} \ln l\left(\underline{y}_i,\underline{z}_i,\underline{w}_i|T_i=k,\beta,\sigma^2_{Pu},\sigma^2_{Pr},(\pi_k)_{k=1}^K,(\theta_k)_{k=1}^K\right) \quad (4.13)$$

First $\pi_k^{(n)}$ is updated such that:

$$\pi_k^{(n+1)} = \frac{\sum_{i=1}^N \pi_{ik}^{(n)}}{\sum_{l=1}^K\sum_{i=1}^N \pi_{il}^{(n)}} \quad (4.14)$$

Second the following three optimization problems are solved separately, thanks to the separability of the conditional completed log-likelihood (see Appendix C).

1. Optimization on the employment equation parameters

$$\max_{\beta^Y,(\theta_k^Y)_{k=1}^K} \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T} \pi_{il}^{(n)}\mathbb{I}_{y_{it}=0}\ln\Phi\left(-X_{it}^Y\beta_Y-\theta_k^Y\right) + \pi_{il}^{(n)}\mathbb{I}_{y_{it}=1}\ln\Phi\left(X_{it}^Y\beta_Y+\theta_k^Y\right),$$

2. Optimization on the sector choice equation parameters

$$\max_{\beta^Z,(\theta_k^Z)_{k=1}^K} \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T} \pi_{il}^{(n)}\mathbb{I}_{y_{it}=1,z_{it}=0}\ln\Phi\left(-X_{it}^Z\beta_Z-\theta_k^Z\right) + \pi_{il}^{(n)}\mathbb{I}_{y_{it}=1,z_{it}=1}\ln\Phi\left(X_{it}^Z\beta_Z+\theta_k^Z\right),$$

3. Optimization on the wage equations parameters

$$\min_{\beta^{Pr},\sigma_{Pr},(\theta_k^{Pr})_{k=1}^K} \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T} \pi_{il}^{(n)}\mathbb{I}_{y_{it}=1,z_{it}=1}\ln\sigma_{Pr} + \frac{\pi_{il}^{(n)}}{2\sigma^2_{Pr}}\mathbb{I}_{y_{it}=1,z_{it}=1}(w_{it}^{Pr}-X_{it}^{Pr}\beta_{Pr}-\theta_k^{Pr})^2,$$

and

$$\min_{\beta^{Pu},\sigma_{Pu},(\theta_k^{Pu})_{k=1}^K} \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T} \pi_{il}^{(n)}\mathbb{I}_{y_{it}=1,z_{it}=0}\ln\sigma_{Pu}$$

$$+ \quad \frac{\pi_{il}^{(n)}}{2\sigma_{Pu}^2}\mathbb{I}_{y_{it}=1,z_{it}=1}(w_{it}^{Pu} - X_{it}^{Pu}\beta_{Pu} - \theta_k^{Pu})^2.$$

Standard errors estimates are obtained by a parametric bootstrap procedure, instead of a non parametric one, since this last method is unstable when applied to the EM algorithm. The parametric bootstrap consists first in obtaining reliable parameter estimates for the whole set of unknown parameters denoted $\widehat{\chi}$. $\widehat{\chi}$ is obtained by replicating the previously described EM algorithm with different random initial values for the parameters. The iteration process is necessary to ensure that a global maximum is obtained. Then, given X and $\widehat{\chi}$, we generate H vectors of the endogenous variables $\left(y_i^h, z_i^h, w_i^h\right)_{h=1\ldots H}$. For each newly generated data set, we estimate the whole set of unknown parameters. Final parameters and standard error estimates are computed as

$$\overline{\beta^*} = \frac{1}{H}\sum_{h=1}^{H}\beta_h^* \tag{4.15}$$

$$\sigma_{\beta^*} = \frac{1}{H-1}\sum_{h=1}^{H}(\beta_h^* - \overline{\beta^*})^2. \tag{4.16}$$

5. Results

5.1. Identification

The model identification does not only rely on functional assumptions, imposed on the residual distributions, but on exclusion restrictions. On the one hand, in the nonemployment/employment equation, we include the number of children between 0 and 3, and the number of children between 3 and 6. These variables are crucial for explaining female participation (see Hyslop (1999) and Edon and Kamionka (2008)). These variables are excluded from the sector choice and the wage equation. The fertility of a woman may influence the sector choice but this component is captured by the unobserved heterogeneity specific to the sector choice equation.

On the other hand, in the sector choice equation, a proxy of the father status as a proxy of the mother status at the end of the individual studies, is included. This is known to be a determinant of civil servant status (Audier (2000)). The status of the parents occupation is not directly observed, a proxy is built from the two-digit classification of their occupation. Hence we consider the father (resp. mother) was a civil servant when he (resp. she) was either *'senior civil servants, information professionals or creative and performing artists'* or *'middle-level health and teaching workers,*

middle-level civil servants' or finally *'middle-level civil servants'*. These variables are excluded from the rest of the model.[9]

5.2. Estimations

5.2.1. Estimation results

Employment

As expected, gender negatively affects employment, as well as children under 3 and 6 do. Women choose to get out of the labor market to bring up their children, and usually wait for their entrance into nursery or primary school to look for a job. In France, some children part-time attend nursery school. The marriage is not determinant for the employment decision. The effect of age on employment is quadratic. The employment probability first increases, and then decreases with age.

Moreover, the higher the degree, the higher the probability to be employed. Graduate degrees are the most rewarding ones in terms of employment probabilities.

Individuals who live in the region around Paris or in Paris have a greater probability to be employed because of job offer opportunities. Finally, as expected, local unemployment rate has a negative effect on the probability to be employed.

Sector Choice

Our theoretical model teaches us that, given risk aversion, a higher unemployment rate diminishes the cost of searching for a public job when unemployed. The return associated to the private sector drops when the unemployment rate rises. This is confirmed by our estimations: the local unemployment rate favors the choice of public sector (Table 3). Fougère and Pouget (2003) also find that the number of candidates for a public job and the macroeconomic cycle go along. Hence the local unemployment rate is a core variable to understand public sector attractiveness.

The father's position directly influences his children sector choice. A son would prefer public services if his father is a civil-servant. This confirms empirical observations: civil servants'

[9] In 1994, among women who work, 13.3% of them have a father who was civil-servant, and 10.9% a mother who was civil-servant. For men, these percentage are respectively 12.4% and 8.1%.

children are over-represented in the public sector (Audier (2000)). But women whose father is a civil-servant would prefer private sector compared to women whose father is not a civil-servant.

Moreover, as expected, women have a higher propensity than men to work in public services. This would be enforced thanks to the wage analysis. This result is in line with Bell, Elliott, and Scott (2005), who find, for UK, that the gains to staying in the public sector are greater for women than for men, and that women tend to gain by joining the public sector, almost irrespective of their position in the earnings distribution. We find similar results for France (see following sections). Fougère and Pouget (2003) also find that the length of queue for public sector is longer for female.

Individuals with lower degree tend to work in the private sector unlike individuals with post-secondary education. Two reasons may explain this effect: first, French public jobs are more qualified on the whole than jobs in private sector (teaching, executive...). And many public jobs require to pass an exam which is conditional on a given degree.

Wages

Usual results are obtained in terms of effects of different variables, such as age, experience, degree... The comparison between private and public wages is more instructive.

Unlike empirical results, the residual variance is slightly higher in the public sector (0.11) than in the private one (0.08). And, the constant associated to public wages is higher (5.82 against 5.08), what people usually call "the public raw wage premium". For most degrees (except for graduates) we can not conclude that education is less rewarded in public services. But for graduates, the returns to education are statistically different and greater in the private sector.

Worth to be pointed out is the worst returns to gender in the private sector. As Gregory and Borland (1999) underlined, public sector wage inefficiency may counterbalance wage discriminations. That seems to be the case in France. Nonetheless, to be a woman has also a negative effect on public wages. Indeed women have lower probabilities to be promoted than men with similar characteristics (Bessière and Pouget (2007)). But in public services, wage increases are closely linked to grade promotions. Thus career differences may explain wage

differences between men and women in public services.

Finally, living around Paris and in Paris gives a positive premium for wages.

We give further details on wages in the following section thanks to simulations.

Table 3. Employment equation

Variables	*Estimates*	*Standard Error*
Constant	-5.515	0.092
Not French	0.137	0.048
French	-	-
Not Married	0.032	0.017
Married	-	-
Women	-0.546	0.015
Men	-	-
Children under 3	-0.237	0.021
Children between 3 and 6	-0.259	0.021
Age/10	4.556	0.567
Age/10 squared	-5.651	0.673
No secondary degree	-	-
Vocational degree	0.217	0.021
High School degree	0.264	0.026
College and Under Graduate	0.335	0.023
Graduate	0.706	0.035
Paris	0.206	0.027
Local unemployment rate	-0.012	0.003

Source: French European Household.

Table 4. Private sector choice

Variables	*Estimates*	*Standard Error*
Intercept	1.067	0.251
Women	-0.604	0.041
Men	-	-
Age/100	3.024	1.483
Age/100 squared	-5.594	1.795
Not Married	0.032	0.042
Vocational degree	0.160	0.050
High School degree	0.317	0.055
College and Under Graduate	-0.122	0.041
Graduate	-0.109	0.063
Women times Father civil servant	0.406	0.080
Men times Father civil servant	-0.361	0.061
Women times Mother civil servant	0.070	0.085
Men times Mother civil servant	-0.109	0.083
Local unemployment rate	-0.047	0.006

Source: French European Household.

Table 5. Public wage equation

Variables	*Estimates*	*Standard Error*
Intercept	5.822	0.041
Women	-0.185	0.007
Paris	0.118	0.009
Age/100	4.919	0.271
Age/100 squared	-4.152	0.327
Age of end of study/100	0.308	0.205
Age of end of study/100 squared	-1.560	0.689
Vocational degree	0.068	0.010
High School degree	0.339	0.016
College and Under Graduate	0.459	0.015
Graduate	0.667	0.015
Part-time	-0.436	0.010
Variance of residuals	0.110	0.002

Source: French European Household.

Table 6. Private wage equation

Variables	*Estimates*	*Standard Error*
Intercept	5.075	0.028
Women	-0.385	0.005
Paris	0.287	0.007
Age/100	8.935	0.188
Age/100 squared	-9.284	0.232
Age of end of study/100	1.806	0.131
Age of end of study/100 squared	-1.787	0.358
Vocational degree	0.106	0.006
High School degree	0.351	0.008
College and Under Graduate	0.428	0.008
Graduate	0.800	0.010
Part-time	-0.494	0.007
Variance of residuals	0.083	0.001

Source: French European Household.

5.2.2. Types

Table 7. Other parameters

Variables	*Estimates*	*Standard Error*
Probability of different types		
Type 1	0.208	0.006
Type 2	0.218	0.006
Type 3	0.225	0.006
Type 4	0.349	0.006
Type-specific heterogeneity parameters		
	Type 1	
θ^{Y}	-1.051	0.027
θ^{Z}	1.972	0.101
$\theta^{W_{Pr}}$	0.014	0.011
$\theta^{W_{Pu}}$	0.184	0.030
	Type 2	
θ^{Y}	-3.788	0.036
θ^{Z}	0.046	0.086
$\theta^{W_{Pr}}$	-1.199	0.013
$\theta^{W_{Pu}}$	-0.555	0.020
	Type 3	
θ^{Y}	-1.235	0.033
θ^{Z}	-2.581	0.087
$\theta^{W_{Pr}}$	-0.247	0.016
$\theta^{W_{Pu}}$	0.007	0.016
	Type 4	
θ^{Y}	-1.924	0.032
θ^{Z}	1.630	0.087
$\theta^{W_{Pr}}$	-0.468	0.011
$\theta^{W_{Pu}}$	-1.031	0.024

Source: French European Household.

Four types were chosen instead of three because of the insufficient fit we had and the peculiar results we got. Three were not sufficient to capture heterogeneity. Why not five? For computational reasons, since four types are already heavy to implement.

In a first step, we compute the individual posterior probability to be of a given type at the initial date 1994. We consider that an individual is of type k when the individual posterior probability associated to type k is the highest. Results show that a clear type-partition exists. At the initial date, type-2 people are mainly nonemployed (85.9%) whereas type-1, type-3 and type-4 individuals are employed. These latter groups differ according to the sector choice. 85.1% of type-3 individuals

are employed in public services in 1994, whereas 88.4% of type-1 individuals and 74.1% of type-4 individuals are employed in the private sector. Thus employment and sectors distinctly partition individuals in our sample.

Considering the values of unobserved heterogeneity for different types, remark that *type-1* and *type-3* individuals have an unobserved ability roughly similar in both sectors (for type 1: $\theta^{W_{Pr}} = 0.184$ and $\theta^{W_{Pu}} = 0.014$; for type 3: $\theta^{W_{Pr}} = -0.247$ and $\theta^{W_{Pu}} = 0.007$), whereas type-4 individuals would have a different unobserved productivity in both sectors. These individuals clearly have chosen the sector in which they are the most productive. Type-4 individuals are clearly more efficient in the private sector than in the public sector ($\theta^{W_{Pr}} = -0.468$ and $\theta^{W_{Pu}} = -1.031$).They may have sorted themselves into the sector that pays them more.

6. Simulations

6.1. Model fit

This section presents a brief fit analysis of the statistical model presented above. Table (8) presents the predicted probabilities for nonemployment, public jobs and private jobs. The frequencies of nonemployment and of employment in either sector are well replicated whatever the date considered.

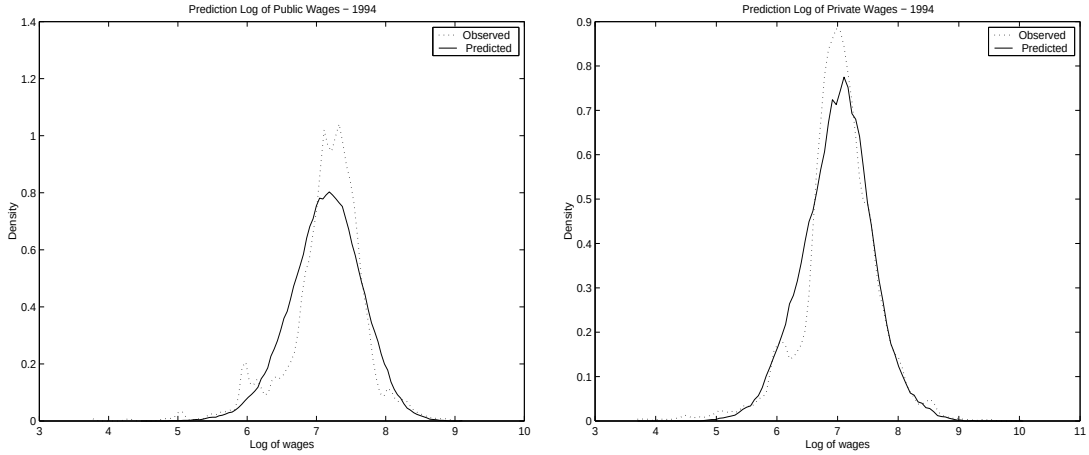
Further the model replicates well the cross sectional distribution. Figure (2) plots the observed and predicted log of monthly wage densities for the two sectors separately, which are quite close.

Table 8. Model Fit

Variables	*Year*	*Predicted probability* %	*Observed probability* %
Nonemployed	1994	34.379	33.229
	1995	33.522	32.463
	1996	33.160	32.659
	1997	32.840	32.816
	1998	32.711	32.502
	1999	32.765	32.148
	2000	33.035	33.229
	2001	33.800	34.171
Public sector	1994	20.105	20.837
	1995	20.356	21.288
	1996	20.605	21.328
	1997	20.647	20.758
	1998	20.586	20.660
	1999	20.355	20.463
	2000	20.033	19.914
	2001	19.804	19.501
Private sector	1994	45.516	45.935
	1995	46.122	46.249
	1996	46.235	46.013
	1997	46.513	46.426
	1998	46.703	46.838
	1999	46.880	47.388
	2000	46.932	46.858
	2001	46.396	46.328

Source: French European Household

Figure 2 Predicted log of monthly wages - 1994



6.2. What private wages would civil servants have? Counterfactual distributions

We derive the counterfactual distribution of the log monthly wages for the individual employed in the public sector. Which log of wages would public servants have if they were employed in the private sector?

We use bootstrapping methods: we draw H independent replicates drawn from the empirical distribution of the explanatory variables of people working in the public sector at date t. For each replicate, we draw a type $t(i)$ in the following posterior distribution

$$\mathbb{P}\left(T_i = k | y_{it} = 1, z_{it} = 0, X_{it}\right) = \frac{\pi_k \mathbb{P}(y_{it} = 1, z_{it} = 0 | T_i = k, X_{it})}{\sum_{l=1}^{K} \pi_l \mathbb{P}(y_{it} = 1, z_{it} = 0 | T_i = l, X_{it})}, \quad (6.17)$$

where T_i is the random variable representing the type of the individual i. Then we compute the corresponding fitted value of the log monthly wages in both sectors:

$$w_{it}^{s,Pr} = X_{it} \widehat{\beta^{Pr}} + \widehat{\theta_{t(i)}^{Pr}}, \quad (6.18)$$

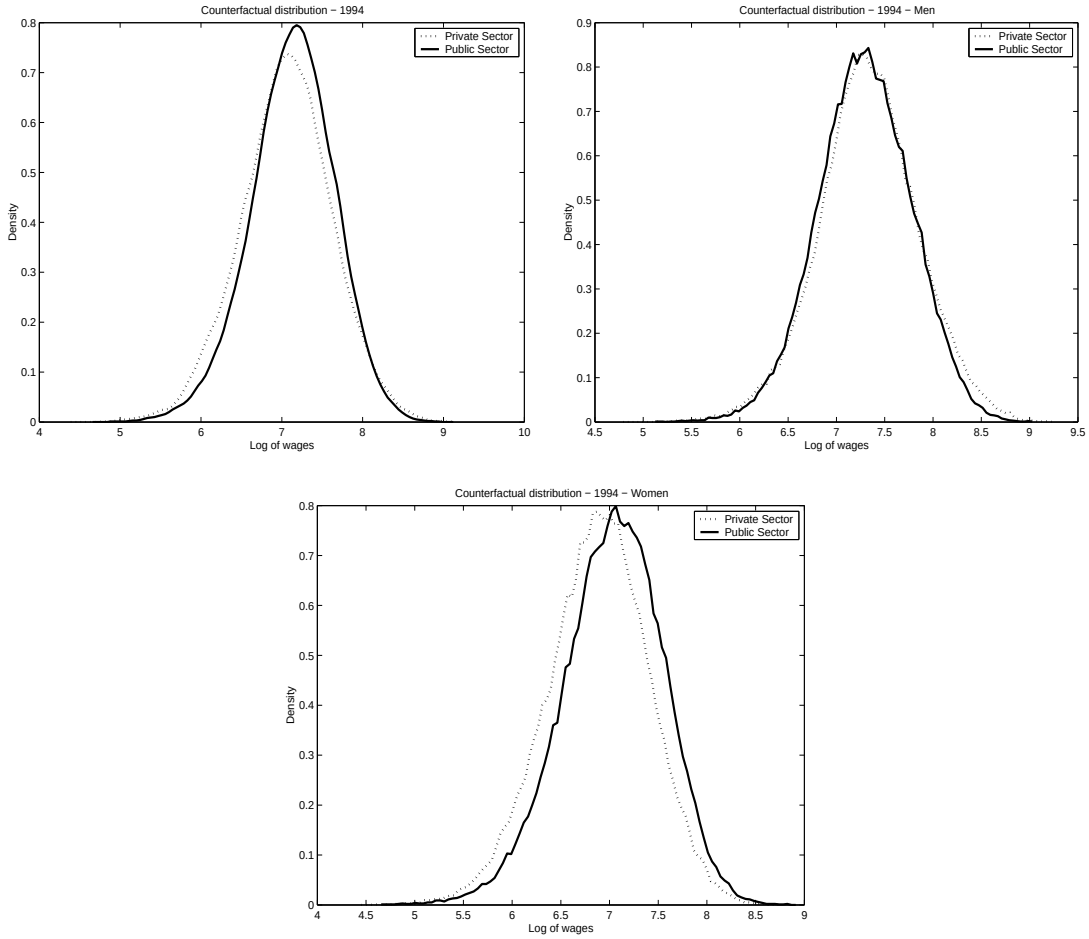
$$w_{it}^{s,Pu} = X_{it} \widehat{\beta^{Pu}} + \widehat{\theta_{t(i)}^{Pu}}. \quad (6.19)$$

Finally we add a sector-specific residual term that is i.i.d normally distributed with a sector-specific variance.

In figure 3, we observe that the counterfactual distribution remains close to the log of wage public distribution. But this hides different effects according to gender and degrees. The counterfactual

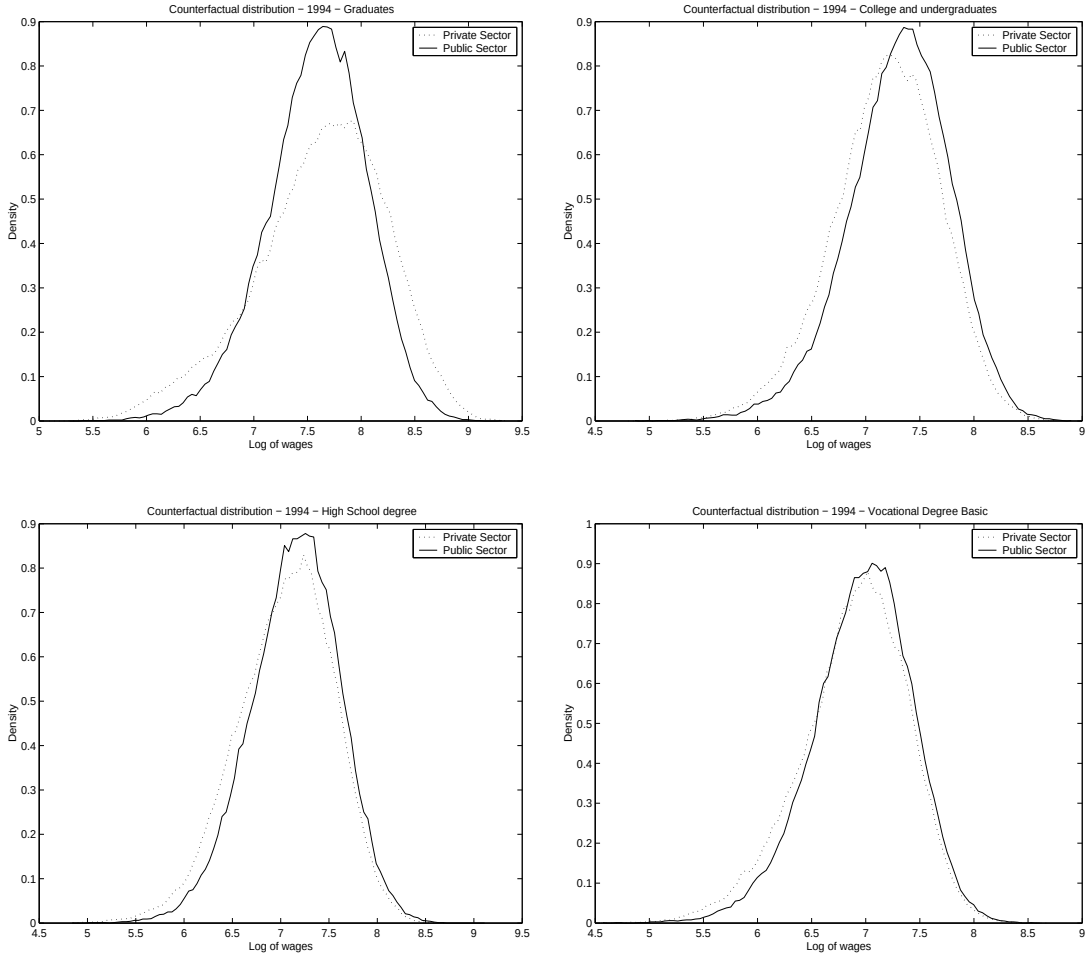
distribution for men is rather stable except for the upper tail, whereas the one for women shifts left.

Figure 3 Counterfactual log of monthly wages for civil servants - 1994



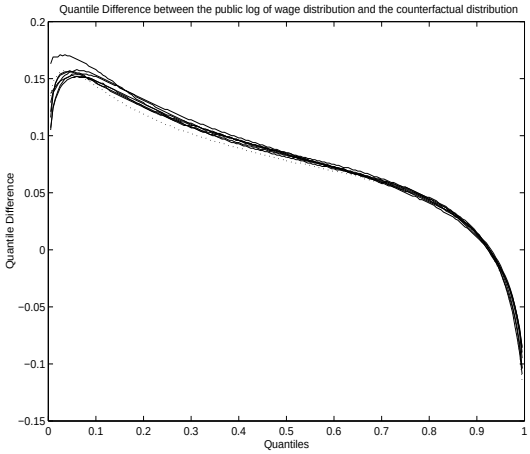
In order to precisely evaluate the public wage premium, let us derive and compare quantiles of the former distributions: the public log of wage distribution and the counterfactual distribution. Further details are given in appendix E. Figures 5 and 8 plot the difference of quantiles between the current and the counterfactual distributions according to gender and degrees. The term "public wage premium" seems to be justified at the bottom of the distribution, whereas it is not anymore for upper wages. But when detailed by degree and gender, we find that the public sector actually gives a premium to women, whereas men would be better paid in the private sector except for those at the lower tail of the distribution. As expected given previous estimations, male graduates would be far better paid in the private sector. So we find similar results as Disney and Gosling (1998) and Bell, Elliott, and Scott (2005): public sector premium is higher for women than for men, and this

Figure 4 Counterfactual log of monthly wages according to degrees



premium differs across the pay distribution. The lower part of the distribution gains from staying in the public sector.

Figure 5 Quantile differences between counterfactual and public log of wage distribution.



7. Conclusion

This paper presents new evidence on the public-private pay gap. The model estimates sector-specific wage equations controlling both for sector choice and employment selection. Thanks to panel data, it captures unobserved heterogeneity which differs between equations. It is modeled using the method of Heckman and Singer (1984) via a discrete distribution.

We find that, when choosing the sector, mimicking the parents has a great importance. But only sons imitate their father's choice. As expected, the public sector attracts more people when the local labor market is depressed. We observe that the wage public premium is effective at the bottom of the public wage distribution. But it turns out to be false at the upper tail. This is in line with empirical observations: low-wage civil-servants are weakly mobile, whereas high-wage civil servants move more frequently from the public to the private sector.

However these results are mitigated when gender is distinguished. The female public sector workers have a comparative advantage in the public sector. Their counterfactual wages in the private sector would be lower. Graduated men would have higher wages in the private sector; in the case of graduated women, the public wage premium is close to zero.

Finally we find that wage differences between the public and the private sectors result from three factors: first the raw mean is greater in the public sector, second the returns to different observable characteristics differ between both sectors, finally the unobserved productivity civil-servants would have in the private sector is roughly equal, though inferior, to the one they have in their current

occupation. Unlike them, a part of workers employed in the private sector seem to acquire specific human capital, that they could not transfer to the public sector.

A. Structural model

From (2.3), differentiating this latter equation with respect to w yields:

$$\begin{aligned} \rho V_{Pr}^{'}(x) &= 1 - \delta V_{Pr}^{'}(x) - \lambda_{Pr,Pr} V_{Pr}^{'}(x)\overline{F}(x) \\ \Rightarrow V_{Pr}(x) - V_{Pr}(\overline{w}) &= -\int_{x}^{\overline{w}_{Pr}} \frac{1}{\rho + \delta + \lambda_{Pr,Pr}\overline{F}_{Pr}(y)} dy \end{aligned}$$

Let us consider both cases.

- Case 1: $d_{Pu} = 0$

  When the individual does not search for a public job, if he receives an offer with a wage w greater than $w^*_{d=0}$, he accepts this offer. The reservation wage $w^*_{d=0}$ checks: $V_n^{d=0} = V_{Pr}(w^*_{d=0})$. Henceforth:

$$\begin{aligned} V_n^{d=0}\rho &= (b-c) + \lambda_{n,Pr} \int_{w^*_{Pr,d=0}}^{\overline{w}} \left(V_{Pr}(x) - V_n\right) dF_{Pr}(x) \\ &= (b-c) + \lambda_{n,Pr} \int_{w^*_{Pr,d=0}}^{\overline{w}} V_{Pr}^{'}(x)\overline{F}_{Pr}(x)dx \\ \Rightarrow & \left(\begin{array}{lll} \rho V_n^{d=0} &= & (b-c) + \lambda_{n,Pr} \int_{w^*_{Pr,d=0}}^{\overline{w}} V_{Pr}^{'}(x)\overline{F}_{Pr}(x)dx \\ \rho V_n^{d=0} &= & w^*_{Pr,d=0} - c + \lambda_{Pr,Pr} \int_{w^*_{Pr,d=0}}^{\overline{w}} V_{Pr}^{'}(x)\overline{F}_{Pr}(x)dx \end{array} \right) \end{aligned}$$

  And:

$$\begin{aligned} w^*_{Pr,d=0} &= b \\ &+ (\lambda_{n,Pr} - \lambda_{Pr,Pr}) \int_{w^*_{Pr,d=0}}^{\overline{w}} \frac{\overline{F}_{Pr}(x)}{\rho + \delta + \lambda_{Pr,Pr}\overline{F}_{Pr}(x)} dx \end{aligned} \tag{A.20}$$

- Case 2: $d_{Pu} = 1$

  Let us proceed the same way thus we get:

$$\rho V_n^{d=1} = (b - c - c_{Pu}) + \lambda_{n,Pr} \int_{w^*_{Pr,d=1}}^{\overline{w}} V_{Pr}^{'}(x)\overline{F}_{Pr}(x)dx\lambda_{n,Pr}$$

$$
\begin{aligned}
&+ \quad p\lambda_{n,Pu}\int_{w^*_{Pu}}^{\overline{w}} V'_{Pu}(x)\overline{F}_{Pu}(x)dx \\
\rho V_n^{d=1} &= w^*_{Pr,d=1} - c + \lambda_{Pr,Pr}\int_{w^*_{Pr,d=1}}^{\overline{w}} V'_{Pr}(x)\overline{F}_{Pr}(x)dx \\
V_n^{d=1} &= V_{Pr}(w^*_{Pr,d=1}) = V_{Pu}(w^*_{Pu}) = \frac{w^*_{Pu}}{\rho}
\end{aligned}
$$

Hence:

$$
\begin{aligned}
w^*_{Pu} &= w^*_{Pr,d=1} - c + \lambda_{Pr,Pr}\int_{w^*_{Pr,d=1}}^{\overline{w}} \frac{\overline{F}_{Pr}}{\rho+\delta+\lambda_{Pr,Pr}\overline{F}_{Pr}(x)}(x)dx \\
w^*_{Pr,d=1} &= (b - c_{Pu}) + (\lambda_{n,Pr} - \lambda_{Pr,Pr})\int_{w^*_{Pr,d=1}}^{\overline{w}} \frac{\overline{F}_{Pr}(x)}{\rho+\delta+\lambda_{Pr,Pr}\overline{F}_{Pr}(x)}dx \\
&+ \quad \frac{\lambda_{n,Pu}p}{\rho}\int_{w^*_{Pu}}^{\overline{w}} \overline{F}_{Pu}(x)dx
\end{aligned}
$$

The values we have to compare to determine whether the individual searches for a public job, are $V_n^{d=0}$ and $V_n^{d=1}$, i.e. $V_{Pr}(w^*_{Pr,d=0})$ and $V_{Pr}(w^*_{Pr,d=1})$:

$$
\begin{aligned}
&V_{Pr}(w^*_{Pr,d=1}) - V_{Pr}(w^*_{Pr,d=0}) \\
= \quad &\int_{w^*_{Pr,d=0}}^{\overline{w}} \frac{1}{\rho+\delta+\lambda_{Pr,Pr}\overline{F}(x)}dx - \int_{w^*_{Pr,d=1}}^{\overline{w}} \frac{1}{\rho+\delta+\lambda_{Pr,Pr}\overline{F}(x)}dx
\end{aligned}
$$

$$
\begin{aligned}
&V_{Pr}(w^*_{Pr,d=1}) = V_{Pr}(w^*_{Pr,d=0}) \Leftrightarrow w^*_{Pr,d=1} = w^*_{Pr,d=0} \\
\text{and} \quad &V_{Pr}(w^*_{Pr,d=1}) \geq V_{Pr}(w^*_{Pr,d=0}) \Leftrightarrow w^*_{Pr,d=1} \geq w^*_{Pr,d=0}
\end{aligned}
$$

The nonemployed person searches for both public and private jobs when $w^*_{Pr,d=1} \geq w^*_{Pr,d=0}$.

From equations ($w^*_{Pr,d=0}$), ($w^*_{Pr,d=1}$) and (w^*_{Pu}), we get:

$$
\begin{aligned}
w^*_{Pr,d=0} &= f(w^*_{Pr,d=0}) \\
w^*_{Pr,d=1} &= f(w^*_{Pr,d=1}) + g(w^*_{Pr,d=1})
\end{aligned}
$$

with

$$\begin{aligned} f(x) &= b + (\lambda_{n,Pr} - \lambda_{Pr,Pr}) \int_x^{\overline{w}} \frac{\overline{F}_{Pr}(x)}{\rho + \delta + \lambda_{Pr,Pr}\overline{F}_{Pr}(x)} dx \\ \text{and } g(x) &= -c_{Pu} + \frac{\lambda_{n,Pu} p}{\rho} \int_{h(x)}^{\overline{w}} \overline{F}_{Pu}(x) dx \\ \text{and } h(x) &= x - c + \lambda_{Pr,Pr} \int_x^{\overline{w}} \frac{\overline{F}_{Pr}(x)}{\rho + \delta + \lambda_{Pr,Pr}\overline{F}_{Pr}(x)} dx \end{aligned}$$

h is an increasing and bounded function with first derivative:

$$h^{'}(x) = \frac{\rho + \delta}{\rho + \delta + \lambda_{Pr,Pr}\overline{F}_{Pr}(x)}$$

Hence g is a decreasing and bounded function and three different cases are possible. Indeed, $\forall x \in [0, \overline{w}]$, $0 \geq f'(x) \geq f'(x) + g'(x)$.

- $f(0) + g(0) \leq f(0)$: it means that a nonemployed person never searches for a public job for $w^*_{d=0} > w^*_{d=1}$.
- $f(0) + g(0) \geq f(0)$ and $f(\overline{w}) + g(\overline{w}) \geq f(\overline{w})$: it means that a nonemployed person always searches for both public and private jobs.
- $f(0) + g(0) \geq f(0)$ and $f(\overline{w}) + g(\overline{w}) < f(\overline{w})$: it means that a nonemployed person may search for a public job depending on $g(w^*_{d=0})$. If $g(w^*_{d=0}) \geq 0$ the nonemployed person will search for a public job, whereas she won't.

  We show that

  $$w^*_{d=0} \leq w^*_{d=1} \quad \Leftrightarrow \quad g(w^*_{d=0}) \geq 0$$

  Indeed,

  $\diamondsuit$ On one hand,

  $$\begin{aligned} & w^*_{d=0} \leq w^*_{d=1} \Rightarrow f(w^*_{d=0}) + g(w^*_{d=0}) \geq f(w^*_{d=1}) + g(w^*_{d=1}) \\ \text{And} \quad & w^*_{d=0} \leq w^*_{d=1} \Rightarrow f(w^*_{d=0}) \leq f(w^*_{d=1}) + g(w^*_{d=1}) \\ \text{Thus} \quad & f(w^*_{d=0}) \leq f(w^*_{d=0}) + g(w^*_{d=0}) \Rightarrow g(w^*_{d=0}) \geq 0 \end{aligned}$$

$\diamondsuit$ On the other hand

$$g(w^*_{d=0}) \geq 0 \Rightarrow f(w^*_{d=0}) \leq f(w^*_{d=0}) + g(w^*_{d=0})$$

Thus $\quad w^*_{d=0} \leq f(w^*_{d=0}) + g(w^*_{d=0})$

And $\quad f(x) + g(x) - x$ is a decreasing function such as

$$\left\{ \begin{array}{l} f(w^*_{d=1}) + g(w^*_{d=1}) - w^*_{d=1} = 0 \\ f(w^*_{d=0}) + g(w^*_{d=0}) - w^*_{d=0} \geq 0 \end{array} \right\}$$

$$\Rightarrow \quad w^*_{d=1} \geq w^*_{d=0}$$

Proposition A.1 *Search for a public job*

The nonemployed agent decides to search for both public and private jobs when:

$$c_{Pu} \quad \leq \quad p_S \frac{\lambda_{n,Pu}}{\rho} \int_{h(w^*_{d=0})}^{\overline{w}} \overline{F}_{Pu}(x)dx$$

where $h(w^*_{d=0})$ *is a function of* $\left(b, \lambda_{n,Pr}, \lambda_{Pr,Pr}, \overline{w}, \rho, \delta, \overline{F}_{Pr}\right)$

with:

– c_{Pu} *the cost to search for a public job*

– $\lambda_{n,Pu}$ *the public job offer arrival rate*

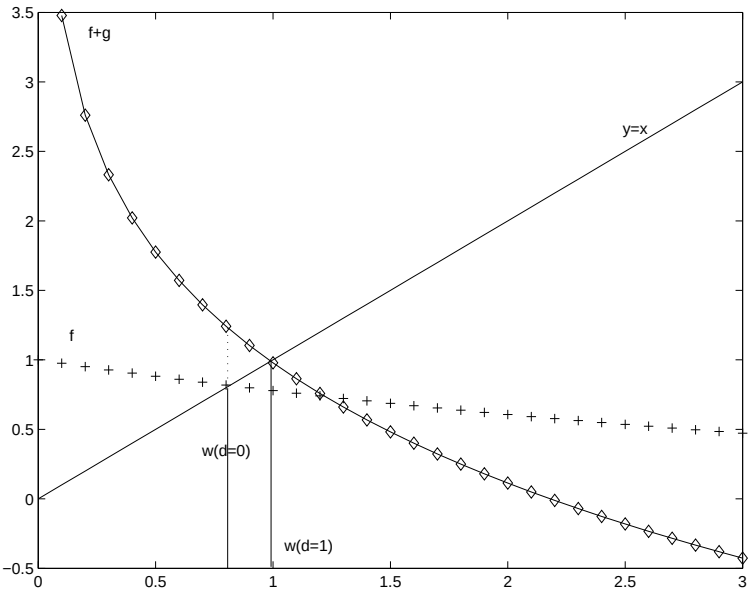
– p_S *the individual probability to succeed public entrance exam*

– $h(w^*_{d=0}) = \frac{\lambda_{n,Pr}}{\lambda_{n,Pr} - \lambda_{Pr,Pr}} w^*_{d=0} - c - \frac{\lambda_{Pr,Pr}}{\lambda_{n,Pr} - \lambda_{Pr,Pr}} b$

He only searches for private jobs when:

$$c_{Pu} \quad > \quad p_S \frac{\lambda_{n,Pu}}{\rho} \int_{h(w^*_{d=0})}^{\overline{w}} \overline{F}_{Pu}(x)dx$$

Figure 6: Function value comparison



B. Descriptive Statistics

C. Implementation

Let me describe precisely the implementation of the M-step.

$$l\left(\underline{y}_i, \underline{z}_i, \underline{w}_i | T_i = k\right) = \prod_{t=1}^{T} \quad \left(\mathbb{P}\left(y_{it} = 0 | T_i = k\right)\right)^{y_{it}=0} \tag{C.21}$$
$$\left(\mathbb{P}\left(y_{it} = 1, z_{it} = 1, w_{it}^{Pr} | T_i = k\right)\right)^{y_{it}=1, z_{it}=1}$$
$$\left(\mathbb{P}\left(y_{it} = 1, z_{it} = 0, w_{it}^{Pu} | T_i = k\right)\right)^{y_{it}=1, z_{it}=0}$$

Given the individual type, for a given date, residuals are independent, hence:

$$l\left(\underline{y}_i, \underline{z}_i, \underline{w}_i | T_i = k\right) = \prod_{t=1}^{T} \quad \left(\mathbb{P}\left(y_{it} = 0 | T_i = k\right)\right)^{y_{it}=0} \left(\mathbb{P}\left(y_{it} = 1 | T_i = k\right)\right)^{y_{it}=1} \tag{C.22}$$
$$\prod_{t=1}^{T} \quad \left(\mathbb{P}\left(z_{it} = 1 | T_i = k\right)\right)^{y_{it}=1, z_{it}=1} \left(\mathbb{P}\left(z_{it} = 0 | T_i = k\right)\right)^{y_{it}=1, z_{it}=0}$$
$$\prod_{t=1}^{T} \quad \left(\mathbb{P}\left(w_{it}^{Pu} | T_i = k\right)\right)^{y_{it}=1, z_{it}=0}$$
$$\prod_{t=1}^{T} \quad \left(\mathbb{P}\left(w_{it}^{Pr} | T_i = k\right)\right)^{y_{it}=1, z_{it}=1}$$

Therefore, we can maximize separately each equation once we condition on the individual type.

D. Cross-section results

Table 9. Public wage equation without selection - 1994

Variables	*Estimates*	*Standard Error*
Intercept	4.568	0.329
Women	-0.183	0.025
Paris	0.133	0.033
Age/100	5.692	1.016
Age/100 squared	-5.162	1.259
Age of end of study/100	11.644	2.508
Age of end of study/100 squared	-27.296	5.949
Vocational degree	-0.019	0.033
High School degree	0.244	0.053
College and Under Graduate	0.323	0.034
Graduate	0.590	0.043
Part-time	-0.609	0.033

Source: French European Household

Table 10. Private wage equation without selection - 1994

Variables	*Estimates*	*Standard Error*
Intercept	4.785	0.178
Women	-0.252	0.020
Paris	0.234	0.026
Age/100	7.903	0.685
Age/100 squared	-8.384	0.887
Age of end of study	4.622	1.207
Age of end of study squared	-8.349	2.792
Vocational degree	0.076	0.021
High School degree	0.307	0.036
College and Under Graduate	0.382	0.033
Graduate	0.698	0.044
Part-time	-0.687	0.027

Source: French European Household

Figure 7 Wages according to degrees in 1994

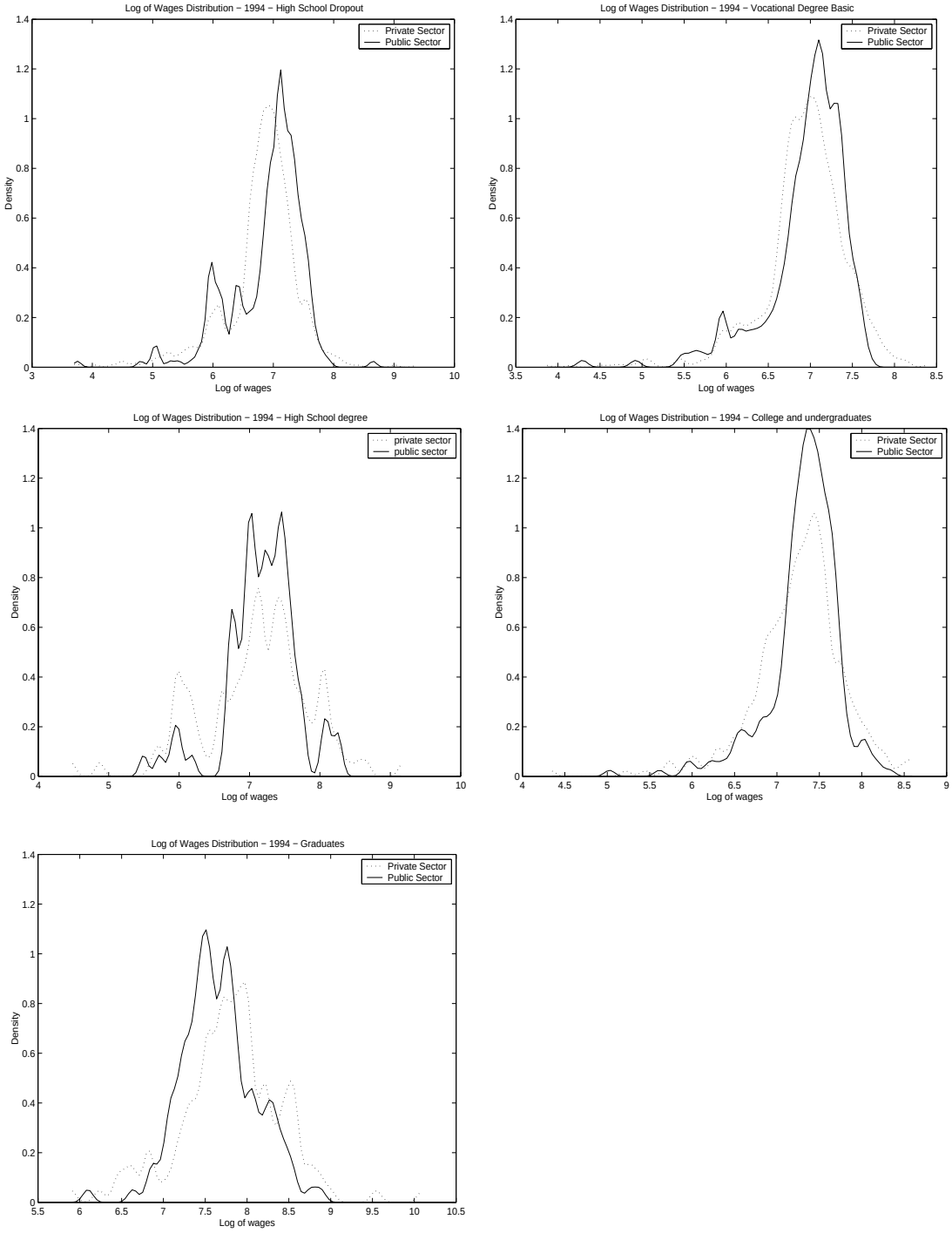


Table 11. Switching regression model with sector selection - 1994

Variables	*Estimates*	*Standard Error*
Private wage equation		
Intercept	4.787	0.178
Women	-0.256	0.023
Paris	0.234	0.026
Age/100	7.872	0.690
Age/100 squared	-8.361	0.887
Experience/100	4.618	1.204
Experience/100 squared	-8.344	2.785
Vocational degree	0.077	0.022
High School degree	0.307	0.036
College and Under Graduate	0.379	0.034
Graduate	0.693	0.046
Part-time	-0.687	0.027
Variance of residuals	0.432	0.006
Public wage equation		
Intercept	5.817	0.322
Women	-0.365	0.031
Paris	0.138	0.030
Age/100	3.092	1.189
Age/100 squared	-2.725	1.482
Experience/100	12.772	2.088
Experience/100 squared	-30.091	4.918
Vocational degree	0.039	0.038
High School degree	0.202	0.062
College and Under Graduate	0.107	0.042
Graduate	0.353	0.053
Part-time	-0.508	0.031
Variance of residuals	0.592	0.021

Source: French European Household

Table 12: Switching regression model with sector selection - 1994 - continued

Variables	*Estimates*	*Standard Error*
Sector choice equation		
Intercept	1.767	0.373
Women	-0.466	0.049
Men	-	-
Age/100	-3.316	1.871
Age/100 squared	2.329	2.345
Married	0.098	0.040
Vocational degree	0.169	0.057
High School degree	0.069	0.095
College and Under Graduate	-0.377	0.066
Graduate	-0.474	0.086
Women times Father civil servant	-0.248	0.073
Men times Father civil servant	-0.229	0.072
Women times Mother civil servant	-0.044	0.079
Men times Mother civil servant	-0.203	0.088
Local unemployment rate	-0.005	0.007

Source: French European Household

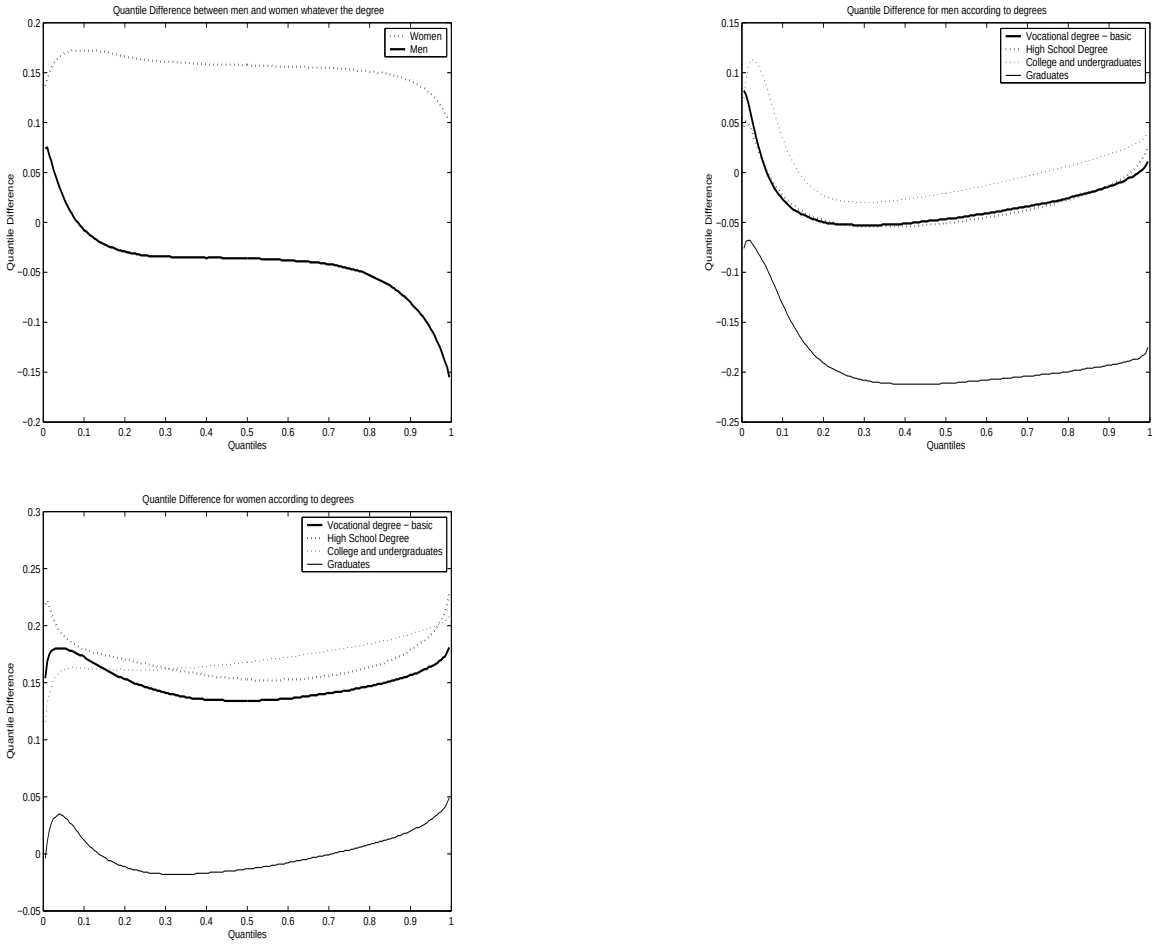
E. Quantile differences between the public log of wage distributions and the counterfactual distribution

Table 13. Quantile differences between the public log of wage and the counterfactual distributions.

	1994	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**
Quantiles				*Estimates*				
				Standard Errors				
0.050	0.156	0.168	0.155	0.155	0.153	0.154	0.155	0.157
	0.018	0.021	0.019	0.019	0.019	0.019	0.019	0.019
0.100	0.144	0.157	0.152	0.150	0.153	0.150	0.146	0.148
	0.017	0.019	0.018	0.017	0.018	0.018	0.018	0.018
0.150	0.130	0.142	0.143	0.138	0.142	0.140	0.136	0.136
	0.019	0.018	0.017	0.017	0.018	0.018	0.018	0.018
0.200	0.118	0.128	0.130	0.126	0.132	0.129	0.127	0.125
	0.019	0.018	0.017	0.017	0.018	0.018	0.017	0.018
0.250	0.109	0.117	0.120	0.116	0.122	0.120	0.119	0.116
	0.019	0.019	0.018	0.017	0.018	0.018	0.017	0.018
0.300	0.102	0.108	0.111	0.108	0.113	0.111	0.111	0.109
	0.019	0.018	0.018	0.017	0.017	0.017	0.017	0.018
0.350	0.095	0.100	0.102	0.100	0.106	0.103	0.104	0.102
	0.018	0.018	0.018	0.017	0.017	0.018	0.018	0.018
0.400	0.090	0.093	0.095	0.094	0.099	0.096	0.097	0.096
	0.019	0.017	0.018	0.018	0.017	0.017	0.018	0.018
0.450	0.084	0.087	0.088	0.088	0.092	0.090	0.091	0.090
	0.018	0.018	0.018	0.018	0.017	0.018	0.018	0.018
0.500	0.079	0.082	0.083	0.083	0.085	0.084	0.085	0.085
	0.018	0.018	0.018	0.018	0.017	0.018	0.018	0.018
0.550	0.074	0.077	0.077	0.077	0.078	0.078	0.079	0.079
	0.018	0.018	0.018	0.018	0.018	0.019	0.018	0.019
0.600	0.070	0.072	0.072	0.072	0.072	0.073	0.073	0.074
	0.018	0.018	0.018	0.019	0.018	0.019	0.018	0.019
0.650	0.064	0.067	0.067	0.067	0.066	0.067	0.066	0.068
	0.018	0.019	0.019	0.019	0.018	0.019	0.019	0.019
0.700	0.059	0.061	0.061	0.061	0.059	0.060	0.059	0.062
	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019
0.750	0.053	0.054	0.053	0.054	0.052	0.052	0.051	0.054
	0.018	0.019	0.019	0.019	0.019	0.020	0.019	0.019
0.800	0.044	0.045	0.044	0.045	0.042	0.044	0.041	0.044
	0.018	0.020	0.020	0.019	0.019	0.020	0.019	0.019
0.850	0.032	0.033	0.033	0.033	0.030	0.032	0.029	0.031
	0.019	0.020	0.019	0.019	0.019	0.019	0.019	0.019
0.900	0.014	0.015	0.015	0.014	0.013	0.014	0.012	0.013
	0.020	0.021	0.020	0.020	0.020	0.019	0.020	0.019
0.950	-0.019	-0.020	-0.017	-0.018	-0.018	-0.015	-0.017	-0.018
	0.021	0.019	0.021	0.022	0.021	0.021	0.021	0.020

Source: French European Household

Figure 8. Quantile differences between the public log of wage and the counterfactual distributions given sex and degrees.



References

ARCIDIACONO, P. (2005): "Affirmative Action in Higher Education: how do admission and financial aid rules affect future earnings?," *Econometrica*, 73, 1477–1524.

AUDIER, F. (2000): "La transmission du statut dans la Fonction publique," *Economie et statistique*, 337-338, 121–133.

BARGAIN, O., AND B. MELLY (2008): "Public Sector Pay Gap in France: New Evidence using Panel Data," IZA Working Paper.

BELL, D., B. ELLIOTT, AND A. SCOTT (2005): "The earnings consequences of moving between the private and public sectors in Great Britain," .

BELLANTE, D., AND A. LINK (1981): "Are Public Sector Workers More Risk Averse than Private Sector Workers?," *Industrial and Labor Relations Review*, 34(3), 408–412.

BESSIÈRE, S., AND J. POUGET (2007): "Les carrières dans la fonction publique d'Etat. Premiers éléments de caractérisation," *Les salaires en France*, pp. 51–66.

CAMERON, S., AND J. HECKMAN (1998): "Life Cycle Schoolong and Dynamic Selection Biais: models and evidence for five cohorts of american males," *Journal of Political Economy*, 106, 262–333.

CAPPELLARI, L. (2002): "Earnings dynamics and uncertainty in Italy: how do they differ between the private and public sectors?," *Labour Economics*, 9, 477–496.

CARD, D. (1996): "The effects of Unions in the structure of wages: a longitudinal analysis," *Econometrica*, (64), 957–979.

DEMPSTER, A., M. LAIRD, AND D. RUBIN (1977): "Maximum Likelihood from Incomplete Data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, 39, 1–38.

DISNEY, R., AND A. GOSLING (1998): "Does it pay to work in the Public Sector?," *Fiscal Studies*, 19(4), 347–374.

——— (2003): "A new method for estimating public sector pay premia: evidence from Britain in 1990s," CEPR Discussion Paper n°3787.

DUSTMAN, C., AND A. VAN SOEST (1998): "Public and Private Sector Wages of Male Workers in Germany," *European Economic Review*, 42, 1417–1441.

EDON, C., AND K. KAMIONKA (2008): "Modélisation dynamique de la participation au marché du travail des femmes en couple," forthcoming in Annales d'Economie et de Statistique.

FOUGÈRE, D., AND J. POUGET (2003): "Les déterminants économiques de l'entrée dans la fonction publique," *Economie et Statistique*, 369, 15–48.

GODDEERIS, J. (1988): "Compensating Differentials and Self-Selection: an application to Lawyers," *The Journal of Political Economy*, 96(2), 411–428.

GREGORY, R., AND J. BORLAND (1999): *Recent Developments in Public Sector Labor Markets*,vol. 3C, chap. 53, pp. 3573–3630. North-Holland, Amsterdam.

GYOURKO, J., AND J. TRACY (1988): "An Analysis of Public- and Private-Sector Wages Allowing for Endogenous Choices of both Government and Union Status," *Journal of Labor Economics*, 6(2).

HARTOG, J., AND H. OOSTERBEEK (1993): "Public and private sector wages in the Netherlands," *European Economic Review*, 37, 97–114.

HECKMAN, J. (1981): *Heterogeneity and State Dependence*,chap. Studies in Labor Market. University of Chicago Press.

HECKMAN, J. J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

HEITMUELLER, A. (2006): "Public-Private Sector Wage Differentials in Scotland: an Endogenous Switching Model," *Journal of Applied Econometrics*, 9, 295–323.

HEYWOOD, J., AND M. MOHANTY (1995): "Estimation of the US Federal Job queue in the presence of an endogeneous queue," *Economica*, 62, 479–493.

HYSLOP, D. (1999): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women," *Econometrica*, (67), 1255–1294.

POSTEL-VINAY, F., AND H. TURON (2007): "The public pay gap in Britain: small differences that (don't?) matter," *Economic Journal*, 117, 1460–1503.

SMITH, S. (1977): *Equal Pay in the Public Sector: Fact or Fantasy*. Princeton, N.J.

VAN OPHEM, H. (1993): "A modified switching regression model for earnings differentials between the public and the private sectors in the Netherlands," *Review of Economics and Statistics*, 75(2), 215–223.

VENTI, S. (1987): *Wages in the Federal and the Private Sectors*chap. in Public Sector Payrolls, pp. 147–182. University of Chicago Press.

WOOLDRIDGE, J. (2005): "Simple Solutions to the initial conditions problem in dynamic, non-linear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, pp. 39–54.

□