DOCUMENT DE TRAVAIL DT/2006-03

Theory and evidence on the glass ceiling effect using matched worker-firm data

Mohamed JELLAL
Christophe J. NORDMAN
François-Charles WOLFF

DIAL • 4, rue d'Enghien • 75010 Paris • Téléphone (33) 01 53 24 14 50 • Fax (33) 01 53 24 14 51
E-mail : dial@dial.prd.fr • Site : www.dial.prd.fr

THEORY AND EVIDENCE ON THE GLASS CEILING EFFECT USING MATCHED WORKER-FIRM DATA[1]

Mohamed Jellal
Université Mohammed V, Rabat
Conseils-Eco, Toulouse
jellalmohamed@yahoo.fr

Christophe J. Nordman
IRD, DIAL, Paris
nordman@dial.prd.fr

François-Charles Wolff
LEN, Université de Nantes
CNAV, INED
wolff@sc-eco.univ-nantes.fr

Document de travail DIAL
Janvier 2006

ABSTRACT

In this paper, we investigate the glass ceiling hypothesis according to which there exists larger gender wage gaps at the upper tail of the wage distribution. We demonstrate that in some circumstances, more qualified women may be offered lower wages than men at the equilibrium. This occurs for instance in a competitive model of wage determination where employers face gender-specific probabilities concerning the stability of their employees in their firms. Then, we focus on the relevance and the magnitude of the glass ceiling effect in France using a representative matched worker-firm data set in 1992 of about 130,000 employees and 14,000 employers. We estimate quantile regressions and use a principal component analysis to summarize information specific to the firms. Our different results show that accounting for firm-related characteristics, in particular firm-specific wage policies, reduces the gender earnings gap at the top of the distribution, but the latter still remains much higher at the top than at the bottom.

Key Words : Gender wage gap; Glass ceiling; Quantile regressions; Matched worker-firm data.

RESUME

Nous analysons l'existence du phénomène de "plafond de verre" selon lequel il existerait des écarts salariaux selon le genre plus importants dans le haut de la distribution des revenus. Nous montrons dans un modèle compétitif de détermination des salaires que, sous certaines hypothèses et à l'équilibre, les femmes les plus qualifiées reçoivent des salaires plus faibles que ceux des hommes de même niveau de qualification. Cela se produit si les employeurs apprécient différemment selon le sexe des employés la stabilité des travailleurs dans leur entreprise. Nous examinons ensuite la pertinence de cette hypothèse et l'étendue de l'effet de plafond de verre à l'aide de données représentatives de l'industrie privée française en 1992 liant quelque 130.000 employés à plus de 14.000 établissements. Nous estimons des régressions de quantiles et utilisons une analyse factorielle pour résumer les informations spécifiques à chaque établissement. Nos différents résultats montrent que prendre en compte les caractéristiques des établissements, en particulier leur politique salariale spécifique, réduit l'estimation de l'écart de revenus entre sexes en haut de la distribution, mais celui-ci n'en demeure pas moins beaucoup plus important en haut qu'en bas de cette distribution des revenus.

Mots clefs: Écart salarial selon le genre; Plafond de verre; Régressions de quantiles; Données appariées employeurs-employés

JEL Code : J16, J31, D80

[1] We thank seminar participants at DIAL, INED and LEST for helpful suggestions.

Table des matières

Liste des tableaux

1. INTRODUCTION

The persistence of wage differentials between men and women with identical productive characteristics is an important stylised fact of labour markets in both industrialised and developing countries. Evidence of a gender wage gap in pay is indeed abundant. Wage differentials across genders that are not compensated by observed socio-economic characteristics were found on numerous occasions in empirical studies (see for instance the review in Blau and Kahn, 2000). Many models have attempted to give a theoretical interpretation to these gender pay gaps. Traditionally, economists have focused on either qualifications or labour market treatment of similarly qualified individuals[2]. Other theories like the insider-outsider or the efficiency wage models have stressed non-competitive mechanisms of wage determination.

More recently, it has been suggested that there exists larger gender wage gaps at the upper tail of the wage distribution, so that it concerns in most cases the more skilled workers. This is the so-called glass ceiling effect above women in the labour market, which can be defined as an invisible barrier that inhibits promotion opportunities for women, but not for men, and prevents them from reaching top positions. Several papers have empirically shed light on the magnitude of the glass ceiling effect in different European countries.

For instance, using data collected in 1998 in Sweden, Albrecht et alii (2003) show that the gender wage gap is increasing throughout the conditional wage distribution and accelerating at the top, and they interpret this result as evidence of a glass ceiling in Sweden. Using data for Spain, De la Rica et alii (2005) stratify their sample by education group and find that the gender wage gap is expanding over the wage distribution only for the group with tertiary education. For less educated groups, the gender wage gap is wider at the bottom than the top. This means that, in Spain, there is a glass ceiling for the more educated, while for the less educated there is not.

Using the European Community Household Panel data set, Arulampalam et alii (2004) find that for most of their ten EU countries, in both the public and private sectors, the average gender wage gap can be broken up into a gap that is typically wider at the top and occasionally also wider at the bottom of the conditional wage distribution. They interpret the gender wage gap at the top of the wage distribution as a glass ceiling evidence, whereby women otherwise identical to men can only advance so far up the pay ladder. At the bottom of the wage distribution in some of their EU countries, they also find that the gender pay gap widens significantly and define this phenomenon as a sticky floor (see also Booth et alii, 2003, Ichino and Filippin, 2005).

To date, there is no clear theoretical argument to rationalize the glass ceiling effect among the various usual explanations which have been suggested for the gender wage gap. According to the beckerian theory, discrimination is due to the discriminatory tastes of employers, co-workers, or customers. Alternatively, in models of statistical discrimination, differences in the treatment of men and women arise from average differences between the two groups in the expected value of productivity or in the reliability with which productivity may be predicted, which leads employers to discriminate on the basis of that average. Discriminatory exclusion of women from 'male' jobs can also result in an excess supply of labour in 'female' occupations, depressing wages there for otherwise equally productive workers. But in these various approaches, there is no reason to expect larger gaps at the upper tail of the wage distribution.

De la Rica et alii (2005) provide an interesting formal analysis in order to rationalize the glass ceiling hypothesis[3]. They suggest that a dead-end argument operate in the upper tail of the distribution.

[2] The former, within the competitive framework, emphasize the existence of compensating wages due, for instance, to differences in human capital accumulation across gender. Because women anticipate shorter and more discontinuous work lives, they have lower incentives to invest in market-oriented formal education and on-the-job training, and their resulting smaller human capital investments will lower their earnings relative to those of men.

[3] At the same time, to rationalize the glass floor observed in Spain for low-educated women, De la Rica et alii (2005) suggest that employers may use statistical discrimination amongst women. But as the job tenure of women expands, women become more reliable and then the female wage is expected to converge to the male wage.

Women are less frequently promoted because their jobs can less easily be promoted[4]. Employers are most often reluctant to invest in women's training, for instance because women have more favourable outside opportunities than men within the household (for domestic work or child care), which again increases the gender wage gap. Therefore, women have to be more productive than men in order to be promoted owing to a higher probability of departure and less training.

Our purpose in this paper is to further investigate the glass ceiling hypothesis, both from a theoretical and empirical viewpoint. Our contribution is twofold. On the one hand, we consider a competitive model of wage determination with uncertainty on the women's productivity and show that more qualified women are offered lower wages at the competitive equilibrium. On the other hand, we test the relevance of the glass ceiling effect using matched employer-employee data. Indeed, the gender wage gap at the upper tail of the wage distribution may be wrongly overstated if firms reward highly educated women differently than men.

We first present a model of gender discrimination at the firm level, which predicts the competitive wage structure within firms. The key assumption is that employers account for conjectures about the future stability of both their male and female workers, but the probability to stay within the firm differs between men and women. While employers admit that male and female workers have an equal productivity, they attach more uncertainty to the women's careers. This is the case if they face greater incertitude towards females' employment duration over time, for instance as a result of their more discontinuous work participation (owing to family events or societal discrimination in domestic tasks). In this setting, we demonstrate that firms are expected to offer lower wages to women as a result of the uncertainty regarding the long-term stability of their female employees. Indeed, firms pass the risk of variability in women's production on female wages and the negative risk premium increases as women are more qualified.

Then, we provide empirical evidence on the gender wage gap in France. One popular way to account for firms and workers' characteristics, including their human capital features, is to base the empirical analysis on matched worker-firm data (see Abowd and Kramarz, 1999). Although differences in productivity across workers could stem from their differences in human capital, it is well acknowledged that some skills or human capital attributed to workers are also specific to the firms in which those workers operate. Thus, part of the returns to human capital for the worker remuneration can be viewed as originating from the firm (human capital externalities)[5]. This means that not controlling for firm specific effects on individual earnings differentials may lead researchers to bias their estimates of the different returns to human capital. But curiously, firm specific effects are most often neglected in the analysis of gender earnings gap, Meng (2004) being a recent and worthwhile exception.

In our paper, we shall avail ourselves of such data by using French worker-firm matched data. To the best of our knowledge, our paper is the first on the glass ceiling effect to control for these specific firm effects. Following previous studies, we first use quantile regressions techniques to assess the existence and the extent to which the glass ceiling phenomenon exists in France. Then, we control for firms' specific wage policies by introducing the firms' features into the different earnings functions. A novelty of our approach is to perform quantile regressions using a preliminary principal component analysis of firms' characteristics, as in Muller and Nordman (2004, 2005). According to the French data, we show that introducing firms-related characteristics into earnings equation significantly reduces the gender earnings gap at the top of the distribution. Nevertheless, the gender wage gap still remains greater at the top than at the bottom.

The remainder of the paper is organized as follows. In Section 2, we describe a competitive model of wage determination with uncertainty on female productivity. The French data that we use for the empirical analysis is described in Section 3. In Section 4, we present the econometric methodology

[4] Conversely, since high-educated women have participation rates which are only slightly lower than male participation rates, women's and men's wages should not be very different in the lower part of the income distribution (De la Rica et al., 2005).

[5] It is also possible that part of what could be interpreted as human capital externalities in the estimates is in fact a consequence of the selection of workers by firms and vice versa. For instance, highly educated workers (i.e. high wage workers) are more likely to match with high wage firms (Abowd et alii, 1999).

and our strategy to account for firm related characteristics. We discuss in section 5 the different results of the quantile regressions. Finally, Section 6 concludes.

2. A COMPETITIVE MODEL OF WAGE DETERMINATION WITH UNCERTAINTY

Despite of the impressive literature on the theoretical foundations of the gender wage gap, it remains difficult to understand how the gap evolves throughout the distribution of wages. As suggested in De la Rica et alii (2005), several explanations may be invoked to rationalize the glass ceiling effect. We believe that basic, albeit realistic, assumptions are sufficient to explain why the gender wage gap is higher at the top of the wage distribution. Specifically, we focus on the consequences of the fact that women are more likely to quit their job than men. So, our aim in this theoretical section is not to propose a structural model to be tested, but instead to better understand the role of factors that can affect the shape of the gender wage profile.

Let us consider a representative firm which produces a good Y^t at date t. This good is sold on a competitive market and it is treated as the numeraire ($p = 1$). To produce that good, the firm hires two types of workers, men and women. We denote respectively by h_1 and h_2 the levels of human capital for a man and for a woman. We have either $h_1 \neq h_2$ or $h_1 = h_2$, meaning that there may exist gender differences in the level of human capital. For the sake of simplicity, there is no on-the-job training in our model and the level of human capital remains constant over time for each employee. Let n_1 and n_2 be respectively the numbers of men and women who are currently working in the representative firm.

We consider that the firm does not really know how long a worker will stay in the firm. This does not seem unrealistic a priori. Assuming that the expected duration of a job in a specific firm is given by $E(T_1)$ and $E(T_2)$ respectively for a man and a woman, then $q_1 = 1/E(T_1)$ and $q_2 = 1/E(T_2)$ are the probabilities respectively for a man and for a woman to quit their current job. We make the following assumption concerning q_1 and q_2.

Assumption 1. *The probability to quit a job is higher for a woman, i.e.* $q_2 > q_1$.

So, we suppose that employers believe that on average, women are characterized by higher quit rates. To justify that crucial assumption, one can rely on gender differences in the labour force participation. For instance, it is well known that women are less likely to have a paid job than men and that they most often interrupt their formal activities. This may occur because of births or other family events, women being more likely to care for their elderly parents or to spend time educating their children. Owing to these more frequent interruptions, we argue that the firm suffers from a higher uncertainty when evaluating the long-term productivity of their female workers with respect to the male workers.

To formalize this uncertainty, we introduce into the definition of the level of employment a random term on the female productivity. As a consequence, the quantity of efficient labour is itself a random variable denoted by $\widetilde{N}^t$, which is given by the following sum:

$$\widetilde{N}^t = h_1 n_1^t + \widetilde{\varepsilon} h_2 n_2^t \tag{1}$$

where $h_1 n_1^t$ and $\widetilde{\varepsilon} h_2 n_2^t$ are respectively the male and female levels of employment. Importantly, we do not assume that the mean level of productivity is different for male and female workers. We just suppose that there is more uncertainty on the female labour force participation. As a consequence, the term $\widetilde{\varepsilon}$ may be described by a random variable whose mean is $E(\widetilde{\varepsilon}) = 1$ and $Var(\widetilde{\varepsilon}) = \sigma_\varepsilon^2$. To find a more explicit result and in order to get closed-form solutions for our problem, we make the following assumption concerning the female productivity parameter.

Assumption 2. *The parameter* $\tilde{\varepsilon}$ *follows a Normal distribution* $N(1,\sigma_\varepsilon^2)$.

Without loss of generality, we neglect the role of the capital factor in the production process[6]. The production function for the representative firm may be expressed as:

$$\tilde{Y}^t = F(h_1 n_1^t + \tilde{\varepsilon} h_2 n_2^t) \tag{2}$$

where the function $F(.)$ is continuous and characterized by decreasing returns, i.e. $F'(.) > 0$ and $F''(.) < 0$. In the above production function, we note that female and male labour force participations are two perfectly substitutable inputs.

Given the possibility of quitting the firm, we account for turnover costs in the model. Let l_1^t and l_2^t be the numbers of men and women who are hired each year by the firm, and we denote by $c_1^t(l_1^t)$ and $c_2^t(l_2^t)$ the corresponding hiring cost functions. As might be expected, we consider that hiring more qualified workers is more costly for the firm and we express these turnover costs as a linear function of the number of hiring. Hence, we have $c_1^t(l_1^t) = c(h_1)l_1^t$ and $c_2^t(l_2^t) = c(h_2)l_2^t$, with $c'(.) > 0$ and $c''(.) < 0$. Finally, as there are entry and exit of workers in our setting, the following equations fully characterize the dynamics of employment within the representative firm respectively for men and women:

$$\frac{dn_1^t}{dt} = l_1^t - q_1 n_1^t \tag{3}$$

$$\frac{dn_2^t}{dt} = l_2^t - q_2 n_2^t \tag{4}$$

meaning that at date t, the total level of employment either for men or women is given by the number of employees at date $t-1$ plus the difference between the number of hiring workers and the number of voluntary exits between $t-1$ and t.

Finally, we define the profit function for the firm at date t. The two variables of interest for our problem are the levels of wage for men and women, respectively denoted by w_1 and w_2. Workers are remunerated at their marginal productivity, so that the firm's expected profit $E\tilde{\Pi}^t$ (given the randomness of the female productivity) is simply:

$$E\tilde{\Pi}^t = E[F(h_1 n_1^t + \tilde{\varepsilon} h_2 n_2^t)] - w_1 n_1^t - w_2 n_2^t - c(h_1)l_1^t - c(h_2)l_2^t \tag{5}$$

So, the problem for the firm is to maximize its expected profit discounted at the interest rate r subject to the constraints which characterize the dynamics of employment over time. The corresponding program may be expressed as:

$$\begin{aligned} &\max_{l_1^t, l_2^t} \int_{\Re+} \left\{E[F(h_1 n_1^t + \tilde{\varepsilon} h_2 n_2^t)] - w_1 n_1^t - w_2 n_2^t - c(h_1)l_1^t - c(h_2)l_2^t\right\} e^{-rt} dt \\ &s.t. \quad \begin{cases} dn_1^t / dt = l_1^t - q_1 n_1^t \\ dn_2^t / dt = l_2^t - q_2 n_2^t \end{cases} \end{aligned} \tag{6}$$

[6] In so doing, we rely on the argument developed in Nickell (1986). When firms plan campaigns of hiring, decisions related to capital investments have most often already been made and thus capital may be seen as pre-determined.

The optimal wage policy may easily be found by solving this problem of optimal control. Let us define the current value of the Hamiltonian H such that:

$$\begin{aligned} H = {} & E[F(h_1 n_1 + \tilde{\varepsilon} h_2 n_2)] - w_1 n_1 - w_2 n_2 - c(h_1) l_1 - c(h_2) l_2 \\ & + \lambda_1 (l_1 - q_1 n_1) + \lambda_2 (l_2 - q_2 n_2) \end{aligned} \tag{7}$$

where λ_1 and λ_2 are the co-state variables associated respectively to the constraints on the levels of employment n_1 and n_2. The conditions of optimality for this problem are $\partial H / \partial l_1 = 0$, $\partial H / \partial l_2 = 0$, $d\lambda_1 / dt = r\lambda_1 - \partial H / \partial n_1$ and $d\lambda_2 / dt = r\lambda_2 - \partial H / \partial n_2$. Hence, we get:

$$-c(h_1) + \lambda_1 = 0 \tag{8}$$

$$-c(h_2) + \lambda_2 = 0 \tag{9}$$

$$\frac{d\lambda_1}{dt} = r\lambda_1 - E[h_1 F'(\tilde{N})] + w_1 + \lambda_1 q_1 \tag{10}$$

$$\frac{d\lambda_2}{dt} = r\lambda_2 - E[\tilde{\varepsilon} h_1 F'(\tilde{N})] + w_2 + \lambda_2 q_2 \tag{11}$$

Let us first briefly interpret these different first-order conditions. According to (8) and (9), firm's decisions with respect to recruitment are such that the marginal cost of hiring an additional worker $c(.)$ is equal to the marginal benefit due to that hiring. Since the marginal cost $c(.)$ is fixed by assumption (remind that education levels do not vary over time), this implies that λ_1 and λ_2 are constant, so that $d\lambda_1 / dt = 0$ and $d\lambda_2 / dt = 0$. From (10) and (11), it follows that $\lambda_1 = (E[h_1 F'(\tilde{N})] - w_1)/(r + q_1)$ and $\lambda_2 = (E[\tilde{\varepsilon} h_2 F'(\tilde{N})] - w_2)/(r + q_2)$. Since $c(h_1) = \lambda_1$ and $c(h_2) = \lambda_2$ from (8) and (9), we obtain the following optimal wages for men and for women:

$$w_1 = h_1 E[F'(\tilde{N})] - (r + q_1) c(h_1) \tag{12}$$

$$w_2 = h_2 E[\tilde{\varepsilon} F'(\tilde{N})] - (r + q_2) c(h_2) \tag{13}$$

It is straightforward to give an interpretation to the previous equalities. At the equilibrium, we find that the optimal male wage is given by the difference between the expected marginal productivity $h_1 E[F'(\tilde{N})]$ and the turnover costs $(r + q_1) c(h_1)$. A similar reasoning applies for (13), but we note that there is an additional random term $\tilde{\varepsilon}$ when defining the woman's marginal productivity $h_2 E[\tilde{\varepsilon} F'(\tilde{N})]$. The normality assumption for the random perturbation $\tilde{\varepsilon}$ allows us to further specify the optimal wage policy for a competitive firm.

Proposition 1. *The optimal wage policy for the firm is such that:*

$$w_1 = h_1 E[F'(\tilde{N})] - (r + q_1) c(h_1) \tag{14}$$

$$w_2 = h_2 E[F'(\tilde{N})] + \upsilon E[F''(\tilde{N})] h_2^2 - (r + q_2) c(h_2) \tag{15}$$

where $\upsilon = Var(\tilde{\varepsilon} n_2) / E(\tilde{\varepsilon} n_2)$ is the coefficient of variation associated to the female productivity.

Proof. Given the normality assumption for $\tilde{\varepsilon}$, we can use the lemma of Stein (Rubinstein, 1976). Let us consider two variables X and Y which are bivariate normally distributed. If the function $f(Y)$ is continuously differentiable, then $\text{cov}(X, f(Y)) = E(f'(Y)) \text{cov}(X, Y)$. By definition, we have $E[\tilde{\varepsilon} F'(\tilde{N})] = E(\tilde{\varepsilon}) E[F'(\tilde{N})] + \text{cov}(\tilde{\varepsilon}, F'(\tilde{N}))$. Now, applying the Stein's lemma to our problem, we get $\text{cov}(\tilde{\varepsilon}, F'(\tilde{N})) = E[F''(\tilde{N})] \text{cov}(\tilde{\varepsilon}, h_1 n_1 + h_2 n_2 \tilde{\varepsilon})$. Provided that $E(\tilde{\varepsilon}) = 1$ and

$Var(\tilde{\varepsilon}) = \sigma_\varepsilon^2$, we deduce $\text{cov}(\tilde{\varepsilon}, h_1 n_1 + h_2 n_2 \tilde{\varepsilon}) = h_2 n_2 \sigma_\varepsilon^2$. Using (13), we finally obtain the following female wage $w_2 = h_2 E[F'(\tilde{N})] + h_2^2 \sigma_\varepsilon^2 n_2 E[F''(\tilde{N})] - (r + q_2) c(h_2)$.

Let us further expand the term $h_2^2 \sigma_\varepsilon^2 n_2 E[F''(\tilde{N})]$. We know that $Var(\tilde{\varepsilon} n_2) = n_2^2 \sigma_\varepsilon^2$ and $E(\tilde{\varepsilon} n_2) = n_2$ since $E(\tilde{\varepsilon}) = 1$. Hence, given the definition of $\upsilon = Var(\tilde{\varepsilon} n_2) / E(\tilde{\varepsilon} n_2)$, it follows that $h_2^2 \sigma_\varepsilon^2 n_2 E[F''(\tilde{N})] = \upsilon h_2^2 E[F''(\tilde{N})]$. **QED**

With respect to our previous interpretation of the optimal wages, there is now an additional term in the definition of the optimal female wage. According to (15), the optimal wage for women is the sum of the marginal expected productivity $h_2 E[F'(\tilde{N})]$ and a negative term $\upsilon E[F''(\tilde{N})] h_2^2$, minus the opportunity cost in terms of turnover $(r + q_2) c(h_2)$. A central feature for our analysis deals with the interpretation of $\upsilon h_2^2 E[F''(\tilde{N})]$. As it stands, this term is a risk premium associated to the uncertainty on female productivity.

In our setting, the optimal wage policy depends on the pattern of qualification (h_1, h_2), on the probability to quit the firm (q_1, q_2), on the interest rate r, on the shape of technology given by F' and F'', and on the coefficient of variation related to the female productivity υ.

Corollary 1. *The gender wage gap is given by:*

$$w_1 - w_2 = (h_1 - h_2) E[F'(\tilde{N})] - \upsilon h_2^2 E[F''(\tilde{N})] - (r + q_1)(c(h_1) - c(h_2)) \qquad (16)$$

Let us briefly comment the different factors which are expected to influence the gender wage gap. Imagine first a situation such that $\upsilon = 0$, meaning that there is no uncertainty. In that case, the difference between w_1 and w_2 depends both on the difference in skill levels respectively for men and for women and on the pattern of turnover costs. In a situation characterized by similar skill levels for male and female workers (meaning that $h_1 = h_2$), the male wage may still be higher than the female wage when the inequality $c(h_1) < c(h_2)$ holds. In that case, the gender wage gap would only be explained by the pattern of turnover costs. But imagine now a situation where uncertainty on the female productivity prevails, i.e. $\upsilon > 0$. Then, our model formally explains the glass ceiling effect.

Corollary 2. *The glass ceiling effect is picked up by the term* $\upsilon h_2^2 E[F''(\tilde{N})]$.

According to the definition of the female wage, the negative risk premium $\upsilon E[F''(\tilde{N})] h_2^2$ depends on the shape of the technology F, on the coefficient of variation for the female productivity, and also on the squared level of the woman's skill level. As the gender wage gap is a convex positive function of h_2 (remind that the risk premium $\upsilon h_2^2 E[F''(\tilde{N})]$ is negative), one expects a significantly higher difference between male and female wages at the top of the income distribution. This is exactly the core of the glass ceiling hypothesis.

So, our theoretical framework sheds light on the role of employer's expectations on the propensity for workers to quit their current jobs. Even in a setting where human capital does not differ between men and women, we demonstrate that uncertainty on the female productivity is a sufficient condition to rationalize the existence of a glass ceiling. At that stage, it seems important to stress the difficulty to empirically assess the relevance of our model. Indeed, the two main implications are that i) the firm sets a lower wage for women given uncertainty and ii) the negative risk premium is higher for high-skilled women. On the one hand, it is straightforward to evidence whether gender differences in wages vary along the income distribution, i.e. to give an answer to the question "does there exist a glass ceiling?". On the other hand, it is much more difficult to provide an accurate measure of uncertainty on productivity.

That employers' subjective expectations on the behaviours of their employees matter to explain the glass ceiling effect is a very important result, which has empirical implications. Indeed, the mobility of workers and their intention to quit their jobs are certainly determined by the firm's environment, even partly. For instance, workers hired by a firm where high wages and good working conditions prevail have less incentive to look for other jobs. Therefore, not accounting for the firm-specific characterization is likely to affect the measurement of the gender wage gap throughout the distribution of wages. So, in what follows, we focus on a non-structural, econometric analysis where we investigate the magnitude of the glass ceiling effect in France and introduce into the earnings regressions some information on firms.

3. THE FRENCH MATCHED WORKER-FIRM DATA

The data we use in this paper are drawn from a unique French survey which matches information of both employers and employees, the 1992 INSEE survey on labour cost and wage structure (*Enquêtes sur le Coût de la Main-d'Oeuvre et la Structure des Salaires en 1992*, ECMOSS thereafter). It is well known that such data sets allow the structure of wages to be modelled while controlling for firm-specific effects (see Abowd et alii, 1999). Specifically, the French survey contains information on 150,000 different workers across 16,000 different workplaces[7]. The sampling population covered by these data is very broad, as all establishments are covered independently of their size and in all industries apart from agriculture, fisheries, non-traded services, and central and local government.

The ECMOSS survey contains a great deal of information. Concerning the employees, data are available on workers' gross annual wage, which is broken down into fixed salary, bonuses, overtime, and data on their gender, age, nationality, tenure, occupation, education level and number of paid hours. There is also some detailed information on the employer, including main economic activity, size, geographical location, management style, work organisation and salary policy. In order to perform our econometric analysis, several additional variables have been constructed and we describe them below.

Concerning the workers, we determine the total number of years of education calculated from the final level reached, total potential experience in the labour market which is given by age minus number of years of education minus six, hourly earnings (gross salary plus payments in kind, all divided by the number of paid hours over the year), and the average number of paid hours of training per worker in the establishment (the number of hours of paid training by worker by occupational category – executive or non-executive – divided by the total number of workers by occupational category)[8]. After deleting observations with missing values or outliers, the worker sample amounts to 137,211 individuals divided into 14,693 establishments. Table 1 provides a description of the characteristics of the employees.

Table 1: Description of the workers' characteristics

Main sample characteristics	Mean	[Min ; Max]	Standard dev.
Number of observed employees per establishment	18.99	[2 ; 152]	15.53
Sex (1 for men, 0 otherwise)	0.60		
Age	37.68	[16.25 ; 65]	10.30
Nationality (1 if French, 0 otherwise)	0.93		
Hourly earnings (gross wage plus payments in kind, all divided by the number of paid hours over the year)	69.48	[29.00 ;395.83]	39.49
Education (number of completed years of schooling)	12.77	[8 ; 18]	1.65
Potential previous experience (Number of years of labour market experience: age - tenure - education – 6)	9.27	[0 ; 48.91]	8.72
Tenure in the current establishment (number of years of tenure)	9.71	[0 ; 46.5]	8.84
Executives (1 if executive, 0 otherwise)	0.11		
Number of hours paid work per year	1671.78	[33 ; 2310]	585.46
Type of contract (1 if fixed duration contract, 0 otherwise)	0.08		
Workplace (1 if Paris, 0 otherwise)	0.19		

Source: Survey INSEE ECMOSS 1992.
The size of the sample is 137211 employees, working in 14693 establishments.

[7] This labour cost survey is concurrently carried out in all European Union countries every four years and aims at providing comparable labour market statistics across EU countries. In the 1992 wave of this survey, INSEE matched the data with those on the wage structure. For previous studies which have estimated earnings functions on the same data, see among others Abowd et alii (2001), Destré and Nordman (2002), and Destré (2003).

[8] The education variable is constructed as follows. For a sub-sample of more than 8000 workers for whom the number of years of education is available (besides the highest paper certificate), we calculate the median number of years of education for each qualification considered. This indirect method for calculating the length of education has the advantage of partially removing the endogeneity of the education variable (see the discussion in Destré, 2003).

To test the glass ceiling hypothesis, one novelty of our approach is to control for firm level variables in the analysis of wage determination. We describe below the information that is utilised for a preliminary multivariate analysis of firm-related characteristics. The definitions and descriptive statistics of these variables appear in Table A1.

First, we make use of twelve sectoral dummies (S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12), six dummies for the size of the establishments (T1, T2, T3, T4, T5, T6)[9], and four dummies for the average number of paid hours of training per worker in the establishment (in increasing order, FF1, FF2, FF3, FF4). Firm level human capital variables are respectively average, standard deviation, maximum and distance to the maximum of workers' education, experience off the current job and tenure in each establishment (respectively for education, experience and tenure: EDUCM, EXPHEM, ANCM, SDEDUC, SDEXPHE, SDANC, SUPEDUC, SUPEXPHE, SUPANC, SE, SEX, SA).

The following variables relate to qualitative aspects of firms' activity: dummies describing the intensity of business the past five years ("strongly growing" or "growing": va1, "stable": va2; "strongly decreasing" or "decreasing": va3), whether activity is usually affected by seasonal movements (d21), whether it is rather regular (d31), or irregular (d32), whether firms have been affected by unusual shocks in 1992 (d4a1) and, if it is the case, whether it was a downturn (d4b1) or an upturn (d4b2). We also make use of qualitative features of intra-firm wage determination such as dummies for the presence of union representatives (ps1), for the existence of wage negotiations in 1992 (d151), and for the use of a formal wage scale system for blue collar workers' wage base (d19a1). If such a formal system is used, the questionnaire provides further information as to whether it is based on the branch's collective agreement (d19b1), on the firm's collective agreement (d19b2) or on another evaluation scheme (evaluation of posts, d19b3).

Further detailed information describes the importance accorded by employers to different criteria in individual wage increases (for both blue collar and white collar workers). In the questionnaire, the answers were ranked according to three different levels of importance: "none", "weak", "medium", and "very strong". In our analysis, we make use of dummies taking into account the answers "very strong": workers' tenure in the job (d3513), increase in workers' performance (d3523), workers' training effort (d3533), accumulation of experience (d3543), acquisition of versatility (d3553), increase in workers' responsibilities (d3563), intra-firm mobility (d3573), and difficulty of workers' eventual replacement (d3583).

Qualitative variables are then utilised to describe the extent to which employers favour individual or general wage increases in their wage policy: whether the base wage progressions "exclusively" (d331), "principally" (d332), "little" (d333) or "never" (d334) depend on individual increases and on general increases (respectively, d341, d342, d343, d344). Dummies regarding individual bonuses according to performances are also introduced in the following way: d3911 indicates whether firms give relative bonuses (the best workers are awarded), d3921 describes whether bonuses are of "absolute" type (the production standards are exceeded). If these two schemes exist in the same firm, d39b1 reports which one is the most important (equals to one if it is relative bonuses). Finally, d411 signals firms having implemented an explicit wage policy characterised by precise objectives.

Firms' organisational features are likely to influence employers' wage settings as well as skill diffusion and acquisition processes (Lindbeck and Snower, 2000; Caroli et alii, 2001; Greenan, 2003). We construct dummies describing the firm's hierarchical structure such as the number of intermediate levels of management between the firm's manager and the blue collar workers assigned to productive lines (zero levels: d250; from 1 to 4: d251; from 5 to 10: d252; 11 levels and above: d253), dummies indicating the existence of job rotation schemes and how they are implemented (whether they are put into practice within production teams: d26a1, and whether they are intended to some versatile workers independently from team working: d26b1), a dummy when direct collaborations between employees of different departments are encouraged (d281), a dummy reporting whether achieved work is "permanently" controlled rather than "intermittently" or "occasionally" (d301), a dummy signalling whether individual performances are "systematically" controlled rather than "occasionally" or "never" (d311), and a dummy for the existence of a formal system to measure individual performances (d321).

[9] The dummies are defined as follows: less than 20 employees (T1), 20 to 49 (T2), 50 to 99 (T3), 100 to 199 (T4), 200 to 499 (T5), and more than 500 employees (T6).

4. ECONOMETRIC STRATEGY

Quantile regression and least absolute deviation estimators have recently become very popular estimation methods (Koenker and Bassett, 1978, Buchinsky, 1998). This technique can be interpreted as using the error distribution in the earnings equation for the definition of different earnings categories, i.e. quantiles, instead of the observed earnings differentials. The popularity of these methods relies on three sets of properties.

First, they provide robust estimates, particularly for misspecification errors related to non-normality and heteroskedasticity, but also for the presence of outliers due to data contamination. Second, they allow the researcher to focus on specific parts of the distribution of interest, which is the conditional distribution of the dependent variable, and to estimate the marginal effect of a covariate on log earnings at various points in the distribution. So, quantile regressions allow to estimate the effect of gender, education or experience on log earnings at the bottom of the log earnings distribution (at the 5th percentile), at the median, and at the top of the distribution (at the 95th percentile). Third, quantile regressions are appropriate when earnings functions contribute only to a small part of the variance of earnings, so that the distribution of earnings and the distribution of errors are close[10].

The ECMOSS survey allows the structure of wages to be modelled while controlling firm-specific effects. With our matched data, we can deal with the firm heterogeneity by introducing firm characteristics into the earnings equation. Nevertheless, a difficulty is that we cannot model unobserved individual heterogeneity in the way of Abowd et alii (1999) as this is a cross-sectional data set. In order to temper the effect of firm heterogeneity, the natural attempt is to estimate firm fixed effects models including firm-specific dummies. Nevertheless, this technique seems to be futile in our case. Indeed, as we estimate quantile regressions, the large number of establishments rules out the possibility of doing this[11].

An alternative approach is described in Muller and Nordman (2004, 2005). It consists of summarising the main information on the firms' characteristics using a multivariate analysis and introducing the computed principal components (factors) stemming from this analysis into the earnings functions. Using factors may be seen as a further step with respect to those studies which have added mean firm variables into earnings functions, individual characteristics being controlled for. By contrast with firms' fixed effects that are introduced in wage regressions, the principal factors suggest qualitative characteristics of the firms. Specifically, we use a principal component analysis (PCA) to summarise the information about the surveyed establishments[12].

This method is based on the calculation of the inertia axes for a cloud of points that represents the data in table format. There are different possible uses of factor analysis in this context. First, factor analyses can be used to elicit hidden characteristics correlated with observable characteristics. Second, PCA results could be utilised as a guide to replace these hidden firm characteristics with observable characteristics correlated with the main factors (as in Muller and Nordman, 2004). Third, and foremost in our case, the PCA is used as a substitute for firm fixed effect regressions. Indeed, the PCA allows us to investigate the determinants of the firm effects in our data. As long as the computed factors account for most of the firm heterogeneity bias, this approach allows us to obtain consistent estimates of the returns to worker characteristics and of the gender wage gap.

For our purpose, the first ten inertia axes (the estimated factors which are linear components of all the firm's characteristics described in the previous section) concentrate a large proportion of the total variance of the original variables (about 40%) and reflect, therefore, a fair amount of the relevant

[10] In our empirical analysis, we rely on bootstrap confidence intervals for quantile regressions in order to avoid the consequences of the slow convergence of classical confidence intervals of estimates (Hahn, 1995). However, given the large size of our sample, the results are only marginally modified.

[11] However, very recently, Koenker (2005) has proposed a new advanced method which allows estimating fixed effects quantile regressions with a large number of fixed effects, but the estimation is far from being straightforward.

[12] In principal component analysis, a set of variables is transformed into orthogonal components, which are linear combinations of the variables and have maximum variance subject to being uncorrelated with one another. Typically, the first few components account for a large proportion of the total variance of the original variables, and hence can be used to summarize the original data. The computed factors were rotated using an oblique rotation. As in Muller and Nordman (2005), we have tried many other techniques of factor analysis, which all lead to similar conclusions.

information about the firm's characteristics. The correlation coefficients of the firms' characteristics with the first ten factors are used for the interpretation of the computed factors. The other factors represent a negligible amount of the statistical information and are dropped from the analysis.

Further details on this rotated PCA can be found in Muller and Nordman (2005) and obtained from the authors upon request. Let us note that the first five factors are closely associated with the size of the firms, their workers' human capital characteristics (mean, variance, minimum and maximum of workers' education and tenure) as well as the firms' training capacity. Therefore, they reflect the characteristics of large-sized firms with high human capital density as compared to small-sized firms frequently associated with low-skilled workers. Factors 5 to 7 are closely correlated with the firms' sectoral belonging and organisational features (firm's hierarchical structure, supervision) while the remaining factors are more strongly related to the various criteria used by employers for defining their implemented wage policy. We are now ready to comment on our econometric analysis.

5. QUANTILE REGRESSIONS ESTIMATES

We turn to the discussion of the different quantile regressions which control for both workers' and firms' observed characteristics, in order to assess the existence and the extent to which the glass ceiling phenomenon exists in France.

We first study the effects of differences in characteristics on the gender earnings gap at different points in the distribution. We carry out a set of quantile regressions on the pooled dataset, i.e. with both men and women, which imposes the restriction that the returns to included labour market characteristics are the same for the two genders. The gender dummies in these regressions are interpreted as the effects of gender on log earnings at the various percentiles once one controls for any differences in these labour market characteristics between genders. Estimates for various specifications are reported in Table 2, where we only focus on the magnitude of the dummy coefficient for female.

Table 2: Gender dummy coefficients using alternative quantile earnings models

Specification	Quantile regressions (percentage of the earnings distribution)							OLS	Firm fixed effects
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean	mean
(1) Gender log earnings gap (*n*= 137,211)	-0.0762***	-0.0968***	-0.1428***	-0.1734***	-0.2223***	-0.3295***	-0.4096***	-0.1912***	-0.2077***
	(0.0018)	(0.0020)	(0.0021)	(0.0026)	(0.0035)	(0.0053)	(0.0081)	(0.0023)	(0.0022)
(2) Basic control variables[a] (*n*= 137,211)	-0.0859***	-0.1006***	-0.1264***	-0.1592***	-0.1948***	-0.2270***	-0.2446***	-0.1779***	-0.1834***
	(0.0022)	(0.0019)	(0.0019)	(0.0021)	(0.0027)	(0.0040)	(0.0061)	(0.0019)	(0.0019)
(3) Extended control variables[b] (*n*= 137,211)	-0.0850***	-0.0978***	-0.1219***	-0.1514***	-0.2051***	-0.2558***	-0.2839***	-0.1803***	-0.1843***
	(0.0021)	(0.0019)	(0.0019)	(0.0018)	(0.0024)	(0.0037)	(0.0052)	(0.0019)	(0.0019)
(4) Extended control variables plus 12 sectoral dummies (*n*= 137,211)	-0.0831***	-0.0956***	-0.1238***	-0.1527***	-0.2089***	-0.2708***	-0.3041***	-0.1881***	_
	(0.0023)	(0.0020)	(0.0019)	(0.0020)	(0.0025)	(0.0043)	(0.0058)	(0.0019)	
(5) Extended control variables plus 10 firms' factor effects[c] (*n*= 137,169)	-0.0921***	-0.1053***	-0.1288***	-0.1570***	-0.2018***	-0.2499***	-0.2757***	-0.1825***	_
	(0.0024)	(0.0020)	(0.0018)	(0.0019)	(0.0026)	(0.0038)	(0.0052)	(0.0018)	
(6) Extended control variables plus 12 sectoral and 29 occupational dummies (*n*= 137,211)	-0.0746***	-0.0814***	-0.0969***	-0.1148***	-0.1381***	-0.1690***	-0.1955***	-0.1267***	_
	(0.0030)	(0.0024)	(0.0020)	(0.0017)	(0.0021)	(0.0030)	(0.0041)	(0.0017)	
(7) Extended control variables, firms' factor effects and 29 occupational dummies (*n*=137,169)	-0.0848***	-0.0904***	-0.1017***	-0.1219***	-0.1447***	-0.1764***	-0.2008***	-0.1326***	-0.1098***
	(0.0030)	(0.0023)	(0.0019)	(0.0017)	(0.0021)	(0.0028)	(0.0039)	(0.0017)	(0.0016)

Survey INSEE ECMOSS 1992.

Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

a: basic control variables include education, experience off the current job, tenure in the current firm, their squared values, a dummy for non-French workers, dummies for the matrimonial status (single, widowed, divorced) and the number of dependent children.

b: extended control variables include the basic control variables mentioned above plus a dummy for the workplace (1 if Paris region), a dummy for the type of work contract (1 if CDD, "*Contrat à durée déterminée*") and the logarithm of the number of hours paid work per year.

c: the variables introduced in the factor analysis are educm, exphem, ancm, sdeduc, sdexphe, sdanc, supeduc, supexphe, supanc, sa, se, sex, s1, s2, s3, s4, s5, s6, s7, s8, s9, s10, s11, s12, t1, t2, t3, t4, t5, t6, ff1, ff2, ff3, ff4, va1, va2, va3, d21, d31, d32, d4a1, d4b1, d4b2, ps1, d151, d19a1, d19b1, d19b2, d19b3, d250, d251, d252, d253, d26a1, d26b1, d281, d301, d311, d321, d331, d332, d333, d341, d342, d3513, d3523, d3533, d3543, d3553, d3563, d3573, d3583, d3911, d3921, d39b1, d411.

The first row presents a series of simple quantile regressions in which we condition the log earnings on gender at the 5th, 10th, 25th, 50th (median), 75th, 90th, and 95th percentiles, without any control variable[13]. We notice that the observed log earnings gap increases as we move up the earnings distribution with a sharp acceleration after the 75th percentile. For instance, at the 75th percentile, we see a raw gender earnings gap of slightly less than 25%. This means that the log earnings of a man at the 75th percentile of the male earnings distribution is a bit more than 22 points above the log earnings of a woman at the 75th percentile of the female earnings distribution.

Interestingly, very similar patterns have emerged in other European countries. First, male and female earnings are closer at the bottom of the earnings distribution. Second, male and female earnings are extremely unequal at the top of the distribution, up to a maximum difference of about 50%. Third, there is a steady increase in the gender log earnings gap as we move up in the earnings distribution. Fourth, there is a sharp acceleration in the increase in the gender log earnings gap starting at about the 75th or 80th percentile in the earnings distribution. Following Albrecht et alii (2003), De la Rica et alii (2005) or Arulampalam et alii (2004), we interpret this last feature of the gender log earnings gap by percentile as a glass ceiling evidence.

Then, we examine various quantile estimates of the gender dummy coefficients when adding both male and female's labour market characteristics. Several specifications have been considered, the list of explanatory variables being further described in Table 2[14]. In what follows, we only focus on the gender dummy coefficient which indicates the extent to which the gender earnings gap remains unexplained at the different quantiles after controlling for individual differences in various combinations of characteristics. We begin by introducing into the earnings equations the covariates commonly used in labour economics, i.e. education, potential experience, tenure, dummies for the matrimonial status, nationality, and the number of dependent children. Then, we add job-specific variables such as the type of work contract, the workplace, the number of hours worked per year, sector of employment, and occupation.

When we control for education, experience off the current job, firm tenure, and other basic socio-economic characteristics (panel 2, Table 2), the gender dummies increase in absolute value *relative to the raw gender dummy* of panel 1 at the 5th and the 10th percentile, but then decrease from about the 20th through the 95th percentiles. The OLS gender dummy coefficient (at the mean) also diminishes. One explanation could be that, in the first quartile of the log hourly earnings distribution, women display more labour market experience than men while this is not the case as for workers belonging to the second, third and fourth quartiles.

In the panel 3 of Table 2, we introduce the extended control variables which include basic control variables plus the type of work contract, the log of hours paid per year, and the location of the firm. The quantile estimates indicate that the gender dummy decreases in absolute value from the 5th to the median percentile *as compared to the preceding model*. Then, from about the 75th to the 95th, however, the gender dummy increases. This might be a first indication that job-related characteristics (working conditions) do matter in explaining why the earnings gap is much greater at the upper tail of the earnings distribution[15].

We next present the estimated gender dummy coefficients after adding 12 sectoral dummies in the quantile regressions (panel 4). The same picture emerges from these estimates, and the gender dummy is reduced only minimally at the bottom of the earnings distribution and slightly increases from about the 20th percentile and more substantially at the top of the earnings distribution. Of course, the sector of employment is to some extent an endogenous characteristic since the choice of sector in which to work is typically made after education is completed.

---

[13] As pointed out in Albrecht et alii (2003), these quantile regression estimates have the advantage to attach standard errors to the estimated gender gaps at the various percentiles. The coefficient estimates for the gender dummy in this panel are identical to the log earnings gaps.

[14] The full set of estimates is in an appendix.

[15] For instance, men are more likely to have a temporary work contract (CDD) than women are as we move up along the earnings distribution: 18.2% for both men and women in the first quartile against 2.5% for men and 4.5% for women in the fourth quartile.

In panel 5, we account for the firms' computed factors stemming from the factor analysis. In so doing, our aim is to substitute a firm fixed-effect regression by a "firm factor effect" regression that may account for qualitative aspects of firms' wage policies and, in particular, for human capital external effects. In a sense, following Muller and Nordman (2004, 2005), we generalize the approach developed in Cardoso (1999) who regresses the firms' fixed effects on different variables[16]. For our purpose, the first ten computed factors concentrate most of the relevant information about the firms' characteristics (see previous section). A Wald test rejects at the 1 percent level the null hypothesis that the coefficients of these ten factors are jointly equal to zero. Since the covariates introduced in the factor analysis include the sectoral dummies, we omit these explanatory variables into the earnings functions.

For the sake of comparison, we have also performed a linear regression with firms' fixed effects. The female dummy coefficient in panel 3 (-0.1843) can be compared with the one estimated at the mean with the firm factor effect, which is equal to -0.1825 (panel 5). So, the female coefficient is slightly reduced as we move from sector fixed effects or firm fixed effects models towards a firm factor effects specification. It may be that the computed factors stemming from our PCA of firms' characteristics add a qualitative aspect of the firms' wage policy to our regressions that fixed effects models, either with sector or firm dummies, may not be able to totally control for.

According to the quantile estimates reported in panel 5, we find that taking into account the firms' factors increases the gender dummy coefficient in absolute value at the lower tail of the log earnings distribution (from the 5^{th} to the median percentile). Conversely, the coefficient is significantly reduced at the upper tail of the distribution, especially above the 75^{th} percentile. Now, the gender earnings gap amounts to about 31% at the 95^{th} percentile while it amounted to more than 35% with the extended and sectoral control variables. Therefore, a first novelty of our analysis with respect to the previous literature is to show that controlling for firms' characteristics reduces the extent of the glass ceiling phenomenon, albeit moderately[17].

Finally, panels 6 and 7 of Table 2 present the quantile log earnings regression estimates adding 29 occupational dummies. We present these estimates separately because there is no clear consensus as to whether occupation (and to some extent industry) should be taken into account to assess the extent of the gender wage gap. If employers differentiate between men and women through their tendency to hire into certain occupations, then occupational assignment is an outcome of employer practices rather than an outcome of individual choice or productivity differences[18]. While panel 6 presents the gender dummy coefficient of a sector fixed effect model, panel 7 accounts for the coefficients of a firm factor effect model. Again, both sets of estimates are very close.

As might be expected, controlling for occupation considerably reduces the gender gap throughout the earnings distribution. In panel 7, the unexplained gender gap falls to 8.8% at the 5^{th} percentile and, more importantly, to 22% at the 95^{th} percentile (compared to 9.4% and 31.7% in panel 5). We would argue that the effect of controlling for occupation on the gender earnings gap reflects the occupational segregation that may be present in France. However, we also note that if the gender earnings gap varies significantly at the upper tail of the earnings distribution from panel 1 to panel 7, it remains remarkably stable at the bottom 5^{th} or 10^{th} percentiles.

To summarize our results, we find that after adjusting for a set of socio-economic control variables as well as for firm-related characteristics, the gender earnings gap is significantly reduced at the top of the distribution. However, it is still greater at the top than at the bottom of the distribution, meaning that the glass ceiling pattern remains in the estimated unexplained gender earnings gaps. As a final

[16] An alternative strategy would be to estimate earnings functions including all the firms' characteristics used in the factor analysis. We tried this out, but our models suffered from severe multicollinearity problems (some important regressors were dropped or insignificant) since firms' characteristics are often highly correlated to each other. By definition, the main components derived from the factor analysis are poorly correlated and also have the advantage to sum up the main statistical information of these firm level variables.

[17] Note that this result is not sensitive to the number of included firms' factors in the earnings functions. In fact, adding more factors (up to a total of 20) does not change significantly the estimated coefficients on the gender dummy at each considered quantile of the earnings distribution.

[18] Conversely, one can argue that analyses that omit occupation and industry may overlook the importance of background and choice-based characteristics on wage outcomes, while analyses that fully control for these variables may undervalue the significance of labour market constraints on wage outcomes (Altonji and Blank, 1999).

step, we have attempted to investigate whether the uncertainty argument makes sense to explain the gender earnings gap. Unfortunately, it appears quite difficult to provide an accurate measurement of the uncertainty assumption, so that we have to rely on proxy variables evaluated on the firm sample. To better understand the factors that influence firms to pay different premia for men and women, we choose to focus on the difference between the male and female firm fixed effects. As pointed out in Meng and Meurs (2004), this difference is an estimate of the within-firm gender wage gap.

We first get the firm unobserved heterogeneity components from the previous male and female fixed effects regressions and compute the difference. Then, we introduce a set of covariates related to observable firm characteristics in order to explain the within-firm gender earnings gap. The difference in fixed effects is estimated using a weighted OLS regression, and the corresponding estimates are in Table 3. Explaining the within-firm gender gap remains difficult, as the R^2 is around 0.03[19]. We observe that the firm sectoral dummies along with the firm size affect this gender gap, but we focus in what follows on the covariates which are more likely to be related to uncertainty.

Table 3: Firm fixed effects regression
(Dependent variable: FFEm-FFEf)

Variables	Weighted OLS	
Constant	0.0620	(0.0392)
S1	0.0022	(0.0361)
S2	0.0000	(0.0000)
S3	-0.0653*	(0.0349)
S4	-0.0230	(0.0347)
S5	-0.0625*	(0.0345)
S6	-0.0417	(0.0383)
S7	-0.0455	(0.0334)
S8	-0.0099	(0.0338)
S9	-0.0383	(0.0325)
S10	-0.1000***	(0.0365)
S11	-0.0172	(0.0422)
S12	-0.0657*	(0.0359)
T1	-0.0279***	(0.0106)
T2	-0.0259**	(0.0103)
T3	-0.0224**	(0.0105)
T4	0.0000	(0.0000)
T5	0.0212**	(0.0100)
T6	0.0121	(0.0129)
PS1	-0.0370***	(0.0106)
D19B2	-0.0144	(0.0149)
D19B1	-0.0246**	(0.0100)
D19B3	-0.0199	(0.0176)
D26A1	-0.0019	(0.0022)
D26B1	-0.0164*	(0.0090)
D281	0.0076	(0.0082)
D301	0.0085	(0.0088)
D311	0.0049	(0.0087)
D321	-0.0185*	(0.0110)
D331	-0.0043	(0.0164)
D332	-0.0213*	(0.0110)
D341	-0.0198*	(0.0103)
D342	-0.0141	(0.0090)
D3583	0.0238*	(0.0142)
D39B1	0.0257*	(0.0141)
Observations	5931	
R-squared	0.0298	

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.
FFEm and FFEf respectively stand for male and female firm fixed effects. The regression also includes the following explanatory variables: VA1, VA2, D21, D32, D4A1, D4B1,D4B2, D252, D253,D131,D151, D3513, D3523, D353, D3543, D3553, D3563, D3573, D411.

[19] Meng and Meurs (2004) obtain a similar result, since they find that less than 4% of the variation of the within-firm gender earnings gap is explained by observable firm characteristics in France and about 7% in Australia.

A very significant effect is the presence of union representatives (PS1), which strongly reduces the within-firm gender gap. There are at least two explanations. On the one hand, it may be that the presence of a trade union prevents the firm from setting up wage discrimination between male and female workers, especially if the objective of the union is to reduce wage inequality within the firm. On the other hand, wages are most often higher in firms when there is a trade union, and the latter also allows for a better job protection. As the probability to be dismissed may be significantly lower in such cases, this reduces uncertainty on the worker's productivity (incentives to quit the job are lessened). A similar interpretation may be given for the use of a formal wage scale system based on the branch's collective agreement (D19B1), which also diminishes the within-firm gender earnings gap.

A second result in favour of the uncertainty argument concerns the variable D3583, which is significant at the 10 percent level. This variable indicates whether there exist some difficulties for the employer with workers' eventual replacement. It may be that such difficulties are the sign of more competitive industries with an increased turnover and less stability of workers. Also, being less confident as to the regularity and the presence of one's employees may entail more fear from employers about the possibility to replace vacant workers. Estimates in Table 4 show that such difficulties increase the within-firm gender gap. Finally, we also note that the variable D39B1 increases the within-firm gap. When firms favour relative bonuses such that the best workers are awarded, the within-firm gender gap is increased.

6. CONCLUDING COMMENTS

In this paper, we have brought both some theory and empirical evidence on the glass ceiling hypothesis, according to which the gender wage gap is more important at the upper tail of the wage distribution. Several studies have recently shed light on this phenomenon in European countries (Albrecht et alii, 2003, De la Rica et alii, 2005, Arulampalam et alii, 2004).

To rationalize the glass ceiling effect, we have introduced some uncertainty on the female productivity in a competitive labour market model. Women are likely to have more frequently interrupted careers (because of birth event for instance), and they may choose to quit the labour force either to spend time with their children, to care for elderly parents, or to move with their husband when the latter is promoted in a new location. We formally prove that this uncertainty entails a risk premium in the wage determination. As the premium is a quadratic function of the female level of human capital, the model predicts a larger wage gap at the top of the distribution.

We assess the relevance of the glass ceiling hypothesis in France using the 1992 ECMOSS data set, which provides information on both employers and employees. This is an empirical novelty with respect to the previous studies on the glass ceiling effect, which have neglected the role of the firm characteristics. Whatever the theoretical explanation behind the glass ceiling, the work environment is likely to affect the magnitude of the gender wage gap along the earnings distribution. Results from quantile regressions indicate that there exists a significant glass ceiling effect in France. While male and female earnings are close at the bottom of the income distribution, there is a strong increase in the gender earnings gap above the 75^{th} percentile of this distribution. Also, it indeed matters to control for the firms' characteristics. Following Muller and Nordman (2004, 2005), we rely on a principal component analysis to extract the most influential factors of the surveyed establishments in order to introduce them into the earnings regressions. This approach allows accounting for qualitative aspects of the firms and, in particular, for their implemented wage policy, which may not be the case when one controls for unobserved heterogeneity through firm fixed effect models.

According to the French data, the gender earnings gap would be overstated at the top of the distribution when the influence of firms' characteristics is omitted. However, despite of the reduction in the earnings gap for the more well-paid workers, there is still a large and significant difference between the male and female earnings. Focusing on the differences in firm fixed effects for male and female, we find that the within-firm gender earnings gap is increased when there are no union representatives in the firm and when employers express fear regarding the recruitment of workers in

case of replacements. In both cases, uncertainty on the employee's stability seems more a concern for employers.

A shortcoming of our analysis relate to the relevance of this uncertainty argument. Indeed, we have proposed a model which explains why women receive significantly lower wages than men only at the upper tail of the earnings distribution. While the consequence of uncertainty on female productivity is to lead to a glass ceiling effect, it may be that there exist other theoretical explanations which could rationalize this phenomenon (see De la Rica et alii, 2005). Clearly, the difficulty for researchers would be to construct formal models leading to testable, differentiated predictions. While there is now a burgeoning literature which evidences the relevance of the glass ceiling effect in industrialised countries, better understanding the origins and causes of this important stylised fact remains an important issue which is left for future research.

BIBLIOGRAPHY

Abowd J.M., Kramarz F. (1999), "The Analysis of Labor Markets Using Matched Employer-Employee Data", in Ashenfelter O. and Card D. (eds), *Handbook of Labor Economics*, vol. 3B, Elsevier, North-Holland, pp. 2629-2710.

Abowd J.M., Kramarz F., Margolis D. N. (1999), "High-Wage Workers and High-Wage Firms", *Econometrica*, vol. 67, pp. 251-333.

Abowd J., Kramarz F., Margolis D.N., Troske K.R. (2001), "The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States", *Journal of the Japanese and International Economies*, vol. 15, pp. 419-436.

Albrecht J., Björklund A., Vroman S. (2003), "Is There a Glass Ceiling in Sweden?", *Journal of Labor Economics*, vol. 21, pp. 145-177.

Altonji J.G., Blank R.M. (1999), "Race and Gender in the Labor Market", in Ashenfelter O.C., Card D. (eds), *Handbook of Labor Economics*, vol. 3C, Elsevier, chapter 48, pp. 3143-3258.

Arulampalam W., Booth A.L., Bryan M.L. (2004), "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wages Distribution", Communication at the 7th Australasian Labour Econometrics Workshop, University of Auckland, 13-14 August.

Blau F., Kahn L. (2000), "Gender Differences in Pay", *Journal of Economic Perspectives*, vol. 14, pp. 75-99.

Booth A., Francesconi M., Frank J. (2003), "A Sticky Floors Model of Promotion, Pay and Gender", *European Economic Review*, vol. 47, pp. 295-322.

Buchinsky M. (1998), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research", *Journal of Human Resources*, vol. 33, pp. 88-126.

Cardoso A. R. (1999), "Firms' Wage Policies and the Rise in Labor Market Inequality: The Case of Portugal", *Industrial and Labor Relations Review*, vol. 53(1), pp. 87-102.

Caroli E., Greenan N., Guellec D. (2001), "Organizational Change and Skill Accumulation", *Industrial and Corporate Change*, vol. 10, pp. 481-506.

De la Rica S., Dolado J.J., Llorens S. (2005), "Ceiling and Floors: Gender Wage Gaps by Education in Spain"; IZA Discussion Paper, n° 1483.

Destré G. (2003), "Fonctions de gains et diffusion du savoir : une estimation sur données françaises appariées", *Economie et Prévision*, vol. 158, pp. 89-104.

Destré G., Nordman C. (2002), "The Impacts of Informal Training on Earnings: Evidence from French, Moroccan and Tunisian Employer-Employee Matched Data", *L'Actualité Économique*, vol. 78, pp. 179-205.

Greenan N. (2003), "Organisational Change, Technology, Employment and Skills: an Empirical Study of French Manufacturing", *Cambridge Journal of Economics*, vol. 27, pp. 287-317.

Hahn J. (1995), "Bootstrapping Quantile Regression Estimators", *Econometric Theory*, vol. 11, pp. 105-121.

Ichino A., Filippin A. (2005), "Gender Wage Gap in Expectations and Realizations", *Labour Economics*, vol. 12, pp. 125-145.

Koenker R., Bassett G. (1978), “Regression Quantiles”, *Econometrica*, vol. 46, pp. 33-50.

Koenker R. (2005), *Quantile Regression*, Econometric Society Monograph Series, Cambridge University Press.

Lindbeck A., Snower D.J. (2000), “Multitask Learning and the Reorganization of Work: From Tayloristic to Holistic Organization”, *Journal of Labor Economics*, vol. 18, pp. 353-376.

Meng X. (2004), “Gender Earnings Gap: the Role of Firm Specific Effects”, *Labour Economics*, vol. 11, pp. 555-573.

Meng X., Meurs D. (2004), “The Gender Earnings Gap: Effects of Institutions and Firms – A Comparative Study of French and Australian Private Firms”, *Oxford Economic Papers*, vol. 56, pp. 189-208.

Muller C., Nordman C. (2004), “Which Human Capital Matters for Rich and Poor’s Wages? Evidence from Matched Worker-Firm Data from Tunisia”, DIAL Document de Travail, DT/2004-09 (www.dial.prd.fr). CREDIT Research Paper 04/08, University of Nottingham.

Muller C., Nordman C. (2005), “Firm Effects or Firm Human Capital? An Investigation of Wages from French Matched Worker-Firm Data”, mimeo, University of Alicante.

Nickell S.J. (1986), “Dynamic Models of Labour Demand”, in Ashenfelter O., Layard R. (eds), *Handbook of Labor Economics*, North-Holland, Elsevier, pp. 473-522.

Rubinstein M. (1976), “The Valuation of Uncertain Income Streams and the Pricing of Options”, *Bell Journal of Economics*, vol. 7, pp. 407-425.

APPENDICES

Table A 1: Descriptive statistics of the firms' characteristics

Variables	Mean	Standard deviation	Min	Max
Firms' human capital characteristics				
EDUCM	12.781	1.081	10	18
EDUCMNEW	12.783	1.116	9.5	18
EXPHEM	9.991	5.228	0	37.39
ANCM	8.544	5.662	0	33.81
SDEDUC	1.009	0.844	0	4.95
SDEDUCNEW	0.990	0.840	0	5.16
SDEXPHE	7.445	3.477	0	28.34
SDANC	5.610	3.577	0	27.28
SUPEDUC	14.269	1.928	10	18
SUPEXPHE	21.881	9.347	0	48.92
SUPANC	17.028	10.335	0	46.50
SA	4.930	6.717	0	42.58
SE	1.511	1.921	0	8
SEX	5.888	7.903	0	43.25
Sectoral dummies				
S1 : INDUSTRIES AGRICOLES ET ALIMENTAIRES	0.028		0	1
S2 : PRODUCTION ET DISTRIBUTION D'ENERGIE	0.011		0	1
S3 : BIENS INTERMEDIAIRES	0.054		0	1
S4 : BIENS D'EQUIPEMENT	0.055		0	1
S5 : BIENS DE CONSOMMATION COURANTE	0.065		0	1
S6 : BATIMENT GENIE CIVIL ET AGRICOLE	0.073		0	1
S7 : COMMERCE	0.173		0	1
S8 : TRANSPORT ET TÉLÉCOMMUNICATION	0.082		0	1
S9 : SERVICES MARCHANDS	0.398		0	1
S10 : LOCATION CREDIT-BAIL	0.009		0	1
S11 : ASSURANCES	0.020		0	1
S12 : ORGANISMES FINANCIERS	0.033		0	1
Firm size				
T1: less than 20 employees	0.313		0	1
T2 : 20 to 49 employees	0.341		0	1
T3 : 50 to 99 employees	0.153		0	1
T4 : 100 to 199 employees	0.066		0	1
T5 : 200 to 499 employees	0.063		0	1
T6 : more than 500 employees	0.065		0	1
Declared average number of paid hours of training per worker and per year in the establishment (dummies)				
FF1 : zero hour or no hours declared	0.498		0	1
FF2 : 1 to 10 hours	0.216		0	1
FF3 : 11 to 40 hours	0.203		0	1
FF4 : more than 40 hours	0.060		0	1
Firms' general and organisational characteristics				
VA1	0.440		0	1
VA2	0.293		0	1
VA3	0.210		0	1
D21	0.360		0	1
D31	2.448	2.225	1	9
D32	1.934	3.463	0	9
D4A1	0.333		0	1
D4B1	0.038		0	1
D4B2	0.308		0	1
PS1	0.252		0	1
D151	0.225		0	1
D19A1	0.769		0	1
D19B1	0.614		0	1
D19B2	0.099		0	1
D19B3	0.065		0	1
D250	0.222		0	1
D251	0.673		0	1
D252	0.035		0	1
D253	0.000		0	1
D26A1	2.258	1.738	1	9
D26B1	0.284		0	1
D281	0.499		0	1
D301	0.618		0	1
D311	0.361		0	1
D321	0.209		0	1
D331	0.057		0	1
D332	0.148		0	1
D333	0.260		0	1
D334	0.256		0	1
D341	0.235		0	1
D342	0.381		0	1
D343	0.125		0	1
D344	0.145		0	1
D3513	0.064		0	1
D3523	0.359		0	1
D3533	0.087		0	1
D3543	0.166		0	1
D3553	0.157		0	1
D3563	0.321		0	1
D3573	0.097		0	1
D3583	0.063		0	1
D3911	0.403		0	1
D3921	0.158		0	1
D39B1	0.076		0	1
D411	0.206		0	1

Source: Survey INSEE ECMOSS 1992. The number of establishments is 14,693.

Table A 2: Earnings functions, with basic control variables

(Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0859***	-0.1006***	-0.1264***	-0.1592***	-0.1948***	-0.2270***	-0.2446***	-0.1779***
	(0.0022)	(0.0019)	(0.0019)	(0.0021)	(0.0027)	(0.0040)	(0.0061)	(0.0019)
Constant	3.0594***	2.9315***	2.6759***	2.4401***	2.3124***	2.2892***	2.3880***	2.4047***
	(0.0084)	(0.0073)	(0.0079)	(0.0098)	(0.0135)	(0.0197)	(0.0289)	(0.0088)
Years of schooling	0.0406***	0.0540***	0.0816***	0.1102***	0.1327***	0.1475***	0.1499***	0.1167***
	(0.0005)	(0.0004)	(0.0005)	(0.0007)	(0.0009)	(0.0013)	(0.0020)	(0.0006)
Years of experience off the current establishment	0.0020***	0.0032***	0.0074***	0.0143***	0.0237***	0.0353***	0.0434***	0.0167***
	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0007)	(0.0011)	(0.0003)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0002***	-0.0003***	-0.0004***	-0.0006***	-0.0007***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0159***	0.0194***	0.0242***	0.0274***	0.0273***	0.0265***	0.0244***	0.0241***
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0005)	(0.0007)	(0.0011)	(0.0003)
(Years of tenure in the current firm)^2	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0001***	-0.0000	0.0001	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.0604***	-0.0758***	-0.1004***	-0.1275***	-0.1402***	-0.1211***	-0.1071***	-0.1351***
	(0.0047)	(0.0041)	(0.0042)	(0.0046)	(0.0059)	(0.0088)	(0.0135)	(0.0042)
Dummy for single	-0.0301***	-0.0262***	-0.0314***	-0.0280***	-0.0257***	-0.0260***	-0.0277***	-0.0357***
	(0.0031)	(0.0026)	(0.0026)	(0.0029)	(0.0037)	(0.0054)	(0.0084)	(0.0026)
Dummy for widowed	-0.0236**	-0.0297***	-0.0423***	-0.0497***	-0.0748***	-0.1240***	-0.0862***	-0.0611***
	(0.0100)	(0.0087)	(0.0088)	(0.0098)	(0.0124)	(0.0182)	(0.0281)	(0.0088)
Dummy for divorced	-0.0027	0.0015	0.0034	0.0049	-0.0035	-0.0140*	-0.0039	-0.0008
	(0.0047)	(0.0041)	(0.0041)	(0.0046)	(0.0058)	(0.0085)	(0.0131)	(0.0041)
Number of dependent children	0.0112***	0.0132***	0.0143***	0.0163***	0.0175***	0.0172***	0.0203***	0.0173***
	(0.0012)	(0.0010)	(0.0010)	(0.0010)	(0.0012)	(0.0018)	(0.0027)	(0.0009)
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table A 3: Earnings functions, with extended control variables
(Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0850***	-0.0978***	-0.1219***	-0.1514***	-0.2051***	-0.2558***	-0.2839***	-0.1803***
	(0.0021)	(0.0019)	(0.0019)	(0.0018)	(0.0024)	(0.0037)	(0.0052)	(0.0019)
Constant	2.8268***	2.7226***	2.5050***	2.3916***	2.5172***	2.9192***	3.3144***	2.4863***
	(0.0147)	(0.0132)	(0.0136)	(0.0137)	(0.0182)	(0.0286)	(0.0405)	(0.0138)
Years of schooling	0.0432***	0.0538***	0.0773***	0.1019***	0.1224***	0.1340***	0.1368***	0.1090***
	(0.0005)	(0.0004)	(0.0005)	(0.0006)	(0.0008)	(0.0013)	(0.0017)	(0.0006)
Years of experience off the current establishment	0.0021***	0.0028***	0.0064***	0.0123***	0.0206***	0.0307***	0.0362***	0.0147***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0006)	(0.0009)	(0.0003)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0002***	-0.0003***	-0.0004***	-0.0005***	-0.0006***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0147***	0.0174***	0.0216***	0.0253***	0.0269***	0.0298***	0.0310***	0.0239***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0007)	(0.0010)	(0.0004)
(Years of tenure in the current firm)^2	-0.0001***	-0.0002***	-0.0002***	-0.0002***	-0.0001***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.1045***	-0.1240***	-0.1553***	-0.1839***	-0.1982***	-0.1906***	-0.1558***	-0.1961***
	(0.0044)	(0.0040)	(0.0041)	(0.0040)	(0.0053)	(0.0082)	(0.0117)	(0.0040)
Dummy for single	-0.0277***	-0.0299***	-0.0340***	-0.0335***	-0.0338***	-0.0400***	-0.0403***	-0.0440***
	(0.0029)	(0.0026)	(0.0026)	(0.0025)	(0.0032)	(0.0050)	(0.0070)	(0.0025)
Dummy for widowed	-0.0257***	-0.0237***	-0.0405***	-0.0390***	-0.0534***	-0.0969***	-0.1154***	-0.0543***
	(0.0094)	(0.0085)	(0.0085)	(0.0084)	(0.0108)	(0.0168)	(0.0237)	(0.0084)
Dummy for divorced	-0.0045	-0.0002	0.0054	0.0002	-0.0041	-0.0138*	-0.0073	-0.0041
	(0.0044)	(0.0040)	(0.0040)	(0.0039)	(0.0051)	(0.0079)	(0.0111)	(0.0039)
Number of dependent children	0.0129***	0.0144***	0.0169***	0.0207***	0.0215***	0.0211***	0.0221***	0.0206***
	(0.0011)	(0.0010)	(0.0010)	(0.0009)	(0.0011)	(0.0016)	(0.0023)	(0.0009)
Dummy for workplace (1: Paris)	0.1204***	0.1550***	0.2028***	0.2366***	0.2787***	0.3195***	0.3419***	0.2500***
	(0.0025)	(0.0023)	(0.0023)	(0.0023)	(0.0029)	(0.0046)	(0.0065)	(0.0023)
Dummy for type of contract	-0.0032	-0.0051	-0.0069*	-0.0186***	-0.0482***	-0.0673***	-0.0992***	-0.0411***
(CDD: « *contrat à durée déterminée* »)	(0.0047)	(0.0042)	(0.0041)	(0.0039)	(0.0050)	(0.0078)	(0.0112)	(0.0039)
Log of hours paid work per year	0.0266***	0.0289***	0.0299***	0.0187***	-0.0138***	-0.0678***	-0.1116***	-0.0011
	(0.0018)	(0.0016)	(0.0016)	(0.0015)	(0.0020)	(0.0032)	(0.0047)	(0.0015)
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table A 4: Earnings functions, with extended control variables and sectoral dummies (Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0831***	-0.0956***	-0.1238***	-0.1527***	-0.2089***	-0.2708***	-0.3041***	-0.1881***
	(0.0023)	(0.0020)	(0.0019)	(0.0020)	(0.0025)	(0.0043)	(0.0058)	(0.0019)
Constant	3.0113***	2.9348***	2.7024***	2.5434***	2.6610***	3.0683***	3.4537***	2.4341***
	(0.0169)	(0.0150)	(0.0151)	(0.0164)	(0.0205)	(0.0354)	(0.0467)	(0.0155)
Years of schooling	0.0411***	0.0496***	0.0725***	0.0988***	0.1196***	0.1314***	0.1338***	0.1059***
	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(0.0008)	(0.0014)	(0.0017)	(0.0006)
Years of experience off the current establishment	0.0021***	0.0030***	0.0071***	0.0133***	0.0215***	0.0304***	0.0360***	0.0157***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0007)	(0.0009)	(0.0003)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0002***	-0.0003***	-0.0004***	-0.0005***	-0.0006***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0159***	0.0173***	0.0203***	0.0230***	0.0253***	0.0282***	0.0309***	0.0227***
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0008)	(0.0010)	(0.0003)
(Years of tenure in the current firm)^2	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0001***	-0.0001***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.1003***	-0.1106***	-0.1380***	-0.1598***	-0.1735***	-0.1701***	-0.1277***	-0.1771***
	(0.0044)	(0.0040)	(0.0039)	(0.0042)	(0.0052)	(0.0089)	(0.0117)	(0.0040)
Dummy for single	-0.0265***	-0.0293***	-0.0350***	-0.0330***	-0.0324***	-0.0417***	-0.0417***	-0.0433***
	(0.0028)	(0.0025)	(0.0025)	(0.0026)	(0.0032)	(0.0054)	(0.0070)	(0.0025)
Dummy for widowed	-0.0275***	-0.0298***	-0.0456***	-0.0413***	-0.0596***	-0.0944***	-0.0957***	-0.0567***
	(0.0093)	(0.0083)	(0.0082)	(0.0086)	(0.0106)	(0.0179)	(0.0233)	(0.0083)
Dummy for divorced	-0.0071	-0.0033	-0.0024	-0.0024	-0.0075	-0.0174**	-0.0125	-0.0070*
	(0.0044)	(0.0039)	(0.0038)	(0.0040)	(0.0049)	(0.0084)	(0.0110)	(0.0039)
Number of dependent children	0.0124***	0.0127***	0.0144***	0.0178***	0.0191***	0.0183***	0.0202***	0.0188***
	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0011)	(0.0018)	(0.0023)	(0.0009)
Dummy for workplace (1: Paris)	0.1176***	0.1431***	0.1895***	0.2320***	0.2797***	0.3148***	0.3337***	0.2444***
	(0.0025)	(0.0023)	(0.0023)	(0.0024)	(0.0029)	(0.0050)	(0.0065)	(0.0023)
Dummy for type of contract (CDD: « *contrat à durée déterminée* »)	-0.0164***	-0.0119***	-0.0136***	-0.0266***	-0.0523***	-0.0764***	-0.1043***	-0.0467***
	(0.0046)	(0.0040)	(0.0039)	(0.0040)	(0.0049)	(0.0084)	(0.0112)	(0.0038)
Log of hours paid work per year	0.0267***	0.0289***	0.0298***	0.0177***	-0.0150***	-0.0683***	-0.1097***	-0.0021
	(0.0017)	(0.0015)	(0.0015)	(0.0016)	(0.0020)	(0.0035)	(0.0046)	(0.0015)
Control for sector	yes	yes	yes	yes	yes	yes	yes	yes
Control for occupation	no	no	no	no	no	no	no	no
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table A 5: Earnings functions, with extended control variables and firms' factor effects
(Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0746***	-0.0814***	-0.0969***	-0.1148***	-0.1381***	-0.1690***	-0.1955***	-0.1267***
	(0.0030)	(0.0024)	(0.0020)	(0.0017)	(0.0021)	(0.0030)	(0.0041)	(0.0017)
Constant	3.5225***	3.5954***	3.7356***	4.0149***	4.4499***	5.0738***	5.7606***	4.0597***
	(0.0531)	(0.0441)	(0.0373)	(0.0338)	(0.0421)	(0.0605)	(0.0835)	(0.0338)
Years of schooling	0.0180***	0.0191***	0.0210***	0.0254***	0.0289***	0.0314***	0.0316***	0.0283***
	(0.0008)	(0.0007)	(0.0006)	(0.0005)	(0.0007)	(0.0010)	(0.0013)	(0.0005)
Years of experience off the current establishment	0.0026***	0.0031***	0.0047***	0.0070***	0.0093***	0.0107***	0.0118***	0.0083***
	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0004)	(0.0006)	(0.0002)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0001***	-0.0001***	-0.0002***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0157***	0.0168***	0.0184***	0.0197***	0.0206***	0.0213***	0.0212***	0.0199***
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0007)	(0.0003)
(Years of tenure in the current firm)^2	-0.0002***	-0.0002***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.0481***	-0.0480***	-0.0473***	-0.0463***	-0.0366***	-0.0267***	-0.0205***	-0.0481***
	(0.0048)	(0.0040)	(0.0034)	(0.0031)	(0.0039)	(0.0056)	(0.0078)	(0.0031)
Dummy for single	-0.0218***	-0.0239***	-0.0228***	-0.0227***	-0.0247***	-0.0237***	-0.0241***	-0.0271***
	(0.0030)	(0.0025)	(0.0021)	(0.0019)	(0.0024)	(0.0034)	(0.0047)	(0.0019)
Dummy for widowed	-0.0175*	-0.0261***	-0.0256***	-0.0295***	-0.0278***	-0.0339***	-0.0364**	-0.0303***
	(0.0099)	(0.0083)	(0.0070)	(0.0063)	(0.0079)	(0.0113)	(0.0156)	(0.0063)
Dummy for divorced	-0.0058	-0.0034	0.0003	0.0007	-0.0013	-0.0016	0.0018	-0.0031
	(0.0047)	(0.0039)	(0.0033)	(0.0030)	(0.0037)	(0.0053)	(0.0073)	(0.0029)
Number of dependent children	0.0109***	0.0108***	0.0111***	0.0120***	0.0122***	0.0138***	0.0116***	0.0129***
	(0.0011)	(0.0009)	(0.0008)	(0.0007)	(0.0008)	(0.0011)	(0.0016)	(0.0007)
Dummy for workplace (1: Paris)	0.0906***	0.1029***	0.1216***	0.1436***	0.1662***	0.1674***	0.1728***	0.1453***
	(0.0029)	(0.0024)	(0.0020)	(0.0018)	(0.0022)	(0.0031)	(0.0043)	(0.0018)
Dummy for type of contract (CDD: « *contrat à durée déterminée* »)	-0.0386***	-0.0330***	-0.0225***	-0.0203***	-0.0379***	-0.0587***	-0.0739***	-0.0422***
	(0.0049)	(0.0040)	(0.0033)	(0.0030)	(0.0037)	(0.0055)	(0.0078)	(0.0029)
Log of hours paid work per year	0.0138***	0.0105***	0.0037***	-0.0097***	-0.0386***	-0.0925***	-0.1419***	-0.0302***
	(0.0020)	(0.0016)	(0.0013)	(0.0012)	(0.0015)	(0.0023)	(0.0032)	(0.0012)
Control for sector	yes	yes	yes	yes	yes	yes	yes	yes
Control for occupation	yes	yes	yes	yes	yes	yes	yes	yes
Firm factor effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table A 6: Earnings functions, with extended control variables and sectoral and occupational effects (Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0921***	-0.1053***	-0.1288***	-0.1570***	-0.2018***	-0.2499***	-0.2757***	-0.1825***
	(0.0024)	(0.0020)	(0.0018)	(0.0019)	(0.0026)	(0.0038)	(0.0052)	(0.0018)
Constant	2.8992***	2.8164***	2.6677***	2.5772***	2.6610***	3.0234***	3.3881***	2.6491***
	(0.0169)	(0.0143)	(0.0133)	(0.0139)	(0.0204)	(0.0298)	(0.0413)	(0.0138)
Years of schooling	0.0390***	0.0480***	0.0682***	0.0915***	0.1130***	0.1273***	0.1319***	0.0987***
	(0.0006)	(0.0005)	(0.0005)	(0.0006)	(0.0009)	(0.0013)	(0.0018)	(0.0006)
Years of experience off the current establishment	0.0028***	0.0042***	0.0072***	0.0128***	0.0208***	0.0306***	0.0352***	0.0153***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0006)	(0.0009)	(0.0003)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0002***	-0.0003***	-0.0004***	-0.0005***	-0.0005***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0161***	0.0180***	0.0209***	0.0234***	0.0261***	0.0295***	0.0307***	0.0234***
	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0007)	(0.0010)	(0.0004)
(Years of tenure in the current firm)^2	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.1006***	-0.1181***	-0.1321***	-0.1628***	-0.1796***	-0.1888***	-0.1487***	-0.1792***
	(0.0049)	(0.0042)	(0.0039)	(0.0040)	(0.0058)	(0.0085)	(0.0116)	(0.0040)
Dummy for single	-0.0329***	-0.0322***	-0.0361***	-0.0358***	-0.0366***	-0.0395***	-0.0392***	-0.0461***
	(0.0032)	(0.0027)	(0.0024)	(0.0025)	(0.0036)	(0.0051)	(0.0069)	(0.0025)
Dummy for widowed	-0.0343***	-0.0303***	-0.0490***	-0.0403***	-0.0562***	-0.1047***	-0.1318***	-0.0547***
	(0.0105)	(0.0089)	(0.0081)	(0.0083)	(0.0119)	(0.0172)	(0.0235)	(0.0083)
Dummy for divorced	-0.0054	-0.0068	0.0023	-0.0033	-0.0033	-0.0250***	-0.0184*	-0.0078**
	(0.0049)	(0.0042)	(0.0038)	(0.0039)	(0.0056)	(0.0080)	(0.0110)	(0.0039)
Number of dependent children	0.0113***	0.0127***	0.0140***	0.0181***	0.0188***	0.0184***	0.0214***	0.0180***
	(0.0012)	(0.0010)	(0.0009)	(0.0009)	(0.0012)	(0.0017)	(0.0023)	(0.0009)
Dummy for workplace (1: Paris)	0.1199***	0.1452***	0.1787***	0.2119***	0.2499***	0.2893***	0.3159***	0.2260***
	(0.0029)	(0.0025)	(0.0023)	(0.0023)	(0.0033)	(0.0048)	(0.0066)	(0.0023)
Dummy for type of contract (CDD: « *contrat à durée déterminée* »)	-0.0262***	-0.0268***	-0.0213***	-0.0324***	-0.0472***	-0.0719***	-0.1003***	-0.0481***
	(0.0051)	(0.0043)	(0.0039)	(0.0039)	(0.0055)	(0.0080)	(0.0111)	(0.0038)
Log of hours paid work per year	0.0238***	0.0258***	0.0240***	0.0119***	-0.0174***	-0.0709***	-0.1139***	-0.0060***
	(0.0020)	(0.0016)	(0.0015)	(0.0015)	(0.0022)	(0.0033)	(0.0046)	(0.0015)
Control for sector	no	no	no	no	no	no	no	no
Control for occupation	no	no	no	no	no	no	no	no
Firm factor effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table A 7: Earnings functions, with extended control variable, firms' factor effects and occupational effects (Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
Dummy for female	-0.0848***	-0.0904***	-0.1017***	-0.1219***	-0.1447***	-0.1764***	-0.2008***	-0.1326***
	(0.0030)	(0.0023)	(0.0019)	(0.0017)	(0.0021)	(0.0028)	(0.0039)	(0.0017)
Constant	3.4573***	3.5584***	3.7292***	4.0518***	4.4540***	5.1596***	5.6814***	4.1618***
	(0.0521)	(0.0404)	(0.0356)	(0.0337)	(0.0413)	(0.0573)	(0.0793)	(0.0335)
Years of schooling	0.0154***	0.0164***	0.0181***	0.0220***	0.0252***	0.0281***	0.0293***	0.0250***
	(0.0008)	(0.0006)	(0.0006)	(0.0005)	(0.0007)	(0.0009)	(0.0013)	(0.0005)
Years of experience off the current establishment	0.0023***	0.0033***	0.0049***	0.0068***	0.0089***	0.0108***	0.0123***	0.0079***
	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0004)	(0.0006)	(0.0002)
(Years of experience off the current firm)^2	-0.0001***	-0.0001***	-0.0001***	-0.0001***	-0.0002***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of tenure in the current firm	0.0159***	0.0172***	0.0185***	0.0201***	0.0207***	0.0213***	0.0219***	0.0203***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0007)	(0.0003)
(Years of tenure in the current firm)^2	-0.0002***	-0.0002***	-0.0002***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Dummy for non-French	-0.0509***	-0.0499***	-0.0485***	-0.0462***	-0.0413***	-0.0326***	-0.0327***	-0.0505***
	(0.0048)	(0.0037)	(0.0033)	(0.0031)	(0.0039)	(0.0054)	(0.0075)	(0.0031)
Dummy for single	-0.0260***	-0.0237***	-0.0248***	-0.0245***	-0.0267***	-0.0270***	-0.0248***	-0.0292***
	(0.0030)	(0.0023)	(0.0020)	(0.0019)	(0.0024)	(0.0033)	(0.0046)	(0.0019)
Dummy for widowed	-0.0211**	-0.0211***	-0.0249***	-0.0237***	-0.0200**	-0.0331***	-0.0312**	-0.0271***
	(0.0098)	(0.0076)	(0.0067)	(0.0064)	(0.0079)	(0.0109)	(0.0150)	(0.0063)
Dummy for divorced	-0.0062	-0.0056	0.0001	-0.0004	0.0012	-0.0024	-0.0006	-0.0033
	(0.0046)	(0.0036)	(0.0032)	(0.0030)	(0.0037)	(0.0051)	(0.0070)	(0.0030)
Number of dependent children	0.0110***	0.0107***	0.0110***	0.0119***	0.0123***	0.0112***	0.0101***	0.0129***
	(0.0011)	(0.0009)	(0.0007)	(0.0007)	(0.0008)	(0.0011)	(0.0015)	(0.0007)
Dummy for workplace (1: Paris)	0.0902***	0.0979***	0.1120***	0.1319***	0.1453***	0.1454***	0.1488***	0.1311***
	(0.0029)	(0.0022)	(0.0020)	(0.0018)	(0.0022)	(0.0031)	(0.0043)	(0.0018)
Dummy for type of contract (CDD: « *contrat à durée déterminée* »)	-0.0414***	-0.0361***	-0.0249***	-0.0229***	-0.0405***	-0.0624***	-0.0739***	-0.0446***
	(0.0049)	(0.0037)	(0.0032)	(0.0030)	(0.0037)	(0.0053)	(0.0075)	(0.0029)
Log of hours paid work per year	0.0131***	0.0100***	0.0027**	-0.0109***	-0.0405***	-0.0946***	-0.1431***	-0.0312***
	(0.0020)	(0.0015)	(0.0013)	(0.0012)	(0.0015)	(0.0022)	(0.0031)	(0.0012)
Control for sector	no	no	no	no	no	no	no	no
Control for occupation	yes	yes	yes	yes	yes	yes	yes	yes
Firm factor effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	137211	137211	137211	137211	137211	137211	137211	137211
Pseudo R-squared	0.1131	0.1402	0.1823	0.2151	0.2347	0.2509	0.2494	0.3534

Survey INSEE ECMOSS 1992.
Standard errors are in parentheses. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.