



**HAL**  
open science

## Evidence on the glass ceiling effect in France using matched worker-firm data

Mohamed Jellal, Christophe Jalil Nordman, Francois-Charles Wolff

► **To cite this version:**

Mohamed Jellal, Christophe Jalil Nordman, Francois-Charles Wolff. Evidence on the glass ceiling effect in France using matched worker-firm data. *Applied Economics*, 2008, 40 (24), pp.3233-3250. 10.1080/00036840600994070 . hal-00582039

**HAL Id: hal-00582039**

**<https://hal.science/hal-00582039>**

Submitted on 1 Apr 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Evidence on the glass ceiling effect in France using matched worker-firm data**

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-06-0005.R1
Journal Selection:	Applied Economics
JEL Code:	J16 - Economics of Gender < J1 - Demographic Economics < J - Labor and Demographic Economics, J31 - Wage Level, Structure; Differentials by Skill, Occupation, etc. < J3 - Wages, Compensation, and Labor Costs < J - Labor and Demographic Economics, J24 - Human Capital Skills Occupational Choice Labor Productivity < J2 - Time Allocation, Work Behavior, and Employment Determination/Creation < J - Labor and Demographic Economics
Keywords:	gender wage gap, glass ceiling, quantile regressions, matched worker-firm data

powered by ScholarOne  
Manuscript Central™

# Evidence on the glass ceiling effect in France using matched worker-firm data <sup>#</sup>

Mohamed JELLAL\*, Christophe NORDMAN\*\* and François-Charles WOLFF\*\*\*

*Applied Economics, first revision*

June 2006

**Abstract:** In this paper, we investigate the relevance of the glass ceiling hypothesis in France, according to which there exist larger gender wage gaps at the upper tail of the wage distribution. Using a matched worker-firm data set of about 130,000 employees and 14,000 employers, we estimate quantile regressions and rely on a principal component analysis to summarize information specific to the firms. Our different results show that accounting for firm-related characteristics 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. Furthermore, a quantile decomposition shows that the gender wage gap is mainly due to differences in the returns to observed characteristics rather than in differences in characteristics between men and women.

**JEL Classification:** J24, J31, J16

**Keywords:** gender wage gap; glass ceiling; quantile regressions; matched worker-firm data

---

<sup>#</sup> We would like to thank seminar participants at DIAL, INED and LEST for helpful suggestions. We have also benefited from the constructive comments and remarks of an anonymous referee. Any remaining errors are ours.

\* Université Mohammed V, Rabat, Morocco ; and Conseils-Eco, 10 impasse de Mansencal, 31500 Toulouse, France. E-mail : [jellalmohamed@yahoo.fr](mailto:jellalmohamed@yahoo.fr)

\*\* Corresponding author. DIAL, IRD, Paris. 4 rue d'Enghien, 75010 Paris, France. E-mail: [nordman@dial.prd.fr](mailto:nordman@dial.prd.fr)

\*\*\* LEN, Université de Nantes, BP 52231 Chemin de la Censive du Tertre, 44322 Nantes Cedex, France ; CNAV and INED, Paris, France. E-mail: [wolff@sc-eco.univ-nantes.fr](mailto:wolff@sc-eco.univ-nantes.fr) <http://www.sc-eco.univ-nantes.fr/~fcwolff>

## 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<sup>1</sup>. 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 exist 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<sup>2</sup>.

For instance, using data collected in 1998 in Sweden, Albrecht *et al.* (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 al.* (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.

---

<sup>1</sup> The former, within the competitive framework, emphasise 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.

1  
2  
3 Using the European Community Household Panel data set, Arulampalam *et al.* (2004)  
4  
5 find that for most of their ten EU countries, in both the public and private sectors, the average  
6  
7 gender wage gap can be broken up into a gap that is typically wider at the top and occasionally  
8  
9 also wider at the bottom of the conditional wage distribution. They interpret the gender wage gap  
10  
11 at the top of the wage distribution as a glass ceiling evidence, whereby women otherwise identical  
12  
13 to men can only advance so far up the pay ladder. At the bottom of the wage distribution in  
14  
15 some of their EU countries, they also find that the gender pay gap widens significantly and define  
16  
17 this phenomenon as a sticky floor (see also Booth *et al.*, 2003; Ichino and Filippin, 2005).  
18  
19  
20  
21

22 Surprisingly, to date, there is no clear theoretical argument to rationalize the glass ceiling  
23  
24 effect among the various usual existing explanations for the gender wage gap. According to the  
25  
26 beckerian theory, discrimination is due to the discriminatory tastes of employers, co-workers, or  
27  
28 customers. Alternatively, in models of statistical discrimination, differences in the treatment of  
29  
30 men and women arise from average differences between the two groups in the expected value of  
31  
32 productivity or in the reliability with which productivity may be predicted, which lead employers  
33  
34 to discriminate on the basis of that average. Discriminatory exclusion of women from ‘male’ jobs  
35  
36 can also result in an excess supply of labour in ‘female’ occupations, depressing wages there for  
37  
38 otherwise equally productive workers. But in these various approaches, there is no reason to  
39  
40 expect larger gaps at the upper tail of the wage distribution.  
41  
42  
43  
44

45 De la Rica *et al.* (2005) suggest that a dead-end argument operate in the upper tail of the  
46  
47 distribution<sup>3</sup>. Women are less frequently promoted because their jobs can less easily be promoted.  
48  
49 Employees are most often reluctant to invest in women’s training, for instance because women  
50  
51 have more favourable outside opportunities than men within the household. Jellal *et al.* (2005)  
52  
53 introduce uncertainty on the female productivity in a competitive labour market model. Women  
54  
55  
56

---

57  
58 <sup>2</sup> Conversely, evidence on gender-earnings differentials in less developed countries remains scarce. See for instance  
59 Sakellariou (2004) for quantile wage regressions in the Philippines, Hinks (2002) in South Africa or Nielsen (2000) in  
60 Zambia.

1  
2  
3 are likely to have more frequently interrupted careers (because of birth event for instance), and  
4 they may choose to quit the labour force either to spend time with children or to care for elderly  
5 parents. Owing to this uncertainty, firms pass the risk of variability in women's production on  
6 female wages and the negative risk premium increases as women are more qualified.  
7  
8  
9

10  
11  
12 In this paper, we wonder whether it matters to control for firms' characteristics when  
13 estimating the gender wage gap along the wage distribution. Our contribution is thus mainly  
14 empirical, as we do not propose any economic hypotheses related to firm characteristics. This is  
15 undoubtedly a shortcoming of the present analysis, but the difficulty is that several explanations  
16 may be invoked to rationalize the glass ceiling effect. Furthermore, it may be excessively difficult  
17 to assess the relevance of these hypotheses<sup>4</sup>. Thus, our paper may be seen as a first step in the  
18 inclusion of firms' characteristics in the gender wage gap and glass ceiling literature, with a focus  
19 on empirical assessment. The next step would be naturally to further investigate the influence of  
20 these firms' factors in order to get a comprehensive understanding on the glass ceiling effect.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

33  
34 So, we primarily assess the relevance of the glass ceiling hypothesis using matched  
35 worker-firm data collected in 1992 in France. Although differences in productivity across workers  
36 could stem from their differences in human capital, it is well acknowledged that some skills or  
37 human capital attributed to workers are also specific to the firms in which those workers operate.  
38 Thus, part of the returns to human capital for the worker remuneration can be viewed as  
39 originating from the firm (see Abowd *et al.*, 1999)<sup>5</sup>. Hence, not controlling for firm specific  
40 effects on individual earnings differentials may lead to biased estimates when focusing on returns  
41 to human capital by gender. Curiously, firm specific effects have been neglected so far in the  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

---

53  
54 <sup>3</sup> Conversely, since high-educated women have participation rates which are only slightly lower than male  
55 participation rates, women's and men's wages should not be very different in the lower part of the income  
56 distribution (De la Rica *et al.*, 2005).

57 <sup>4</sup> This is the case for instance for the uncertainty argument suggested by Jellal *et al.* (2005). The French data provides  
58 very poor proxy on the measure of uncertainty. Also, finding some results in favor of the uncertainty argument does  
59 not preclude that the motivation for the glass ceiling effect is due to other arguments not specifically related to  
60 uncertainty considerations.

<sup>5</sup> 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 al.*, 1999).

1  
2  
3 analysis of gender earnings gap, Bayard *et al.* (2003) and Meng (2004) being worthwhile  
4  
5 exceptions.  
6

7  
8 As a consequence, the gender wage gap at the upper tail of the wage distribution may be  
9  
10 wrongly overstated if firms reward highly educated women differently than men. Thus, our main  
11  
12 contribution is to propose for the first time an empirical investigation of the glass ceiling effect in  
13  
14 a context where specific firm effects are controlled for. Following previous studies, we first use  
15  
16 quantile regressions techniques to assess the extent to which the glass ceiling phenomenon exists  
17  
18 in France. Then, we control for firms' specific wage policies by introducing the firms' features  
19  
20 into the different earnings functions. A novelty of our approach is to perform quantile regressions  
21  
22 using a preliminary principal component analysis of firms' characteristics, as in Muller and  
23  
24 Nordman (2004, 2005). We also carry out a quantile decomposition analysis with the inclusion of  
25  
26 firms' factors. We follow the method of Machado and Mata (2005) and examine whether gender  
27  
28 wage differences stem from differentiated returns to observable characteristics.  
29  
30  
31  
32

33  
34 In France, we find that introducing firm-related characteristics into earnings equation  
35  
36 significantly reduces the gender earnings gap at the top of the distribution. However, the gender  
37  
38 wage gap still remains greater at the top than at the bottom as in other European countries. The  
39  
40 rest of the paper is organized as follows. In Section 2, we describe the French ECMOSS survey  
41  
42 conducted in 1992 by INSEE. In Section 3, we present the econometric methodology and our  
43  
44 strategy to account for firm related characteristics. We discuss in Section 4 the different results of  
45  
46 the quantile regressions along with results from the quantile decomposition. Section 5 concludes.  
47  
48  
49  
50

## 51 52 **2. The French matched worker-firm data**

53  
54 The data we use in this paper are drawn from a unique French survey which matches  
55  
56 information of both employers and employees, the 1992 INSEE survey on labour cost and wage  
57  
58 structure (*Enquêtes sur le Coût de la Main-d'Oeuvre et la Structure des Salaires en 1992*, ECMOSS  
59  
60 thereafter). It is well known that such data sets allow the structure of wages to be modelled while

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
controlling for firm-specific effects (see Abowd *et al.*, 1999). Specifically, the French survey  
contains information on 150,000 different workers across 16,000 different workplaces<sup>6</sup>. 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.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
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.

31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
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)<sup>7</sup>.

47  
48  
49  
50  
51  
52  
53  
54  
By definition, we have only information on persons having a paid job in the French  
matched data set. This is undoubtedly a shortcoming as we are unable to account for gender  
differences in the labour force participation. It is well known that women have a lower probability

---

55  
56  
57  
58  
59  
60  
<sup>6</sup> 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 al.* (2001), Destré and Nordman (2002), Destré (2003) and Meng and Meurs (2004).

<sup>7</sup> 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).



1  
2  
3 to take part in the labour market, which raises some selectivity issue. Albeit the expected positive  
4 self-selection for women that may affect the magnitude of the gender wage gap, the problem is  
5 certainly not too severe. The selection issue has been recently addressed within a quantile  
6 decomposition framework by Albrecht *et al.* (2006). In the Netherlands, these authors evidence a  
7 positive self-selection of women into full-time work, but still find that the bulk of the gender  
8 wage gap is due to differences between men and women in the return to individual characteristics.  
9

10  
11 In the same way, again owing to data limitation, we focus on workers currently employed  
12 in the private sector. Again, this restriction leads to a selection bias as women tend to be more  
13 present in the public sector. Nevertheless, in the French context, the problem is certainly less  
14 severe than it seems owing to the fact that, in the public sector, wages are mainly fixed by law and  
15 then cannot really respond to productivity or discrimination strategies. After deleting  
16 observations with missing values or outliers, the worker sample amounts to 137,211 individuals  
17 divided into 14,693 establishments. Table 1 provides a description of the characteristics of the  
18 employees.  
19

20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
*Insert Table 1 about here*

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 A of the Appendix.

First, we make use of twelve sectoral dummies (s1, s2, s3, s4, s5, s6, s7, s8, s9, s10, s11, s12), the size of the establishments (A4)<sup>8</sup>, and four dummies for the average number of paid hours of training per worker in the establishment (in increasing order, FF1, FF2, FF3, FF4). 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

1  
2  
3 (D21), whether it is rather regular (D31), or irregular (D32), whether firms have been affected by  
4 unusual shocks in 1992 (D4A1) and, if it is the case, whether it was a downturn (D4B1) or an  
5 upturn (D4B2). We also make use of qualitative features of intra-firm wage determination such as  
6 dummies for the presence of union representatives (PS1), for the existence of wage negotiations in  
7 1992 (D151), and for the use of a formal wage scale system for blue collar workers' wage base  
8 (D19A1). If such a formal system is used, the questionnaire provides further information as to  
9 whether it is based on the branch's collective agreement (D19B1), on the firm's collective  
10 agreement (D19B2) or on another evaluation scheme (evaluation of posts, D19B3).

11  
12 Further detailed information describes the importance accorded by employers to different  
13 criteria in individual wage increases (for both blue collar and white collar workers). In the  
14 questionnaire, the answers were ranked according to three different levels of importance: "none",  
15 "weak", "medium", and "very strong". In our analysis, we make use of dummies taking into  
16 account the answers "very strong": workers' tenure in the job (D3513), increase in workers'  
17 performance (D3523), workers' training effort (D3533), accumulation of experience (D3543),  
18 acquisition of versatility (D3553), increase in workers' responsibilities (D3563), intra-firm mobility  
19 (D3573), and difficulty of workers' eventual replacement (D3583).

20  
21 Qualitative variables are then used to describe the extent to which employers favour  
22 individual or general wage increases in their wage policy: whether the base wage progressions  
23 "exclusively" (D331), "principally" (D332), "little" (D333) or "never" (D334) depend on individual  
24 increases and on general increases (respectively, D341, D342, D343, D344). Dummies regarding  
25 individual bonuses according to performances are also introduced in the following way: D3911  
26 indicates whether firms give relative bonuses (the best workers are awarded), D3921 describes  
27 whether bonuses are of "absolute" type (the production standards are exceeded). If these two  
28 schemes exist in the same firm, D39B1 reports which one is the most important (equals to one if it

---

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
<sup>8</sup> For the econometric analysis, dummies are also 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).

1  
2  
3 is relative bonuses). Finally, D411 signals firms having implemented an explicit wage policy  
4 characterised by precise objectives.  
5  
6

7  
8 Firms' organisational features are likely to influence employers' wage settings as well as  
9 skill diffusion and acquisition processes (Lindbeck and Snower, 2000; Caroli *et al.*, 2001; Greenan,  
10 2003). We construct dummies describing the firm's hierarchical structure such as the number of  
11 intermediate levels of management between the firm's manager and the blue collar workers  
12 assigned to productive lines (zero levels: D250; from 1 to 4: D251; from 5 to 10: D252; 11 levels  
13 and above: D253), dummies indicating the existence of job rotation schemes and how they are  
14 implemented (whether they are put into practice within production teams: D26A1, and whether  
15 they are intended to some versatile workers independently from team working: D26B1), a dummy  
16 when direct collaborations between employees of different departments are encouraged (D281), a  
17 dummy reporting whether achieved work is "permanently" controlled rather than "intermittently"  
18 or "occasionally" (D301), a dummy signalling whether individual performances are  
19 "systematically" controlled rather than "occasionally" or "never" (D311), and a dummy for the  
20 existence of a formal system to measure individual performances (D321).  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39

### 40 **3. Econometric strategy**

41  
42 Quantile regression and least absolute deviation estimators are now popular estimation  
43 methods (Koenker and Bassett, 1978; Buchinsky, 1998). This technique can be interpreted as  
44 using the error distribution in the earnings equation for the definition of different earnings  
45 categories, i.e. quantiles, instead of the observed earnings differentials. The popularity of these  
46 methods relies on three sets of properties.  
47  
48  
49  
50  
51  
52

53  
54 First, they provide robust estimates, particularly for misspecification errors related to non-  
55 normality and heteroskedasticity, but also for the presence of outliers due to data contamination.  
56  
57 Second, they allow the researcher to focus on specific parts of the distribution of interest, which  
58 is the conditional distribution of the dependent variable, and to estimate the marginal effect of a  
59  
60

1  
2  
3 covariate on log earnings at various points in the distribution. So, quantile regressions allow to  
4 estimate the effect of gender, education or experience on log earnings at the bottom of the log  
5 earnings distribution, at the median, and at the top of the distribution. Third, quantile regressions  
6 are appropriate when earnings functions contribute only to a small part of the variance of  
7 earnings, so that the distribution of earnings and the distribution of errors are close<sup>9</sup>.  
8  
9

10 Let us briefly describe the underlying econometric specification. We denote by  $w_{ij}$  the log  
11 earnings of individual  $i$  working in firm  $j$ , and  $x_{ij}$  a vector of explanatory variables excluding  
12 gender. We define  $f_{ij}$  as a dummy variable being equal to one when the employee is a woman,  
13 and equal to zero otherwise. Under the assumption of a linear specification, the model that we  
14 seek to estimate is given by:  
15  
16

$$17 \quad q_{\theta}(w_{ij}|x_{ij}) = x_{ij}'\beta(\theta) + f_{ij}\gamma(\theta) \quad (1)$$

18 where  $q_{\theta}(w_{ij}|x_{ij})$  is the  $\theta^{\text{th}}$  conditional quantile of  $w_{ij}$ . In a quantile regression, the distribution  
19 of the error term is left unspecified (Koenker and Bassett, 1978). In (1), the set of coefficients  
20  $\beta(\theta)$  provides the estimated rates of return to the covariates at the  $\theta^{\text{th}}$  quantile of the log  
21 earnings distribution, and the coefficient  $\gamma(\theta)$  measures the intercept shift due to gender  
22 differences. Two comments are in order.  
23  
24

25 First, we begin by estimating the magnitude of the gender earnings gap on the whole  
26 sample which includes both male and female employees. However, when pooling the data, we  
27 implicitly assume that the returns to the labour market characteristics are the same at various  
28 quantiles for men and women. As this assumption is not necessarily satisfied, we will relax it  
29 latter on.  
30  
31

32 Second, the ECMOSS survey allows the structure of wages to be modelled while  
33 controlling firm-specific effects. With our matched data, we can deal with the firm heterogeneity  
34  
35

---

36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
<sup>9</sup> 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

1  
2  
3 by introducing firm characteristics into the earnings equation. Nevertheless, a difficulty is that we  
4 cannot model unobserved individual heterogeneity in the way of Abowd *et al.* (1999) as this is a  
5 cross-sectional data set. In order to temper the effect of firm heterogeneity, the natural attempt is  
6 to estimate firm fixed effects models including firm-specific dummies. Nevertheless, this  
7 technique seems to be futile in our case. Indeed, as we estimate quantile regressions, the large  
8 number of establishments (more than 10000) rules out the possibility of doing this<sup>10</sup>.  
9

10  
11 An alternative approach is described in Muller and Nordman (2004, 2005). It consists of  
12 summarising the main information on the firms' characteristics using a multivariate analysis and  
13 introducing the computed principal components (factors) stemming from this analysis into the  
14 earnings functions. Using factors may be seen as a further step with respect to those studies  
15 which have added mean firm variables into earnings functions, individual characteristics being  
16 controlled for. By contrast with firms' fixed effects that are introduced in wage regressions, the  
17 principal factors suggest qualitative characteristics of the firms. Specifically, we use a principal  
18 component analysis (PCA) to summarise the information about the surveyed establishments<sup>11</sup>.  
19

20  
21 This method is based on the calculation of the inertia axes for a cloud of points that  
22 represents the data in table format. There are different possible uses of factor analysis in this  
23 context. First, factor analyses can be used to elicit hidden characteristics correlated with  
24 observable characteristics. Second, PCA results could be used as a guide to replace these hidden  
25 firm characteristics with observable characteristics correlated with the main factors (as in Muller  
26 and Nordman, 2004). Third, and foremost in our case, the PCA is used as a substitute for firm  
27 fixed effect regressions. Indeed, the PCA allows us to investigate the determinants of the firm  
28 effects in our data. As long as the computed factors account for most of the firm heterogeneity  
29

30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

---

the large size of our sample, the results are only marginally modified.

<sup>10</sup> 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.

<sup>11</sup> 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.

1  
2  
3 bias, this approach allows us to obtain consistent estimates of the returns to worker  
4 characteristics and of the gender wage gap.  
5  
6

7  
8 For our purpose, the first ten inertia axes (the estimated factors which are linear  
9 components of all the firm's characteristics described in the previous section) concentrate a large  
10 proportion of the total variance of the original variables (about 40%) and reflect, therefore, a fair  
11 amount of the relevant information about the firm's characteristics. The correlation coefficients  
12 of the firms' characteristics with the first ten factors are used for the interpretation of the  
13 computed factors. The other factors represent a negligible amount of the statistical information  
14 and are dropped from the analysis.  
15  
16  
17  
18  
19  
20  
21  
22  
23

24 Further details on this rotated PCA can be found in Muller and Nordman (2005) and  
25 obtained from the authors upon request. Let us note that the ten factors are closely associated  
26 with the firms' sectoral belonging (factor 1), the various criteria used by employers for defining  
27 their implemented wage policy (factors 2, 4 and 9), their organisational features (hierarchical  
28 structure and supervision; factors 3, 6 and 7) and the firms' general features such as the state of  
29 business in 1992 (factors 5 and 10) and the firm size and training capacity (factor 8). The ten  
30 factors therefore reflect a wide range of firm characteristics and account for the existence of very  
31 different types of firms in the French private sector in 1992. These firms can mainly be described  
32 by their sector affiliation, size, organisational features (including firm wage policy) and human  
33 capital density. We are now ready to comment on our econometric analysis.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

#### 50 **4. Econometric results**

##### 51 *4.1. Quantile regressions estimates*

52  
53 Under the assumption that the returns to included labour market characteristics are the  
54 same for the two genders, the gender dummies in the quantile regressions are interpreted as the  
55 effects of gender on log earnings at the various percentiles once one controls for any differences  
56 in these labour market characteristics between genders. Estimates for the gender coefficient on  
57  
58  
59  
60

1  
2  
3 the pooled dataset are reported in Table 2 for various specifications, the full set of estimates  
4  
5 being reported in Table 3<sup>12</sup>.  
6

7  
8 *Insert Table 2 about here*  
9

10  
11 *Insert Table 3 about here*  
12

13 The first row presents a series of quantile regressions in which we condition the log  
14 earnings on gender at the 5<sup>th</sup>, 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, 90<sup>th</sup>, and 95<sup>th</sup> percentiles, without any control  
15 variable<sup>13</sup>. We notice that the observed log earnings gap increases as we move up the earnings  
16 distribution, with a sharp acceleration after the 75<sup>th</sup> percentile. For instance, at the 75<sup>th</sup> percentile,  
17 we see a raw gender earnings gap of slightly less than 25%. This means that the log earnings of a  
18 man at the 75<sup>th</sup> percentile of the male earnings distribution is a bit more than 22 points above the  
19 log earnings of a woman at the 75<sup>th</sup> percentile of the female earnings distribution.  
20  
21  
22  
23  
24  
25  
26  
27

28 Interestingly, very similar patterns have emerged in other European countries. First, male  
29 and female earnings are closer at the bottom of the earnings distribution. Second, male and  
30 female earnings are extremely unequal at the top of the distribution, up to a maximum difference  
31 of about 50%. Third, there is a steady increase in the gender log earnings gap as we move up in  
32 the earnings distribution. Fourth, there is a sharp acceleration in the increase in the gender log  
33 earnings gap starting at about the 75<sup>th</sup> or 80<sup>th</sup> percentile in the earnings distribution. Following  
34 Albrecht *et al.* (2003), De la Rica *et al.* (2005) or Arulampalam *et al.* (2004), we interpret this last  
35 feature of the gender log earnings gap by percentile as evidence of a glass ceiling.  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 Then, we examine various quantile estimates of the gender dummy coefficients when  
48 adding both male and female's labour market characteristics. Several specifications have been  
49 considered, the list of explanatory variables being further described in Table 2. In what follows,  
50 we only focus on the gender dummy coefficient which indicates the extent to which the gender  
51 earnings gap remains unexplained at the different quantiles after controlling for individual  
52 differences in various combinations of characteristics. We begin by introducing into the earnings  
53  
54  
55  
56  
57  
58  
59  
60

---

<sup>12</sup> In Table 3, we report the full set of estimates for the regression with firm factor effects.

1  
2  
3 equations the covariates commonly used in labour economics, i.e. education, potential experience,  
4 tenure, dummies for the matrimonial status, nationality, and the number of dependent children.  
5  
6 Then, we add job-specific variables such as the type of work contract, the workplace, the number  
7  
8 of hours worked per year, sector of employment, and occupation.  
9  
10

11  
12 When we control for education, experience off the current job, firm tenure, and other  
13 basic socio-economic characteristics (panel 2, Table 2), the gender dummies increase in absolute  
14 value relative to the raw gender dummy at the 5<sup>th</sup> and the 10<sup>th</sup> percentile, but then decrease from  
15 about the 20<sup>th</sup> through the 95<sup>th</sup> percentiles. The OLS gender dummy coefficient (at the mean) also  
16 diminishes. One explanation could be that in the first quartile of the log hourly earnings  
17 distribution, women display more labour market experience than men while this is not the case as  
18 for workers belonging to the second, third and fourth quartiles.  
19  
20  
21  
22  
23  
24  
25  
26  
27

28  
29 In the panel 3 of Table 2, we introduce the extended control variables which include basic  
30 control variables plus the type of work contract, the log of hours paid per year, and the location  
31 of the firm. The quantile estimates indicate that the gender dummy decreases in absolute value  
32 from the 5<sup>th</sup> to the median percentile as compared to the preceding model. Then, from about the  
33 75<sup>th</sup> to the 95<sup>th</sup>, however, the gender dummy increases. This might be a first indication that job-  
34 related characteristics (working conditions) do matter in explaining why the earnings gap is much  
35 greater at the upper tail of the earnings distribution<sup>14</sup>.  
36  
37  
38  
39  
40  
41  
42  
43  
44

45 We next present the estimated gender dummy coefficients after adding 12 sectoral  
46 dummies in the quantile regressions (panel 4). The same picture emerges from these estimates,  
47 and the gender dummy is reduced only minimally at the bottom of the earnings distribution and  
48 slightly increases from about the 20<sup>th</sup> percentile and more substantially at the top of the earnings  
49 distribution. Of course, the sector of employment is to some extent an endogenous characteristic  
50 since the choice of sector in which to work is typically made after education is completed.  
51  
52  
53  
54  
55  
56  
57  
58  
59

---

60 <sup>13</sup> The coefficient estimates for the gender dummy in this panel are identical to the statistical log earnings gaps.



1  
2  
3 In panel 5, we account for the firms' computed factors stemming from the factor analysis.  
4  
5 In so doing, our aim is to substitute a firm fixed-effect regression by a "firm factor effect"  
6  
7 regression that may account for qualitative aspects of firms' wage policies, human capital and  
8  
9 organisational features. In a sense, following Muller and Nordman (2004, 2005), we generalise the  
10  
11 approach developed in Cardoso (1999) who regresses the firms' fixed effects on different  
12  
13 variables. For our purpose, the first ten computed factors concentrate most of the relevant  
14  
15 information about the firms' characteristics. A Wald test rejects at the 1 percent level the null  
16  
17 hypothesis that the coefficients of these ten factors are jointly equal to zero. Since the covariates  
18  
19 introduced in the factor analysis include the sectoral dummies, we omit these explanatory  
20  
21 variables in the earnings functions.  
22  
23  
24  
25

26 For the sake of comparison, we have also performed a linear regression with firms' fixed  
27  
28 effects. The female dummy coefficient in panel 3 (-0.184) can be compared with the one  
29  
30 estimated at the mean with the firm factor effect, which is equal to -0.177 (panel 5). So, the  
31  
32 female coefficient is slightly reduced as we move from sector fixed effects or firm fixed effects  
33  
34 models towards a firm factor effects specification. It may be that the computed factors stemming  
35  
36 from our PCA of firms' characteristics add a qualitative aspect of the firms' wage policy to our  
37  
38 regressions that fixed effects models, either with sector or firm dummies, may not be able to  
39  
40 totally control for.  
41  
42  
43  
44

45 According to the quantile estimates reported in panel 5, we find that taking into account  
46  
47 the firms' factors slightly increases the gender dummy coefficient in absolute value at the lower  
48  
49 tail of the conditional log earnings distribution (from the 5<sup>th</sup> to the median percentile).  
50  
51 Conversely, the coefficient is significantly reduced at the upper tail of the distribution, especially  
52  
53 above the 75<sup>th</sup> percentile. Now, the gender earnings gap amounts to about 27% at the 95<sup>th</sup>  
54  
55 percentile while it amounted to more than 30% with the extended and sectoral control variables.  
56  
57  
58  
59

---

60 <sup>14</sup> 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.

1  
2  
3 With respect to the existing literature, our results show that controlling for firms' characteristics is  
4 likely to have a reducing effect on the extent of the glass ceiling phenomenon, albeit moderately<sup>15</sup>.  
5  
6

7  
8 It is of interest to compare the magnitude of the gender wage gap obtained respectively  
9 with inclusion of firms' factors (derived from the PCA) and firms' characteristics. If we find that  
10 the various estimates lead to very similar results, this would be the sign that the PCA method may  
11 be useful to account for the firm's environment with a minimum number of factors, thereby  
12 avoiding potential problems of multicollinearity when multiple firm-specific variables are  
13 controlled for in wage regressions. Indeed, by definition, the main components derived from the  
14 factor analysis are poorly correlated and also have the advantage to sum up the main statistical  
15 information of the firm level variables.  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

26 According to the different estimates described in Table 2, we observe that the results  
27 from these two specifications are very similar. On the one hand, the coefficients associated with  
28 the gender dummy are increasing (in absolute value) along the conditional earnings distribution,  
29 especially at the 75<sup>th</sup> percentile and above. On the other hand, once observed firms' characteristics  
30 are controlled for, we note that the magnitude of the coefficient on the gender variable is slightly  
31 lower with respect to the regression with firms' factors. At the top of the distribution, the gender  
32 earnings gap is equal to 24.9% with firms' characteristics, 27.1% with firms' factors, while it  
33 amounts to 30.4% with extended controls and sectoral dummies.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

45 Finally, panels 6 and 7 of Table 2 present the quantile log earnings regression estimates  
46 adding 29 occupational dummies. We present these estimates separately because there is no clear  
47 consensus as to whether occupation (and to some extent industry) should be taken into account  
48 to assess the extent of the gender wage gap. If employers differentiate between men and women  
49 through their tendency to hire into certain occupations, then occupational assignment is an  
50 outcome of employer practices rather than an outcome of individual choice or productivity  
51  
52  
53  
54  
55  
56  
57  
58  
59

---

60 <sup>15</sup> 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.

1  
2  
3 differences<sup>16</sup>. While panel 6 presents the gender dummy coefficient of a sector fixed effect model,  
4  
5 panel 7 accounts for the coefficients of a firm factor effect model. Again, both sets of estimates  
6  
7 are very close.  
8  
9

10 As might be expected, controlling for occupation considerably reduces the gender gap  
11  
12 throughout the conditional earnings distribution. In panel 7, the unexplained gender gap falls to  
13  
14 8.8% at the 5<sup>th</sup> percentile and, more importantly, to 22% at the 95<sup>th</sup> percentile (compared to 9.4%  
15  
16 and 31.7% in panel 5). We would argue that the effect of controlling for occupation on the  
17  
18 gender earnings gap reflects the occupational segregation that may be present in France.  
19  
20 However, we also note that if the gender earnings gap varies significantly at the upper tail of the  
21  
22 conditional earnings distribution from panel 1 to panel 7, it remains remarkably stable at the  
23  
24 bottom 5<sup>th</sup> or 10<sup>th</sup> percentiles.  
25  
26  
27  
28  
29  
30

#### 31 *4.2. Quantile regressions by gender*

32

33 In the previous section, we have estimated the magnitude of the gender earnings gap  
34  
35 conditional on the characteristics of the pooled sample of male and female workers at different  
36  
37 points of the earnings distribution, thereby implicitly assuming that the returns to those  
38  
39 characteristics were the same at various quantiles for men and women. However, this assumption  
40  
41 seems *a priori* unrealistic. We now test its relevance by introducing into the earnings functions a  
42  
43 set of the same covariates crossed with the gender dummy. Estimates of these quantile  
44  
45 regressions are shown in Table 4, which can be compared to pooled estimates in Table 3.  
46  
47  
48  
49

50 *Insert Table 4 about here*

51

52 We use a Wald test to assess the joint significativeness of the crossed variables all along  
53  
54 the conditional earnings distribution. The values of the statistics at each considered percentile  
55  
56 reveal that we have to reject the hypothesis of joint nullity of the crossed variables at the 1%  
57  
58  
59

---

60 <sup>16</sup> 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).

1  
2  
3 level. This rejection of the pooling assumption therefore implies that gender specific equations  
4  
5 (i.e. for men and for women) have to be estimated instead. More specifically, Table 4 is indicative  
6  
7 of which return of the introduced labour market characteristics significantly differs across  
8  
9 genders at each considered quantile.  
10

11  
12 We notice that almost all returns vary across the sexes whatever the workers' relative  
13  
14 position in the conditional distribution. This is all the more true for the returns to human capital,  
15  
16 i.e. schooling and experience, which are always significantly lower for women. The only exception  
17  
18 is for the return to tenure at the midpoint of the distribution (50<sup>th</sup> percentile). Interestingly, some  
19  
20 demographic characteristics, such as the number of dependent children or a dummy for being  
21  
22 widowed (the reference is being married), have differentiated impacts between men and women  
23  
24 depending on their relative position in the earnings distribution. For instance, while the wage  
25  
26 premium for the number of children is not significantly different between men and women in the  
27  
28 lower tail of the conditional distribution (percentiles 5<sup>th</sup> to 50<sup>th</sup>), this premium differs increasingly  
29  
30 across gender as we move up in the distribution.  
31  
32  
33  
34

35  
36 *Insert Table 5 about here*  
37

38 Differentiated regressions by gender are then displayed in Table 5. Let us briefly describe  
39  
40 the influence of a few explanatory variables. As previously noticed, the returns to human capital  
41  
42 are always lower for women. For instance, the return to schooling increases, respectively for men  
43  
44 and women, from 4.2% versus 3.3% at the 5<sup>th</sup> percentile to 13.4% versus 12.5% at the 95<sup>th</sup>  
45  
46 percentile. Also, while the marginal return to potential experience off the current job is  
47  
48 decreasing for both genders (the quadratic term is negative), it diminishes more rapidly for males  
49  
50 than for females. Other results are worth noting. First, being divorced is detrimental to men, but  
51  
52 not to women. Second, the wage premium for working in the Paris area is higher for men in the  
53  
54 lower tail of the distribution but, then, in the upper tail, the reverse is true: women have certainly  
55  
56 access to better jobs there. These results provide therefore incentives to perform earnings  
57  
58 decompositions at quantiles instead of only looking at decompositions at the sample mean.  
59  
60

### 4.3. Quantile decomposition

After having analysed the extent to which returns to exogenous factors differ between men and women, we now perform a quantile decomposition of the gender gap. As in Albrecht *et al.* (2003) and following the recent approach described in Machado and Mata (2005), we decompose the difference between the male and female log earnings distributions into two components<sup>17</sup>. The first one is due to differences in labour market characteristics between male and female employees. The second one is due to differences in the rewards that both men and women receive for their (observable) characteristics.

Instead of relying on the Oaxaca-Blinder decomposition technique whose purpose is to identify the sources of differences between the means of two distributions, we implement the decomposition at each quantile of the earnings distribution. Let us briefly describe the approach developed by Machado and Mata (2005). For the presentation, let  $\beta^m$  and  $\beta^f$  denote respectively men and women's returns to labour market characteristics  $x^m$  and  $x^f$  respectively. The decomposition of the difference between the male and female earnings densities is:

$$x^m \beta^m(\theta) - x^f \beta^f(\theta) = (x^m - x^f) \beta^f(\theta) + x^m (\beta^m(\theta) - \beta^f(\theta)) \quad (2)$$

In the above equation, the first term on the right hand side indicates the magnitude of the gap which is due to dissimilarities in labour market characteristics. The second term indicates the magnitude of the gap which is due to differences in the rewards to these characteristics. As two counterfactual densities may be constructed, we choose to generate the density that would arise if women were endowed on the basis of men's labour market characteristics and paid like women.

To construct the counterfactual density, we rely on the three following steps. First, we draw a sample of 150 numbers from a standard uniform distribution. Second, using these different numbers  $\theta_j$  ( $j=1, \dots, 150$ ), we estimate the quantile regressions coefficient vectors

---

<sup>17</sup> For further discussion on decomposition methods and the gender wage gap, see Silber and Weber (1999).

1  
2  
3  
4  $\beta^f(\theta_j)$  for the various  $j$  using the female subsample. Third, for each value of  $j$ , we take a  
5  
6 draw with replacement from the male data set and generate the predicted wage  $x^m \beta^f(\theta_j)$ . Two  
7  
8 additional comments are in order. First, in order to get standard errors for the counterfactual  
9  
10 density, we have replicated the whole procedure exactly 40 times. Second, as the procedure is  
11  
12 excessively time-consuming with very large samples, we perform the quantile decomposition on  
13  
14 the basis of a random sample of 34 303 employees (i.e. the sample rate is 1 employee out of 4).  
15  
16

17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
*Insert Table 6 about here*

When applying the quantile decomposition, we again wonder whether the inclusion of firms' factors or characteristics may affect the underlying conclusions. Specifically, we provide the decomposition results for three specifications, i.e. with no firms variables, with firms' factors, and with firms' characteristics. Results from the decompositions are given in Table 6.

In the first, third and fifth columns, we report the observed gender gaps at respectively the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentile of the distribution for the various specifications. The gender gaps reported in columns 2, 4 and 6 are constructed using women's returns only, and then assuming that these women have the male distribution of labour market characteristics. In any cases, we observe that the gender gap strongly increases throughout the earnings distribution. Interestingly, the gap due to differences in the returns to observable characteristics is of very similar magnitude independently on the inclusion of firm factors or firm characteristics.

Our results are in fact very similar to those obtained in Sweden by Albrecht *et al.* (2003). We clearly observe that gender differences in labour market characteristics do not really explain the larger gap observed at the top of the distribution. Instead, the gender gap is mainly due to the differential rewards that women receive for their own characteristics. It is clear from the results of Table 6 that women would receive much lower wages were they paid as women while being endowed with male characteristics. Hence, in France, our main finding is that the glass ceiling

1  
2  
3 effect is mainly due to differences in returns to observed labour market characteristics across  
4  
5 genders at the top of the distribution rather than to differences in those characteristics.  
6  
7  
8  
9

## 10 **5. Discussion and concluding comments**

11  
12 In this paper, we have brought empirical evidence on the glass ceiling effect according to  
13  
14 which the gender wage gap is more important at the upper tail of the wage distribution. While  
15  
16 several studies have recently shed light on this phenomenon in European countries (Albrecht *et*  
17  
18 *al.*, 2003; De la Rica *et al.*, 2005; Arulampalam *et al.*, 2004), our contribution is the first one to  
19  
20 account for both characteristics of employers and employees. It is only recently that researchers  
21  
22 have acknowledged that the use of matched employer-employee data for studying labour market  
23  
24 discrimination can deepen the understanding of sex segregation in the professional environment  
25  
26 (Hellerstein and Neumark, 2005). Interestingly, while the role of the firm characteristics has been  
27  
28 neglected so far in studies dealing with the gender gap, we believe that the work environment is  
29  
30 likely to affect the magnitude of the gender wage gap along the earnings distribution.  
31  
32  
33  
34

35  
36 Specifically, we assess the relevance of the glass ceiling hypothesis using the French 1992  
37  
38 ECMOSS data, which provides rich information on about 130,000 employees and 14,000  
39  
40 employers. Econometric results from quantile regressions show that there exists a significant glass  
41  
42 ceiling effect in France. While male and female earnings are close at the bottom of the income  
43  
44 distribution, there is a strong increase in the gender earnings gap above the 75<sup>th</sup> percentile of this  
45  
46 distribution. In order to control for the firms' characteristics, we follow Muller and Nordman  
47  
48 (2004, 2005) and rely on a principal component analysis to extract the most influential factors of  
49  
50 the surveyed establishments. This approach allows accounting for qualitative aspects of the firms  
51  
52 including their implemented wage policy, which may not be the case when one controls for  
53  
54 unobserved heterogeneity through firm fixed effect models.  
55  
56  
57  
58

59  
60 According to the French data, the gender earnings gap would be overstated at the top of  
the distribution if the influence of firms' characteristics were omitted. Hence, accounting for job-

1  
2  
3 related characteristics as well as for characteristics of the workers' environment has a reducing  
4 effect on the extent of the glass ceiling phenomenon. The reduction of the observed gender wage  
5 gap once firm characteristics are controlled for may be understood in a context of sorting among  
6 firms, where productive firms primarily form relationships with productive workers. In this  
7 setting, the gender wage gap may arise as a result of sorting of male and female workers across  
8 firms that pay different wages. If this hypothesis is true, accounting for the firm characteristics  
9 should reduce the magnitude of the gender wage gap, if not clear it up totally. In other words, the  
10 gender gap may be non-existent within firms if the gender pay gap is only linked to sorting.  
11 However, this does not provide any convincing argument as for the persistence of larger wage  
12 gaps at the upper tail of the wage distribution.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

26 Other forces that would rationalize the existence of the glass ceiling effect may also be at  
27 work. Despite of the reduction in the earnings gap for the most paid workers, there is still a large  
28 and significant difference between the male and female earnings. Furthermore, our counterfactual  
29 decomposition shows that the glass ceiling effect is mostly due to differences in the returns to  
30 labour market characteristics across genders at the top of the distribution rather than to average  
31 differences in these characteristics. Hence, our results are very similar to those reported in recent  
32 studies conducted in Europe. As there is now robust evidence on the magnitude of the glass  
33 ceiling effect in industrialised countries, it would be worthwhile to deepen the understanding of  
34 the origins and causes of this stylised fact. We leave this issue for future research.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## References

- Abowd, J.M. and Kramarz, F. (1999) The analysis of labor markets using matched employer-employee data, in *Handbook of Labor Economics* (Eds) O.C. Ashenfelter and D. Card, vol. 3B, Elsevier, North-Holland, pp. 2629-2710.
- Abowd, J.M., Kramarz, F. and Margolis, D. N. (1999) High-wage workers and high-wage firms, *Econometrica*, **67**, 251-333.
- Abowd, J., Kramarz, F., Margolis, D.N. and 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*, **15**, 419-436.
- Albrecht, J., Björklund, A. and Vroman, S. (2003) Is there a glass ceiling in Sweden?, *Journal of Labor Economics*, **21**, 145-177.
- Albrecht, J., van Vuuren, A. and Vroman, S. (2006) Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands, *mimeo*, University of Georgetown.
- Altonji, J.G. and Blank, R.M. (1999) Race and gender in the labor market, in *Handbook of Labor Economics* (Eds) O.C. Ashenfelter and D. Card, vol. 3C, Elsevier, North-Holland, pp. 3143-3258.
- Arulampalam, W., Booth, A.L. and Bryan, M.L. (2004) Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution, *mimeo*, University of Auckland.
- Bayard, K., Hellerstein, J., Neumark, D. and Troske, K. (2003) New evidence on sex segregation and sex differences in wages from matched employer-employee data, *Journal of Labor Economics*, **21**, 887-922.
- Blau, F. and Kahn, L. (2000) Gender differences in pay, *Journal of Economic Perspectives*, **14**, 75-99.
- Booth, A., Francesconi, M. and Frank, J. (2003) A sticky floors model of promotion, pay and gender, *European Economic Review*, **47**, 295-322.
- Buchinsky, M. (1998) Recent advances in quantile regression models: A practical guideline for empirical research, *Journal of Human Resources*, **33**, 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*, **53**, 87-102.
- Caroli, E., Greenan, N. and Guellec, D. (2001) Organizational change and skill accumulation, *Industrial and Corporate Change*, **10**, 481-506.
- De la Rica, S., Dolado, J.J. and 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*, **158**, 89-104.
- Destré, G. and Nordman, C. (2002) The impacts of informal training on earnings: Evidence from French, Moroccan and Tunisian employer-employee matched data, *L'Actualité Économique*, **78**, 179-205.
- Greenan, N. (2003), Organisational change, technology, employment and skills: An empirical study of French manufacturing, *Cambridge Journal of Economics*, **27**, 287-317.
- Hahn, J. (1995) Bootstrapping quantile regression estimators, *Econometric Theory*, **11**, 105-121.

- 1  
2  
3 Hellerstein, J. and Neumark, D. (2005) Using matched employer-employee data to study labor  
4 market discrimination, *IZA Discussion Paper*, n°1555.  
5  
6 Hinks, T. (2002) Gender wage differentials and discrimination in the New South Africa, *Applied*  
7 *Economics*, **34**, 2043-2052.  
8  
9 Ichino, A. and Filippin, A. (2005) Gender wage gap in expectations and realizations, *Labour*  
10 *Economics*, **12**, 125-145.  
11  
12 Jellal, M., Nordman, C. and Wolff, F-C. (2005) Gender wage gap under uncertainty, *mimeo*, DIAL,  
13 Paris.  
14  
15 Koenker, R. (2005) *Quantile Regression*, Econometric Society Monograph Series, Cambridge  
16 University Press.  
17  
18 Koenker, R. and Bassett, G. (1978) Regression quantiles, *Econometrica*, **46**, 33-50.  
19  
20 Lindbeck, A. and Snower, D.J. (2000) Multitask learning and the reorganization of work: From  
21 tayloristic to holistic organization, *Journal of Labor Economics*, **18**, 353-376.  
22  
23 Machado, J.A. and Mata, J. (2005) Counterfactual decomposition of changes in wage distributions  
24 using quantile regression, *Journal of Applied Econometrics*, **20**, 445-465.  
25  
26 Meng, X. (2004) Gender earnings gap: The role of firm specific effects, *Labour Economics*, **11**, 555-  
27 573.  
28  
29 Meng, X. and Meurs, D. (2004) The gender earnings gap: Effects of institutions and firms – A  
30 comparative study of French and Australian private firms, *Oxford Economic Papers*, **56**, 189-  
31 208.  
32  
33 Muller, C. and Nordman, C. (2004) Which human capital matters for rich and poor's wages?  
34 Evidence from matched worker-firm data from Tunisia, *mimeo*, CREDIT Research Paper  
35 04/08, University of Nottingham.  
36  
37 Muller, C. and Nordman, C. (2005) Firm effects or firm human capital? An investigation of  
38 wages from French matched worker-firm data, *mimeo*, University of Alicante.  
39  
40 Nielsen, H. S. (2000) Wage discrimination in Zambia: An extension of the Oaxaca-Blinder  
41 decomposition, *Applied Economics Letters*, **7**, 405-408.  
42  
43 Sakellariou, C. (2004) The use of quantile regressions in estimating gender wage differentials: A  
44 case study of the Philippines, *Applied Economics*, **34**, 1001-1007.  
45  
46 Silber, J. and Weber, M. (1999) Labour market discrimination: Are there significant differences  
47 between the various decomposition procedures?, *Applied Economics*, **31**, 359-365.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**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		

Survey INSEE ECMOSS 1992.

The size of the sample is 137 211 employees, working in 14 693 establishments.

**Table 2. Gender dummy coefficients using alternative quantile earnings models**  
(Dependent variable: log hourly earnings)

Specification	Quantile regressions (percentage of the conditional 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 ( $n= 137,211$ )	-0.076*** (42.44)	-0.097*** (49.37)	-0.143*** (68.15)	-0.173*** (66.30)	-0.222*** (62.84)	-0.330*** (62.74)	-0.410*** (50.74)	-0.191*** (83.16)	-0.208*** (93.11)
(2) Basic control variables <sup>a</sup> ( $n= 137,211$ )	-0.086*** (38.45)	-0.101*** (52.09)	-0.126*** (65.40)	-0.159*** (74.17)	-0.195*** (71.41)	-0.227*** (56.97)	-0.245*** (39.81)	-0.178*** (92.42)	-0.183*** (95.89)
(3) Extended control variables <sup>b</sup> ( $n= 137,211$ )	-0.085*** (40.34)	-0.098*** (51.72)	-0.122*** (64.29)	-0.151*** (81.95)	-0.205*** (85.74)	-0.256*** (68.98)	-0.284*** (54.26)	-0.180*** (97.16)	-0.184*** (96.33)
(4) Extended control variables plus 12 sectoral dummies ( $n= 137,211$ )	-0.083*** (36.93)	-0.096*** (47.99)	-0.124*** (63.86)	-0.153*** (75.42)	-0.209*** (83.71)	-0.271*** (62.46)	-0.304*** (52.38)	-0.188*** (96.53)	-
(5) Extended control variables plus 10 firms' factor effects <sup>c</sup> ( $n= 137,169$ )	-0.085*** (38.68)	-0.100*** (48.53)	-0.123*** (66.49)	-0.149*** (75.11)	-0.197*** (78.66)	-0.243*** (65.79)	-0.271*** (53.46)	-0.177*** (95.33)	-
(6) Extended control variables plus firms' characteristics <sup>d</sup> ( $n= 137,169$ )	-0.081*** (37.16)	-0.095*** (46.71)	-0.118*** (59.82)	-0.146*** (77.67)	-0.185*** (75.93)	-0.230*** (63.58)	-0.249*** (48.86)	-0.170*** (91.23)	-
(7) Extended control variables plus 12 sectoral and 29 occupational dummies ( $n= 137,211$ )	-0.075*** (25.18)	-0.081*** (33.67)	-0.097*** (48.93)	-0.115*** (66.40)	-0.138*** (65.81)	-0.169*** (56.83)	-0.195*** (47.45)	-0.127*** (73.79)	-
(8) Extended control variables, firms' factor effects and 29 occupational dummies ( $n=137,169$ )	-0.084*** (26.42)	-0.090*** (38.34)	-0.101*** (56.73)	-0.119*** (66.02)	-0.140*** (60.51)	-0.173*** (61.59)	-0.194*** (48.47)	-0.130*** (76.29)	-

Survey INSEE ECMOSS 1992.

Absolute values of t statistics 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 a4, a44, s1, s2, s3, s4, s5, s6, s7, s8, s9, s10, s11, s12, 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.

*d*: the firm characteristics introduced in the regressions are s1, s2, s3, s4, s5, s6, s7, s8, s9, s10, s11, T2, T3, T4, T5, T6, FF2, FF3, FF4, VA1, VA2, D21, D32, D4A1, D4B1, D4B2, PS1, D151, D19B1, D19B2, D19B3, D252, D253, D26A1, D26B1, D281, D301, D311, D321, D331, D332, D341, D342, D331, D332, D3513, D3523, D3533, D3543, D3553, D3563, D3573, D3583, D39B1, D411.

**Table 3. Quantile earnings functions with firms' factor effects, pooled regressions**  
(Dependent variable: Log hourly earnings)

Variables	Quantile regressions							OLS mean
	0.05	0.10	0.25	0.5	0.75	0.90	0.95	
Constant	2.829*** (186.02)	2.729*** (191.42)	2.515*** (191.00)	2.432*** (165.86)	2.549*** (134.37)	2.974*** (104.69)	3.362*** (85.07)	2.524*** (184.04)
Dummy for female	-0.085*** (38.68)	-0.100*** (48.53)	-0.123*** (66.49)	-0.149*** (75.11)	-0.197*** (78.65)	-0.243*** (65.79)	-0.271*** (53.46)	-0.177*** (95.33)
Years of schooling	0.043*** (84.82)	0.054*** (110.52)	0.077*** (159.48)	0.101*** (167.29)	0.121*** (143.35)	0.131*** (103.53)	0.133*** (78.06)	0.107*** (190.18)
Years of experience off the current firm	0.002*** (5.87)	0.003*** (8.76)	0.007*** (21.04)	0.013*** (37.06)	0.020*** (47.58)	0.030*** (47.03)	0.034*** (39.80)	0.015*** (46.51)
(Years of experience off the current firm) <sup>2</sup>	-0.008*** (6.98)	-0.010*** (9.38)	-0.017*** (17.49)	-0.027*** (25.23)	-0.038*** (27.33)	-0.049*** (23.69)	-0.051*** (17.77)	-0.028*** (27.87)
Years of tenure in the current firm	0.015*** (35.90)	0.018*** (45.28)	0.022*** (62.38)	0.025*** (67.81)	0.027*** (56.98)	0.030*** (41.80)	0.032*** (32.28)	0.024*** (69.35)
(Years of tenure in the current firm) <sup>2</sup>	-0.017*** (12.43)	-0.018*** (14.14)	-0.019*** (16.65)	-0.019*** (16.04)	-0.015*** (9.96)	-0.016*** (6.77)	-0.019*** (5.61)	-0.017*** (15.20)
Dummy for non-French	-0.104*** (22.87)	-0.122*** (28.59)	-0.150*** (38.33)	-0.178*** (41.50)	-0.194*** (35.26)	-0.192*** (23.41)	-0.149*** (13.16)	-0.191*** (47.67)
Dummy for single	-0.029*** (9.66)	-0.031*** (11.17)	-0.034*** (13.83)	-0.034*** (12.65)	-0.039*** (11.63)	-0.037*** (7.43)	-0.039*** (5.79)	-0.044*** (17.65)
Dummy for widowed	-0.023** (2.35)	-0.029*** (3.24)	-0.044*** (5.39)	-0.047*** (5.30)	-0.061*** (5.44)	-0.085*** (5.11)	-0.117*** (5.11)	-0.055*** (6.55)
Dummy for divorced	-0.005 (1.15)	-0.002 (0.54)	0.004 (1.07)	-0.003 (0.75)	-0.006 (1.09)	-0.021*** (2.74)	-0.021* (1.96)	-0.005 (1.40)
Number of dependent children	0.013*** (11.14)	0.014*** (13.00)	0.016*** (17.73)	0.021*** (22.52)	0.020*** (17.78)	0.022*** (13.40)	0.021*** (9.51)	0.021*** (23.64)
Dummy for workplace (1: Paris)	0.119*** (43.71)	0.151*** (59.14)	0.196*** (84.78)	0.220*** (88.49)	0.252*** (80.57)	0.286*** (61.72)	0.304*** (47.76)	0.230*** (99.16)
Dummy for type of contract (CDD: « <i>contrat à durée déterminée</i> »)	-0.006 (1.33)	-0.007* (1.66)	-0.006 (1.41)	-0.021*** (5.18)	-0.040*** (7.71)	-0.068*** (8.73)	-0.088*** (8.06)	-0.039*** (10.03)
Log of hours paid work per year	0.027*** (14.45)	0.028*** (16.56)	0.028*** (18.43)	0.015*** (8.92)	-0.015*** (7.27)	-0.070*** (21.79)	-0.112*** (24.72)	-0.004** (2.56)
Control for firms' factors	yes	yes	Yes	yes	yes	yes	yes	yes
Observations	137169	137169	137169	137169	137169	137169	137169	137169
Pseudo R <sup>2</sup> / R <sup>2</sup>	0.135	0.168	0.217	0.255	0.280	0.301	0.306	0.416

Survey INSEE ECMOSS 1992.

Absolute values of t statistics are in parentheses. \*\*\*, \*\* and \* mean respectively significant at the 1%, 5% and 10% levels.

**Table 4. Quantile earnings functions with firms' factor effects, pooled regressions and crossed effects**  
(Dependent variable: Log hourly earnings)

Variables	Quantile regressions														OLS mean	
	0.05		0.10		0.25		0.5		0.75		0.90		0.95		mean	
	X	X*Female	X	X*Female	X	X*Female	X	X*Female	X	X*Female	X	X*Female	X	X*Female	X	X*Female
Constant	2.786** (120.10)		2.717** (132.36)		2.573** (132.18)		2.499** (130.25)		2.603** (96.95)		2.966** (72.10)		3.309** (58.05)		2.566** (135.96)	
Dummy for female	0.155** (4.88)		0.140** (4.93)		0.130** (4.75)		0.097** (3.51)		0.072* (1.86)		0.077 (1.30)		0.114 (1.39)		0.132** (4.87)	
Years of schooling	0.043** (58.88)	-0.011** (9.50)	0.055** (82.49)	-0.016** (14.96)	0.077** (113.35)	-0.020** (19.33)	0.101** (136.15)	-0.024** (20.64)	0.124** (111.51)	-0.028** (16.37)	0.132** (80.27)	-0.019** (7.18)	0.135** (60.65)	-0.014** (3.76)	0.107** (145.94)	-0.024** (20.88)
Years of experience off the current firm	0.004** (6.74)	-0.001* (1.67)	0.005** (9.86)	-0.001** (2.00)	0.009** (19.99)	-0.004** (5.30)	0.016** (36.90)	-0.006** (8.99)	0.025** (43.20)	-0.009** (9.88)	0.034** (40.27)	-0.010** (7.13)	0.039** (32.80)	-0.008** (4.54)	0.017** (41.94)	-0.005** (7.98)
(Years of experience off the current firm) <sup>2</sup>	-0.010** (6.53)	0.004 (1.64)	-0.012** (8.44)	0.003 (1.37)	-0.020** (14.70)	0.007** (3.28)	-0.031** (23.17)	0.009** (4.34)	-0.042** (22.62)	0.008** (2.78)	-0.049** (17.67)	-0.002 (0.42)	-0.048** (12.53)	-0.012* (1.87)	-0.030** (22.71)	0.003 (1.50)
Years of tenure in the current firm	0.018** (32.13)	-0.005** (5.28)	0.020** (39.30)	-0.005** (5.82)	0.022** (46.40)	-0.002** (3.08)	0.024** (53.96)	-0.001* (1.95)	0.028** (44.86)	-0.004** (4.25)	0.031** (32.89)	-0.005** (3.30)	0.033** (25.37)	-0.005** (2.40)	0.024** (54.80)	-0.002** (3.16)
(Years of tenure in the current firm) <sup>2</sup>	-0.024** (13.93)	0.008** (2.76)	-0.025** (15.99)	0.012** (4.44)	-0.021** (14.52)	0.005** (1.98)	-0.019** (13.55)	0.001 (0.39)	-0.020** (10.17)	0.008** (2.29)	-0.021** (6.87)	0.011** (2.05)	-0.022** (5.15)	0.008 (1.06)	-0.019** (13.94)	0.004* (1.74)
Dummy for non-French	-0.115** (20.65)	0.033** (3.17)	-0.134** (26.63)	0.047** (5.03)	-0.154** (31.91)	0.052** (5.76)	-0.179** (37.39)	0.047** (5.34)	-0.200** (29.89)	0.062** (5.09)	-0.207** (20.21)	0.077** (4.17)	-0.180** (12.86)	0.069** (2.71)	-0.201** (42.81)	0.068** (7.86)
Dummy for single	-0.056** (13.46)	0.061** (9.83)	-0.064** (17.40)	0.069** (12.60)	-0.071** (20.68)	0.080** (15.37)	-0.078** (23.59)	0.093** (18.34)	-0.083** (18.61)	0.095** (13.57)	-0.088** (13.29)	0.096** (9.09)	-0.089** (9.93)	0.098** (6.76)	-0.088** (27.11)	0.101** (20.19)
Dummy for widowed	-0.041** (2.15)	0.032 (1.43)	-0.038** (2.22)	0.025 (1.22)	0.012 (0.74)	-0.056** (2.90)	0.015 (0.93)	-0.052** (2.76)	0.010 (0.46)	-0.054** (2.10)	-0.009 (0.26)	-0.062 (1.61)	0.012 (0.28)	-0.100* (1.91)	0.006 (0.40)	-0.053** (2.86)
Dummy for divorced	-0.023** (3.38)	0.035** (3.68)	-0.029** (4.76)	0.044** (5.25)	-0.022** (3.72)	0.046** (5.68)	-0.024** (4.14)	0.047** (5.92)	-0.011 (1.42)	0.030** (2.78)	-0.022* (1.85)	0.030* (1.82)	0.001 (0.07)	-0.008 (0.35)	-0.024** (4.21)	0.044** (5.74)
Number of dependent children	0.010** (7.00)	-0.003 (1.08)	0.011** (8.76)	-0.002 (0.81)	0.012** (10.44)	0.000 (0.03)	0.016** (15.01)	-0.001 (0.29)	0.017** (12.57)	-0.004* (1.82)	0.020** (9.89)	-0.012** (3.19)	0.023** (8.46)	-0.014** (2.79)	0.017** (16.33)	-0.004** (2.04)
Dummy for workplace (1: Paris)	0.133** (36.95)	-0.028** (4.86)	0.157** (48.08)	-0.030** (5.85)	0.179** (57.99)	-0.005 (1.02)	0.202** (66.96)	0.014** (3.02)	0.229** (55.36)	0.034** (5.29)	0.271** (43.45)	0.027** (2.78)	0.295** (34.49)	0.029** (2.17)	0.221** (74.39)	0.006 (1.21)
Dummy for type of contract (CDD: « <i>contrat à durée déterminée</i> »)	-0.023** (3.17)	-0.010 (1.01)	-0.035** (5.59)	0.011 (1.26)	-0.040** (6.99)	0.031** (3.86)	-0.061** (11.21)	0.046** (5.90)	-0.091** (12.17)	0.062** (5.80)	-0.109** (9.39)	0.048** (2.94)	-0.130** (8.11)	0.044** (1.96)	-0.075** (13.92)	0.041** (5.40)
Log of hours paid work per year	0.032** (11.40)	-0.013** (3.30)	0.028** (11.35)	-0.005 (1.59)	0.022** (9.77)	-0.000 (0.13)	0.006** (2.64)	0.007** (2.31)	-0.030** (10.04)	0.017** (4.02)	-0.076** (15.94)	-0.002 (0.37)	-0.114** (17.14)	-0.018** (1.98)	-0.009** (4.43)	0.001 (0.37)
Control for firms' factors	yes		yes		yes		yes		yes		yes		yes		yes	
Wald test: value, prob.	34.00; 0.000		45.18; 0.000		51.86; 0.000		69.56; 0.000		63.83; 0.000		47.41; 0.000		35.99; 0.000		90.30; 0.000	
Number of observations	137169		137169		137169		137169		137169		137169		137169		137169	
Pseudo R <sup>2</sup> / R <sup>2</sup>	0.154		0.189		0.239		0.272		0.293		0.309		0.311		0.435	

Survey INSEE ECMOSS 1992.

Absolute values of t statistics are in parentheses. \*\*\*, \*\* and \* mean respectively significant at the 1%, 5% and 10% levels.

Table 5. Quantile earnings functions with firms' factor effects, by gender

(Dependent variable: Log hourly earnings)

Variables	Quantile regressions														OLS mean	
	0.05		0.10		0.25		0.5		0.75		0.90		0.95		Male	Female
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Constant	2.747*** (114.01)	2.880*** (176.95)	2.626*** (119.03)	2.762*** (181.88)	2.442*** (121.66)	2.542*** (144.31)	2.387*** (106.17)	2.396*** (129.28)	2.510*** (97.24)	2.536*** (104.09)	2.921*** (69.33)	2.951*** (77.77)	3.319*** (58.22)	3.327*** (62.54)	2.466*** (122.17)	2.536*** (143.31)
Years of schooling	0.047*** (67.18)	0.036*** (56.20)	0.061*** (91.77)	0.046*** (75.67)	0.086*** (129.22)	0.065*** (88.34)	0.111*** (128.69)	0.088*** (105.71)	0.131*** (123.28)	0.104*** (88.95)	0.137*** (81.53)	0.119*** (63.06)	0.136*** (61.95)	0.126*** (48.16)	0.115*** (148.86)	0.093*** (116.84)
Years of experience off the current firm	0.003*** (5.43)	0.002*** (3.72)	0.004*** (8.42)	0.002*** (5.10)	0.008*** (18.03)	0.005*** (11.40)	0.015*** (30.85)	0.010*** (22.00)	0.025*** (44.66)	0.016*** (27.27)	0.033*** (38.03)	0.025*** (27.87)	0.038*** (32.11)	0.030*** (23.32)	0.017*** (37.77)	0.012*** (27.98)
(Years of experience off the current firm) <sup>2</sup>	-0.010*** (6.09)	-0.006*** (4.51)	-0.013*** (8.11)	-0.007*** (6.09)	-0.020*** (13.98)	-0.015*** (10.58)	-0.032*** (20.07)	-0.024*** (16.42)	-0.042*** (23.93)	-0.034*** (17.93)	-0.047*** (16.68)	-0.052*** (17.18)	-0.048*** (12.59)	-0.058*** (13.30)	-0.030*** (20.88)	-0.027*** (19.76)
Years of tenure in the current firm	0.018*** (31.46)	0.012*** (22.14)	0.021*** (37.93)	0.015*** (29.46)	0.023*** (48.30)	0.020*** (37.30)	0.026*** (49.25)	0.026*** (48.26)	0.029*** (48.64)	0.025*** (36.81)	0.032*** (32.78)	0.027*** (25.61)	0.033*** (25.58)	0.029*** (19.32)	0.025*** (53.28)	0.023*** (45.58)
(Years of tenure in the current firm) <sup>2</sup>	-0.024*** (13.37)	-0.011*** (6.31)	-0.025*** (15.03)	-0.010*** (6.00)	-0.022*** (15.00)	-0.015*** (8.06)	-0.020*** (12.33)	-0.022*** (12.25)	-0.020*** (10.55)	-0.013*** (5.75)	-0.020*** (6.49)	-0.010*** (2.88)	-0.021*** (4.87)	-0.014*** (2.75)	-0.020*** (13.12)	-0.015*** (9.08)
Dummy for non-French	-0.120*** (20.92)	-0.079*** (11.90)	-0.144*** (26.44)	-0.094*** (15.45)	-0.169*** (34.06)	-0.123*** (17.90)	-	-0.147*** (21.25)	-0.215*** (33.41)	-0.154*** (17.26)	-0.206*** (19.78)	-0.127*** (9.32)	-0.178*** (12.72)	-0.113*** (5.94)	-0.212*** (42.11)	-0.146*** (22.11)
Dummy for single	-0.055*** (12.66)	0.004 (1.23)	-0.060*** (14.90)	0.004 (1.29)	-0.069*** (19.35)	0.013*** (3.65)	-0.075*** (19.34)	0.017*** (4.72)	-0.081*** (18.87)	0.013*** (2.86)	-0.085*** (12.39)	0.012* (1.71)	-0.086*** (9.60)	0.009 (0.85)	-0.086*** (24.64)	0.015*** (4.43)
Dummy for widowed	-0.038* (1.92)	-0.008 (0.92)	-0.030 (1.58)	-0.012 (1.50)	0.020 (1.17)	-0.037*** (3.98)	0.016 (0.84)	-0.043*** (4.69)	0.001 (0.03)	-0.041*** (3.55)	-0.004 (0.11)	-0.062*** (3.51)	-0.006 (0.14)	-0.084*** (3.38)	0.007 (0.39)	-0.045*** (5.20)
Dummy for divorced	-0.028*** (3.92)	0.013*** (2.66)	-0.029*** (4.31)	0.020*** (4.44)	-0.016*** (2.68)	0.028*** (5.59)	-0.021*** (3.13)	0.022*** (4.39)	-0.016*** (2.09)	0.024*** (3.84)	-0.025** (2.09)	0.008 (0.86)	0.006 (0.36)	-0.001 (0.10)	-0.023*** (3.83)	0.025*** (5.26)
Number of dependent children	0.011*** (7.34)	0.007*** (4.72)	0.013*** (8.94)	0.011*** (8.08)	0.015*** (12.16)	0.016*** (10.56)	0.019*** (15.31)	0.019*** (13.68)	0.020*** (14.90)	0.015*** (8.75)	0.022*** (10.67)	0.012*** (4.63)	0.024*** (8.67)	0.009** (2.39)	0.020*** (17.32)	0.017*** (12.38)
Dummy for workplace (1: Paris)	0.133*** (35.69)	0.101*** (30.30)	0.165*** (46.50)	0.132*** (42.85)	0.197*** (61.23)	0.193*** (56.39)	0.206*** (57.75)	0.235*** (69.67)	0.228*** (56.89)	0.275*** (64.54)	0.260*** (40.54)	0.318*** (48.81)	0.276*** (32.30)	0.336*** (36.79)	0.221*** (68.81)	0.241*** (74.83)
Dummy for type of contract (CDD: « <i>contrat à durée déterminée</i> »)	-0.008 (1.03)	-0.010** (1.99)	-0.015** (2.15)	-0.003 (0.70)	-0.026*** (4.33)	0.004 (0.74)	-0.054*** (8.48)	0.002 (0.32)	-0.077*** (10.66)	-0.028*** (4.26)	-0.107*** (9.07)	-0.060*** (5.89)	-0.127*** (7.94)	-0.077*** (5.41)	-0.065*** (11.27)	-0.024*** (4.94)
Log of hours paid work per year	0.031*** (10.30)	0.023*** (11.60)	0.029*** (10.82)	0.026*** (14.52)	0.023*** (9.78)	0.029*** (14.29)	0.004* (1.67)	0.021*** (10.37)	-0.029*** (9.96)	-0.009*** (3.40)	-0.075*** (15.36)	-0.076*** (18.32)	-0.113*** (16.99)	-0.127*** (21.58)	-0.010*** (4.20)	-0.003* (1.67)
Control for firms' factors	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Number of observations	82376	54793	82376	54793	82376	54793	82376	54793	82376	54793	82376	54793	82376	54793	82376	54793
Pseudo R <sup>2</sup> / R <sup>2</sup>	0.138	0.112	0.166	0.141	0.209	0.193	0.239	0.245	0.271	0.256	0.289	0.254	0.292	0.265	0.399	0.374

Survey INSEE ECMOSS 1992.

Absolute values of t statistics are in parentheses. \*\*\*, \*\* and \* mean respectively significant at the 1%, 5% and 10% levels.

**Table 6. Quantile decomposition and counterfactual gender gap**

Distribution	With no firm controls		With firms' factors		With firms' characteristics	
	Observed	Counterfactual	Observed	Counterfactual	Observed	Counterfactual
10 <sup>th</sup> percentile	-0.098	-0.102 (0.022)	-0.100	-0.113 (0.015)	-0.095	-0.105 (0.014)
25 <sup>th</sup> percentile	-0.122	-0.129 (0.018)	-0.123	-0.137 (0.005)	-0.118	-0.124 (0.005)
50 <sup>th</sup> percentile	-0.151	-0.168 (0.023)	-0.149	-0.158 (0.009)	-0.146	-0.159 (0.019)
75 <sup>th</sup> percentile	-0.205	-0.224 (0.023)	-0.197	-0.215 (0.019)	-0.185	-0.196 (0.014)
90 <sup>th</sup> percentile	-0.256	-0.291 (0.035)	-0.243	-0.233 (0.010)	-0.230	-0.233 (0.013)

Survey INSEE ECMOSS 1992.

Counterfactuals are constructed using the characteristics of male employees and returns to these characteristics of female employees. Standard errors in parentheses are obtained with 40 replications of the decomposition. The decomposition is performed using a random sample of 34 303 employees.



## Appendix

Table A. Descriptive statistics of the firms' characteristics

Variables	Mean	Standard deviation	Min	Max
<b>Firms' manpower structure</b>				
Share of production blue collar workers in the firm (A41)	0.593	0.27	0.01	1
Share of non-production blue collar workers in the firm (A42)	0.454	0.32	0.00	1
Share of technicians and supervisors in the firm (A43)	0.320	0.19	0.01	1
Share of executives in the firm (A44)	0.133	0.16	0.00	1
<b>Sectoral dummies</b>				
Food and agriculture (s1)	0.028	0.166	0	1
Production and distribution of energy (s2)	0.011	0.105	0	1
Intermediary goods (s3)	0.054	0.225	0	1
Equipment goods (s4)	0.055	0.227	0	1
Current consumption goods (s5)	0.065	0.246	0	1
Construction (s6)	0.073	0.261	0	1
Trade (s7)	0.173	0.379	0	1
Transports and telecommunications (s8)	0.082	0.274	0	1
Tradable services (s9)	0.398	0.489	0	1
Real-estate services (s10)	0.009	0.092	0	1
Insurances (s11)	0.020	0.138	0	1
Financial organizations (s12)	0.033	0.180	0	1
<b>Firm size</b>				
Total number of employees in the firm (A4)	373.688	1008.66	1.00	22174
Less than 20 employees (T1)	0.313	0.464	0	1
20 to 49 employees (T2)	0.341	0.474	0	1
50 to 99 employees (T3)	0.153	0.360	0	1
100 to 199 employees (T4)	0.066	0.248	0	1
200 to 499 employees (T5)	0.063	0.243	0	1
More than 500 employees (T6)	0.065	0.246	0	1
<b>Declared average number of paid hours of training per worker and per year in the establishment (dummies)</b>				
Zero hour or no hours declared (FF1)	0.498	0.500	0	1
1 to 10 hours (FF2)	0.216	0.411	0	1
11 to 40 hours (FF3)	0.203	0.402	0	1
More than 40 hours (FF4)	0.060	0.237	0	1
<b>Firms' general and organisational characteristics</b>				
State of business: Strongly growing or growing (VA1)	0.440	0.496	0	1
State of business: Stable (VA2)	0.293	0.455	0	1
State of business: Strongly decreasing or decreasing (VA3)	0.210	0.407	0	1
Activity is usually affected by seasonal movements (D21)	0.360	0.480	0	1
Activity is rather regular (D31)	2.448	2.225	1	9
Activity is rather irregular (D32)	1.934	3.463	0	9
Firm has been affected by unusual shocks in 1992 (D4A1)	0.333	0.471	0	1
Downward shock in 1992 (D4B1)	0.038	0.190	0	1
Upward shock in 1992 (D4B2)	0.308	0.462	0	1
Presence of union representatives in the firm (PS1)	0.252	0.434	0	1
Existence of wage negotiations in 1992 (D151)	0.225	0.418	0	1
Existence of a formal wage scale system for blue collar workers (D19A1)	0.769	0.421	0	1

Based on the branch's collective agreement (D19B1)	0.614	0.487	0	1
Based on the firm's collective agreement (D19B2)	0.099	0.299	0	1
Based on another evaluation scheme (D19B3)	0.065	0.246	0	1
Number of intermediate levels of management between the firm's manager and blue collar workers:				
Zero (D250)	0.222	0.416	0	1
From 1 to 4 (D251)	0.673	0.469	0	1
From 5 to 10 (D252)	0.035	0.185	0	1
From 11 and above (D253)	0.000	0.022	0	1
Existence of job rotation schemes:				
Put into practice within production teams (D26A1)	2.258	1.738	1	9
Intended to some versatile workers only (D26B1)	0.284	0.451	0	1
Direct collaborations between employees encouraged (D281)	0.499	0.500	0	1
Achieved work is 'permanently' controlled (D301)	0.618	0.486	0	1
Individual performances are 'systematically' controlled (D311)	0.361	0.480	0	1
Existence of a formal system to measure individual performances (D321)	0.209	0.406	0	1
Base wage progressions 'exclusively' depend on individual increases (D331)	0.057	0.232	0	1
Base wage progressions 'principally' depend on individual increases (D332)	0.148	0.356	0	1
Base wage progressions 'little' depend on individual increases (D333)	0.260	0.439	0	1
Base wage progressions 'never' depend on individual increases (D334)	0.256	0.436	0	1
Base wage progressions 'exclusively' depend on general increases (D341)	0.235	0.424	0	1
Base wage progressions 'principally' depend on general increases (D342)	0.381	0.486	0	1
Base wage progressions 'little' depend on general increases (D343)	0.125	0.331	0	1
Base wage progressions 'never' depend on general increases (D344)	0.145	0.352	0	1
Importance accorded by employers to the following criteria in individual wage increases:				
Workers' tenure in the job (D3513)	0.064	0.244	0	1
Increase in workers' performance (D3523)	0.359	0.480	0	1
Workers' training effort (D3533)	0.087	0.282	0	1
Accumulation of experience (D3543)	0.166	0.372	0	1
Acquisition of versatility (D3553)	0.157	0.364	0	1
Increase in workers' responsibilities (D3563)	0.321	0.467	0	1
Intra-firm mobility (D3573)	0.097	0.296	0	1
Difficulty of workers' eventual replacement (D3583)	0.063	0.244	0	1
Firm gives relative bonuses (the best workers are awarded) (D3911)	0.403	0.491	0	1
Firm gives absolute bonuses (production standards are exceeded) (D3921)	0.158	0.365	0	1
Firm favours relative bonuses (D39B1)	0.076	0.264	0	1
Firm has an explicit wage policy characterized by precise objectives (D411)	0.206	0.405	0	1

Survey INSEE ECMOSS 1992.

The number of establishments is 14 693.