



# Are union representatives badly paid? Evidence from France

Thomas Breda

► To cite this version:

| Thomas Breda. Are union representatives badly paid? Evidence from France. 2010. halshs-00564905

**HAL Id: halshs-00564905**

**<https://halshs.archives-ouvertes.fr/halshs-00564905>**

Submitted on 10 Feb 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.



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2010 - 26**

**Are union representatives badly paid?**

**Evidence from France**

**Thomas Breda**

**JEL Codes: J31, C49, J51, J53, J71**

**Keywords: Unions, wage differentials, representative,  
probability-based estimator**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# Are Union Representatives Badly Paid? Evidence From France\*

Thomas Breda<sup>†</sup>

October 5, 2010

## **Abstract:**

In this paper, I study the wage differential between firms' union representatives and their coworkers using a linked employer-employee dataset. On the employee side of the data, the surveyed workers are asked if they are unionized but we do not know which unionized workers are union representatives. On the employer side of the data, I have access to the number of union representatives and unionized workers in each firm. I use this information to construct an indicator of the firm-level probability for a randomly drawn unionized worker to be union representative. This indicator is then used to split the directly observable wage differential between unionized and non-unionized workers into two differentials: one between union representatives and non-unionized workers and another one between unionized workers who are not a union representative and non-unionized workers. Estimates that control for individual characteristics and firm-level fixed effects show that union representatives' wages are 10% lower than those of other unionized workers and non-unionized workers. Additional tests suggest that this gap can be understood as the result of a non-cooperative strategic interaction between employers and union representatives.

JEL codes: J31 C49 J51 J53 J71

Keywords: Unions, Wage Differentials, Representative, Probability-based Estimator

---

\*The author would like to thank the Direction de l'Animation de la Recherche, des Etudes et des Statistiques (DARES) for providing the data used in this study as well as Philippe Askenazy, Luc Behagel and Thomas Piketty for their helpful comments.

<sup>†</sup>Paris School of Economics-Ecole Normale Supérieure. Email: thomas.breda@ens.fr.

## Introduction

France is a country of open shop unionism in which only few designated workers within firms or workplaces represent all their coworkers, even those who are not unionized. In this context, the strategic interaction at play between the employer and the few workers representing all the others is unclear: on the one hand the employer can try to buy the union representatives<sup>1</sup>, according them individual benefits in exchange for a less tough bargaining; on the other hand, the employer can have the opposite strategy and discriminate against union representatives in order to discourage further attempts to organize.

This paper examines this strategic interaction mostly from an empirical point of view. I use a linked employer-employee dataset from the French private sector to show that union representatives' wages are lower than their coworkers' wages. On the employee side of the data, I only know if the surveyed workers are unionized or not. An empirical difficulty is the impossibility to distinguish, among the unionized workers, those who are a union representative from those who are only a member of a union without being a union representative. To sidestep this problem, I use available information in the employer part of the data on the number of union representatives (there can be more than one) and the number of unionized workers in each workplace. These variables enable to construct the workplace-level probability for a randomly drawn worker to be a union representative. This probability is equal to 0 for the workers declaring that they are not unionized (since union representatives have to be unionized) and is equal to the proportion of union representatives among unionized workers in their workplace for the workers declaring that they are unionized. I use this probability variable to split the directly observable wage differential between unionized and non-unionized workers into two differentials: one between union representatives and non-unionized workers and another one between unionized workers who are not a union representative and non-unionized workers. Assuming that these two differentials do not vary with the probability variable (or vary according to a parametrized curve), I show that they can be estimated consistently, conditional to workers characteristics and firm-level fixed effects. The estimation relies on the fact that the observed wage differences between unionized and non-unionized workers are more likely to be wage differences between union representatives and non-unionized workers in workplaces in which the proportion of union representatives among unionized workers is higher.

---

<sup>1</sup>Also called shop stewards in UK.

Estimating a series of standard wage determination models that control for individual and firm-level characteristics, I find that unionized workers taken as a whole are paid 2 to 3% less than non unionized ones. When the technique described above is used to split this directly observable wage differential, I find that union representatives are paid around 10% less than non-unionized workers whereas the other unionized workers are paid equivalently or slightly more than non-unionized workers.

In open shop countries such as France or in right-to-work states in United States, workers do not have to be unionized to be covered by firm-level union contracts. In this case, why do some workers are unionized and pay union dues whereas they could simply be free-riders and benefit from the union contracts without getting involved? Workers have indeed incentives to be covered nonmembers and act as freeriders unless unions provide benefits exclusive to their members. According to the classical Olson explanation (Olson, 1965), if a large group exists, it must have formed either because membership is compulsory or because the group provides private goods and services accessible only to its members, with ancillary provision of the collective good as a “byproduct”. An empirical litterature tries to solve this classical Olson free-riding problem by showing that union members enjoy non official wage premiums. In the United States, studies by Blakemore et al. (1986), Schumacher (1999), Budd and Na (2000) and Eren (2008) find a membership wage premium around 10%. In contrast, Hildreth (2000) and Booth and Bryan (2004) in Britain as well as Bunel and Raveaud (2008) in France do not find any wage premium for union members. This paper confirms only partly the results by Bunel and Raveaud (2008) for France and shows that the membership wage premium is small but yet statistically significant in certain empirical specifications. The main improvment on these authors’ seminal work is to consider separetly the union members who are union representatives and those who are not. Doing so, I obtain a higher wage premium for members who are not representative, which is more in line with the theoretical prediction of the free-riding model.

The real novelty of this paper is to focuss mainly on union representatives: to my knowledge, it is the first piece of work in the economic litterature that looks at union representatives’ wages. There are two explanations for this lack of interest. First, data on union representatives’ wage seem to be scarce or even nonexistent. Second, the situation of union representatives does not always appear as a valuable research question. In the U.S., there is a lot of evidence that anti-union policies and discrimination take place when

a union tries to organize a firm (see Bronfenbrenner, 2009, for a recent study). Employers also discriminate against prounion job applicants (Leap et al, 1990). However, unions in the US are majority unions and once the union has won the right to represent the workforce, it would usually have the power to prevent a company from paying lower wages to a representative. If a company were to pay a union representative less, the union would use the law and the contractual grievance procedure to fight back<sup>2</sup>. The legal settings in the U.S. imply *de facto* a collective bargaining between a majority union and an employer.

In France, the five largest unions are officially recognized in a specific firm for collective bargaining as long as they find a worker in this firm who is willing to be a union representative for them. This is a very weak legal requirement: union recognition works on a voluntary basis and there is no majority election. This weak requirement implies that union representative can be isolated on the field since they represent workers that have not elected them. An other implication of the French legal settings is that firm-level collective bargaining between an employer and the workers takes naturally the form of an “individual bargaining” between the employer and the union representative who defends the interest of a community of workers to which he is not necessarily closely related. In such a case, one can expect a strategic interaction to take place specifically between the employer and the union representative.

Union representatives enjoy non monetary rewards due to their situation: they have a work discharge (about half a day per work), they have a direct access to important information about their firm, they participate to work councils and are consulted about the important decisions their firm has to take and they have sometimes possibilities to begin a political career. They might also enjoy a form of social reward by getting the esteem of their coworkers and a higher social status. However, unions are present in only 36% of the private sector firms with more than 20 employees (Breda 2010). This is not a very high recognition rate since the main unions only have to find one worker who accepts to be their representative to be recognized. As a matter of fact, it appears that only 1 out of 125 workers is a union representative. The workers’ willingness to become union representative is in fact very small, which is a surprising result regarding the apparent advantages that representatives can get. This paper also attempts to discuss and propose an explanation for this apparent paradox.

The paper is organized as follows. Section 1 describes briefly the French institutional

---

<sup>2</sup>This point derives directly from informal discussions with Chris Tilly, from UCLA.

setting. Section 2 presents the econometric strategy I use to measure the “union representatives wage gap”<sup>3</sup>. Section 3 briefly presents the data. Section 4 gives the main empirical results. Section 5 discusses the different possible causal interpretations of the results and provides additional econometric tests in order to identify the mechanisms at play. Section 6 briefly concludes.

## 1 Institutional settings

The legal settings of union representation in France have been slightly modified on the 4th of May 2004 and more recently on the 20th of August 2008. As this study focuses on year 2004, I describe the functioning of industrial relations before these two laws were passed.

At first sight, France shares with most continental Europe countries characteristics of a regulated industrial relation system with multi-level bargaining. First, industry wide agreements negotiated by unions and employer associations cover most of the workforce. Second, individual employers can sign firm specific agreements with unions when unions are recognized at the firm level. According to the Statistics Department of the French Ministry of Labor (DARES), 97.7% of the workforce was covered by a collective agreement in 2004. With a union density around 8%, France is the OECD country with both the highest coverage rate and the lowest union density (OECD Employment Outlook, 2004).

In 1982, the Loi Auroux (August 4, 1982) encouraged decentralized bargaining. As a consequence, industry-level bargaining became less significant (Barrat et al 1996). Some of the existing agreements are even outdated because they have been rarely renegotiated in the past two decades and are now weaker than national standards in many sectors and regarding many topics. As a matter of fact, in 2006, exactly 50% of the 160 branches covering more than 5,000 employees had a branch minimum wage which was below the national minimum wage and was consequently useless (Breda 2010). To summarize, almost all workers are covered by industry-level agreements but a lot of these contracts are weak or even outdated, which gives an even more central role to firm-level bargaining in the French multi-level industrial relations.

---

<sup>3</sup>Note that this presentation is not completely standard: since the measure of the wage differential between union representatives and non-unionized workers (the union representatives wage gap) is subject to technical difficulties, I begin in section 2 with a description of the strategy I use to get this measure. Discussions about interpretation or identification of any causal effect will follow later in section 5.

Firm-level agreements can be signed between unions and employers as soon as unions have been recognized within firms. Concerning wages, these agreements can only improve the industry minimum wage and must be above the national minimum wage. Three key features differentiate France from Anglo-Saxon countries concerning union coverage at the firm level: first, there is no certification election, second, many unions can be present in the same firm and represent workers collectively and third, unionism is completely “open shop”.

*There is no certification election:*

To be recognized in a firm with more than 50 employees, the main unions almost only need to find one worker who accepts to officially represent the union in the firm. Such a worker is called a union representative. 95% of union representatives belong to only 5 large national “historical” unions. These “historical” unions are recognized as legal bargaining units within firms as soon as a worker accepts to be their representative. This is a fundamental feature of the French industrial relations: there is no certification election required for historical unions to organize larger firms. In firms with size between 10 and 50 employees, unions have to choose their representatives among workers who have already been elected, the so-called “firm delegates”. These “firm delegates” are legally recognized non-union representatives acting as the voice of the workers in their day-to-day relationship with the employer (they are generally also members of the work councils). They are elected every four years by workers in firms with more than 10 employees among voluntary candidates in a simple majority rule voting (the winning candidates are simply those that have collected the larger number of votes). The process of union recognition is more binding in firms with size between 10 and 50 employees, but even in these firms, union recognition remains less binding than the U.S. certification process which requires a majority of workers to be pro-union. The very weak legal constraints bearing on firm-level union recognition makes it easier for unions to legally organize firms. However, the low organizational cost paid by unions in these firms and the fact that they are not necessarily supported by the majority of the workforce should limit their bargaining power and leave the union representatives isolated on the field.

*Different unions can organize the same firm:*

The recognition process described above applies to each union, which makes in theory unlimited the total number of unions that can cover the workers of a given firm. In practice, around 36% of the private sector workplaces with more than 20 employees are



organized, which represents 64% of the workforce in these workplaces. This discrepancy is explained by the fact that the firm's probability to be organized increases considerably with its size. A particular union can also be represented in a working establishment by up to five representatives, depending on the size of the establishment<sup>4</sup>.

*Unionism is completely "open shop":*

When one or more unions are recognized in a firm, in place and newly hired employees do not have the duty to become union members, neither to participate in strikes. Finally, union contracts negotiated by union representatives must apply to all workers in the firm. Finally, the institutional settings concerning industrial relations and bargaining at the establishment level are exactly identical to the institutional settings at the firm level which is described above. Latter on, I use establishment-level variables on the number of unionized workers and union representatives to estimate the union representatives wage gap.

Figure 1 provides a visual illustration of the way professional relations work in France. The blue area represents a workplace. The yellow and red areas are respectively the unionized workers and the union representatives. Notice that proportions on the figure correspond to the estimated share of unionized workers and union representative on the data sample used for this study (in each workplace, around 1 worker out of 10 is unionized and around 1 unionized worker out of 10 is union representative). In what follows, I compare as precisely as possible, the wages of the union representatives and unionized workers to the wage of their coworkers.

## 2 Empirical specifications and estimation strategy

### 2.1 General framework with constant wage premia

I first provide a precise estimation of the wage differential between unionized and non unionized workers that controls for individual-level observable characteristics and establishment-level fixed effects. To do so, I present a series of regression models of the

---

<sup>4</sup>The maximal number of representatives that can be legally recognized is 1 for workplaces with less than 1000 employees, 2 for workplaces having between 1000 and 1999 employees, 3 for workplaces having between 2000 and 3999 employees, 4 for workplaces having between 4000 and 9999 employees and 5 for workplaces having 10000 employees or more.

type:

$$\ln(w_{ij}) = \alpha U_{ij} + \beta X_i + \eta_j + u_{ij} \quad (1)$$

where  $w_{ij}$  represents the hourly wage of individual  $i$  in establishment  $j$ ,  $X_i$  is a set of observed skill characteristics (such as age and education) of worker  $i$ ,  $\eta_j$  an establishment-level fixed effect and  $U_{ij}$  an indicator equal to 1 if worker  $i$  in establishment  $j$  is unionized. In some specifications, the fixed effect  $\eta_j$  will be replaced by a vector  $Z_j$  of establishment-level covariates.

In equation 1,  $\alpha$  can be interpreted in log-points as the within-establishment wage premium for unionized workers conditional on their observable skill characteristics. These unionized workers can be split in two groups: the workers who are Union Representatives (UR) and the other ones who are “Unionized Only” (UO). The wage premia for union representatives and workers who are “unionized only”, conditional on their characteristics and on establishments fixed-effects, are defined similarly as the coefficients  $\alpha_1$  and  $\alpha_2$  in the following regression model:

$$\ln(w_{ij}) = \alpha_1 UR_{ij} + \alpha_2 UO_{ij} + \beta X_i + \eta_j + u_{ij} \quad (2)$$

Let us assume that the standard identification assumption  $\mathbb{E}[u_{ij}|UR_{ij}, UO_{ij}] = 0$  is true. In this case,  $\alpha_1$  and  $\alpha_2$  can be estimated consistently by conventional OLS regression applied to 2. The problem is that the variables  $UR_{ij}$  and  $UO_{ij}$  are not observable directly in the data (see next section). The goal of this section is to recover the wage premia  $\alpha_1$  and  $\alpha_2$  using proxy variables for  $UR_{ij}$  and  $UO_{ij}$  that are available in the data. In other words, we assume that the wage premia for union representatives and workers who are “unionized only” would be identified if these variables were observable directly and we try to recover an estimate of these wage premia using an indirect estimation strategy.

Let us define by  $p_j$  the probability for a surveyed unionized worker in workplace  $j$  to be a union representative. Figure 2 provides an illustration of the link between the variables  $U_{ij}$ ,  $UR_{ij}$  and  $p_j$ . Notice that if workers are sampled randomly -as it is the case in the data I use (see next section)-,  $p_j$  is simply equal to the proportion of union representatives among unionized workers in workplace  $j$ <sup>5</sup>.

---

<sup>5</sup>If instead of being sampled randomly in each workplace, workers were selected according to some of their observable characteristics (age, gender, etc), the probability for a surveyed unionized worker in workplace  $j$  to be a union representative could be different from the proportion of union representatives among

The key result is that the unobservable variable  $UR_{ij}$  (resp.  $UO_{ij}$ ) in equation 2 can be replaced by the proxy variables  $p_j U_{ij}$  (resp.  $(1 - p_j)U_{ij}$ ). In other words, we will still have an estimation of the desired wage premium if we replace the dummy variables for being a union representative by the probability to be a union representative. This probability is equal to the individual indicator of being unionized ( $U_{ij}$ , which is observable) times the establishment-level probability of being a union representative conditional on being a unionized worker ( $p_j$ ). Formally, we have the following propositions:

**Proposition 1:** Let us write  $e_{ij} = UR_{ij} - p_j U_{ij}$ . If  $e_{ij}$  is not correlated with the error term  $u_{ij}$  in equation 2 then the wage premia  $\alpha_1$  and  $\alpha_2$  can be consistently estimated by OLS regression applied to:

$$\ln(w_{ij}) = \alpha'_1(p_j U_{ij}) + \alpha'_2((1 - p_j)U_{ij}) + \beta' X_i + \eta'_j + v_{ij} \quad (3)$$

Mathematically, this means that  $\mathbb{E}[\alpha'_1] = \alpha_1$  and  $\mathbb{E}[\alpha'_2] = \alpha_2$ .

**Proposition 2:** If  $e_{ij}$  is not correlated with the error term  $u_{ij}$  in equation 2 then the variances  $\sigma_u^2$ ,  $\sigma_v^2$  and  $\sigma_e^2$  of  $u$ ,  $v$  and  $e$  are related as follows:

$$\sigma_v^2 = \sigma_u^2 + (\alpha_1 - \alpha_2)^2 \sigma_e^2 \quad (4)$$

Propositions 1 and 2 are proved in mathematical appendix A. In the empirical section, I will estimate equation 3 by conventional (OLS). I will also correct the standard errors of the OLS estimates and provide a maximum-likelihood estimator of the desired wage differentials. These procedures are quickly described below:

### Correction of standard errors and tests:

Calling  $\hat{\sigma}_{\alpha_1}$ ,  $\hat{\sigma}_{\alpha'_1}$  the usual finite distance consistent estimators of the standard errors of  $\alpha_1$  and  $\alpha'_1$ , we have from proposition 2:

$$\hat{\sigma}_{\alpha_1}^2 = \hat{\sigma}_{\alpha'_1}^2 - (X'X)_{11}^{-1}(\alpha_1 - \alpha_2)^2 \sigma_e^2$$

---

unionized workers. For example, if union representatives are older than the average unionized worker and if the workers' sampling strategy over-represent older workers, then the probability for a sampled unionized worker to be representative is higher than the proportion of union representatives among unionized workers.

where  $(X'X)_{11}^{-1}$  designates the first diagonal coefficient (the one corresponding to  $\alpha_1$ ) of the variance covariance matrix of the regressors. This formula will be used in the empirical analysis to correct the estimated standard errors and run the appropriate Student's tests.

### Maximum likelihood estimator:

There is *a priori* no reason why the OLS “indirect estimation” (IE) procedure proposed above would provide, as in the case of standard OLS, the most efficient estimators of  $\alpha_1$  and  $\alpha_2$ . I thus also compute the log-likelihood of the sample under the hypothesis of normality of the residuals and show that the maximum (log)likelihood estimator (ML) does not correspond to the IE estimator.

Let us consider that the data obey to the following linear model (illustrated on figure 2):

$$\begin{cases} \ln(w_{ij}) = \beta X_{ij} + v_{ij0} & \text{if } U_{ij} = 0 \\ \ln(w_{ij}) = \alpha_1 + \beta X_{ij} + v_{ij1} \text{ with probability } p_j & \text{if } U_{ij} = 1 \\ \ln(w_{ij}) = \alpha_2 + \beta X_{ij} + v_{ij2} \text{ with probability } 1 - p_j & \text{if } U_{ij} = 1 \end{cases}$$

with  $v_{ij0}, v_{ij1}, v_{ij2} \sim \mathcal{N}(0, \sigma^2)$

For simplicity, I have included individual, establishment-level characteristics and the constant term in the unique vector  $X_{ij}$ . Doing this modeling, I have made two non obvious assumptions. First, the return to observable characteristics  $\beta$  is identical for non-unionized workers, only unionized workers and union representatives. Otherwise, I would be doing a kind of Oaxaca-Blinder decomposition estimated by maximum likelihood, which is not our goal here since we would like to estimate an equivalent of equation 2. Second, the standard deviation of the residuals  $\sigma$  is also identical accross groups of workers, as it is the case in the OLS estimation. I will show that under this hypothesis the ML estimator is not identical to the ME estimator. But later on, in the empirical analysis, I will allow the standard deviations to be different accross groups.

We denote by  $w$ ,  $X$ ,  $U$  and  $p$  respectively the vectors of the  $N$  observable variables  $w_{ij}$ ,  $X_{ij}$ ,  $U_{ij}$  and  $p_j$  and by  $\phi$ , the standardized normal density. The log-likelihood

$\mathcal{L}(w, X, U, p, \alpha_1, \alpha_2, \beta, \sigma^2) = \ln(\prod_{i=1}^N (P(w_{ij}, p_j, X_{ij}, U_{ij} | \beta, \alpha_1, \alpha_2, \sigma^2)))$  can be written:

$$\mathcal{L} = \sum_{i=1}^N \ln[(1-U_i)\phi((w_{ij}-\beta X_{ij})/\sigma) + p_j U_i \phi((w_{ij}-\beta X_{ij}-\alpha_1)/\sigma) + (1-p_j) U_i \phi((w_{ij}-\beta X_{ij}-\alpha_2)/\sigma)] \quad (5)$$

It is easy to check that differentiating equation 5 relative to  $\beta$ ,  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  does not simplify as in the case of OLS. Indeed, in the case of OLS, the ln functions have as argument

only one normal density function which equals  $e^{-(u_i^2/2)}/\sqrt{(2\pi)}$ . Consequently, maximizing the likelihood is equivalent to minimizing the sum of the square of the residuals. In equation 5, this is not true anymore because the ln functions have as arguments a sum of 3 density functions that does not simplify. Consequently, the IE and ML estimators have no reasons to be equal and estimates of the parameters that minimize the log-likelihood defined in equation 5 will be presented in the next section as a robustness check.

## 2.2 Allowing for non-constant wage premia:

There are good reasons to think that the wage differential between union representatives and their coworkers can vary with the proportion of union representatives among unionized workers. If there are a lot of unionized workers for example, the representatives will be less isolated on the field and it might be harder for the employer to either discriminate or bribe them. However, we have assumed in proposition 1 that the measure of the wage penalties  $\alpha_1$  and  $\alpha_2$  are constant and independent of  $p_j$ . This is a usual and implicit assumption that one makes when estimating a linear model. It offers a convenient way to get the effect of a variable on the mean of another one. But the fact that  $\alpha_1$  and  $\alpha_2$  do not depend on  $p_j$  is crucial when replacing the observed variable  $UR_{ij}$  by the proxy  $p_j U_{ij}$ . Figure 3 illustrates this point in the simpler case with no control variables. On the left chart, the wage differentials  $\Delta w_{UR} = \alpha_1$  between union representatives and non unionized workers and  $\Delta w_{UO} = \alpha_2$  between workers who are “unionized only” and non-unionized workers remain constant when  $p_j$  varies. The observed within workplaces wage differential  $\Delta w_U = \alpha$  between unionized and non-unionized workers has for equation  $\Delta w_U = [p_j \Delta w_{UR} + (1 - p_j) \Delta w_{UO}]$ . It is a straight line. In this case, the estimation strategy resulting from proposition 1 simply consists in estimating the slope and the y-intercept of this straight line (which can be obtained by a simple orthogonal projection of  $\Delta w_U$  on the lines  $y = p_j$  and  $y = 1 - p_j$  represented by dashed lines on the chart). However, if  $\Delta w_{UR}$  and  $\Delta w_{UO}$  are not constant with  $p_j$ ,  $\Delta w_U$  is not a straight line any more. This is illustrated on the right side of the figure. In this case, the estimation strategy resulting from proposition 1 fails to identify the wage premia for union representatives and workers who are “unionized only”. In the case of standard OLS estimation, a non-constant parameter would be estimated to be equal to its mean. But in our case, the possibility to estimate a wage premium explicitly relies on the fact that this wage premium is constant. If instead  $\alpha_1$  and  $\alpha_2$  vary with  $p_j$ , their estimation through equation 3 can be completely

erroneous. To overcome this problem, I propose 2 solutions:

- When  $p_j = 1$ ,  $\alpha_1 = \Delta w_{UR} = \Delta w_U$  and is thus observable directly (when a sampled worker declares itself as unionized in a workplace where only the union representatives are unionized, we know with certainty that he is a union representative). I will thus estimate  $\alpha_1$  on the subsample of firms with  $p_j = 1$ .

- I will plot the observable wage differential between unionized and non-unionized workers for different values of  $p_j$  and see if it varies linearly such as in the left chart of figure 3.

### 3 The data

The dataset I use is the 2004 French Workplace Employment Relations Survey (RE-  
PONSE04) conducted by the Ministry of Labor towards up to 10 employees randomly  
drawn in each of 2929 business establishments with more than 20 employees. RE-  
PONSE04 contains extensive information on industrial relations at the workplace level and on the  
firms' organizational and technological structure<sup>6</sup>. In each surveyed workplace, union den-  
sity, the name of the unions that are present and the number of their representatives are  
available. I will divide the total number of union representatives by the number of union-  
ized workers (which is equal to union density at the workplace-level times the number of  
employees in the workplace) to get the proportion of union representatives among union-  
ized workers. Net hourly wages in December 2003 have been retrieved from Social Security  
records (the Déclaration Annuelles de Données Sociales, DADS) by the Ministry of Labor  
for the workers surveyed in RE-  
PONSE04 and have been matched with the dataset. The  
RE-  
PONSE04 survey covers mainly the private sector but some public companies are also  
present. After cleaning, the employee survey contains 7836 workers for whom we have  
the usual observable characteristics (education, gender, occupation, age, tenure, working  
hours) and we know if they are unionized or if they have been unionized in the past and  
are not unionized anymore.

---

<sup>6</sup>RE-  
PONSE follows the same design than WERS in the U.K. See Bryson et al. (2009) for a study that  
uses both RE-  
PONSE and WERS to study unions and workplace performance or Blanchflower and Bryson  
(2008).

## 4 Results

The data appendix provides a table of summary statistics on currently unionized workers and past unionized workers who are not unionized anymore and a figure showing which observable characteristics in the data affects the most the probability to be unionized. It can be seen in particular that unionized workers are more often men than women and that the probability to be unionized increases with age and increases sharply with tenure. The last table of the data appendix gives the distribution of the data workplaces in terms of the number of unionized and non-unionized workers. Estimations of the wage penalty for unionized workers that include workplace fixed effects rely on the 658 workplaces with at least one unionized and one non-unionized worker. Even if the sample size is relatively small, this number is large enough to run fixed effects estimations.

I start with the estimation of the wage differential between unionized workers (taken as a whole) and non-unionized ones such as defined in equation 1. Table 1 presents OLS and workplace fixed effects estimations of this differential with 3 different sets of control variables. In the REPOSE04 survey, the workers are also asked if they were unionized in the past. In all specification I add a dummy variable for workers who have been unionized and are not unionized anymore. This implies that the omitted group to which the unionized workers are compared is the group composed by the workers that have never been unionized. The first specification, in columns 1 and 2, is simply a raw differential (with no control variables). We can see that unionized workers are paid in average 6 to 7 percent more than workers that have never been unionized. However, within workplaces, this difference vanishes (second column). In columns 3 and 4 (specification 2), the usual exogenous wage determinants have been added as control variables (age, age square, tenure, tenure square, education, gender). Unsurprisingly, the OLS gap found in column 2 disappears: the fact that unionized workers are older, have longer tenure and are more often men explained entirely the raw difference. More strikingly, when workplace fixed effects are added in this second specification, unionized workers are found to be paid about 5% less. To check if this observation is robust, I have tried in specification 3 (last 2 columns) to add additionnal controls for occupation and working hours on the ground that unionized workers can be segregated in low skill occupations and might work less. The estimated within workplace wage penalty for unionized workers is lower but remains very significant. The roughly 2.5% wage penalty that we find in estimates that control for

workplace fixed-effects and detailed individual skill characteristics appears to be robust to a wide range of other specifications that control more finely for the observable individual characteristics. This result differs from the one of Bunel and Raveaud (2008) who found a non significant wage differential between unionized and non-unionized workers. Table 1 also present the wage penalty for workers who were unionized in the past and are not unionized anymore. In specifications that include workplace-level fixed effects, these workers appear to be paid between the workers who are unionized and those who have never been unionized. This result will be discussed in the next section which deals with interpretation.

I now turn to the key estimation of the paper: the 2.5% negative wage differential between unionized workers and their coworkers found in table 1 will be splitted in two differentials, one for union representatives and one for the other unionized workers. It has been possible to construct a variable  $p_j$  for 2081 workplaces in the sample. In 860 of these workplaces, unions are not recognized (there are no union representatives). For those workplaces,  $p_j$  is equal to 0. Figure 4 shows the distribution of the proportion of unionized workers and union representatives as well as the distribution of the constructed variable  $p_j$  accross the other workplaces (those having at least one union representative). In almost 10% of these workplaces,  $p_j$  is higher than 0,9. The proxy variable  $p_j U_{ij}$  has been constructed for 7597 employees. 28 of these employees have declared to be unionized and work in a workplace in which  $p_j$  is equal to 1. These workers are identified with certainty as union representatives.  $p_j U_{ij}$  is the individual probability to be union representative. Taking the sum of the variable on the sample gives the expectation of the total number of union representatives in the sample. It is equal to 128.

Table 2 shows the results obtained when estimating equation 3 by OLS (the so-called IE estimation strategy) and when maximizing the log-likelihood function 5 (the so-called ML estimation strategy). All specifications include a detailed set of individual characteristics as control variables (age, age square, tenure, tenure square, 8 education dummies, a dummy for gender, 4 dummies for occupation and working hours). The results are striking: in the specification that includes workplace-level fixed effects (column 1), the wage penalty is entirely beared by union representatives and gets as large as 8% for these workers as compared to those who are not unionized. In the second specification, the workplace fixed effects have been replaced by a set of workplace observable characteristics



(6 dummies for workplace size, 5 dummies for age, 16 for industry and 8 for region)<sup>7</sup>. The penalty for union representatives is even larger in this second estimation. However, workers who are “unionized only” appear to be actually paid a slightly better wage than non-unionized workers, which is consistent with the Olson explanation of collective action (Olson, 1965). I provide standard errors that have been corrected according to proposition 2 for all estimates obtained via the IE estimation strategy. I also provide standard errors clustered by couple (*workplace, U<sub>i</sub>*) since the variables of interest  $p_j U_i$  and  $(1 - p_j) U_i$  are identical for all unionized workers or non-unionized workers in the same workplace. These clustered standard errors probably slightly overestimate the true standard errors of the estimators since they have not been corrected. The third column of table 2 gives the estimates without fixed effects obtained via the ML strategy<sup>8</sup>. They appear to be close and to confirm the results obtained via the IE strategy. In these specifications, I have authorized the variance of the error term to be different for non-unionized workers, union representatives and only unionized workers. Finally, the results in table 5 are robust to modifications of the individual set of control variables, providing that the usual Mincer equation controls for age, education and gender remain in this set)<sup>9</sup>

As discussed in section 2, the results in the 3 first columns of table 5 would be biased if the wage penalty for union representatives vary with  $p_j$ . Columns (4) and (5) of table 5 provide a first robustness check of the identification strategy. In these specifications, I have excluded from the sample the workers having an unclear union status, that is, those who declare to be unionized and working in a workplace in which the probability to be a union representative is neither 0 nor 1. After this cleaning, the sample (in column 5) comprises 6416 non unionized workers (none of these workers have been removed from the sample) and 114 unionized workers. 27 of these unionized workers are identified with certainty as union representatives and the remaining 87 as workers who are “only unionized”. Point estimates in these specifications are a bit lower (gaps are larger)

---

<sup>7</sup>Of course, the regression model with no fixed effects leads to a more imprecise measure of the wage differential between union representatives and their workplace’s coworkers. However, it is less demanding for the data. Since the dataset I use is relatively small and since my estimation is indirect, I believe that it is important to estimate this model too.

<sup>8</sup>the ML estimator with workplace fixed effects is computationally too intensive to be calculated with a standard optimization algorithm on a standard computer (after several months, it still failed to converge and I had to stop the process).

<sup>9</sup>No results are not reported but they are available on demand.

and standard deviations a bit higher, but the results of the previous specifications are corroborated. Figure 5 plots the observable wage differential between unionized and non-unionized workers for different values of  $p_j$ . More precisely, I have reported the estimated wage differential between unionized and non-unionized workers (conditional on observable characteristics) in the 2 groups of establishments with  $p_j = 0$  and  $p_j = 1$  (those are identical to the estimates in the last columns of table 5). I have then divided the workplaces having  $p_j$  strictly between 0 and 1 in 4 quartiles and reported on the plot the average  $p_j$  and the average conditional wage differential in each of these groups. In each case, the wage differential increases regularly from virtually 0 in establishments where  $p_j = 0$  to 10% or more in those where  $p_j = 1$ . Even though the estimated functions  $\alpha(p_j) = \delta w_u(p_j)$  are not perfectly linear as in the theoretical case exposed on the left panel of figure 3, they do not exhibit any abnormal point.

## 5 Discussion and causal tests

Why are union representatives so badly paid? Two usual explanations are possible: discrimination and adverse selection. Keeping the previous notations (without the  $j$  subscript) and denoting by  $\theta_i$  the productivity of worker  $i$ , we can give a mathematical definition of these 2 usual statistical notions in the context of this study:

$$\left\{ \begin{array}{ll} \text{- Discrimination (taste based):} & \mathbb{E}[w_i|\theta_i, UR_i = 1] < \mathbb{E}[w_i|\theta_i, UR_i = 0] \\ \text{- Adverse selection:} & \mathbb{E}[\theta_i|X_i, UR_i = 1] < \mathbb{E}[\theta_i|X_i, UR_i = 0] \end{array} \right.$$

Statistical discrimination on wages is impossible in the long run. Indeed, there is statistical discrimination if the employer does not observe a worker productivity and lower this worker's wage regarding an observable non productive characteristic (such as gender, race or union status) because he knows or thinks that this unproductive characteristic is correlated with some unobservable component of productivity (for example lazyness is unobservable and affects productivity and the employer thinks that unionized workers are more lazy). In the long run, unionized workers average productivity is observed by the employer and lower wages can only reflect adverse selection (unionized workers are indeed

more lazy) or actual taste based discrimination (unionized workers are not more lazy, the employer knows it but still he pays them a lower wage).

The legal settings in France tend to isolate the union representatives on the field and to turn collective bargaining into a more individual bargaining between the employer and the representative. Consequently, a specific strategic interaction can then take place between the employer and the few union representatives in the firm. A careful look at the incentives of the employers in this strategic game shows that they can rationally have interest to hurt the representative to discourage other workers to become representatives and to discourage further attempts to organize. In that sense, the term “taste based discrimination” is misleading even though the employer consciously pays to the union representatives wages that are lower than their productivity. From a statistical point of view, we can speak about discrimination. However, from an industrial organization point of view, the potential discrimination should probably be seen as the result of a non-cooperative strategic interaction between the employer and the union representatives.

I do not have an experimental design that enables to identify with certainty the causal explanation of the wage penalty for union representatives. An informal argument support the “discrimination interpretation”. The union representatives are drawn among the unionized workers and they probably share with these unionized workers a lot of socio-economic characteristics. This implies that they are not very likely to be far less productive than the other unionized workers. Hence, the high wage differential between union representatives and the workers who are “unionized only” is unlikely to reflect a selection process because such a selection process would concern only the representatives but not the other unionized workers who do not appear to be paid less than their non-unionized counterparts. However, the law gives to union representatives a roughly 3 hours a week work discharge for their union activities. From a legal point of view, the employer should not pay these workers less because of their work discharge. These work discharges are indeed a legal duty whose the cost has in theory to be entirely beared by the employer. But from an economic perspective, if union representatives work less, they might well be paid less as a consequence. I now present 2 pieces of evidence that reinforce the discrimination interpretation and invalidates the idea that the representatives are paid less because of selection or because of their work discharges.

The first test I provide consists in separating the representatives according to their

tenure. If discrimination is at play, it cannot happen instantly. In practice, it can take the form of a lower rate of promotions and pay raises for the representatives. However, if the “bad workers” select themselves among the union representatives, this means that the representatives are drawn from the bottom of the wage distribution (conditional on their observable characteristics). In that case, we should already observe a negative wage differential between the representatives with short tenure and their coworkers. Denoting by  $ST_i$  a dummy variable equal to 1 for workers having less than 5 years of tenure, I estimate in table 3 the following equation<sup>10</sup>:

$$\ln(w_{ij}) = \alpha_1^{ST} (p_j U_{ij} ST_i) + \alpha_1^{LT} (p_j U_{ij} (1 - ST_i)) + \alpha_2 ((1 - p_j) U_{ij}) + ST_i + \beta X_i + \eta'_j + v_{ij}$$

In the two first columns, the usual detailed set of individual characteristics has been added as control. In the two last columns, only the exogenous characteristics of the Mincer wage equation are included. In particular, I do not control for occupation which can be endogenous: occupations are affected by promotions, which themselves depend on the tenure of workers and on potential discrimination against them. In all specifications, the wage penalty for union representatives is beared by those having more than 5 years of tenure. This is even stronger in specifications that do not control for occupation. Since union representatives with short tenure are given the same work discharge than the ones with longer tenure, these results show that work discharge do not explain the wage penalty for union representatives.

My second approach to suggest that strategic interaction is at play is to look at the wage penalty for representatives from different unions and to correlate the results to the behavior of each particular union. In the REPNONSE survey, the managers are asked the number of union representatives of each of the French main unions. It is thus possible to compute the proportion  $p_j$  of union representatives among unionized workers for each main union and to apply the IE and ML techniques to obtain consistent (IE) or most powerful (ML) estimators of the wage penalty for the union representatives of these different unions. Due to the relatively small sample size of the data, I have done it only for the 3 largest unions and grouped together the other unions. These largest unions are the “Confédération Générale du Travail” (CGT), the “Confédération Française Démocratique du Travail” (CFDT) and “Force Ouvrière” (FO). CGT and CFDT are almost the same size and FO is

---

<sup>10</sup>Unionized workers with less than 5 years of tenure represent exactly 25% of the sample of unionized workers.

slightly smaller. Almost 70% of the union representatives belong to these unions (about 27% belong to CGT, 27% belong to CFDT and 15% belong to FO). CGT is historically a communist union. Even though since the mid '90s communism no longer stands out as the dominant ideology driving the organization, its concrete counterpart of class struggle still characterizes the action of CGT today. According to a large sociological literature, CGT can be seen as more aggressive and less willing to make concessions than CFDT, its more direct rival. As a matter of fact, statistics from the Ministry of Labor show that, in 2002, 26% of the strikes in the firms were initially conducted by CGT alone and 33% by an association of CGT and other unions, whereas only 10% of the strikes were generated by CFDT alone and 24% by CFDT and other unions. Finally, Breda (2008) has studied the union wage premium in France, that is, the average wage premium obtained by unions in firms in which they are recognized. He found that CGT is the union associated with the largest premia.

Table 4 shows the results. CGT union representatives appear to be paid around 20% less than non-unionized workers. CFDT union representatives seem to be paid roughly 10% less in the IE specification with no fixed effects but this result is not very stable and not statistically significant in other specifications. FO or the union representatives representing other unions are not paid differently in comparison with non-unionized workers. The fact that union representatives from CGT, who are the most fighting and the more able to bargain better wages for the workers in their firm, are less paid than union representatives from other unions reinforces the idea that they play a non-cooperative game with the employers which leads the employer to pay them poorly<sup>11</sup>. The results in table 4 also strongly contradict the fact that union representatives are paid less because they have a weekly work discharge. Indeed, if this were the only explanation, we should observe an identical wage penalty for the representatives of each union since they are granted the same work discharge.

---

<sup>11</sup>It should be kept in mind that these last results rely on a small number of "observations". The expectation of the number of CGT, CFDT and FO union representatives on the sample is respectively 38, 41 and 26 individuals. In addition, those workers union status is not observable directly. But yet results for CGT are very significant. This probably means that, conditional on observables, almost all the few workers identified as likely to be CGT union representatives experience far lower wages. Of course, standard errors and confidence intervals are large too and direct data on the workers' union status would be necessary to assess precisely what is the size of the gap. The 95% confidence interval for the wage penalty experienced by CGT union representatives is [-25%,-15%].

Interpreting the wage penalty for union representatives as the result of a non-cooperative interaction between employers and them enables to solve an apparent paradox: even though union recognition at the workplace level only depends on the presence of a volunteer worker to be a union representative, unions are only present in 36% of private sector workplaces with more than 20 employees. How comes so few workers accept to become representatives even though it apparently looks to be a privileged position (union representatives have a direct access to important information about their firm, they participate to work councils, they get working discharges and they might also enjoy a form of social reward by getting the esteem of their coworkers and a higher social status)? If true, the fact that workers are discriminated and have to renounce partly to their professional career when they become a union representative could be the hidden factor that makes such a decision difficult to take.

## 6 Conclusion

This research has shown that unionized workers in France are slightly less paid than their non-unionized coworkers. When this wage gap is broken apart between union representatives, who bargain for all the employees in their workplace with the employer, and the workers who are only unionized, a clear pattern appears: only unionized workers earn as much or even slightly more than non-unionized ones whereas union representatives are paid 8 to 11% less, even in specifications that control for workplaces fixed effects.

A non-cooperative game probably takes place between employers and union representatives, which leads the employer to discriminate against the representatives. Such an exclusive interaction which does not comprise the other workers is made easier by the French legal context in which union representatives are not democratic representants of their coworkers (they are not elected). Empirical results reinforce the idea that a non-cooperative game takes place: the most penalized union representatives are precisely those from the least cooperative union and those with the longest tenure.

To my knowledge, this research is the first on union representatives. It has the virtue to reveal an unknown important statistical fact that concern a lot of workers: in France, even if unionization rates are low, there are still more than 1 million unionized workers in the private sector and more than 100,000 of them are union representatives. This

research thus shed light on the problematic situation of professional relations: worker representatives should not be paid so badly as compared to the workers they represent, especially because one of their main role is to bargain wages.

Finally, union representatives are not observable directly in the data and their potential number (128) is relatively small. I call for further research on this topic and direct data collection on union representatives in order to get more precise estimates of the exact value of the wage differential between union representatives and the workers they represent.

## References

- BLAKEMORE, A., J. HUNT, AND B. KIKER (1986): “Collective bargaining and union membership effects on the wages of male youths,” *Journal of Labor Economics*, 4(2), 193–211.
- BLANCHFLOWER, D., AND A. BRYSON (2010): “The wage impact of trade unions in the UK public and private sectors,” *Economica*, 77(305), 92–109.
- BOOTH, A. L., AND M. L. BRYAN. (2004): “The Union Membership Wage Premium Puzzle: Is there a Free Rider Problem?,” *Industrial and Labor Relations Review*, 57(3), 402–421.
- BREDA, T. (2008): “Union Wage Differentials: What About France?,” in *European Association of Labour Economics Conference Paper, Amsterdam*.
- (2010): “Firms’ rents, workers’ bargaining power and the union wage premium in France,” *Working Paper Paris School of Economics*.
- BRONFENBRENNER, K. (2009): “No holds barred: The intensification of employer opposition to organizing,” *Kate Bronfenbrenner*, p. 40.
- BRYSON, A., J. FORTH, AND P. LAROCHE (2009): “Unions and workplace performance in Britain and France,” *NIESR Discussion Papers*.
- BUDD, J., AND I. NA (2000): “The union membership wage premium for employees covered by collective bargaining agreements,” *Journal of Labor Economics*, 18(4), 783–807.
- BUNEL, M., AND G. RAVEAUD (2008): “Unionism does not pay: evidence from France,” in *European Association of Labour Economics Conference Paper, Amsterdam*.
- DE OLIVERA, O. (2006): “Les demandes de licenciement de salariés protégés : Une baisse de 10% en 2004,” *Premières synthèses DARES*.
- DE OLIVERA, O., R. MERLIER, AND S. ZILBERMAN (2005): “Les licenciement de salariés protégés de 2001 à 2003 : le motif économique prédominant, toujours et encore,” *Premières synthèses DARES*.



- EREN, O. (2008): “Does Membership Pay Off for Covered Workers-A Distributional Analysis of the Free Rider Problem,” *Indus. & Lab. Rel. Rev.*, 62, 367.
- HILDRETH, A. (2000): “Union wage differentials for covered members and nonmembers in Great Britain,” *Journal of Labor Research*, 21(1), 133–147.
- LEAP, T., W. HENDRIX, R. CANTRELL, AND G. TAYLOR (1990): “Discrimination Against Prounion Job Applicants,” *Industrial Relations*, 29(3), 469–478.
- OLSON, M. (1971): *The logic of collective action: Public goods and the theory of groups*. Harvard Univ Pr.
- SCHUMACHER, E. (1999): “What Explains Wage Differences Between Union Members and Covered Nonmembers?,” *Southern Economic Journal*, 65(3), 493–512.

## Mathematical appendix: Proof of propositions 1 and 2.

The term  $e_{ij} = UR_{ij} - p_j U_{ij}$  can be seen as a measurement error: the difference between the fact and the probability to be a union representative. By construction, this term verifies 2 properties enonced in the following lemmas:

*Lemma 1:*  $\mathbb{E}[e_{ij}] = 0$

*Proof:* We have  $\mathbb{E}[e_{ij}|U_{ij} = 0] = 0$  (because non-unionized workers cannot be union representatives) and  $\mathbb{E}[e_{ij}|U_{ij} = 1] = P(UR_{ij} = 1)(1 - p_j) + P(UR_{ij} = 0)(-p_j) = p_j * (1 - p_j) + (1 - p_j) * (-p_j) = 0$ . This implies Lemma 1.

*Lemma 2:*  $Cov(p_j U_{ij}, e_{ij}) = 0$

*Proof:* First,  $Cov(p_j U_{ij}, e_{ij}) = \mathbb{E}[(p_j U_{ij} - \mathbb{E}[p_j U_{ij}])(e_{ij} - \mathbb{E}[e_{ij}])] = \mathbb{E}[p_j U_{ij} e_{ij}]$ . Next, we have:

$$\begin{aligned} \mathbb{E}[p_j U_{ij} e_{ij} | U_{ij} = 0] &= 0 \\ \mathbb{E}[p_j U_{ij} e_{ij} | U_{ij} = 1] &= \mathbb{E}[\mathbb{E}[p_j e_{ij} | UR_{ij} = 1] P(UR_{ij} = 1) + \mathbb{E}[p_j e_{ij} | UR_{ij} = 0] P(UR_{ij} = 0) | U_{ij} = 1] \\ &= \mathbb{E}[p_j(1 - p_j) * p_j - p_j(-p_j) * (1 - p_j)] \\ &= 0 \end{aligned}$$

Consequently,  $Cov(p_j U_{ij}, e_{ij}) = 0$ <sup>12</sup>.

Noticing that  $UO_{ij} = U_{ij} - UR_{ij} = (1 - p_j) * U_{ij} - e_{ij}$  and plugging  $p_j U_{ij}$  and  $(1 - p_j) U_{ij}$  in equation 2. We get:

$$\ln(w_{ij}) = \alpha_1(p_j * U_{ij}) + \alpha_2((1 - p_j) * U_{ij}) + \beta X_i + \eta_j + u_{ij} + (\alpha_1 - \alpha_2)e_{ij}.$$

$\mathbb{E}[u_{ij}|p_j U_{ij}] = \mathbb{E}[u_{ij}|UR_{ij} - e_{ij}] = 0$  because we have assumed that  $\mathbb{E}[u_{ij}|UR_{ij}] = 0$  and supposed that  $u_{ij}$  and  $e_{ij}$  are not correlated.  $\mathbb{E}[e_{ij}|p_j U_{ij}] = 0$  also follows from Lemmas 1 and 2.

Denoting  $v_{ij} = u_{ij} + (\alpha_1 - \alpha_2)e_{ij}$ .  $v_j$  the residual in the econometric equation above, we finally have  $\mathbb{E}[v_{ij}|p_j U_{ij}] = 0$ , which is a sufficient condition to prove that the OLS estimation of 3 provides consistent estimates of  $\alpha_1$  and  $\alpha_2$ .<sup>13</sup>

<sup>12</sup>Note that  $e_{ij}$  should not be seen as a classical measurement error. Indeed, the classical error in variable assumption that econometricians would have in mind when dealing with measurement errors would be  $Cov(UR_{ij}, e_j) = 0$ , that is, the measurement error is not correlated with the true value of the considered variable. This assumption is obviously wrong here since  $\mathbb{E}[e_{ij}|UR_{ij} = 0] < 0$  and  $\mathbb{E}[e_{ij}|UR_{ij} = 1] > 0$ .

<sup>13</sup>Notice that the estimates of  $\beta$  and  $\eta_j$  could be biased if  $\mathbb{E}[v_{ij}|X_i, \eta_j] \neq 0$ . That will occur if  $\mathbb{E}[e_{ij}|X_i, \eta_j] \neq 0$ . The measurement error  $e_{ij}$  actually plays the role of an omitted variable: not having it in the regression biases the estimation for the variables that are correlated with it. As, by construction  $p_j U_{ij}$  is not correlated with  $e_{ij}$ , it follows that the estimates of  $\alpha_1$  and  $\alpha_2$  are unbiased.

Also, if  $u_{ij}$  is uncorrelated with  $e_{ij}$ , we immediately have that  $\sigma_v^2 = \sigma_u^2 + (\alpha_1 - \alpha_2)^2 \sigma_e^2$ .  
**QED.**

## Data Appendix:

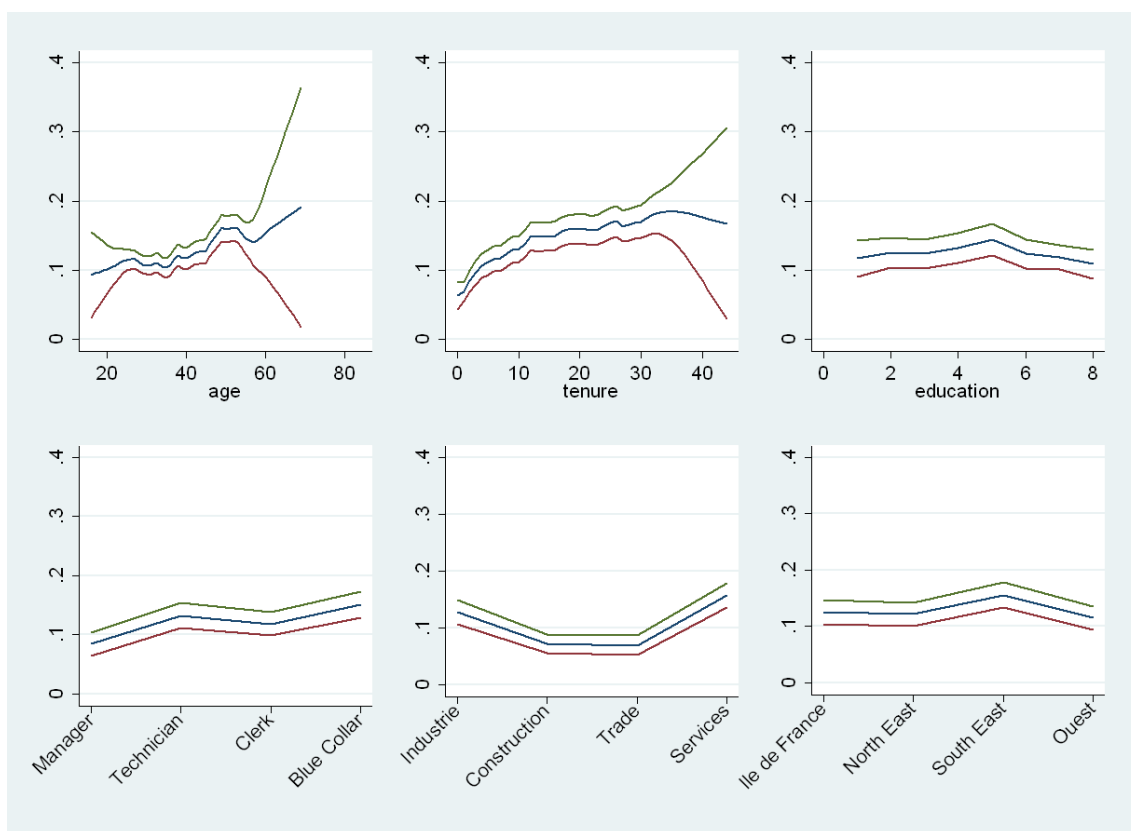
Descriptive statistics: distribution of current and past unionized workers in terms of their observable characteristics (weighted).

	Currently Unionized	Previously unionized
Average	7,4%	13,3%
<b>Gender:</b>		
Men	8,6%	15,1%
Women	5,6%	10,3%
<b>Occupation:</b>		
Blue collar	8,2%	16,4%
Clerk	6,0%	9,6%
Technician	8,9%	14,1%
White collar/manager	5,5%	9,8%
<b>Diploma:</b>		
None	8,2%	14,2%
Less than Bac	8,5%	17,1%
Bac	7,6%	9,7%
More than bac	5,2%	7,5%
<b>Working time:</b>		
Part time	6,2%	10,7%
Full time	7,5%	13,6%
<b>Sector:</b>		
Industry	9,4%	17,4%
Construction	3,7%	13,5%
Trade	3,6%	9,6%
Services	8,9%	13,4%
<b>Region:</b>		
Ile de France (Paris)	8,1%	10,0%
North East	7,3%	16,5%
South East	10,2%	14,7%
Ouest	7,0%	12,9%

Distribution of workplaces regarding the number of unionized and non-unionized workers in the survey.

		Number of unionized workers						Total
		0	1	2	3	4	6	
Number of non unionized workers	0	258	57	25	8	3	0	351
	1	432	122	37	10	4	1	606
	2	541	145	49	9	1	0	745
	3	420	126	31	5	1	0	583
	4	288	56	11	1	0	0	356
	5	159	31	5	0	0	0	195
	6	53	11	0	0	0	0	64
	7	19	0	0	0	0	0	19
	8	7	1	0	0	0	0	8
	9	0	1	0	0	0	0	1
	11	1	0	0	0	0	0	1
Total		2,178	550	158	33	9	1	2,929

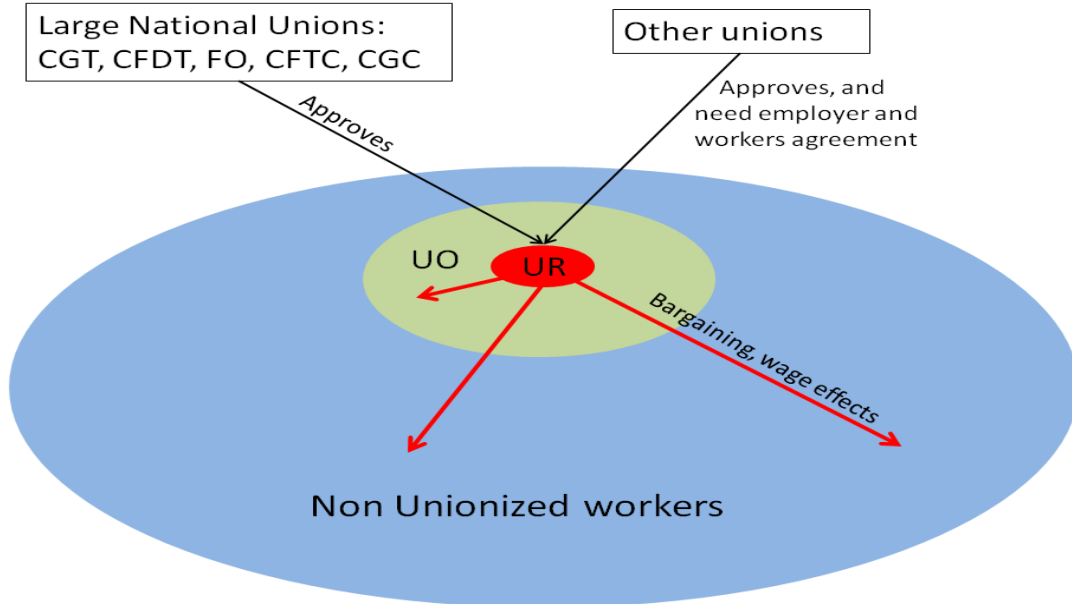
Probability to be unionized as a function of observed characteristics (estimate and confidence interval)



Obtained via simultaneous smoothing using non parametric estimation.

# Tables and figures:

Figure 1: Within firm industrial relations



UO: Unionized Only, UR: Union Representative.

Notice that proportions have been respected: the relative size of the bloc corresponds exactly to the relative share of each type of worker in the data sample.

Figure 2: Link between the variables  $U_{ij}$  and  $UR_{ij}$

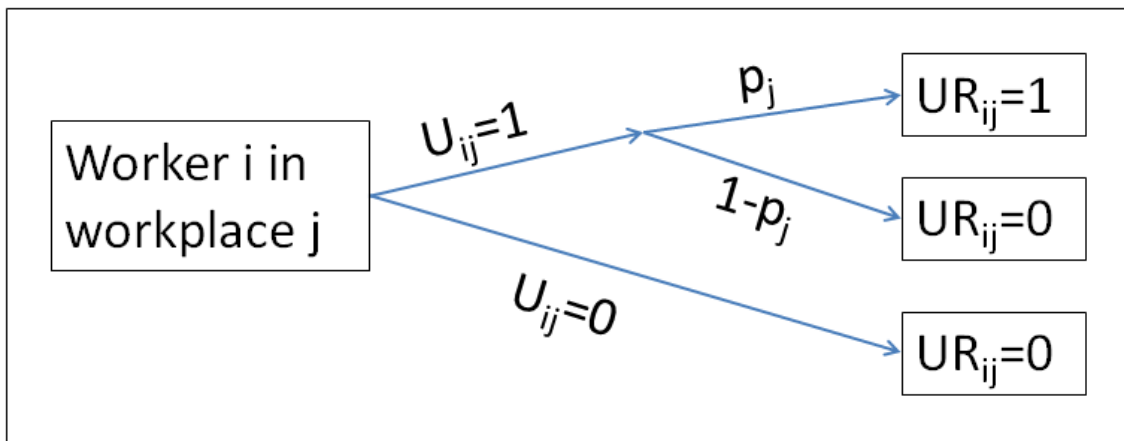
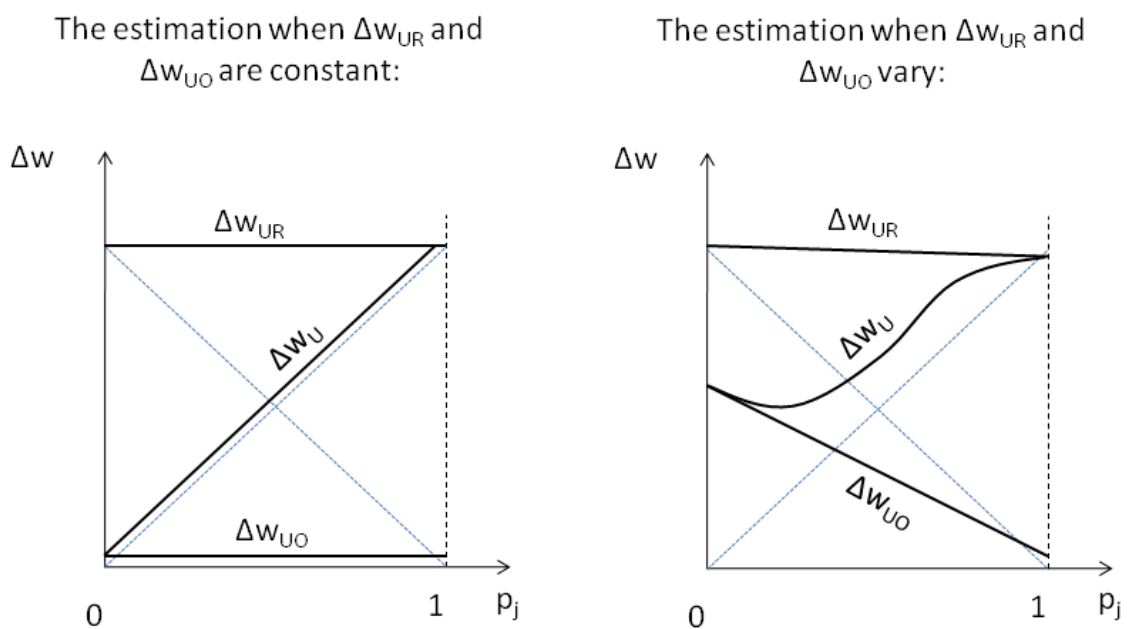


Figure 3: Link between the different wage differentials when  $p_j$  varies:



$\Delta w_U = [p_j \Delta w_{UR} + (1 - p_j) \Delta w_{UO}]$  with  $\Delta w_U$ ,  $\Delta w_{UR}$  and  $\Delta w_{UO}$  denoting respectively the wage penalty for all unionized workers, union representatives and only unionized workers.

Table 1: Wage differential between non unionized and unionized or previously unionized workers, various controls.

	1:FE	1:No FE	2:FE	2:No FE	3:FE	3:No FE
Unionized	-.024*	.068***	-.052***	-.010	-.026***	.014
	(.014)	(.015)	(.011)	(.011)	(.010)	(.009)
Previously unionized	.009	.057***	-.046***	-.010	-.017*	.011
	(.013)	(.015)	(.011)	(.011)	(.010)	(.010)
<b>Control variables:</b>						
a) Age,tenure,educ,gender	No	No	Yes	Yes	Yes	Yes
b) Occup, work. hours	No	No	No	No	Yes	Yes
c) Firms characteristics	-	No	-	Yes	-	Yes
Observations	7836	7836	7741	7689	7356	7309
R2	.64	.004	.79	.54	.85	.66

Control variables in a) are age, age square, tenure, tenure square, 8 education dummies and gender. They are all highly significant in regressions. Controlling more finely for age or tenure (for example a dummy per age and tenure year) only modifies the results marginally.

Control variables in b) are 4 dummies for blue collars, clerks, technicians and managers, a dummy for part time workers and the number of hours worked per week. Controlling more finely for occupation (24 groups) only modifies the results marginally.

Control variables in c) are 8 region dummies, 16 sector dummies, 6 dummies for the workplace size, 5 dummies for the workplace age and a dummy for the presence of a union representative. Controlling more finely for size and age only affects the results marginally.

\*: significant at the 10% level. \*\*: significant at the 5% level. \*\*\*: significant at the 1% level.



Figure 4: Distribution of the proportion of unionized workers, of the number and proportion of union representatives and of the variable  $p_j$  across workplaces where unions are recognized (e.g. with at least one union representative).

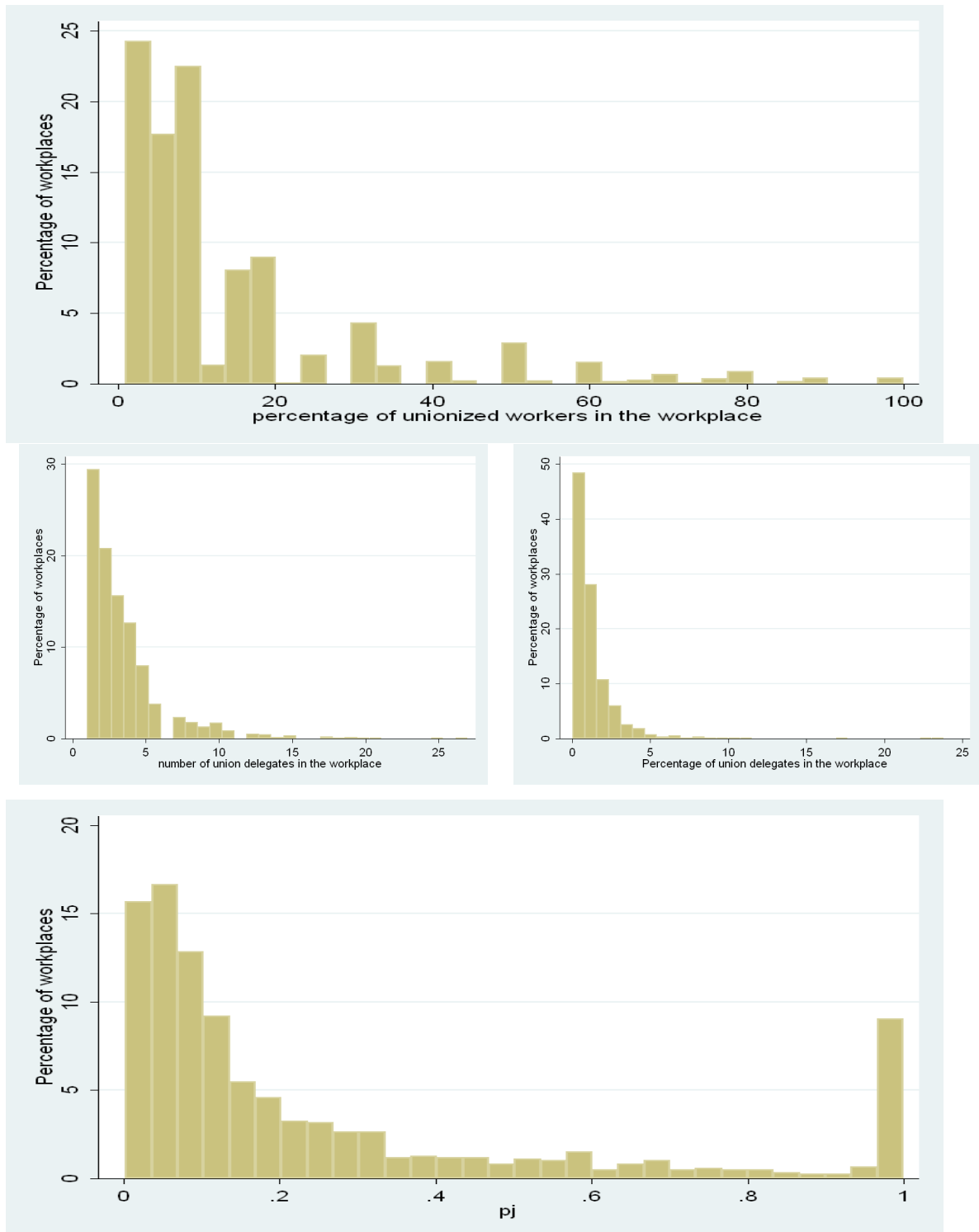


Table 2: IE and ML estimators of the wage differentials between union representatives or only unionized workers and non-unionized workers.

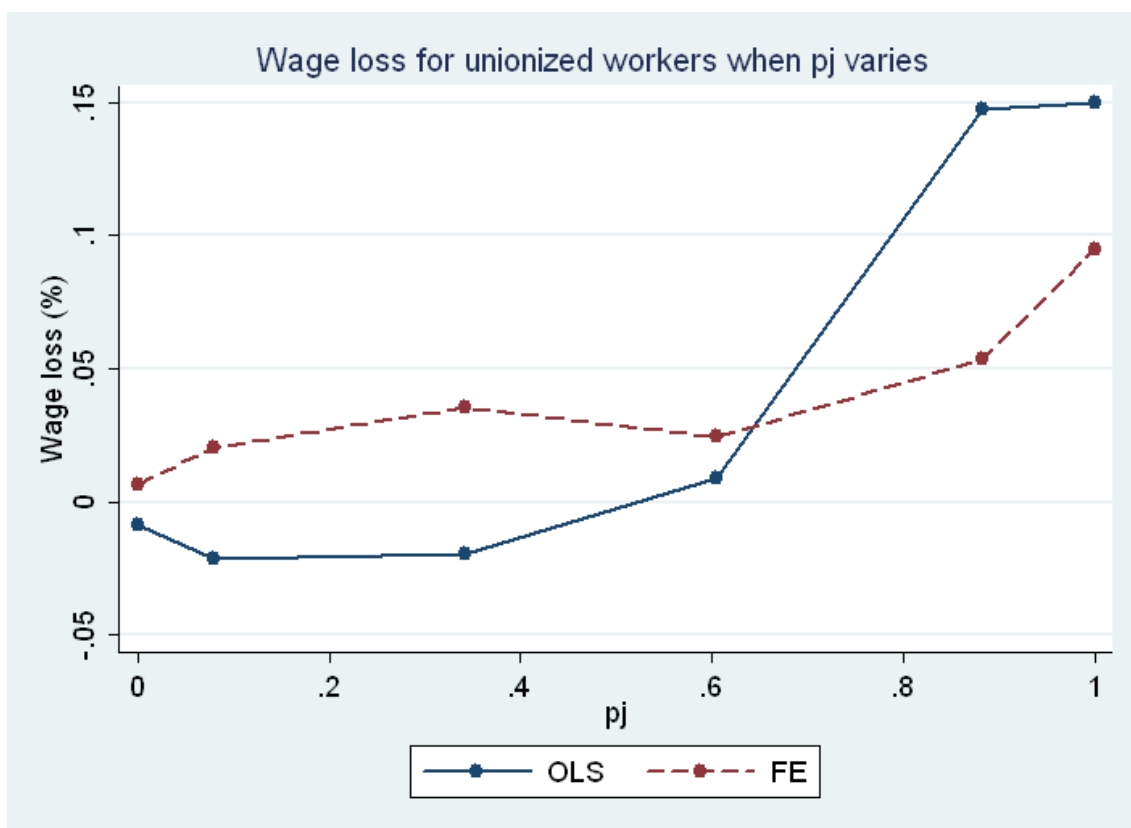
	IE		ML	OLS: $p_j \in \{0.1\}$	
	1:FE	2:No FE	4:No FE	5:FE	6:No FE
<b>Union Representative</b>	-0.079**	-0.113***	-0.101***	-0.095*	-0.149***
<i>st dev</i> <sup>a</sup>	(0.034)	(0.034)	(0.028)	(0.060)	(0.048)
<i>Corrected st dev</i>	(0.036)	(0.034)			
<b>Unionized Only</b>	-0.012	0.036**	0.030**	-0.005	0.010
<i>st dev</i> <sup>a</sup>	(0.013)	(0.015)	(0.013)	(0.030)	(0.028)
<i>Corrected st dev</i>	(0.014)	(0.013)			
Indiv. controls	Detailed	Detailed	Detailed	Detailed	Detailed
Firm controls	No	Yes	Yes	No	Yes
Observations	7123	6729	6729	6530	6144
R2	0.86	0.66	-	0.87	0.66

<sup>a</sup>: In specifications (1) and (2), standard errors have been clustered by groups of workers with the same observable union status (unionized or not) in the same workplace. In specifications (5) and (6), standard errors are robust.

Individual control variables correspond to specification 3 in table 1. They include controls for age, tenure, education, occupation, gender and hours worked. Workplace control variables in non FE estimations are size, age, sector and region dummies.

\*: significant at the 10% level. \*\*: significant at the 5% level. \*\*\*: significant at the 1% level.

Figure 5: Estimation of the wage differential between unionized and non unionized workers when the probability for unionized workers to be union delegate increases



The estimations are obtained by running on 6 subsamples of the dataset OLS regressions that include detailed controls for individual characteristics and either controls for workplaces characteristics or workplaces fixed effects.

Table 3: Estimation of the wage differentials between union representatives and non-unionized workers for workers with long and short tenure

	Indirect Estimation (IE)			
	FE	No FE	FE	No FE
Union Rep. with tenure < 5 years	0.013 (0.068)	-0.029 (0.053)	0.041 (0.119)	-0.050 (0.069)
Union Rep. with tenure $\geq$ 5 years	-0.108*** (0.035)	-0.141*** (0.039)	-0.134*** (0.042)	-0.186*** (0.044)
Indiv. controls	Detailed	Detailed	Standard	Standard
Firm controls	No	Yes	No	Yes
Observations	7123	6729	7494	7085
R2	0.86	0.66	0.79	0.54

Standard errors have been clustered by groups of workers with the same observable union and tenure status (unionized or not, less or more than 5 years of tenure) in the same workplace. A dummy variable for workers with more than 5 years of tenure (not interacted with other variables) has been added as an additional control.

Standard individual control variables are age, age square, tenure, tenure square, education (8 groups) and gender. Detailed individual control variables also include occupation (4 groups), a dummy for part time positions and hours worked. Workplace control variables in non FE estimations are size (6 groups), age (5 groups), 16 industry dummies and 9 region dummies.

\*: significant at the 10% level. \*\*: significant at the 5% level. \*\*\*: significant at the 1% level.

Table 4: IE and ML estimators of the wage differentials between union representatives or only unionized workers and non-unionized workers

	IE		ML
	1:FE	1:OLS	No FE
<b>UR CGT</b>	-0,222***	-0,206***	-0,200***
<i>st dev</i> <sup>a</sup>	(0,063)	(0,076)	(0,056)
<b>UR CFDT</b>	-0,012	-0,123**	-0,066
<i>st dev</i> <sup>a</sup>	(0,077)	(0,056)	(0,052)
<b>UR FO</b>	0,014	0,011	-0,034
<i>st dev</i> <sup>a</sup>	(0,078)	(0,081)	(0,081)
<b>UR other unions</b>	-0,028	-0,025	-0,045
<i>st dev</i> <sup>a</sup>	(0,052)	(0,065)	(0,054)
Indiv. controls	Detailed	Detailed	Detailed
Firm controls	No	Yes	Yes
Observations	7133	6738	6738

<sup>a</sup>: In IE estimations, standard errors have been clustered by groups of workers with the same observable union status (unionized or not) in the same workplace.

Individual control variables are controls for age, age square, tenure, tenure square, education (8 groups), occupation (4 groups), gender, part time, and hours worked. Workplace control variables in non FE estimations are size (6 groups), age (5 groups), 16 industry dummies and 9 region dummies.

\*: significant at the 10% level. \*\*: significant at the 5% level.

\*\*\*: significant at the 1% level.