Discussion Papers No. 535, April 2008
Statistics Norway, Research Department

Erling Barth, Bernt Bratsberg,
Torbjørn Hægeland and Oddbjørn Raaum

Performance Pay and Within-Firm Wage Inequality

Abstract:
This paper examines the impact of performance-related pay on wage differentials within firms. Our theoretical framework predicts that, compared to a fixed pay system, pay schemes based on individual effort increase within-firm wage inequality, while group-based bonuses have minor effects on wage dispersion. Theory also predicts an interaction between performance-related pay and union bargaining, where union power reduces the impact of performance pay on wage dispersion. The empirical contribution utilizes two recent Norwegian employer surveys, linked to a full set of employee records. A longitudinal sub-sample allows for identification based on fixed establishment effects. Introduction of performance-related pay is shown to raise residual wage inequality in nonunion firms, but not in firms with high union density. Our findings suggest that even though performance-related pay appears to be on the rise, the overall impact on wage dispersion is likely to be small, particularly in European countries with strong unions.

Keywords: Performance related pay, wage inequality, union bargaining

JEL classification: J31, J33

Acknowledgement: We are grateful to Carlo Dell'Aringa, Alex Bryson, Donna Ginther and Espen Moen for comments on earlier versions of this paper. We also thank participants at the 2005 SOLE conference in San Francisco, the 2006 workshop on the Use and Analysis of Employer Employee Data in Oslo, and the 2006 Lower Conference in Esbjerg, as well as seminar participants at the Catholic University of Milan and University of Oslo, for useful comments. This work has received financial support from the Norwegian Research Council (grants # 156035/S20 and 173591/S20), the Norwegian Ministry of Labour and Social Inclusion (Frisch project #1391), the EDWIN project of the European Commission IHP project (grant #HPSE-CT-2002-00108), and is part of the research at the ESOP Centre at the Department of Economics, University of Oslo.

Address: Oddbjørn Raaum, Frisch Centre, Gaustadalléen 21, N-0349 Oslo, Norway
oddbjorn.raaum@frisch.uio.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1. Introduction

Performance-related pay and individualized wage setting appears to be on the rise, even in economies with strong unions. Following the publication of the seminal article of Holmström and Milgrom (1987), performance-based pay became a predominant method of rewarding executive managers. But individualized pay and performance-related pay seem to be spreading to lower ranks of employees within organizations, and throughout the labor market as well, see Lemieux et al. (2006) for recent evidence from the US and Kersley et al. (2006) for the UK. Particularly in Europe, where union negotiated, fixed wages have been predominant, this development represents an important break with traditional forms of remuneration. The implications of institutions and pay schemes on pay, productivity, and the distribution of effort are not thoroughly discussed and well understood. This paper develops a theoretical framework for analyzing the effects of performance-related pay on within-firm wage dispersion, and confronts theory with evidence from a rich employer-employee data set. In particular, our focus is on the interaction between performance pay and collective bargaining.

Wage differences among workers with similar qualifications are on the rise as well. Autor et al. (2005) conclude that increased within-group inequality accounts for much of the increase in the US wage dispersion over the past 30 years. Barth and Lucifora (2005), using comparable micro data from 15 European countries over 20 years, find evidence of similar trends in Europe. Different explanations have been put forward, but the relative impact of demand and supply factors versus institutional factors both on levels and trends in wage inequality across countries remains to be determined. DiNardo et al. (1996) conclude that a decline in unionism contributed to a considerable portion of the increase in wage inequality in the United States. Kahn (2000) argues that differences in collective coverage explain a large part of the observed differences in wage dispersion across countries.

An expansion of individualized and performance-related pay systems is yet another possible explanation for a widening wage distribution. In its simplest form, the argument is that wages become more dispersed when pay more closely reflects individual performance heterogeneity, stemming from differences in talent, skills, effort, and even luck. Indeed, Lemieux et al. (2006) conclude that 'growing incidence of performance pay accounts for 25 percent of the growth in [the US] male wage inequality between the late 1970s and early 1990s."

Performance pay may have very different effects depending on the institutional environment in which it is implemented. In particular, the presence of unions and collective wage bargaining may change the conditions for and the design of performance-related pay schemes. Understanding the interplay between pay systems and union bargaining is thus crucial in order to understand the development of pay setting regimes, particularly in European labor markets where unions are prevalent, as well as to understand differences in levels and changes in within-group wage dispersion across economies. The theoretical literature on performance pay typically considers employer strategic behavior in the context of incomplete information about worker characteristics and/or behavior. Our theoretical model of performance pay includes a regime with union bargaining in order to derive hypotheses on the interaction effects between these two types of wage setting. Three features of unions are particularly important in this context.

First, unions appropriate rent. Blanchflower and Bryson (2003) provide comprehensive evidence on the evolution of union-nonunion wage gaps in the UK and the US. This effect is usually thought of as a negative monopoly face of unionism (Freeman and Medoff, 1984). The existence of a union wage premium also implies, however, that workers may obtain part of the productivity gain caused by performance-based incentives. Furthermore, in the presence of rent, the wage policy of the firm is not necessarily constrained by outside options (i.e., a participation constraint), but rather limited by the bargaining power of the union. This feature of the employment relationship, that unions appropriate rent, has important implications for the design of the optimal pay scheme of the firm.

Second, unions tend to compress wages. Wage dispersion is typically lower in unionized environments, see, e.g., Freeman and Medoff (1984) and Card et al. (2004) for evidence. Since performance pay may induce differences in effort and productivity, as well as pay, it is important to understand how the wage compressing mechanisms of unions operate. Below, we consider the wage compressing effects of a utilitarian union that represents the interests of risk-averse employees. Risk aversion implies that the union places more weight on low-paid workers than on high-paid workers, providing collective preferences that favor pay equality.

Third, unions influence information flows. Freeman and Medoff (1984) argue that unions provide workers with a "collective voice" that improves the information flow within the firm. In particular, unions may voice grievances or preferences about workplace conditions that individual workers are reluctant to put forward. In the agency literature, however, the important information problem is lack of observability (or at least verifiability) of effort. In the present paper, we argue that a union may operate as a monitoring device.[1] It is easier (or less costly) for workers than for management to observe each others' daily work. A higher level of trust among workers facilitates information sharing across employees. Consequently, the union may be able to enforce a given job standard, without substantial monitoring costs. The problem for the firm, then, is how to provide the union with incentives to keep a certain job standard. A group bonus may serve this purpose. The advantage of a bonus based on collective, rather than individual, outcomes is that risk is shared among workers. Below we analyze both individual and group-based performance pay schemes. It turns out that in the union bargaining case, the trade-off between individual and group-based pay schemes depends on the relative importance of risk sharing within the group versus the loss associated with enforcing a common job standard.

Our empirical contribution identifies the effect of introducing performance-related pay on the conditional within-establishment wage distribution, using a representative panel of two employer surveys from Norway, matched with the full set of employee records. The data allow us to use a comprehensive set of controls, both at the worker and at the firm level. Furthermore, we obtain identification by the use of establishment fixed effects. Controlling for establishment effects allows us to sweep out the permanent effect of worker sorting across firms as well as firm-specific heterogeneity, two issues that are of great concern in analyses of the relationship between pay and remuneration schemes.

There exists only a rather small literature analyzing the effects of performance-related pay on productivity and wage inequality. A few important results stem from case studies. For example, Lazear (2000) finds that the variance in output across individuals rose when the Safelite Glass Corporation shifted to piece rates. Among large scale analyses based on representative data sets, Belfield and Marsden (2003) find a positive association between the use of performance related pay and a crude measure of pay inequality. Lemieux et al. (2007) use US data, drawn from the PSID and the NLSY, to sort out the effect of performance pay on recent developments of the US wage structure. They find that growth in performance pay can explain a significant fraction of the increase in male wage inequality that occurred between the late 1970s and the early 1990s. Few studies link unionism and performance pay, and the empirical evidence from these studies appears inconclusive. While Brown (1990) and

[1] See Pencavel (1977) for an early discussion of such a mechanism. According to Metcalf (2003), the Donovan Commission (1968) study of UK shop stewards emphasized the shop stewards' role in communication, information, and discipline of workers. See also Vroman (1990), who uses such a mechanism to explain the cyclical behavior of the union-nonunion wage differential in an efficiency wage model.

Heywood et al. (1997) find less use of performance-related pay in unionized establishments, Booth and Frank (1999) and Booth and Francesconi (2000) uncover positive associations between union status and performance pay.

In the next section we provide a theoretical analysis of the relationship between wage inequality and remuneration schemes under bargaining and unilateral decisions on part of the employer. In the following section, we detail our data. Section 4 presents the empirical analyses, and section 5 concludes.

2. A Theoretical Model

Our point of departure is a simple principal agent model in which performance pay is introduced in order to induce effort from the worker.[2] To analyze the effects of payment schemes on wage dispersion within firms, we introduce individual heterogeneity along two dimensions. Workers differ in observable productivity (α_i) as well as 'efficiency,' defined as the ability to transform effort (z_i) into valuable output. Heterogeneous efficiency across workers is represented by p_i which is an observable exogenously given parameter. The total value of a single worker's output is given by

(1) $$Y_i = \alpha_i + p_i(z_i + \varepsilon_i), \qquad \varepsilon_i \sim N(0, \sigma^2)$$

where luck is measured by ε_i. The agency problem arises because it is hard to distinguish whether changes in output is caused by effort (z_i) or luck (ε_i). Unless costly monitoring devices are implemented, the employer only observes the sum of the effort-induced contribution and the random component. Effort z_i is chosen optimally by the worker, subject to constraints that vary across pay schemes. We assume that p_i has a log-normal distribution across workers;

(2) $$\ln p_i \sim N\left(-\gamma^2/2, \gamma^2\right)$$

where the normalization ensures that the expectation of the efficiency distribution across workers is unity;

(3) $$E(p_i) = 1 \quad , V(p_i) = \exp(\gamma^2) - 1 .$$

The parameter γ captures the heterogeneity of effort efficiency across workers, and an increase in γ represents a mean-preserving spread of the efficiency distribution. We assume

[2] See Holmström and Milgrom (1987). For textbook expositions, see, e.g., Lazear (1995) and Cahuc and Zylberberg (2005).

that employers observe productivity (α_i) and effort efficiency (p_i) of each worker. Thus, wage contracts can be specified conditional on these characteristics.[3]

The employer may observe individual effort (or equivalently, ε_i) by means of a costly monitoring scheme implemented through a set of rules or conditions related to the job, rather than to particular individuals, taking the form of a (common) standard for all workers. With monitoring, the employer is able to enforce a common effort standard ($\bar{z}$) at a cost, M, where[4]

$$(4) \qquad M(\bar{z}) = n\lambda\bar{z} .$$

The parameter n denotes the number of workers at the firm and $\lambda \geq 0$ is the marginal cost of measuring (in a verifiable way) effort. A positive λ implies that it is more costly to monitor and verify workers' effort the higher is the level of effort. For example, it may be inexpensive to verify a worker's presence at work, but costly to verify the work intensity or concentration provided during this presence.

We consider the following simple linear pay scheme:

$$(5) \qquad W_i = \omega_i + b_i p_i (z_i + \varepsilon_i)$$

where ω_i is a fixed component of the wage, and b_i is the piece rate (or the worker's share of the effort-induced productivity). We allow the share to vary across workers, but we will show shortly that it is optimal to have a common value for all. A fixed pay regime (*FP*) means that $b_i = 0$, while a performance pay regime (*PP*) is characterized by $0 < b_i \leq 1$.

Workers are risk averse and care about both wages and effort. Individual utility is given by

$$(6) \qquad U_i = -\exp\left[-a\left(W_i - C(z_i)\right)\right]$$

where $a>0$ is a measure of risk aversion, and $C(z_i)$ is the cost of providing effort, z_i. We assume convex effort costs, given by:

$$(7) \qquad C(z_i) = \frac{1}{2} c z_i^2$$

where c is a positive shift parameter of the marginal effort cost function. With performance pay, worker (ex post) utility depends on the normally distributed random component ε_i.

---

[3] As we show below, this means that the firm will set the wage so that the expected utility of workers equals their reservation level, thus implying that sorting of workers across firms and pay regimes may be considered to be random. This is in contrast to a model where pay for performance is introduced in order to attract more efficient workers (see, e.g., Lazear, 1995).

[4] In principle, it may be optimal for the firm to let the effort requirement vary across workers. At the same time, monitoring technology will make it less costly to impose a common standard. We assume that the latter effect dominates.

Substituting in from equations (5) and (7), and taking the expectation of equation (6), the expected utility of an individual worker is given by

$$EU_i = -\exp(-a\Psi_i)$$

(8) $$\Psi_i = \omega_i + b_i z_i - \frac{1}{2}cz_i^2 - \frac{1}{2}b_i^2 p_i^2 a\sigma^2$$

Note that $E[\exp(-ab_i p_i \varepsilon_i)] = \exp[-(1/2)ap_i^2 b_i^2 \sigma^2]$ since $\varepsilon_i \sim N(0,\sigma^2)$

2.1 The Nonunion Firm

The firm maximizes expected profits (π), given by

(9) $$\pi = \sum_{i=1}^{n}[\alpha_i + (1-b_i)p_i(z_i + \varepsilon_i) - \omega_i] - \lambda n\bar{z},$$

and chooses the optimal remuneration scheme subject to a set of participation and incentive constraints. Participation requires that the expected utility of each worker matches her outside option, X_i, (i.e., $EU_i = -\exp(-a\Psi_i) \geq -\exp(-aX_i)$). In general, the outside option is increasing in the observed productivity component, α_i, but we do not need to specify the relationship in detail in order to proceed.

The incentive constraints differ by pay regime. When the firm does not monitor effort (and sets $\bar{z} = 0$), the worker chooses effort to maximize expected utility, subject to the remuneration scheme. The optimal effort for the worker is given by $z_i = b_i p_i / c$. (We return to the determination of b_i below.) If the firm chooses a common effort standard $(\bar{z} > 0)$ that is verified by the monitoring technology, a fixed pay contract will be preferred.[5]

Case 1: Fixed Pay (FP)

In the fixed pay regime ($b_i = 0$), effort is determined by a common standard and the firm chooses a set of individual-specific fixed wages, ω_i, to match the workers' outside options. An optimal effort standard ($\bar{z}$) maximizes profits subject to the participation (outside option) constraint. The optimal effort standard follows from equalizing marginal gains and marginal costs, which in the Appendix is shown to yield

(10) $$\bar{z} = \frac{(1-\lambda)}{c}.$$

[5] This follows from the fact that a positive b_i requires some wage compensation for risk, while it does not contribute to effort, which in this case is determined by the common job standard.

The marginal revenue per worker, net of monitoring costs, from raising the job standard is equal to (1-λ). When the firm raises the job standard, it must also increase wages to compensate for higher effort. The average marginal cost per worker of increasing effort is simply given by $c\bar{z}$. Note that the optimal effort standard is decreasing in both monitoring costs and average effort costs. The fixed pay is given by

(11) $$W_i = X_i + \frac{1}{2c}(1-\lambda)^2$$

With fixed pay, workers receive a wage that matches their outside option plus a compensation for effort which is common to all workers. Thus, the within-firm wage dispersion (as measured by the variance of *W*) reflects worker heterogeneity in outside options only:

(12) $$Var(W) = Var(X).$$

Under fixed pay there is no impact of the luck component, simply because the firm bears all of the risk.

Case 2: Performance Pay (PP)

Instead of monitoring workers, the firm may choose b_i and ω_i to maximize profits, subject to the incentive and participation constraints. The optimal piece rate, as shown in the Appendix, is given by

(13) $$b_i = b = \frac{1}{1+ac\sigma^2}.$$

In line with the literature, the optimal piece rate is decreasing in risk aversion and the marginal cost of effort. It may seem surprising that b is independent of the efficiency of workers (p_i). It is well known, however, that the optimal power of the incentive (b) arises from a trade off between risk compensation and the value of the effort induced. In our model the efficiency of a worker (p_i) increases both the necessary compensation for risk and the value of increased effort proportionally, and thus it does not affect the trade-off between the two. The individual wage is given by

(14) $$W_i = X_i + \frac{1}{2c}bp_i^2 + bp_i\varepsilon_i$$

where the deterministic part includes the outside option and a term which represents the share of the effort-induced output received by the worker, minus effort costs plus the compensation for risk since $\frac{(bp_i)^2}{c} - \frac{c}{2}\left(\frac{bp_i}{c}\right)^2 + \frac{1}{2}b^2p_i^2a\sigma^2 = \frac{1}{2c}bp_i^2$

In contrast to the fixed pay regime, there is i) a risk to be compensated and ii) the worker receives a fraction of the luck component. Note also that pay is increasing in individual efficiency. More efficient workers provide more effort for which they receive a larger compensation. The stochastic component in (14) depends on the chosen piece rate and individual efficiency. If we let $(1/2c)p_i^2 + p_i\varepsilon_i = q_i$, the wage dispersion within the nonunion firm under performance pay can then be written

(15) $$Var(W) = Var(X) + b^2 Var(q) + 2b\,\mathrm{cov}(X, q)$$

First, outside option variation is reflected in the wage structure. Second, efficiency heterogeneity implies that wages differ within firms because workers respond differently to incentives and do not provide the same effort. Finally, when the most efficient workers also have better outside options, the covariance term in (15) is positive and contributes to larger within firm wage inequality.

Compared to the fixed pay regime, three differences from introducing performance pay stand out. First, wage differentials arise from effort heterogeneity. More efficient workers work harder and they are paid more. Second, wage differentials are reinforced under performance pay due to the positive correlation between efficiency and the outside option. Finally, variability in the unobserved luck component contributes to wage dispersion. The higher the share (*b*) of productivity that is captured by the worker, the larger is the wage dispersion within the firm. Wage inequality is greater under high-powered incentives.

2.2 The Unionized Firm with Collective Bargaining

The workers bargain collectively (by means of a union) with the firm over parameters of the pay structure and in the case of fixed pay, the effort standard. We consider Pareto optimal outcomes only, which would be the outcome if we assumed Nash bargaining. As long as the relative bargaining power remains the same over different bargaining issues covered by the collective agreement, the outcome is typically efficient. Thus, the outcome may be characterized by maximizing profits for a given level of utility on part of the union, and must satisfy the first order conditions of the Lagrangian (*L*)

(16) $$L = \sum_{i=1}^{n}\left(\alpha_i + (1-b_i)p_i(z_i+\varepsilon_i) - \omega_i - \lambda\bar{z}\right) - \chi\sum_{i=1}^{n}\left(-exp(-a\Psi_i) - \bar{u}\right)$$

where χ is the Lagrange multiplier and $\bar{u}$ defines the average given worker utility level.

Case 3: Union Fixed Pay (UFP)

Like without unions, fixed pay contracts have $b = 0$ and common effort standard. From the first order condition of (16), shown in the Appendix, the optimal effort standard is the same as without union presence, $\bar{z} = (1-\lambda)/c$, unaffected by union power, reflected in ($\bar{u}$). The lack of influence of collective bargaining on the optimal common effort follows from the efficient bargaining assumption, in which the parties agree on an effort standard that maximizes the pie (revenue net of effort and monitoring costs). The surplus shares are then determined by negotiations over the fixed component of the pay contract, ω_i. The fixed-pay wage structure under collective bargaining is given by:

$$\begin{aligned} &W_i = E(W) = \bar{\Psi} + \frac{1}{2c}(1-\lambda)^2 \\ &Var(W) = 0 \end{aligned} \tag{17}$$

where $\bar{\Psi} = -\ln(-\bar{u})/a$. To maximize the collective utility of union members, wages are equated for all workers. The individual participation constraints are now replaced by a union utility constraint. Compared to the nonunion case, the union ensures that potential differences arising from outside option heterogeneity are equalized within the firm. The equalizing effect of the union arises from fact that, with equal effort, a risk-averse union will place more weight on low-pay groups. The most efficient way of satisfying the union utility constraint is thus to provide equal pay. Clearly, this result hinges on the assumption that $\bar{\Psi} > X_i \;\forall\, i$, which means that the union is strong enough to ensure that all workers receive a positive rent.[6]

Case 4: Union Individualized Performance Pay (UPP)

With individualized performance pay, there is no monitoring and individual workers choose effort optimally, $z_i=b_ip_i/c$. Individual worker utility is given by

$$EU_i = -\exp(a\Psi_i),\ where\ \Psi_i = \omega_i + (1/2c)b_i^2 p_i^2 \left(1 - a\sigma^2 c\right).$$

The Lagrangian then takes the form

$$L = \sum_{i=1}^{n}\left(\alpha_i + (1-b_i)(p_i^2 b_i / c) - \omega_i\right) - \mu \sum_{i=1}^{n}\left(-\exp(-a\Psi_i) - \bar{u}\right) \tag{17}$$

which is maximized with respect to b_i and ω_i *(i=1,..n)*. The first order conditions imply that the optimal piece rate is the same for all workers and equal to that of the nonunion case, see

[6] Alternatively, one might consider a situation where the binding constraint is the outside options of the highest paid workers, i.e., where low-paid workers rely on coordinated action, while the most productive workers are paid according to their market value. In this case a kinked wage schedule in X_i, with wage compression in all but the upper part of the distribution, would result.

Appendix. Again, the average productivity level is unaffected by the bargaining power of the two parties, just as in the fixed pay regime, because of the efficient bargaining assumption. The wage structure is given by

$$\begin{aligned} &W_i = \overline{\Psi} + \frac{1}{2c} b p_i^2 + b p_i \varepsilon_i == \overline{\Psi} + b q_i \\ &Var(W) = b^2 Var(q) \end{aligned} \tag{18}$$

The wage has two components that are common to all workers in the firm, determined by the bargaining power of the union and the compensation for risk. More efficient workers provide greater effort and are therefore compensated so as to equalize marginal utility across workers. Finally, the wage captures part of the luck component.

Wage dispersion is lower than in the nonunion case with performance pay, because variation in the outside option does not spill over to wage differentials within the firm. Moreover, because outside options are non-binding, there is no positive association between effort compensation and wage-driving outside factors.

Case 5: Group Performance Pay (GPP)

As discussed in the introduction, one possible role of unions may be to provide a form of self-monitoring mechanism. We assume that the union is willing to monitor effort only if the standard is the same for all workers, $z_i = z$, and provided that the reward system leaves workers with the same expected utility, $\Psi_i = \widetilde{\Psi}$. We justify these assumptions on the grounds that the unions cannot be expected to treat its members differently, particularly when cooperation among workers is required. Since the verification problem facing the employer is the same as before, the effort standard that the union is willing to enforce has to be collectively optimal for the workers. In order to induce the union to enforce a higher level of effort, the employer may use a group piece rate (denoted β to distinguish the notation from the individual piece rate, b).[7] With a group bonus, the pay is the sum of a fixed wage and the share of the common effort as well as the average luck; $W_i = \omega_i + \beta(z + \overline{\varepsilon})$.

(Note that $\frac{1}{n}\sum_{i=1}^{n} p_i z = z$ and $\frac{1}{n}\sum_{i=1}^{n} p_i \varepsilon_i$ is labeled $\overline{\varepsilon}$.)

Expected utility is a function of

[7] Note that the group bonus could have been an option in the nonunion case as well. However, in the absence of union monitoring, the free rider problem implies less effort by each individual with group rates than with individual piece rates.

$$\Psi_i = \omega_i + \beta z - \frac{z^2}{2c} - \frac{1}{2n}\beta^2 ac\sigma^2 .$$

The optimal effort standard, according to union preferences, is determined by the condition

$$(19) \qquad \frac{\partial \sum -e^{-a\Psi_i}}{\partial z} = \sum ae^{-a\Psi_i}(\beta - cz) = 0$$

implying that

$$(20) \qquad z = \beta / c \text{ , when } \Psi_i = \tilde{\Psi} .$$

Note that the optimal effort standard is independent of the fixed wage component, ω_i. This result follows from mechanisms similar to those of the traditional performance pay model. It enables the parties to agree on a bonus rate which elicits effort, while the fixed component can be used to smooth utility across workers. The common effort is increasing in the bonus parameter and decreasing in the marginal cost of effort, as with the traditional performance pay schemes.

With a group bonus, the bargaining problem maximizes the Lagrangian

$$(21) \qquad L = \sum_{i=1}^{n}\left(\alpha_i + \frac{p_i\beta}{c} - \omega_i - \frac{p_i\beta^2}{c}\right) - \mu\left(\frac{1}{n}\sum -\exp(-a\Psi_i) - \bar{u}\right)$$

and it follows from the first order conditions shown in Appendix that the optimal group bonus rate is

$$(22) \qquad \beta = \frac{1}{1 + ac\sigma^2 / n}$$

The group rate is increasing in *n*, providing higher powered incentives in the group bonus case than in the case with individualized performance pay. This result is a reflection of the risk pooling across union members. The wage structure is given by:

$$(23) \qquad \begin{aligned} &W_i = \omega_i + \beta(z + \bar{\varepsilon}) = E(W) = \bar{\Psi} + \beta(\frac{1}{2c} + \bar{\varepsilon}) \\ &Var(W) = 0 \end{aligned}$$

With a union bargaining group bonus, individual productivity as well as outside option heterogeneity, are leveled. The outside options are assumed to be non-binding and all workers supply a common effort standard. Finally, the average luck shock is the basis for bonuses which are equal for all workers. Thus, all workers are paid the same.

2.3 Summary of Wage Inequality across Regimes

Wages may differ within firms because of variation in outside options (partly due to observable productivity), heterogeneity in effort efficiency, and luck. Individual performance pay increases the within-firm wage dispersion since individual output luck affects wages. Incentives at the individual worker level also creates wage differentials under performance pay because more efficient workers are compensated for higher effort. In a nonunion environment, different outside options will generate wage dispersions across workers in the same firm, even in a fixed pay regime. With performance pay, the effort compensation is positively correlated with outside options, contributing to additional wage variability compared to fixed pay. Thus, when the firm has the power to unilaterally set wages, within firm dispersion is higher with performance pay.

A union with power to deliver rents to its members contributes to more compressed wages, appearing in our model as an equalization of expected utility across individual workers. With individualized performance pay, union bargaining will also cut the association between efficiency and the outside option as drivers of wages. Thus, our framework unambiguously predicts that (i) performance pay will reduce wage dispersion under union bargaining, but (ii) the impact of introducing performance pay is smaller in the union case than in the nonunion case. Finally, wage dispersion within firms with group-based performance pay systems (where an effort standard is monitored by the union) is likely to be similar to that found under fixed-pay collective bargaining.

2.4 Profits across Regimes

When the pay components are set unilaterally by the firm, as in the nonunion (subscript *nu*) case, expected profits are equal to

$$\pi_{nu}^{PP} = \overline{\alpha} - \overline{X} + \frac{\exp(\gamma^2)}{2c}\frac{1}{1+ac\sigma^2} \tag{24}$$

and

$$\pi_{nu}^{FP} = \overline{\alpha} - \overline{X} + \frac{(1-\lambda)^2}{2c} \tag{25}$$

under performance and fixed pay, respectively. Thus, profits are proportional to average output arising from worker effort. In the fixed pay case, profits are net of monitoring costs What determines the preferred regime for the firm? First, if monitoring costs are sufficiently large, profits are higher with performance pay. Second, if effort efficiency becomes more dispersed, the average effort is raised (even if the optimal share, *b*, is independent of γ). Thus, with greater worker heterogeneity, performance pay is more likely to be preferred by the firm. Finally, factors that reduce the optimal bonus share (b), such as risk aversion (a) and dispersion of random shocks (σ^2), also make performance pay less attractive because average effort and thereby productivity is reduced.

With union presence and efficient bargaining, the ranking of profits is the same (for a given level of union utility), since

$$(26) \qquad \pi_u^{PP} = \overline{\alpha} - \overline{\Psi} + \frac{(1-\lambda)^2}{2c}$$

$$(27) \qquad \pi_u^{PP} = \overline{\alpha} - \overline{\Psi} + \frac{\exp(\gamma^2)}{2c}\frac{1}{1+ac\sigma^2}$$

Even in the presence of unions, individual performance pay is likely to be preferred over fixed pay if monitoring costs are sufficiently large, workers differ in terms of efficiency, risk aversion is low, and if individual productivity is not too heavily influenced by random events.

With a group-based bonus, expected profits are again proportional to average productivity which is equal to the common output standard monitored by the union;

$$(28) \qquad \pi_u^{GP} = \overline{\alpha} - \overline{\Psi} + \frac{1}{2c}\left(\frac{1}{1+ac\sigma^2/n}\right)$$

Profits are highest with group-based performance pay if and only if average effort is higher than in the other regimes. Compared to fixed pay, a group-based bonus is likely to be more profitable if monitoring costs are large, risk aversion is low, and if individual productivity is not too heavily influenced by luck. Efficiency dispersion (γ) has no impact in a group-based bonus regime. Consequently, a more heterogeneous workforce in terms of effort efficiency will raise profits under individual performance pay and make group-based bonuses relatively less attractive.

2.5 The Impact of Performance Pay: Empirical Identification

Our theoretical model clearly suggests that the firm's choice of pay system is endogenous. Firms self-select into pay regimes partly on the basis of worker and firm characteristics which themselves affect within-firm wage dispersion. For example, the optimal pay system depends on the monitoring technology available to the firm. Furthermore, firms with heterogeneity in worker efficiency will tend to implement individual performance pay. This means that firm-specific heterogeneity can be a predictor of performance pay rather than a result of it. A careful assessment of the impact of performance pay on wage variability thus requires linked employer-employee data. It obviously also requires a strategy for identification, since cross-sectional comparisons of firms or individuals may confound the heterogeneity of firms and workers with the effects of performance pay.

Another complexity is that workers may self-select into firms with different pay schemes. In our model, the firm is able to compensate workers with different efficiencies exactly through the fixed component of the wage. This observation may justify the empirical assumption of random sorting of workers across firms. In practice, however, this may be problematic, as individualized compensation may be less than perfect, inducing self-selection of more efficient workers into firms with performance-based pay systems (see, e.g., the discussion in Lazear, 1995).

In our empirical analyses, we rely on fixed-effects estimators for identification. Because we have panel data of both firms and workers, we are able to estimate the effect of introducing performance-related pay from observing changes in wage dispersion in firms that switch pay regime, both with and without restricting the analysis to workers who stay with the firm. In this way we control for time invariant heterogeneity across firms, and at the same time we are able to adjust for the impact of self-selection of workers.

Ideally, we would like to model the switching across pay regimes, and instrument for the performance pay indicator. According to our model, plausible instruments could come from such factors as monitoring technology, idiosyncratic risk, risk aversion, or effort heterogeneity. However, we have not been able to find convincing exclusion restrictions, and thus rely on the fixed-effects methodology. One way to think about the introduction of performance pay schemes, then, is that the switchers represent the marginal firms who happen to change their pay practice during our observational window, for instance as a result of influences such as a change in the management group, labor relations, or even developments in the management literature. In contrast, a cross-sectional analysis would look across the total population of firms with and without performance pay schemes, comparing firms with different incentives for their particular pay regime.

3. Data Sources, Samples, and Variable Construction

The core of our data material consists of the Norwegian Flexibility Survey 1997 and the 2003 Norwegian Work and Establishment Survey. Both surveys were carried out as computer assisted telephone interviews with either the manager or the chief personnel officer of the establishment. In both surveys, random, but stratified (with respect to establishment size, age, and sector), samples were drawn from the population of Norwegian establishments with more than 10 employees. Questions concerning employees typically related to the "main occupational group" at the establishment. In addition, the survey data were matched with detailed data about the establishment and each of its employees taken from various administrative registers. The register data are annual and cover the period 1995-2003.

The response rates of the surveys were 76 percent in 1997 and 77 percent in 2003. In the present study, we focus on private sector establishments. Employees working fewer than 30 hours per week are excluded from the analyses, as are firms with less than ten individual observations in the full-time worker data set (as the within firm differentials are poorly defined in these firms). This leaves us with 2,406 observations of 1,751 establishments with valid data in 1997 or 2003. Of these, 655 establishments are represented in both surveys.

Performance-related Pay

Both establishment surveys contained questions about performance-related pay. Unfortunately, these questions were not identical in the two surveys. In 1997, respondents were asked whether or not "the main occupational group receives any pay through incentive pay systems, bonuses, or profit sharing?" In 2003, the survey instrument instead included separate questions about six different forms of performance-related pay; (a) Individual and group piece rates, (b) Commissions, (c) Group bonuses, (d) Profit sharing, (e) Individual bonuses, and (f) Individual performance assessments.

Respondents were also asked to estimate the share of total wages that were associated with each type of performance pay. It seems reasonable to assume that managers who in 2003 answered affirmative on the use of at least one the five former pay types (a)-(e) would have answered "yes" to the 1997 question. It is not obvious, however, how establishments with type (f) "individual performance assessments," would have interpreted the 1997 question. In addition, it is not clear whether the answers refer to permanent or variable elements of compensation. In the empirical analyses, we classify the wage policy at the firm as performance-related pay in 2003 if the manager answered "yes" on at least one of the types (a)-(f). If "yes" on (f) only, its share of total wages must be at least 3 percent.[8] Performance-related pay is found in about half of the firms, see Table 1.

---

[8] The core results in this paper are, however, not sensitive to whether we instead use a "strict" definition of performance pay in 2003, including firms that ticked one of (a)-(e) only, or a "wide" definition also including all firms with (f), individual performance assessments.

Wage Differentials

Information about the firm-level wage structure is taken from administrative register data. Our individual wage measure is constructed by dividing the firm's report of total wages paid to each individual worker through the year, by the number of days of the employment relationship. This daily wage rate is then adjusted by means of a regression of individual log wages on educational attainment (seven levels); age and its square; interactions between each education level and the age polynomial; immigrant status; union membership; gender and interactions between gender and all other characteristics; an indicator variable for year of observation; interactions between the year variable and all other regressors as well as establishment-year fixed effects. This first-step regression provides 292,690 log wage residuals. From the distribution of log wage residuals of individual workers, we compute the 10th, 50th, and 90th percentiles of the residual log wage distribution within each establishment. Average wages are compared across establishments by means of the set of year-specific fixed effects.

As Table 1 shows, the average 90-10 log wage differential in the sample is .697 and it drops to .607 after controlling for individual observed characteristics of the workforce. As expected, the wage differentials are larger in firms with performance related pay. The wage differentials in the balanced panel are similar to what is observed in the combined cross sections.

Union Density and Control Variables

Union density is the proportion of workers at the establishment who are members of a union. We collect this information from the manager surveys. If not available in the survey data, we computed the union density from data on individual payments of union membership dues identified through registers and aggregated to the establishment level. Union density reflects the ability of the workers to coordinate industrial actions, see, e.g., Barth et al. (2001), and to promote collective interests with respect to both the level and dispersion of wages. Bargaining structure and the presence of collective agreements are typically correlated with union membership. The sample average union density is close to 60 percent, and membership is lower in firms with performance related pay. Again, the balanced sample appears to be fairly representative.

We include firm size and age as well as industry affiliation (12 major industries) among the firm-level control variables. From the individual data, we calculate the fraction of female workers and the skill composition (i.e., the fractions of workers with low and high educational attainment) at the firm level.

Our sample is restricted to the private sector. Due to reorganization of former government monopolies, establishments within postal services and the national telecommunications company (*Telenor*) were classified as belonging to the public sector in 1997 and to the private sector in 2003.

4. Empirical Results

This section reports our empirical estimates of how performance-related pay systems affect wage inequality *within* firms, but we first look at pay differentials *between* firms.

4.1 Performance Pay Raises the Mean Wage, but not by much

From the theoretical model, we expect firms with performance related pay to have higher wages. The effect arises from the need to compensate workers for the exposure to risk. Note that we do not necessarily expect productivity to be higher with performance-related pay. The reason is that, when firms can monitor worker effort at low costs, they will fix standards at a higher level of effort than the average effort level induced by the performance-pay scheme. Moreover, imperfect differentiation with respect to the outside option creates a sorting incentive (see, e.g., Lazear, 1995), in such a way that more productive workers tend to sort into firms with performance pay. This sorting would induce a positive correlation between performance pay and wages, even though it was not a causal relation. In order to sweep out the effect of such sorting, it is necessary to control carefully for individual characteristics.

Table 2 shows a positive effect of performance-pay schemes on the firm average wage level. Workers in firms with performance pay have a pay advantage of 9 percent over workers in firms without performance-related pay, see column (1). The wage premium reflects in part better observed individual characteristics of workers in performance-pay firms; when we account for differences in education, age, and other individual characteristics, the estimated pay advantage drops to 6.6 percent (col. 2). When we also account for differences in observed characteristics of firms, such as degree of unionization, firm size, employee composition, and industry, the pay advantage falls even further—to 2.5 percent (col. 3). When we take advantage of the panel dimension of the data and estimate the wage effect of performance-pay schemes including firm fixed effects, the estimate drops even more to 1.8 percent (col. 5).

Note that the panel estimates without fixed effects (col. 4) are very similar to those for the full sample, suggesting that the balanced panel is fairly representative.

A possible concern is that these estimates are influenced by selective mobility among workers. For example, high-performing workers may be attracted to firms with performance-pay schemes, in which case the observed pay advantage of performance-pay firms might reflect sorting of workers rather than a true effect of the pay scheme. A simple test is to re-estimate the model in the sub-sample of employees that worked at the same establishment both in 1997 and 2003, see cols. (6) and (7). Results based on the restricted sample reveal slightly larger effects of performance pay than those from the overall sample. Of particular interest is the estimated coefficient reported in column (7), which is identified from differential wage growth between employees that experienced the introduction (or elimination) of performance-related pay schemes during the sample period and employees that did not. The estimated effect of performance pay based on this sample is 2.8 percent.[9] The lower longitudinal estimates of the performance-pay effect may reflect correlation between performance pay and unobserved firm characteristics with positive wage effects, such as superior organization or management. However, any attenuation bias arising from measurement error in the performance-pay variable will be amplified by the fixed effect specification. Overall, our preferred estimate of the firm wage effect is 2.8 percent.

In Table 3, we examine whether union power and bargaining regime influence the effect of performance pay on firm-level wages. Using union density to summarize the wage setting environment, column (1) suggests that such an interaction does exist. The estimated effect of performance pay in non-unionized firms (i.e., firms with a zero union density) is a 5.4 percent boost in the average wage, while the average wage is largely unaffected by performance pay in fully unionized firms (i.e., 0.6 percent; .054-.048). The empirical evidence on wage differentials *between* firms in Table 3 is, however, not very strong. When we consider the panel sample, introduce firm fixed effects, or when the sample is restricted to employees at the same workplace both years, the results fail to uncover a significant union interaction on the performance-pay effect on the average firm wage. Overall, Tables 2 and 3 reveal that firms with performance-related pay have slightly higher wages than those with fixed-pay schemes only.

---

[9] The restricted sample consists of 79,068 individuals for whom we observe wage payments from the same employer in both survey years. Among these employees, 24.6 percent experienced the introduction and 5.1 percent the elimination of performance pay during the sample period. When we allow for asymmetric impacts of introducing and eliminating performance pay, the estimated coefficient of *"Add performance pay"* is .029 (se=.012) and the coefficient of *"Drop performance pay"* is -.027 (se=.018).

4.2 Performance Pay Raises Within-firm Wage Inequality, but less so in Unionized Firms

Table 4 displays the results from regressions of within-firm wage inequality on firm characteristics, including whether or not the firm has a performance-pay scheme. In column (1) the dependent variable is the 90^{th} -10^{th} percentile differential in the observed log wage of employees at the establishment, while in columns (2)-(6) the differential is based on the log wage residual. As Table 4 shows, wage dispersion is greater in firms with performance pay; the 90-10 differential is .077 log point higher in performance-pay firms according to the estimate in column (1). Some of that is attributable to variation in worker and firm characteristics. When we account for worker characteristics, the coefficient of the performance-pay indicator drops to .056 (col. 2), and when we add firm characteristics the association between performance pay and within-firm dispersion is reduced to about half of that in observed wages (col. 3). Yet, the association remains statistically significant.

Our theory predicts that performance pay has a modest impact on wage dispersion in firms with union bargaining. When we interact performance pay and union density, the effect seems modest in firms with high union membership although the interaction term is only significant at the 10 percent level (see column 4). In firms without union presence, however, performance pay is associated with considerably higher wage dispersion.

In the balanced panel, we find even stronger empirical support for the effect of performance related pay on within-firm wage inequality (col. 5). In firms without unions (in terms of membership), performance-related pay has a significant and positive effect on wage differentials. In the opposite end with full membership, wage differentials are unaffected by the introduction of performance related pay. One might suspect that these results are driven by unobserved firm characteristics correlated with pay system and organization. However, introducing firm fixed effects in the balanced panel sample has only a negligible impact on coefficient estimates of the performance-pay variables, suggesting that unobserved, time-invariant heterogeneity is not the driving force behind these empirical patterns.

Between 1997 and 2003, the sample average of the dependent variable—the within-firm 90-10 log wage differential conditional on observed worker characteristics—rose by 0.069 log point (not reported in tables). During the same period, there was a 12.6 percentage point increase in firms' use of performance pay (Barth et al., 2007).[10] According to the estimates in Table 4, performance pay raises the within-firm 90-10 differential by 0.041 log

[10] This figure agrees with the observed increase in use of bonuses as identified by official wage statistics (Lunde and Grini, 2007).

points. The implication is that the expansion of performance-based pay systems can explain about 7.5 percent of the observed growth in within-firm wage inequality.[11] Although performance pay schemes lead to greater within-firm wage dispersion, only a small share of the increase in wage inequality over the sample period can be attributed to more widespread use of performance pay.

4.3 Performance Pay Affects Wages more in the Top of the Within-firm Wage Distribution—and more so in Nonunion Firms than in Union Firms

Further inspections into the impact of performance pay across the wage distribution are provided in Tables 5 and 6. First, Table 5 reveals that the effects are equally strong at the top (90-50) and at the bottom (50-10) of the distribution. Moreover, in unionized firms, introduction of performance-related pay does not seem to affect the wage structure at all, as the coefficient of the interaction term between union density and performance pay takes the opposite sign and is of similar magnitude to the main effect listed in the top row of the table. Cols. (3) and (6) also reveal that time-invariant firm heterogeneity does not explain this pattern as the coefficient estimates of the performance pay variables from the firm-fixed effects models are basically identical to those from models without fixed effects.

In table 6, we examine the impact of performance pay at various percentiles of the within-firm residual wage distribution.[12] Columns (1)-(3) present analyses of determinants of the 90th percentile wage residual within the firm. Results confirm that the impact of performance-related pay at the top of the wage distribution depends on the degree of unionization within the firm. In nonunion firms the estimated effect is in the range of five to six percent depending on whether or not the regression includes firm fixed effects. Importantly, the effect on top-level wages declines with the unionization rate of the firm. In firms where all workers are union members, the estimate of the performance-pay effect on the 90th percentile wage is close to zero and statistically insignificant. The cross-sectional estimates suggest a significant and negative main effect of union density, supporting the idea that union power compress wages at the top.

The effects on wages in the lower end of the within-firm wage distribution are shown in columns (4)-(6). Although estimated with larger standard errors, a striking pattern is that signs of coefficients from the 10th percentile regressions are opposite those found for the top

---

[11] Based on the average effect in column (3), the calculation is 0.041*(0.126/0.069) = 0.075.

[12] To save space, we only present results based on the balanced panel. Analyses based on the full sample of firms yielded very similar results and are available upon request.

of the wage distribution. Indeed, the estimated coefficients suggest that, in nonunion firms, performance pay has a negative effect on the bottom wages in the firm. Results also show a positive interaction with degree of unionization which is clearly significant in the balanced sample, even with establishment fixed effect. Again, the main effect of higher unionization is to raise the level for the lowest paid, although the evidence is less clear when identified by means of changes in union density (i.e. with fixed effects). As far as the left tail of the within firm wage distribution is concerned, unions seem to raise the wage of those with lowest pay and also prevent any negative effect that performance-pay schemes may have on for these workers.

4.4 Why Is the Effect of Performance Pay lower in Unionized Firms?

Our theoretical model predicts that the design of a performance-related pay scheme differs according to union presence. Utilitarian unions have preferences for equality and can assist the firm in monitoring (and sanctioning lack of) individual effort to sustain group-based incentives. Our data offer a simple distinction between individual (e.g., individual bonuses based on performance) and group-based (e.g., group bonus) performance pay. Unfortunately, this information is not available from the 1997 survey, but the 2003 survey contained detailed questions on various pay components for the major occupational group at the establishment. Here, we define group-based schemes as group bonuses, while other forms of performance pay are classified as individual performance pay. [13]

In Table 7, columns (1) and (2), we report results pertaining to firms' introduction of performance-pay schemes over the sample period. Column (1) addresses the probability of having (any kind of) performance pay in 2003 among firms which did not have it six years earlier. Firms with strong unions are less likely to introduce performance pay; the coefficient on union density is significantly negative and shows that, among firms without performance pay in 1997, fully unionized firms were 21 percentage points less likely to adopt a performance pay schedule over the sample period compared to nonunion firms. Moreover, among those firms that introduced performance pay, group-based schemes were more likely in the presence of strong unions. In column (2), the coefficient on union density indicates that, in the event of an introduction of a performance pay system, fully unionized firms were 24 percentage points more likely to adopt a group-based bonuses compared to firms without union membership,.

[13] The results are very similar if we include profit sharing as a group-based scheme.

Turning next to within-firm wage dispersion, columns (3) and (4) display the coefficient estimates from equations where the 90-10 residual wage differential is regressed on performance pay and the interaction with group-based systems as well as other firm characteristics. According to the coefficient estimates, individualized performance pay raises within-firm wage inequality by about 7 percent, while a group-based system raises inequality by 3 percent. Importantly, the estimated effect of group-based systems is significantly lower than that of individualized pay.

Column (5) represents the quasi-fixed effects model, where the within-firm change in the wage differential from 1997 to 2003 is estimated as a function of the introduction of performance pay, an interaction between introduction and group-based schemes, as well as change in other time-varying firm characteristics. Firms that introduced individualized performance pay experienced a .102 log point increase in wage differentials. The point estimates suggest a considerably more moderate rise in wage differentials within firms that introduced a group-based system, although the interaction effect is only significant at the 10 percent level. Taken together, the evidence in Table 7 shows that the bulk of the effect of performance pay on within-firm wage dispersion stems from individualized pay systems. Group-based bonuses for the main occupational group has less, if any, impact on wage variability within the firm.

5. Conclusions

Theory predicts that wage differentials *between* and *within* firms are linked to the choice of pay system. First, we do not necessarily expect to find higher productivity in firms with performance-related pay. As frequently stated in the literature, monitoring firms with fixed pay systems are likely to implement effort standards higher than the ones chosen by workers in performance-pay firms. Second, theory suggests that the choice of pay system (i.e., pay for performance or fixed pay) is independent of the bargaining power of the union. Third, the introduction of individual performance-related pay is likely to have a much stronger impact on wage inequality in nonunion firms than in firms where union influence is strong. Fourth, wage dispersion in firms with group-based bonuses is likely to be similar to that in firms with fixed pay.

According to our theory, the main effects of performance pay on within-firm wage dispersion arise from two sources: The first effect stems from a change in the distribution of effort within the firm. In a fixed-pay regime with a common effort standard, wages only vary according to observed productivity and outside options. With performance pay, effort and output of the most efficient workers is enhanced, and consequently so is their pay in order to compensate higher effort cost and risk. As the more efficient workers typically have better outside options, the introduction of performance pay also leads to an increase in wages of workers with high pay in the first place. The second effect relates to the distribution of the luck component of the value of output which in the case of individual performance pay is partly passed on to the workers.

In a union bargaining setting, similar mechanisms operate, but they are modified in important ways. First, the union wage compression breaks the link between the distribution of outside options and the internal distribution of wages under performance pay. If a firm introduces individual performance pay, the more efficient workers obtain higher pay, but the wage inequality is not amplified in a union setting since all are equally paid (above their outside option). Secondly, in the union setting, group-performance pay is more likely because the union may act as a monitoring device for the firm, and thus provide a partial solution to the underlying agency problem. Unions may contribute to the solution of the asymmetric information problem as they possess the ability to use collective action to sustain a high effort level and back their claim on their share of the productivity gain. Theory, then, suggests that group-based performance pay schemes have minor effects on the wage distribution since both effort and total compensation is more evenly distributed. Finally, with group bonuses, the stochastic output component has no effect on the within-firm wage distribution, since the risk is pooled across workers.

Empirically, we do find higher wage levels in firms with performance related pay. However, after controlling for individual worker and firm characteristics, the wage premium of performance pay firms drops to about 2 percent. We are not ready to take this as evidence of productivity effects of pay for performance, simply because it may well be compensation for higher risk. Performance pay is less prevalent in highly unionized firms. This observation seems to be at odds with our predictions, since according to the model, bargaining strength does not affect the motivation to deviate from fixed pay. However, our conjecture is that some of the union's opposition is directed more against payment schemes that leave a lot of discretion in the hands of management. In this case, our bargaining model does not apply straightforwardly, as bargaining power is unevenly distributed across the various elements of the contract, and the parties may generally end up with Pareto-inferior outcomes. Another potentially important reason why performance pay schemes are less prevalent in unionized environments is that union bargaining in itself has similar qualities to that of a group bonus, since unions through bargaining appropriate a share of the rents associated with higher performance. In this sense, union bargaining mimics a group-based pay scheme, and may provide similar incentives for self-monitoring and higher effort.

The observed impact of performance pay on within-firm wage dispersion is consistent with our theoretical predictions. Performance-related pay raises the within-establishment wage dispersion. Furthermore, the increase in wage dispersion is significantly reduced by union presence in the establishment. These results remain after controlling for composition of workers and establishment fixed effects.

The impact of performance pay on wage dispersion is significantly lower with group-based pay schemes than with individualized performance-based pay schemes. Firms that introduce performance pay schemes are more likely to choose group-based pay if there is a strong union present. This evidence is consistent with our model, based on the assumption that the union may serve as a self-monitoring device and thus contributes to solve the free-rider problem associated with group-based pay.

There appears to be an underlying trend in the direction of higher within-firm wage inequality in our data, even after controlling for the use and introduction of performance pay. All in all, we conclude that even though pay for performance is on the rise and does contribute to within-firm wage inequality, the introduction of such types of payment schemes is unlikely to be a major contributor to the increase in the wage dispersion in the highly unionized European labor markets.

References

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2005), ”Trends in U.S. Wage Inequality: Re-Assessing the Revisionists” Discussion Paper Number 2095, Harvard Institute of Economic Research, Cambridge: Harvard University.

Barth, Erling, Bernt Bratsberg, Torbjørn Hægeland, and Oddbjørn Raaum (2006), “Who Pays for Performance?” IZA DP 2142, forthcoming *International Journal of Manpower*.

Barth, Erling, and Claudio Lucifora (2006) “Wage Dispersion, Markets and Institutions, The Effects of the Boom in Education on the Wage Structure” IZA Discussion Paper No. 2181.

Barth, E. , Naylor, R. and O. Raaum (2001), Union Wage Effects: Does Membership Matter?, *Manchester School* Vol. 68 (3) 2001, 259-275.

Belfield, R., and D. Marsden (2003), "Performance pay, monitoring environments, and establishment performance," *International Journal of Manpower,* vol. 24(4): 452-471.

Blanchflower, D. and Bryson, A. (2003) 'Changes over time in union relative wage effects in the UK and the US Revisited', chapter 7 in *International Handbook of Trade Unions*, John T. Addison and Claus Schnabel (eds.), Edward Elgar.

Booth, Alison, and Jeff Frank (1999), "Earnings, Productivity, and Performance-Related Pay," *Journal of Labor Economics,* Vol. 17(3): 447-463.

Booth, Alison L., and Marco Francesconi, (2000), "Collectivism versus Individualism: Performance-related Pay and Union Coverage for Non-standard Workers in Britain," ILR working papers 061, Institute for Labour Research.

Brown, Charles (1990), "Firm's Choice of Method of Pay," *Industrial and Labor Relations Review,* Vol. 43(3): 165S-182S.

Brown, Michelle, and John S. Heywood (2002), *Paying for Performance: An international comparison.* Armonk NY: M.E. Sharpe.

Cahuc, Pierre and André Zylberberg (2004) *Labor Economics*, Cambridge, Mass.: The MIT Press.

Card, David, Thomas Lemieux, and W. Craig Riddell (2004), "Unions and Wage Inequality," *Journal of Labor Research* 25:4 (Fall).

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages 1973-1992, A Semiparametric Approach," *Econometrica* 64(5):1001-1044.

Foss, Nicolai J., and Keld Laursen (2005), "Performance pay, delegation and multitasking under uncertainty and innovativeness: An empirical investigation," *Journal of Economic Behavior and Organization,* Vol. 58: 246-276.

Freeman, Richard B, and James Medoff. *What Do Unions Do?* New York: Basic Books, 1984.

Heywood, John, W.S. Siebert, and X. Wei (1997), "Payment by Results Systems: British Evidence," *British Journal of Industrial Relations,* 35: 1-22.

Holmström, Bengt, and Paul Milgrom (1987), "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica* 55(2): 303-328.

Kahn, Lawrence (2000), "Wage Inequality, Collective Bargaining and Relative Employment 1985-94: Evidence from 15 OECD Countries," *The Review of Economics and Statistics*, 82, No. 4.

Kersley, B, C. Alpin, J. Forth, A. Bryson, H. Bewley, G Dix, and S Oxenbridge (2006), *Inside the Workplace: Findings from the 2004 Workplace Employment Relations Survey*, London and New York: Routledge.

Lazear, Edward P. (1995), *Personnel Economics,* MIT Press.

Lazear, Edward P. (2000), "The Use of Performance Measures in Incentive Contracting," *American Economic Review,* Vol. 90(2): 415-420.

Lazear, Edward P. (2002), "Performance Pay and Productivity," *American Economic Review,* Vol. 90(5):1346-1361.

Lemieux, Thomas, Macleod, W. Bentley and Daniel Parent (2007), Performance pay and Wage Inequality, NBER Working Paper 13128.

Lunde, H. and K.,H.,Grini (2007), "Bonus – hvor mye og til hvem?" SSB Rapport 2007/18.

Metcalf, David (2003), "Unions and productivity, financial performance and investment: International evidence." In *International handbook of trade unions*. Edited by Addison, J.; Schnabel, C. Edward Elgar, 2003, pp. 118-171.

Pencavel, J. (1977), "The Distributional and Efficiency Effects of Trade Unions in Britain", *British Journal of Industrial Relations*, 15 (2), July 1977, 137-56.

Vroman, Susan (1990), "The Union-Nonunion Wage Differential and Monitoring Costs," *Economics Letters* 32, 405-409.

Table 1: Firm Characteristics

	Full sample			Balanced panel		
	All firms	Firms without perform-ance pay	Firms with perform-ance pay	All firms	Firms without perform-ance pay	Firms with perform-ance pay
Within-firm 90-10 log wage differential						
Unadjusted	.697	.655	.737	.679	.628	.730
Adjusted	.609	.576	.639	.590	.552	.629
Explanatory variables:						
Performance pay	.513	0	1	.498	0	1
Share union members	.592	.670	.517	.610	.683	.535
Log(Firm size)	4.492	4.498	4.486	4.522	4.556	4.488
Firm age/100	.440	.476	.406	.495	.531	.459
Share high education	.221	.195	.245	.211	.196	.225
Share low education	.148	.162	.135	.154	.166	.142
Share female	.296	.308	.285	.288	.302	.275
Industry:						
Oil, mining, energy	.056	.038	.072	.047	.037	.057
Nondurables (omitted)	.194	.229	.162	.238	.275	.201
Durables	.209	.230	.189	.253	.263	.243
Construction	.074	.046	.100	.069	.039	.099
Wholesale	.113	.066	.158	.113	.061	.166
Retail	.047	.048	.046	.044	.044	.045
Transportation	.073	.108	.039	.059	.090	.028
Post & telecom	.015	.018	.011	0	0	0
Finance	.036	.030	.041	.029	.032	.026
Business services	.111	.070	.150	.080	.048	.113
Health	.030	.051	.010	.029	.051	.007
Education, pers services	.038	.057	.021	.038	.060	.015
Observations		2,406			1,310	

Note: The adjusted 90-10 differential is based on the residuals from a regression of individual log wages on educational attainment (seven levels); age and its square; interactions between each education level and the age polynomial; immigrant status; union membership; gender and interactions between gender and all other characteristics; an indicator variable for year of observation and interactions between the year variable and all other regressors; and firm-by-year fixed effects. The first-step regression has 292,690 observations.

Table 2: Firm Log Wage Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Performance pay	.090*** (.011)	.066*** (.009)	.025*** (.007)	.022** (.009)	.018** (.008)	.035*** (.009)	.028*** (.009)
Share union			-.010 (.012)	-.001 (.014)	.040* (.022)	-.029** (.014)	.016 (.023)
log(Firm size)			.035*** (.003)	.034*** (.004)	-.014* (.008)	.038*** (.004)	.004 (.009)
Firm age/100			.005 (.009)	.016 (.011)		-.010 (.011)	
Share high educ			.364*** (.022)	.359*** (.029)	.066 (.076)	.303*** (.029)	.007 (.080)
Share low educ			-.232*** (.038)	-.273*** (.050)	-.175** (.073)	-.225*** (.049)	-.127 (.078)
Share female			-.161*** (.020)	-.130*** (.024)	-.087 (.067)	-.107*** (.024)	-.077 (.071)
Controls:							
Individual charact.	No	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects step 2	No	No	No	No	Yes	No	Yes
Observations	2,406	2,406	2,406	1,310	1,310	1,385	1,310
Note				Balanced panel		Restricted individual sample	

***/**/* Significant at 1/5/10 percent level.
Note: Standard errors are reported in parentheses. The dependent variable in column (1) is the mean establishment log wage. Dependent variable in columns (2)-(7) is the year-specific establishment fixed effect from the first-step regression described in the note to Table 1. Second-step regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. The sample in columns (4) and (5) is restricted to the balanced panel of firms. In columns (6) and (7), the firm wage effect is computed from the subsample of individuals working at the same establishment both years. Second-step regressions also control for year of observation and industry (12 major industries).

Table 3: Firm Log Wage Regressions with Union Interaction

	(1)	(2)	(3)	(4)	(5)
Performance pay	.054***	.023	.004	.044**	.033*
	(.014)	(.018)	(.017)	(.018)	(.019)
Share union * Performance pay	-.048**	-.001	.021	-.014	-.008
	(.020)	(.025)	(.023)	(.025)	(.025)
Share union	.016	.001	.031	-.021	.019
	(.016)	(.020)	(.024)	(.019)	(.026)
log(Firm size)	.035***	.034***	-.015*	.038***	.004
	(.003)	(.004)	(.008)	(.004)	(.009)
Firm age/100	.005	.016		-.010	
	(.009)	(.011)		(.011)	
Share high educ	.364***	.359***	.065	.303***	.008
	(.022)	(.029)	(.076)	(.029)	(.080)
Share low educ	-.233***	-.273***	-.173**	-.225***	-.128
	(.037)	(.050)	(.073)	(.049)	(.078)
Share female	-.165***	-.130***	-.086	-.108***	-.078
	(.020)	(.024)	(.067)	(.024)	(.071)
Firm fixed effects	No	No	Yes	No	Yes
Observations	2,406	1,310	1,310	1,385	1,310
Note		Balanced panel of firms		Restricted individual sample	

Note: Standard errors are reported in parentheses. The dependent variable is the year-specific establishment fixed effect from the individual regression described in the note to Table 1. Second-step regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. The sample in columns (2), (3), and (5) is restricted to balanced panel of firms. In columns (4) and (5), the firm wage effect is computed from the subsample of individuals working at the same establishment both years. Second-step regressions also control for year of observation and major industry.

Table 4: Within-Firm 90-10 Log Wage Differential Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Performance pay	.077***	.056***	.041***	.070***	.136***	.104***
	(.012)	(.010)	(.010)	(.020)	(.026)	(.034)
Share union * Performance pay				-.047*	-.131***	-.134***
				(.028)	(.036)	(.046)
Share union			-.163***	-.137***	-.101***	-.003
			(.016)	(.022)	(.029)	(.047)
log(Firm size)			.015***	.015***	.006	-.024
			(.004)	(.004)	(.006)	(.016)
Firm age/100			-.059***	-.059***	-.084***	
			(.012)	(.012)	(.016)	
Share high educ			-.048	-.049	-.024	.064
			(.031)	(.031)	(.042)	(.149)
Share low educ			.219***	.217***	.239***	.266*
			(.052)	(.052)	(.073)	(.145)
Share female			.323***	.319***	.278***	.271**
			(.027)	(.027)	(.035)	(.132)
Year=2003	.037***	.048***	.043***	.043***	.050***	.060***
	(.012)	(.010)	(.010)	(.010)	(.012)	(.010)
Constant	.638***	.555***	.486***	.471***	.499***	
	(.010)	(.009)	(.028)	(.029)	(.038)	
Firm fixed effects	No	No	No	No	No	Yes
Observations	2,406	2,406	2,406	2,406	1,310	1,310
					Balanced panel	

Note: Standard errors are reported in parentheses. The dependent variable is the within-establishment difference between the 90th and the 10th percentile log wage residual from the step-one regression described in the note to Table 1, except for in column (1) where the dependent variable is the observed 90-10 log wage differential at the establishment. Regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. The sample in columns (5) and (6) is restricted to balanced panel of firms. Regressions in columns (3)-(6) also control for 12 major industries.

Table 5: Within-Firm 90-50 and 50-10 Log Wage Differential Regressions

	90-50			50-10		
	(1)	(2)	(3)	(4)	(5)	(6)
Performance pay	.040***	.060***	.048***	.030*	.075***	.056**
	(.010)	(.013)	(.016)	(.016)	(.020)	(.028)
Share union * Performance pay	-.038***	-.054***	-.057***	-.009	-.077***	-.077**
	(.0134	(.018)	(.021)	(.022)	(.028)	(.037)
Share union	-.075***	-.071***	.027	-.062***	-.031	-.030
	(.010)	(.014)	(.022)	(.018)	(.022)	(.039)
log(Firm size)	.013***	.007***	-.011	.002	-.001	-.013
	(.002)	(.003)	(.007)	(.003)	(.004)	(.013)
Firm age/100	-.012**	-.017**		-.047***	-.067***	
	(.006)	(.008)		(.010)	(.013)	
Share high educ	.092***	.081***	.028	-.141***	-.105***	.037
	(.015)	(.021)	(.069)	(.025)	(.033)	(.122)
Share low educ	.095***	.141***	.089	.129***	.099*	.177
	(.025)	(.036)	(.067)	(.040)	(.057)	(.118)
Share female	.065***	.053***	.009	.255***	.225***	.261**
	(.013)	(.017)	(.061)	(.022)	(.027)	(.107)
Year=2003	.011**	.021***	.023***	.032***	.029***	.037***
	(.005)	(.006)	(.005)	(.008)	(.010)	(.008)
Constant	.215***	.229***		.256***	.270***	
	(.014)	(.019)		(.023)	(.030)	
Firm fixed effects	No	No	Yes	No	No	Yes
Observations	2,406	1,310	1,310	2,406	1,310	1,310
		Balanced panel			Balanced panel	

Note: Standard errors are reported in parentheses. The dependent variable in columns (1) – (3) is the within-establishment difference between the 90th and the 50th percentile log wage residual from the step one regression described in note to Table 1. Columns (4)–(6) display results for the 50-10 log wage differential. Second-step regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. The sample in columns (2), (3), (5) and (6) is restricted to the balanced panel of firms. Second-step regressions also control for 12 major industries.

Table 6: 90th and 10th Percentile Firm Log Wage Regressions

	90th Percentile			10th Percentile		
	(1)	(2)	(3)	(4)	(5)	(6)
Performance pay	.039***	.060***	.052***	-.031**	-.075***	-.052**
	(.009)	(.012)	(.015)	(.013)	(.017)	(.023)
Share union * Performance pay	-.031**	-.053***	-.064***	.016	.077***	.070**
	(.013)	(.016)	(.020)	(.018)	(.024)	(.031)
Share union	-.071***	-.063***	.007	.066***	.039**	.011
	(.010)	(.013)	(.020)	(.015)	(.019)	(.032)
log(Firm size)	.012***	.006**	-.011	-.003	.000	.013
	(.002)	(.002)	(.007)	(.003)	(.004)	(.011)
Firm age/100	-.021***	-.028***		.038***	.056***	
	(.005)	(.007)		(.008)	(.010)	
Share high educ	.026**	.021	-.013	.075***	.045*	-.077
	(.014)	(.019)	(.064)	(.020)	(.028)	(.102)
Share low educ	.107***	.134***	.107*	-.110***	-.105**	-.159
	(.023)	(.033)	(.062)	(.034)	(.047)	(.099)
Share female	.115***	.100***	.055	-.110***	-.118***	-.215**
	(.012)	(.016)	(.056)	(.034)	(.023)	(.090)
Firm fixed effects	No	No	Yes	No	No	Yes
Observations	2,406	1,310	1,310	2,406	1,310	1,310
		Balanced panel			Balanced panel	

Note: Standard errors are reported in parentheses. The dependent variable is the within-establishment 90th or 10th percentile log wage residual from the step one regression described in note to Table 1. Second-step regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. Second-step regressions also control for year of observation and major industry.

Table 7: Individual and Group-based Performance Pay

	Probit regressions Probability of introducing performance pay during 1997-2003 (1)	Probability of group-bonuses, given introduction (2)	Wage differentials (90-10) (3)	(4)	(5)
Share union	-.210*** (.075)	.243** (.108)	-.176*** (.027)	-.209*** (.033)	-.140*** (.042)
Performance pay			.069*** (.017)	.077*** (.024)	.102*** (.027)
Performance Pay* Group-bonuses			-.034* (.019)	-.054** (.025)	-.063* (.038)
log(Firm size)			.016** (.006)	.009 (.009)	-.028* (.016)
Firm age/100			-.050*** (.018)	-.087*** (.025)	
Share high educ			-.030 (.044)	-.003 (.065)	.141 (.149)
Share low educ			.378*** (.078)	.456*** (.110)	.287** (.138)
Share female			.283*** (.039)	.234*** (.052)	.489*** (.116)
Sample	Firms without performance pay in 1997	Firms that introduced performance pay during 1997-2003	Cross-section 2003	Balanced panel 2003 obs	Balanced panel 1997 and 2003 obs
Outcome			90-10 in 2003	90-10 in 2003	Change 90-10, 1997-03
Observations	400	179	1233	655	655

Note: Standard errors are reported in parentheses. Estimates in columns (1) and (2) are based on probit models; reported are marginal effects (dP/dX). In columns (3)-(5), the dependent variable is the within-establishment 90-10 log wage differential, where individual wages are residuals from the step one regression described in the note to Table 1. Regressions are weighted by the observation count of the firm-by-year cell in the first-step regression sample multiplied by the sampling weight for the firm. Inclusion in the second-step regression requires at least ten observations in the firm-by-year cell. Regressions in cols. (3)-(5) also control for major industry.

Appendix

Pay determination within the firm is modeled as the outcome of constrained profit maximization, subject to various constraints that differ across pay regimes.

The Nonunion Firm

In the case of a firm without union representation, the Lagrangian is given by

$$
\begin{aligned}
L = &\sum_{i=1}^{n} \left(\alpha_i - \omega_i + (1-b_i) p_i z_i \right) - \lambda n z \\
&-\sum_{i=1}^{n} \mu_i \left(\omega_i + b_i p_i (z_i + \varepsilon_i) - \frac{c}{2} z_i^2 - \frac{p_i^2 b_i^2}{2} a\sigma^2 - X_i \right) \\
&-\sum_{i=1}^{n} \xi_i (z_i - \frac{b_i p_i}{c}) \\
&-\sum_{i=1}^{n} \tau_i (z_i - z)
\end{aligned}
$$

Here, the outside option constraint is always binding, while only one of the incentive constraints is binding, depending on the pay regime.

Case 1: Fixed Pay (FP); $b_i = \xi_i = 0, \quad z_i = z$

Inserting the expression for individual effort, $z_i = z$, the first order conditions are

$$
\begin{aligned}
&\frac{\partial L}{\partial \omega_i} = -1 - \mu_i = 0 \quad \Rightarrow \quad \mu_i = 1 \\
&\frac{\partial L}{\partial z} = \sum_{i=1}^{n} p_i - \lambda n - \sum_{i=1}^{n} \mu_i c z = 0 \qquad , \\
&\Rightarrow \quad z = (1-\lambda)/c \quad as \quad \frac{1}{n}\sum_{i=1}^{n} p_i = 1
\end{aligned}
$$

Case 2: Individual Performance Pay (PP); $z = \tau_i = 0$

Inserting $z_i = b_i p_i / c$, the first order conditions read

$$
\begin{aligned}
&\frac{\partial L}{\partial \omega_i} = -1 - \mu_i = 0 \quad \Rightarrow \quad \mu_i = 1 \\
&\frac{\partial L}{\partial b_i} = \frac{p_i^2}{c} - 2b_i \frac{p_i^2}{c} + \mu_i b_i \frac{p_i^2}{c} + \mu_i b_i p_i^2 a\sigma^2 = 0 \qquad , \\
&\Rightarrow \quad b_i = \frac{1}{1 + ac\sigma^2} = b
\end{aligned}
$$

The Union Bargaining Case

For the unionized firm, the Lagrangian is given by

$$L = \sum_{i=1}^{n}\left(\alpha_i - \omega_i + (1-b_i)p_i z_i\right) - \lambda nz$$
$$-\chi(\sum_{i=1}^{n}(-e^{-a\psi_i}) - \overline{u})$$
$$-\sum_{i=1}^{n}\xi_i(z_i - \frac{b_i p_i}{c})$$
$$-\sum_{i=1}^{n}\tau_i(z_i - z)$$

Relative to the nonunion case, the individual outside option constraints are replaced by the (given) union utility constraint.

Case 3: Union Fixed Pay (UFP); $b_i = 0, \quad z_i = z$

Inserting $z_i = z$, the first order conditions are

$$\frac{\partial L}{\partial z} = \sum_{i=1}^{n} p_i - \lambda n - \chi\sum_{i=1}^{n} a\exp(-a\psi_i)cz = 0$$
$$\frac{\partial L}{\partial \omega_i} = -1 - \chi a\exp(-a\psi_i) = 0 \qquad ,$$
$$\Rightarrow \quad z = (1-\lambda)/c \quad as \quad \frac{1}{n}\sum_{i=1}^{n} p_i = 1$$

Note also that, when $\psi_i = \overline{\psi}, \sum_{i=1}^{n} -a\exp(-a\psi_i) = n\overline{u} \Rightarrow \overline{\psi} = -\ln(\overline{u})/a$.

Case 4: Union Individualized Performance Pay (UPP); $z = 0$, $b_i > 0$

Inserting $z_i = bp_i / c$, the first order conditions read

$$\frac{\partial L}{\partial \omega_i} = -1 - \chi a\exp(-a\psi_i) = 0 \Rightarrow \chi a\exp(-a\psi_i) = -1$$
$$\frac{\partial L}{\partial b_i} = \frac{p_i^2}{c} - 2b_i\frac{p_i^2}{c} + \chi\sum_{i=1}^{n} a\exp(-a\psi_i)\left(\frac{p_i^2}{c}(2b_i - b_i(1+ac\sigma^2)\right) = 0 \qquad ,$$
$$\Rightarrow \quad b_i = \frac{1}{1+ac\sigma^2} = b$$

and $\overline{\psi} = -\ln(\overline{u})/a$.

Case 5: Group Performance Pay (GPP);

The optimal effort standard is given by

$$\frac{\partial \sum -e^{-a\Psi_i}}{\partial z} = \sum a e^{-a\Psi_i}(\beta - cz) = 0 \text{ subject to } \Psi_i = \overline{\Psi} \quad \Rightarrow \quad z = \beta / c$$

Note that in this case the piece rate is labeled *β*, instead of *b.* As the Lagrangian reads

$$L = \sum_{i=1}^{n} \left(\alpha_i + (1-\beta)\frac{p_i \beta}{c} - \omega_i \right) - \vartheta \left(\frac{1}{n} \Sigma - \exp(-a\Psi_i) - \overline{u} \right),$$

and the first order conditions are

$$\frac{\partial L}{\partial \omega_i} = -1 - \chi a \exp(-a\psi_i) = 0 \quad \forall i$$

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{p_i}{c} \right) - 2\beta \sum_{i=1}^{n} \left(\frac{p_i}{c} \right) - \beta\vartheta \sum_{i=1}^{n} a \exp(-a\psi_i) \left(\frac{2}{c} - \frac{1}{c}(1 + ac\sigma^2 / n \right).$$

When combined, the first order conditions give the expression for the optimal group rate,

$$\beta = \frac{1}{1 + ac\sigma^2 / n} ,$$

since

$$\frac{1}{n} \sum_{i=1}^{n} p_i = 1 .$$