Monitoring and Pay: General results

Magnus Allgulin*

SSE/EFI Working Paper Series in Economics and Finance
No. 340

October, 1999

Abstract

This paper considers the optimal incentives for motivating a risk neutral, wealth constrained agent. In particular, monitoring and pay are shown to be complementary instruments under very general conditions, extending earlier results by Allgulin and Ellingsen (1998). The paper also proves that linear incentive schemes are strictly sub-optimal in this setting.

JEL CLASSIFICATION: J31, J41.
KEYWORDS: Monitoring, efficiency wages, incentive pay.

*Address: Department of Economics, Stockholm School of Economics, Box 6501, S—113 83 Stockholm, Sweden. Email: magnus.allgulin@post.harvard.edu

I am grateful to Tore Ellingsen for helpful comments and discussions. Financial support from Handelsbanken, the Stockholm School of Economics and the Royal Swedish Academy of Sciences is gratefully acknowledged.

1 Introduction

Standard analysis of the efficiency wage model argues that monitoring and pay are substitute instruments for motivating workers.[1] However, Allgulin and Ellingsen have recently demonstrated that this result hinges on unduly restrictive assumptions regarding workers' choice of effort - for example that there are only two possible effort levels. They extend the simple shirking model to a continuum of effort levels and characterize the profit maximizing levels of monitoring and pay. The equilibrium levels of monitoring and pay are characterized by two first-order conditions containing three general functions: the principal's benefit function of the agent's effort, the agent's cost of effort and the principal's cost of monitoring. In the analysis it is calculated how monitoring and pay vary according to multiplicative shifts to three functions: the profitability $B(e)$ (i.e. the principal's benefit from the agent's effort), the agent's cost of effort $C(e)$, and the principal's cost of monitoring, $M(p)$. It is shown that under reasonable assumptions, monitoring and pay are complementary instruments, i.e. a parameter change which causes an increase in the accuracy of monitoring also causes an increase in the level of pay and vice versa. This paper will demonstrate that the result is true for any monotonic shift, not just the multiplicative.

Two extensions of the model are also investigated. The first is the introduction of an ex ante individual rationality constraint for the agent in addition to the previous ex post limited liability constraint. This will typically be of interest if unemployed workers bid over each other with entrance fee offers for a job opportunity. The implication is that a binding ex ante individual rationality constraint generalizes the result even further.

The second extension allows the principal to credibly use a mixed strategy when monitoring the agent. In that case, any concave monitoring cost function can be replaced by another monitoring cost function that is cheaper and linear. Hence, the monitoring technology used by the principal is always weakly convex, implying that even this extension will make the result more general.

Finally, the optimal non-linear incentive scheme solution in the paper is compared to the best linear incentive scheme solution. The finding is that, for the principal, there is always a non-linear wage contract which strictly dominates any linear wage contract. This result may not be surprising, but the main contribution of this section is to provide a common framework for efficiency wages and linear incentive wages, in which their different implications can be analyzed.

[1] See e.g. Shapiro and Stiglitz (1984) and Milgrom and Roberts (1992).

The paper is organized as follows. Section 2 sets up a more general version of the model proposed by Allgulin and Ellingsen (1998). Section 3 contains a complete analysis of this model. Section 4 introduces an ex ante individual rationality constraint into the model and discusses its implications when it is binding. In Section 5, the principal is allowed to use mixed strategies, and this is shown to strengthen the previous results. Section 6 shows that the optimal incentive scheme used in the paper strictly dominates a linear incentive scheme. Section 7 concludes.

2 Model

A risk neutral principal employs a single risk neutral agent. The agent can exert effort $e \in \mathbf{R}_+$ which affects the principal's benefit, $B(e)$, at some cost to the agent, $C(e)$. The principal motivates the agent through a compensation contract $w(e)$, where w is the wage that the agent receives if the principal observes e.

The principal can observe and verify the agent's effort with probability p. This probability is affected by the principal's choice of monitoring technology. In order to attain probability p of observing the agent's effort, the principal has to pay $M(p)$. The following assumption regarding functional forms is made.

Assumption 1

(i) $B'(e) > 0$, $B''(e) \leq 0$,

(ii) $C(0) = 0$, $C'(e) > 0$, $C''(e) > 0$,

(iii) $M'(p) > 0$.

The ex post utility of the principal can now be written

$$U = B(e) - w - M(p), \tag{1}$$

and the ex post utility of the agent is

$$V = w - C(e). \tag{2}$$

A key assumption of the efficiency wage model is that there is a lower limit, $w_0 \in \mathbf{R}_+$, to the payment. The limit may be due either to legal rules or to a wealth constraint.

Since effort is not always observed, the compensation contract also needs to specify some payment $\bar{w} \in \mathbf{R}_+$ that the agent is to receive in this case. The agent is assumed to maximize the expected utility

$$\mathrm{E}[V] = pw(e) + (1-p)\bar{w} - C(e). \tag{3}$$

Suppose now that the principal will induce the agent to take the level of effort $\hat{e} \in \mathbf{R}_+$. Then the following incentive compatibility constraint must be satisfied for all e:

$$pw(\hat{e}) + (1-p)\bar{w} - C(\hat{e}) \geq pw(e) + (1-p)\bar{w} - C(e).$$

We see that any incentive compatible contract that implements $\hat{e}$ can, without loss to the principal, be replicated by a step function of the form $w(e) = w_0$ for $e < \hat{e}$ and $w(e) = w$ for $e \geq \hat{e}$. In other words, the principal sets an effort target, $\hat{e}$. The agent gets w if he meets or exceeds the target and the minimum payment w_0 otherwise.[2] Later, it is shown that this kind of contract strictly dominates a linear incentive scheme. Given the above contract, if an agent ever wants to deviate, he will deviate to $e = 0$. Thus, the incentive compatibility constraint becomes

$$p(w - w_0) \geq C(\hat{e}). \tag{4}$$

Finally, the assumption that an indifferent agent takes the action that the principal favors enables the principal to lower the wage down to the level where the incentive compatibility constraint becomes an equality. Inverting this equality, we obtain an expression for the actual effort which the agent will exert,

$$e(p, w) := C^{-1}((w - w_0)p). \tag{5}$$

Note that $\bar{w}$, the wage in the case that the effort level is not observed, is irrelevant for the agent's incentives. It is common to assume that $\bar{w} = w(\hat{e})$, and we follow this practice. One reason why the assumption is plausible is that if $\bar{w} < w(\hat{e})$, the principal would have had an incentive not to monitor (or to falsely claim that he has not monitored).

The principal's problem is to find a probability p and a wage w to maximize

$$U(p, w) = B(e(p, w)) - w - M(p) \tag{6}$$

subject to the constraints $w \geq w_0$ and $p \in [0, 1]$. This is a straightforward maximization problem in two variables. Let a solution to this problem be denoted (p^*, w^*), and let $e^* := e(p^*, w^*)$ denote the associated effort level. The first–order conditions for the solution can then be written

$$p^* \frac{B'(e^*)}{C'(e^*)} - 1 \leq 0, \tag{7}$$

[2] As shown by Demougin and Fluet (1997), the optimality of effort targets is quite general. It is not an artefact of the simple monitoring technology.

with equality if $w^* > w_0$, and

$$(w^* - w_0)\frac{B'(e^*)}{C'(e^*)} - M'(p^*) \geq 0, \tag{8}$$

with equality if $p^* < 1$.

Equation (7) tells us that the marginal benefit from increased effort will be larger than the marginal cost whenever the principal chooses to monitor imperfectly. The reason for the distortion away from the socially optimal effort level is that the agent must be paid a rent in order not to shirk. Also, equation (7) confirms that there must be a positive level of monitoring in order to induce any effort. (If $p^* = 0$, then $w^* = w_0$.)

Let U_{ij} denote the second derivative of U and let U^*_{ij} denote the second derivative of U evaluated at a solution (p^*, w^*). The second–order conditions are then $U^*_{ww} < 0, U^*_{pp} < 0$ and $U^*_{pp}U^*_{ww} - (U^*_{pw})^2 > 0$. To state the conditions in full, note that

$$U_{ww} = p^2 h(e), \tag{9}$$

$$U_{pp} = (w - w_0)^2 h(e) - M''(p), \tag{10}$$

$$U_{pw} = p(w - w_0)h(e) + \frac{B'(e)}{C'(e)}, \tag{11}$$

where

$$h(e) := \frac{B''(e)C'(e) - C''(e)B'(e)}{(C'(e))^3}.$$

3 Analysis

3.1 Method

The method to analyze the effects of a shift in one of the general functions is as follows. The first-order conditions (7) and (8) are used to create two equations characterizing the equilibrium values of w and p where, A) equation 1 is lacking one of the control variables and, B) equation 2 is lacking any form of the general function that is about to be shifted.

The trick is that we can now look at the shift as though it created a chain reaction. The equilibrium value of the remaining control variable in equation 1 is solely determined by this equation. A shift in the general function in equation 1 may change the equilibrium value of this control variable. Then, since equation 2 is lacking any form of the shifted function, the entire effect of the shift on the other control variable is forwarded by the first control variable according to equation 2. Hence we only have to differentiate equation 2 in order to find out the condition for the control variables to be positively related.

3.2 Main results

(7) and (8) together tell us that

$$(w^* - w_0) = pM'(p^*). \tag{12}$$

This equation can be used to substitute for w^* in (7) to get the equation

$$p^* \frac{B'(C^{-1}((p^*)^2 M(p^*)))}{C'(C^{-1}((p^*)^2 M(p^*)))} = 1, \tag{13}$$

which solely determines p^*. The equations (12) and (13) together characterize the equilibrium values of w and p. We see that neither the principal's benefit function, $B(e)$, nor the agent's cost function, $C(e)$, appears in any form in equation (12). Therefore, any effect that any change in these functions will have on w^* will be forwarded by p^* according to equation (12). By differentiating (12) we obtain

$$\frac{dw^*}{dp^*} = p^* M''(p^*) + M'(p^*),$$

which is positive if and only if

$$-p^* \frac{M''(p^*)}{M'(p^*)} < 1.$$

Hence, we can conclude that monitoring and pay are complementary instruments according to *any* shift in the principal's benefit function, $B(e)$, or the agent's cost of effort function, $C(e)$, as long as the condition above is fulfilled.

To analyze the effects of a shift in the principal's monitoring cost function, first define

$$f(p^*) = p^* M'(p^*).$$

Using this and inverting (12) now yields

$$p^* = f^{-1}(w^* - w_0),$$

which we can use to substitute for p^* back in equation (7) to get the equation

$$f^{-1}(w^* - w_0) \frac{B'(C^{-1}((w^* - w_0) f^{-1}(w^* - w_0)))}{C'(C^{-1}((w^* - w_0) f^{-1}(w^* - w_0)))} = 1 \tag{14}$$

which solely determines w^*. This equation together with (7) characterizes the equilibrium values of w and p and (7) is lacking any form of the monitoring cost function, $M(p)$. Differentiation of (7) yields

$$\frac{dw^*}{dp^*} = -\frac{U^*_{pw}}{U^*_{ww}}$$

which is positive if and only if

$$U^*_{pw} > 0.$$

Thus we can conclude that monitoring and pay are complementary instruments according to *any* shift in the principal's monitoring cost function, $M(p)$, as long as monitoring and pay are Edgeworth complements.

The above analysis can be summarized in the following two propositions,

Proposition 1 *According to any shift in the principal's benefit function, $B(e)$, or the agent's cost of effort function, $C(e)$, monitoring and pay are complementary instruments if and only if $-pM''(p^*)/M'(p^*) < 1$.*

Proposition 2 *According to any shift in the principal's monitoring cost function, $M(p)$, monitoring and pay are complementary instruments if and only if $U^*_{pw} > 0$.*

4 Binding individual rationality constraint

So far it has implicitly been assumed that the agent is willing to be hired by the principal, given the equilibrium solution. The agent will receive a higher wage than w_0 in any solution, and the interpretation of w_0 as the market price of labor has justified the previous ignorance of a participation constraint.

However, there are two important objections against this reasoning. Firstly, one of the foundations of efficiency wages is that there is incomplete competition in the labor market, due to for example the non-homogeneity of firms, or the transaction costs of being fired for the agent. Hence, w_0 may be lower than the agent's outside option before the relationship starts. Secondly, it has frequently been argued that the agent's (ex ante) individual rationality constraint *must* be binding because otherwise the agent would just offer to pay an entrance fee for the job opportunity, such that the individual rationality constraint is binding anyway.[3]

Below, in addition to the limited liability constraint, w_0, on the maximum punishment that the principal can impose on the agent ex post, an ex ante individual rationality constraint is introduced into the model,

$$w - C(e) \geq \bar{V}, \tag{15}$$

where the ex ante reservation utility, $\bar{V}$, is assumed to be higher than the ex post limited liability constraint, w_0. Clearly, there are two distinct cases. If

[3] See for example Carmichael (1985) and (1990).

the solution in the above analysis yields utility for the agent that is higher than the reservation utility, i.e. if

$$w^* - C(e^*) \geq \bar{V}, \tag{16}$$

then the participation constraint is not binding and the analysis remains intact. On the other hand, if this is not the case, the following analysis will apply.

Rearranging and assuming that the participation constraint, (15), is binding yields the expression for the lowest possible monitoring accuracy,

$$p = \frac{w - \bar{V}}{w - w_0}. \tag{17}$$

Using this to substitute for the monitoring accuracy and the incentive constraint, (5), to substitute for the effort level, yields the following maximization problem for the principal.

$$\underset{w}{Max} \;\; U(w) = B(C^{-1}(w - \bar{V})) - w - M(\frac{w - \bar{V}}{w - w_0}) \tag{18}$$

subject to the constraints $w \geq w_0$ and $\bar{V} > w_0$. Let the solution to this problem be denoted w^*, and let $e^* := e(w^*)$, $p^* := p(w^*)$ denote the associated effort level and monitoring accuracy respectively. Assuming an interior solution, the first–order condition for the solution can be written

$$\frac{B'(e^*)}{C'(e^*)} - 1 - M'(p^*)\frac{\bar{V} - w_0}{(w^* - w_0)^2} = 0. \tag{19}$$

The optimal wage is solely determined by this condition and neither the principal's benefit function, $B(e)$, the principal's monitoring cost function, $M(p)$, or the agent's cost function, $C(e)$, appears in any form in the expression (17). Hence, differentiation of (17) in equilibrium will reveal how the wage and the monitoring accuracy vary together according to any shift in these functions,

$$\frac{dp^*}{dw^*} = \frac{\bar{V} - w_0}{(w^* - w_0)^2}.$$

This is clearly always positive.

Proposition 3 *If the agent's individual rationality constraint is binding, then according to any shift in the principal's benefit function, $B(e)$, the principal's monitoring cost function, $M(p)$, or the agent's cost of effort function, $C(e)$, monitoring and pay are complementary instruments.*

5 Mixed strategies

With a non-binding individual rationality constraint, the analysis section demonstrated that monitoring and pay are complements according to *any* shift in the principal's benefit function, $B(e)$, or the agent's cost function, $C(e)$, as long as the condition

$$-p^* \frac{M''(p^*)}{M'(p^*)} < 1 \tag{20}$$

is fulfilled. Thus a sufficient, but not necessary, condition is that $M(p)$ is convex. Generally, some increasing returns to monitoring may be allowed for, but when the returns to monitoring increase sufficiently fast the result may be overturned.

However, if the principal can use mixed strategies there will never be increased returns to monitoring. More precisely, if the principal is able to randomize between the monitoring accuracy $p = 0$ and $p = 1$, he will optimally do so whenever the monitoring cost function, $M(p)$, is concave. As illustrated in Figure 1, by this randomization he will create a new and cheaper monitoring cost function,

$$\dot{M}(q) = (1-q)M(0) + qM(1), \tag{21}$$

where $q \in [0,1]$, i.e. is linear in q.

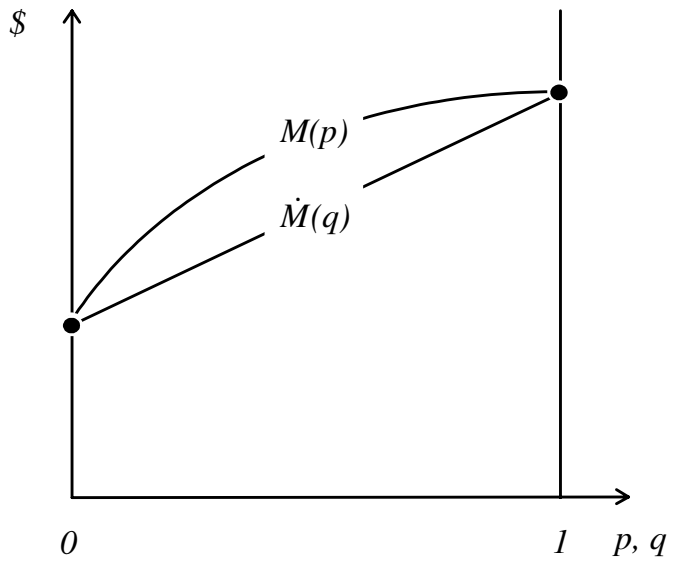

Figure 1

The principal is able to do this because the agent's expected utility (3)

$$\mathrm{E}[V] = pw(e) + (1-p)\bar{w} - C(e) \tag{22}$$

is linear in the detection probability p. The agent's expected utility when the principal randomizes in this manner is

$$\mathrm{E}[V] = q(1 * w(e) + 0 * \bar{w}) + (1 - q)(0 * w(e) + 1 * \bar{w}) - C(e), \tag{23}$$

which can be simplified to

$$\mathrm{E}[V] = qw(e) + (1 - q)\bar{w} - C(e). \tag{24}$$

Hence, the agent will react in exactly the same way towards the randomization variable q as he did towards the detection probability p.

Of course, one could argue that there may be credibility problems for the principal; it could be tempting for the principal to pretend that the randomization variable q is high when it is not. Technically, this problem should arise even with the original monitoring function, but this can be solved by e.g. the assumption that the monitoring technology able to detect at probability p^* is observable and installed before the agent makes his effort. However, this objection against the use of mixed strategies for the principal falls if he is able to create a mechanism that commits him to a certain q. Thus in a final proposition it can be concluded that:

Proposition 4 *According to any shift in the principal's benefit function, $B(e)$, or the agent's cost of effort function, $C(e)$, monitoring and pay are complementary instruments if the principal can credibly use mixed strategies.*

6 The suboptimality of linear incentive schemes

Assume that the principal is interested in implementing a certain effort level, $\hat{e}$. Below, we shall now see that implementation of $\hat{e}$ is (strictly) more costly using a linear incentive scheme than using the optimal non-linear scheme.[4]

If the principal is free to use a non-linear incentive scheme, he will simply solve the minimization problem

$$\underset{w,p}{Min} \;\; w + M(p) \tag{25}$$

subject to the incentive compatibility constraint

$$p(w - w_0) = C(\hat{e}). \tag{26}$$

[4]This result contrasts with the conventional result in many principal-agent models, where linear incentive schemes are optimal. The difference is accounted for by limited liability.

Substitution of the constraint into the minimization problem yields a problem with only one variable, for example

$$\underset{p}{Min} \;\; w_0 + \frac{C(\hat{e})}{p} + M(p) \tag{27}$$

with the first order condition

$$(p^*)^2 M'(p^*) = C(\hat{e}) \tag{28}$$

if an interior solution is assumed.

If the principal is restricted to using a linear incentive scheme, the new restriction is introduced that the wage must be of the form $\alpha_0 + \alpha_1 e$ if the agent's effort is observed. If the agent's effort is not observed, the agent receives the wage $\alpha_0 + \alpha\hat{e}$. Furthermore, the limited liability constraint on the agent restricts the principal to setting α_0 to at least w_0: the wage which the agent gets if he makes no effort and this is detected. The principal does not give anything away for free so he will optimally set $\alpha_0 = w_0$. Hence, the agent's ex ante wage will now be

$$w(e) = w_0 + \alpha(pe + (1-p)\hat{e}). \tag{29}$$

Thus the agent will face the following maximization problem:

$$\underset{e}{Max} \;\; V(e) = w_0 + \alpha(pe + (1-p)\hat{e}) - C(e) \tag{30}$$

with the first-order condition

$$\alpha p = C'(e). \tag{31}$$

Thus, the principal solves the minimization problem (25) subject to the constraints

$$w = w_0 + \alpha\hat{e} \tag{32}$$

and

$$\alpha p = C'(\hat{e}). \tag{33}$$

Substitution of the constraints into the minimization problem yields a problem with only one variable, for example

$$\underset{p}{Min} \;\; w_0 + \frac{C'(\hat{e})\hat{e}}{p} + M(p) \tag{34}$$

with the first order condition

$$(p^{**})^2 M'(p^{**}) = C'(\hat{e})\hat{e} \tag{35}$$

for an interior solution.

By the assumptions $C(0) = 0$ and $C''(e) > 0$ we know that $C(e) < eC'(e)$. This can easily be seen in Figure 2. There, $C(\hat{e}) = \int_0^{\hat{e}} C'(e)de$ is represented by the area B, and $\hat{e}C'(\hat{e})$ is represented by the strictly larger area $A + B$.

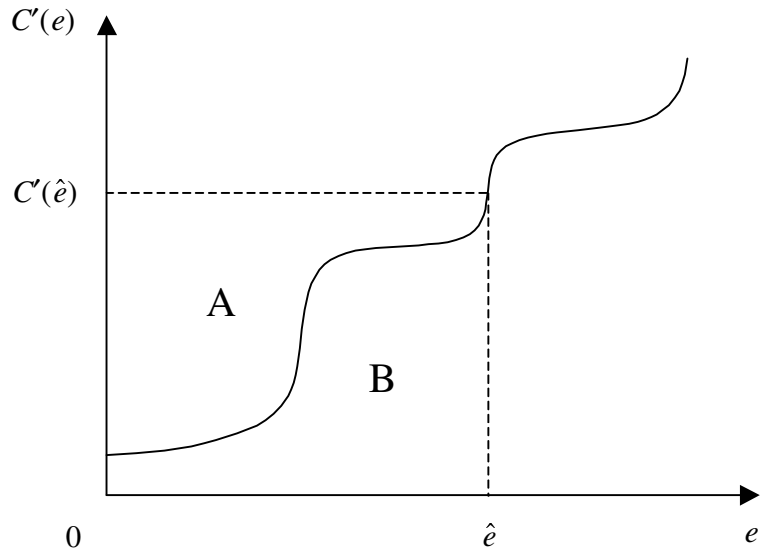

Figure 2

Thus (28) and (35) together tell us that

$$(p^*)^2 M'(p^*) < (p^{**})^2 M'(p^{**}), \tag{36}$$

implying that for any level of effort the principal wants to implement he will monitor more accurately under a linear incentive scheme if $M''(p) \geq 0$.

It is now easy to compare the least resources required to implement the effort level, $\hat{e}$, using a non-linear contract and a linear contract respectively. If one substitutes for the equilibrium wage, it is clear that

$$w_0 + p^* M'(p^*) + M(p^*) < w_0 + p^{**} M'(p^{**}) + M(p^{**}). \tag{37}$$

Thus, the principal can implement any level of effort under the best non-linear incentive scheme using less resources on incentives for the agent than under the best linear incentive scheme.

Proposition 5 *For any level of effort the principal wants to implement, he will monitor more accurately if he is restricted to using linear incentive schemes than if he is not, if* $M''(p) \geq 0$ *or if the principal can credibly use mixed strategies.*

If $M''(p) < 0$, the optimal monitoring accuracy under a non-linear incentive scheme, p^*, may be higher than the optimal monitoring accuracy under a linear incentive scheme, p^{**}. But even if this is the case, it is more costly for the principal to implement $\hat{e}$ if the wage contract is restricted to be linear. To see this, first note that $p^* > p^{**}$ together with (36) imply that

$$p^* M'(p^*) < p^{**} M'(p^{**}), \tag{38}$$

or alternatively expressed,

$$\int_{p^{**}}^{p^*} (M'(p) + pM''(p))\, dp < 0. \tag{39}$$

Assume now that the inequality (37) does not hold, i.e. that

$$\int_{p^{**}}^{p^*} (2M'(p) + pM''(p))\, dp \geq 0. \tag{40}$$

It is easily seen that the following must then be true:

$$\int_{p^{**}}^{p^*} M'(p) > 0. \tag{41}$$

But this is impossible since it contradicts the assumption that $M''(p) < 0$. Evidently, the principal is always harmed by restricting the wage contract to be linear. Thus a final proposition can be stated:

Proposition 6 *The set of optimal linear contracts and the set of optimal non-linear contracts are disjunct.*

7 Concluding remarks

Allgulin and Ellingsen (1998) demonstrate that a natural generalization of the conventional shirking model of efficiency wages completely overturns previous intuitions. The current paper has demonstrated the robustness of this result to a variety of generalizations. In addition, it has shown that linear incentive schemes are suboptimal.

The most obvious argument against the latter result is that it hinges on the assumption of risk neutral workers. In practice, linear incentive schemes may be preferable. But the contribution here is not merely the result; it

is rather the framework. Previously, efficiency wages and linear wage contracts have been described in different settings. Standard shirking models of efficiency wages, such as Shapiro and Stiglitz (1984), assume imperfect monitoring and limited liability but no risk aversion. On the other hand, models of linear wage contracts, such as Holmström (1994) (or Milgrom and Roberts (1992, p226-7)), assume imperfect monitoring, risk aversion and unlimited liability. To be able to compare the two different types of wage contracts, the underlying assumptions for them must of course be the same. A first step towards a common framework is taken here, as both the conventional efficiency wage contract and the linear wage contract are modelled in a world of imperfect monitoring, limited liability and no risk aversion. The next step, which clearly merits an investigation, is a common framework with risk averse workers. A conjecture is that, when both limited liability and risk aversion are present, the optimal wage contract is neither an efficiency wage step function nor a linear wage, but rather a smooth non-linear wage contract.

References

Allgulin, Magnus and Tore Ellingsen (1998): Monitoring and Pay, *Working Paper Series in Economics and Finance* No. 245, Stockholm School of Economics.

Carmichael, H. Lorne (1985): Can Unemployment be Involuntary?: Comment, *American Economic Review* December, 75, 1213-1214.

Carmichael, H. Lorne (1990): Efficiency Wage Models of Unemployment - One View, *Economic Inquiry* April, 28, 269-295.

Demougin, Dominique and Claude Fluet (1997): Monitoring versus Incentives: Substitutes or Complements?, Mimeo, Université du Québec à Montréal.

Holmström, Bengt and Paul Milgrom (1994): The Firm as an Incentive System, *American Economic Review* 84, 972–991.

Milgrom, Paul and John Roberts (1992): *Economics, Organization and Management,* (Englewood Cliffs, NJ: Prentice–Hall).

Shapiro, Carl and Joseph E. Stiglitz (1984): Involuntary Unemployment as a Worker Discipline Device, *American Economic Review* 74, 433–444.