Limited Liability and Dynamic Incentives

Magnus Allgulin*

SSE/EFI Working Paper Series in Economics and Finance
No. 341

October, 1999

Abstract

If efficiency wages really exist, as proposed by Shapiro and Stiglitz (1984), why do we not see more job purchases? A conventional answer is that with multiple periods, low pay in initial periods serves as an implicit payment (Lazear (1981)). This paper presents a formal analysis of this issue. A major result is that the per period worker rents associated with efficiency wages are inversely related to the number of periods, but are never zero. The paper also discusses how remaining worker rents can be eliminated by implicit bonds, such as firm-specific human capital investments.

JEL CLASSIFICATION: J31, J41.
KEYWORDS: Monitoring, efficiency wages, implicit bond, entrance fee.

*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

The objection against the shirking version of the efficiency wage model that most frequently recurs in the literature is that unemployed workers should offer to pay entrance fees for job opportunities, and such fees are very seldom observed. The best known paper which presents a shirking efficiency wage model is Shapiro and Stiglitz (1984). Indeed, Shapiro and Stiglitz were aware of this weakness and had already defended their model against the entrance fee criticism in the original paper. They pointed out the two most obvious arguments against job purchases. The first is that workers may simply be unable to pay the entrance fee. The idea is that workers are wealth constrained due to imperfect capital markets. The second is that the resulting employment contract is not self enforcing. The firm would have an incentive to claim that a worker shirks, fire him once the fee is paid, and sell the employment opportunity again.

However, many economists have not been convinced by these explanations of the lack of entrance fee observations. Carmichael (1985), in his reply to Shapiro and Stiglitz and Carmichael (1990), argues that the presence of an imperfect capital market, or the moral hazard problem, is not sufficient to prevent the emergence of entrance fees. For example, the moral hazard problem can be solved by the introduction of a third party. Moreover, in the Shapiro Stiglitz model, the firm and the worker do not know when the relationship is to end. Bull (1985, 1987) argues that this is unrealistic and continues to show that unemployment will not be sufficient to generate the surplus required for a self-enforcing contract in a model with a last period. As a consequence, the efficiency wage theory has lost advocates to a related theory, referred to as delayed payment/bonding contracts. In these contracts, firms initially pay wages below the alternative wage and later pay wages above the alternative wage to discourage shirking when monitoring is imperfect. The important point is that they are protected against the entrance fee criticism since they do not alter the present value of compensation from the first-best, full-information level.

The main purpose of this paper is to show how a firm should adjust the level and timing of compensation in the best possible way. Incentives can be constructed by both efficiency wages and back-loaded compensation, and it turns out that a rational firm will use both of them. It will be demonstrated that in a self-enforcing efficiency wage model with a finite number of periods and wealth constrained workers, the shape of the optimal wage path may look very much like the one derived from a delayed payment/bonding contract, such as for example Lazear (1981) and Becker and Stigler (1974). It will typically be constituted by a high wage in the last period and a lower and

stationary wage for all preceding periods. The worker rents associated with efficiency wages still exist, but they are strongly diminished by the length of the relationship. This result is perhaps most remarkable for not fully eliminating worker rents; Lazear (1995, page 71) and others have claimed that delayed payment is a perfect substitute for entrance fees.

Bonding contracts and entrance fees might in principle eliminate the remaining worker rents. However, both bonds and entrance fees are inferior means of extracting worker rents. Investment in firm-specific human capital solves the moral hazard problem more efficiently. (With capital market imperfection, the firm might also be able to extract a greater portion of worker rents in this way.) Shapiro and Stiglitz briefly mention that when the cost of losing job-specific human capital is substantial, workers may have an incentive to exert effort even under conditions of full employment. That is true even in this paper. The point is that a firm-specific human capital investment increases the worker's liability and hence increases the firm's profit at the investing worker's expense. Unemployed workers or workers employed by other firms bid over each other for a job opportunity with higher education levels. This may go on until the hired worker is totally extracted. The reason why bids are in education levels and not in money (bond or entrance fee) is simply that it solves the moral hazard problem in the simplest possible way: if the employer fires the worker he will lose the firm-specific human capital investment together with the worker.

The model is based on the static efficiency wage model first presented in Allgulin and Ellingsen (1998), and later revised in Allgulin (1999), which contrary to Shapiro and Stiglitz's model allows the workers' effort levels to be adjusted in a continuous fashion. A secondary purpose of the paper has been to investigate whether the result from this model, which will be referred to as the static model throughout the paper, carries over to the dynamic case. The finding is that it does; the result is even strengthened by an increase in the number of periods that the relationship between the worker and the firm will last.

Two extensions are also discussed. The first makes the firm able to commit itself to future control variables. The solution could be interpreted as the optimal explicit bonding contract, and the wage path is simply as much payment as possible postponed to the last period and, as a consequence, the lowest possible wage in all preceding periods. The second briefly discusses how a worker's commitment to saving or educating himself at work can be an alternative to bonds or pre-work education, if he is too wealth-constrained for the latter.

The paper is organized as follows. Section 2 sets up the general model, derives the optimal contract and analyzes how the results from the static

model carry over to a dynamic extension. Welfare effects are discussed in Section 3. In Section 4 it is explained how bonding contracts are captured by the model and how firm specific human capital investments can act as implicit bonds. Section 5 briefly discusses some extensions and finally, Section 6 concludes.

2 Model

A risk-neutral employer employs a single risk-neutral employee. The employee can exert effort $e_t \in \mathbf{R}_+$ yielding the benefit $B(e_t)$ to the employer at some cost $C(e_t)$ to the employee, where $t \in \{1, 2, .., T-1, T\}$ and T is the potential number of periods the relationship between the employer and the employee can last. The employer motivates the employee through a compensation contract $w(e_t)$. The employer has the ability to observe and verify the exerted level of effort e_t with probability p_t, and he can indirectly choose this detection probability through his choice of resources devoted to monitoring, $M(p_t)$. The benefit, cost and monitoring cost functions are assumed to have the plausible properties: $B'(e_t) > 0$, $B''(e_t) < 0$, $C(0) = 0$, $C'(e_t) > 0$, $C''(e_t) > 0$ and $M'(p_t) > 0$. Both the employer and the employee discount future utilities using a common discount factor δ.

Define the employer's ex post utility as

$$U = \sum_{t=0}^{T-1} \delta^t (B(e_t) - w_t - M(p_t)). \tag{1}$$

The employee's ex post utility in the last period T is

$$V_T = w_T - C(e_T), \tag{2}$$

and the total value of the present and future periods for him at any other time is

$$V_t = w_t - C(e_t) + \delta V_{t+1} \quad \text{for } t \in \{1, .., T-1\}. \tag{3}$$

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. Legal rules also make it impossible for the employer to fire the employee if he fulfills his task, and to exchange him for another employee. Since effort is not always observed, the compensation contract also needs to specify some payment $\bar{w}_t \in \mathbf{R}_+$ that the employee is to receive in this case. Furthermore, the employer cannot commit himself not to change his control variables in future periods. On the other hand, he can commit not to rehire the employee once he has been fired. This is a crucial assumption and it implies that the employee's decisions, and

hence the equilibrium values of the employer's control variables in different periods, are different. Without this assumption the result would just be the static model's outcome repeated T times.

Suppose the employer wants to induce some level of effort $\hat{e}_t$. The employee's only interest is to maximize his expected utility; therefore for all $t \in \{1, ..T-1\}$ the following incentive compatibility (IC) constraints must be satisfied for all e_t:

$$p_t(w_t(\hat{e}_t) + \delta V^*_{t+1}) + (1-p_t)(\bar{w}_t + \delta V^*_{t+1}) - C(\hat{e}_t) \geq$$

$$p_t(w_t(e_t) + \sum_{i=1}^{T-t} \delta^i w_0) + (1-p_t)(\bar{w}_t + \delta V^*_{t+1}) - C(e_t), \tag{4}$$

where V^*_{t+1} denotes the value of future periods for the worker given the (optimal) incentive scheme. For $t = T$, the incentive compatibility constraint becomes

$$p_T w_T(\hat{e}_T) + (1-p_T)\bar{w}_T - C(\hat{e}_T) \geq$$
$$p_T w_T(e_T) + (1-p_T)\bar{w}_T - C(e_T). \tag{5}$$

As shown in Demougin and Fluet (1997), any incentive-compatible contract that implements $\hat{e}_t$ can, without loss to the employer, be replicated by a step function of the form $w_t(e_t) = w_0$ for $e_t < \hat{e}_t$ and $w_t(e_t) = w_t$ for $e_t \geq \hat{e}_t$. In other words, the employer sets an effort target $\hat{e}_t$. The employee gets w_t if he meets or exceeds the target and the minimum payment w_0 otherwise. If the employee deviates under this contract he will choose the effort level $e_t = 0$.

Finally, we impose the standard assumption that an indifferent employee takes the action that the employer favors. Thus, the employer can lower the wage down to the level where the incentive compatibility constraints become equalities,

$$p_t(w_t - w_0 + \delta V^*_{t+i} - \sum_{i=1}^{T-t} \delta^i w_0) = C(\hat{e}_t) \tag{6}$$

for $t \in \{1, .., T-1\}$ and

$$p_T(w_T - w_0) = C(\hat{e}_T). \tag{7}$$

Note that $\bar{w}_t$, is irrelevant for the employee's incentives. It is common to assume that $\bar{w}_t = w_t(\hat{e}_t)$, and this practice is followed. One reason why the assumption is plausible is that if $\bar{w}_t < w(\hat{e}_t)$, the employer would have had an incentive not to monitor (or to falsely claim that he has not monitored).

Using the incentive compatibility constraints (6) and (7), starting at period T, the effort level in each period can now be computed. The highest

possible effort level the employer is able to implement as a function of monitoring and pay in the last period is naturally the same as in the static case,

$$e_T = C^{-1}\left(p_T(w_T - w_0)\right). \tag{8}$$

But for all other periods it is cheaper for the employer to implement a given effort level, since the employee is not only threatened by being fired but also by never being rehired by the employer. Thus we have the following expressions for the highest possible effort levels in each of these periods as a function of monitoring and pay:

$$e_t = C^{-1}\left(p_t(w_t - w_0 + \delta V^*_{t+1} - \sum_{i=1}^{T-t} \delta^i w_0)\right). \tag{9}$$

These expressions will be used to substitute for e_t in the employer's utility function. Hence he will face T maximization problems to choose probabilities p_t and transfers w_t that solve

$$\underset{p_t, w_t}{Max} \;\; B(e_t) - w_t - M(p_t) + \sum_{i=1}^{T-t} \delta^i (B(e_{t+i}) - w_{t+i} - M(p_{t+i}))$$

subject to the constraints $w_t \geq w_0$ and $p_t \in [0,1]$ for all $t \in \{1,..,T\}$. Denote the solutions to these problems (p^*_t, w^*_t), and let $e^*_t := e_t(p^*_t, w^*_t)$ be the corresponding effort.

Assuming an interior solution, first order conditions are

$$\frac{B'(e^*_t)}{C'(e^*_t)} p^*_t - 1 = 0 \tag{10}$$

for $t \in \{1,..,T\}$,

$$\frac{B'(e^*_T)}{C'(e^*_T)}(w^*_T - w_0) - M'(p^*_T) = 0, \tag{11}$$

and

$$\frac{B'(e^*_t)}{C'(e^*_t)}(w^*_t - w_0 + \delta V^*_{t+1} - \sum_{i=1}^{T-t} \delta^i w_0) - M'(p^*_t) = 0 \tag{12}$$

for $t \in \{1,..,T-1\}$.

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

$$U_{tw_tw_t} = (p_t)^2 h(e_t) \;\; \text{for } t \in \{1,..,T\}, \tag{13}$$

$$U_{Tp_Tp_T} = (w_T - w_0)^2 h(e_T) - M''(p_T), \tag{14}$$

$$U_{Tp_Tw_T} = p_T(w_T - w_0)h(e_T) + \frac{B'(e_T)}{C'(e_T)}, \tag{15}$$

$$U_{tp_tp_t} = (w_t - \sum_{i=0}^{T-t} \delta^i w_0 + \delta V_{t+1}^*)^2 h(e_t) - M''(p_t) \tag{16}$$

for $t \in \{1, .., T-1\}$, and

$$U_{tp_tw_t} = p_t(w_t - \sum_{i=0}^{T-t} \delta^i w_0 + \delta V_{t+1}^*)h(e_t) + \frac{B'(e_t)}{C'(e_t)} \tag{17}$$

for $t \in \{1, .., T-1\}$, where

$$h(e_t) := \frac{B''(e_t)C'(e_t) - C''(e_t)B'(e_t)}{(C'(e_t))^3}. \tag{18}$$

2.1 Determination of the optimal contract

The first order-conditions (10), (11) and (12) together tell us that

$$w_T^* = w_0 + p_T^* M'(p_T^*) \tag{19}$$

and

$$w_t^* = w_0 + p_t^* M'(p_t^*) - \delta V_{t+1}^* + \sum_{i=1}^{T-t} \delta^i w_0 \tag{20}$$

for $t \in \{1, .., T-1\}$. Using these expressions to substitute for w_t^* in the incentive constraints (8) and (9) yields

$$e_t^* = C^{-1}\left((p_t^*)^2 M'(p_t^*)\right) \tag{21}$$

for $t \in \{1, .., T\}$. With these expressions, it is now possible to substitute for e_t^* in the first-order condition (10) to get an equation that solely determines p_t^* for $t \in \{1, .., T\}$,

$$\frac{B'\left(C^{-1}\left((p_t^*)^2 M'(p_t^*)\right)\right)}{C'\left(C^{-1}\left((p_t^*)^2 M'(p_t^*)\right)\right)} p_t^* - 1 = 0. \tag{22}$$

Evidently, the equation characterizing the optimal monitoring accuracy is time independent and we have

$$p_t^* = p_{t+1}^* = p^*, \tag{23}$$

where p^* is the optimal monitoring accuracy in the static model. This in turn leads directly to another result. If we drop the time index on the monitoring accuracy (which we are now allowed to do), and look back at the expression (21), we can easily conclude that even the optimal effort level is time independent. Thus

$$e_t^* = e_{t+1}^* = e^*, \tag{24}$$

where e^* is the optimal effort level in the static model.

The time independency of the monitoring accuracy (23), together with the expression for the last period's wage (19), implies that the optimal wage in the last period is the same as the optimal wage in the static model, w^*, i.e. that

$$w_T^* = w^*. \tag{25}$$

To investigate the wage levels in other periods, first note that from the employee's ex post utility and the incentive constraint respectively in the last period, (2) and (7), together with the time independency of the monitoring accuracy, (23), we have:

$$V_T^* = (1 - p^*)(w_T^* - w_0) + w_0. \tag{26}$$

This utility can be used together with the employee's ex post utilities and incentive-compatibility constraints for other periods, (3) and (9), to solve backwards for an expression of V_{t+1}:

$$V_{t+1}^* = (1 - p^*)(w_T^* - w_0) + \sum_{i=0}^{T-t-1} \delta^i w_0. \tag{27}$$

This in turn substituted into the expression for the wage in other periods than the last period, (20), together with the expression for the wage in the last period, (19), finally yield the following expression for the employee's wage in all periods but the last one:

$$w_t^* = (1 - \delta(1 - p^*))w_T^* + \delta(1 - p^*)w_0 \tag{28}$$

for $t \in \{1, .., T-1\}$. The above results can be summarized in the following proposition:

Proposition 1 *The optimal wage path is constituted by the equilibrium wage of the static model in the last period and a lower and stationary wage for all preceding periods. Both wage levels are independent of the length of the relationship. The optimal effort level and the optimal monitoring accuracy are stationary and have the same values as in the static model.*

The wage path is illustrated in Figure 1.

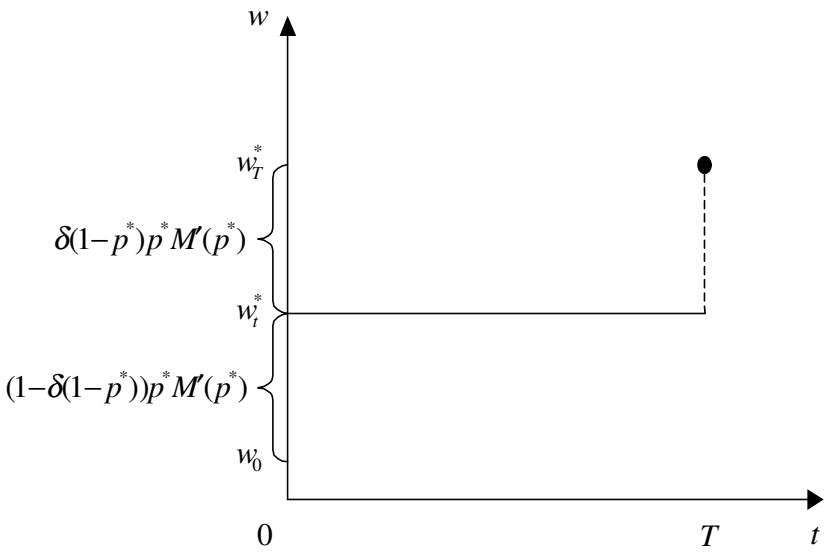

Figure 1

Here one can clearly see that, with no discounting of future periods ($\delta = 1$), all monetary incentives lie in the last period. The intuition is that worker rents in a preceding period exist only to compensate for the depreciation of future rents during that period.

The result implies that all second-order conditions are fulfilled if and only if the second-order conditions for the last period are fulfilled. Thus it is sufficient that

$$U^*_{Tw_Tw_T} < 0, U^*_{Tp_Tp_T} < 0$$

and

$$U^*_{Tp_Tp_T}U^*_{Tw_Tw_T} - (U^*_{Tp_Tw_T})^2 > 0,$$

where

$$U_{Tw_Tw_T} = (p_T)^2 h(e_T), \tag{29}$$

$$U_{Tp_Tp_T} = (w_T - w_0)^2 h(e_T) - M''(p_T) \tag{30}$$

and

$$U_{Tp_Tw_T} = p_T(w_T - w_0)h(e_T) + \frac{B'(e_T)}{C'(e_T)}, \tag{31}$$

to ensure that the interior solution is a max point. The second–order conditions for the last period are the same as are required in the static model; hence the second-order conditions will not put further restrictions on the solution by the above T-period extension of the static model.

2.2 The correlation between monitoring and pay

A quick glimpse at equation (28) also reveals that the main results from the static model remain intact. Clearly the monitoring accuracy and the wage in the last period are complementary instruments for motivating workers in this T-period extension if and only if the monitoring accuracy and the wage are complements in the static model; they are characterized by the same equations and thus have the same values. Thus, the following propositions can be stated:

Proposition 2 *According to any shift in the employer's benefit function, $B(e)$, or the employee's cost of effort function, $C(e)$, monitoring and the last period's wage are complementary instruments if and only if $-p^*M''(p^*)/M'(p^*) < 1$.*

Proposition 3 *According to any shift in the employer's monitoring cost function, $M(p)$, monitoring and the last period's wage are complementary instruments if and only if $U^*_{pw} > 0$.*

Equation (28) tells us that the optimal wage in other periods than the last period is a linear combination between the optimal wage in the last period and the minimum payment. The weight on the minimum payment decreases as the monitoring accuracy increases and vice versa, if the above conditions for complementarity hold. Hence, if monitoring accuracy and the wage in the last period are complementary instruments for motivating workers, then monitoring accuracy and the wage in other periods are as well. More formally, by differentiating the expression for the optimal wage in other periods than the last period, (28), with respect to the optimal monitoring accuracy,

$$\frac{dw_t^*}{dp^*} = (w_T^* - w_0) + (1 - \delta(1 - p^*))\frac{dw_T^*}{dp^*},$$

it is easily seen that if monitoring and the last period's wage are positively correlated, monitoring and the wage in other periods must also be positively correlated. However, the opposite is not true! The condition for complementarity between w_t^* and p^* is weaker than the condition for complementarity between w_T^* and p^*.

Proposition 4 *According to any shift in the employer's benefit function, $B(e)$, or the employee's cost of effort function, $C(e)$, monitoring and pay are complementary instruments if $-p^*M''(p^*)/M'(p^*) < 1$.*

Proposition 5 *According to any shift in the employer's monitoring cost function, $M(p)$, monitoring and the last period's wage are complementary instruments if $U^*_{pw} > 0$.*

The equation (28) also gives an insight into how the employer optimally should respond to changes in the discount factor. Clearly, the optimal monitoring accuracy, the optimal effort level and the optimal wage in the last period are all independent of the discount factor since they are the same as in the static case. On the other hand, the weight on the minimum payment in the linear combination determining the optimal wage increases as the discount factor increases and vice versa. Hence, we have that for all periods but the last one, the optimal wage is negatively correlated with the discount factor.

Proposition 6 *The optimal wage in all periods but the last period is negatively correlated with the discount factor, while the optimal wage in the last period, the optimal monitoring accuracy and the optimal effort level are independent of the discount factor.*

3 Welfare implications

Since both the exerted effort level e_t and the monitoring accuracy p_t are stationary, the length of the relationship does not affect the degree of social efficiency of the contract as long as the principal earns enough to start the project. The contract length will merely have distributional implications. It is easy to rewrite the expression (27) to get the following expression for the present value of getting hired by the firm:

$$V_1^*(T) = p^*(1-p^*)M'(p^*) + \sum_{i=0}^{T-1} \delta^i w_0. \tag{32}$$

Thus, the employee's per period compensation is decreasing with the length of the relationship. Conversely, the employer benefits from a longer relationship and will only offer the longest contract which is allowed for by the nature of the project or other exogenous circumstances.

In the absence of the legal rules making it impossible for the employee to fire a non-shirking employee, the employer is forced to repeat the static outcome T times. In this case, the present value for the employee becomes

$$\sum_{i=0}^{T-1} \delta^i V_1(1) = \sum_{i=0}^{T-1} \delta^i (p^*(1-p^*)M'(p^*) + w_0). \tag{33}$$

This is larger than the present value from the long-term relationship (32). Hence the following propositions can be stated:

Proposition 7 *The degree of social efficiency of the contract is not affected by the length of the relationship.*

Proposition 8 *The length of the relationship is a matter for dispute: the employer prefers longer and the employee prefers shorter relationships.*

Proposition 9 *Legal rules protecting employees from being fired lower the employees' present value of compensation.*

The welfare implications from changes in the employee's liability are almost trivial. It is easy to check that the equations (21) and (22) imply that the equilibrium monitoring accuracy, p^*, and the equilibrium level of effort, e^*, are both independent of the limited liability constraint, w_0. Thus the degree of social efficiency is independent of the employee's liability. This should be said with one important reservation though. The equations (19) and (28) reveal that

$$\frac{dw_t^*}{dw_0} = \frac{dw_T^*}{dw_0} = 1. \tag{34}$$

In other words, an increase in the limited liability parameter redistributes wealth from the employer to the employee. If the limited liability parameter becomes high enough, it is no longer profitable for the employer to hire the employee, and the degree of social efficiency will of course be affected. Hence, it can be stated that:

Proposition 10 *So long as the employer can afford to hire the employee, the degree of social efficiency is not affected by changes in the employee's liability.*

4 Implicit bonding contracts

The above version of the efficiency wage model still implies rents to employees, even if they diminish with the length of the relationship. This common feature of the shirking model of efficiency wages has been criticized in many articles. Carmichael (1985), for example, argues that these rents should be eliminated by entrance fees or bonds. The idea of a bond is found in for example Lazear (1981) and Becker and Stigler (1974). There the employee posts a bond initially, is paid interest on it, and gets it back when the relationship is over. The bond transactions are included in the employee's wage path in their papers. One result is a negative compensation in the first period, and a constant compensation in intermediate periods which is lower than the higher compensation in the last period.

To understand this result in the framework of the model in this paper, an important clarification is called for. A factor closely related to the limited liability constraint is the employee's initial wealth. If contractible, i.e. if it can be taken away from the employee in the case where he does not meet or exceed his effort target, the wealth will add liability and hence lower the liability constraint. Wealthier employees would then be more attractive to hire since, as we learned from the previous section, all periods' equilibrium wages will be lower. On the other hand, if the employee cannot commit to keeping his wealth contractible, as soon as he is hired, he will try to make his wealth non-contractible or simply get rid of it by for example consuming it. The commitment is something that can be accomplished by a bond. This paper will make a clear distinction between the wage and the bond transaction by simply assuming that the wealth that is used for a bond never leaves the employee. A bond is just a contract which says that some of the employee's initial wealth, that can be taken away from him if he does not meet or exceed his effort target, must remain intact during the employment relationship. This is entirely captured by a decrease in the limited liability parameter equivalent to the size of the bond.

The idea of employees purchasing job opportunities, which is sometimes mentioned in this context, is something completely different if the relationship between the employer and the employee lasts longer than one period. If the employer sells a job and the employee uses hi contractible wealth to pay for it, the employer will regret selling the job as soon as he realizes that the equilibrium wage path has shifted upwards and, according to equation (34), he will indirectly repay the employee the whole price during each period of their relationship. It is easy to check that if the employee instead uses non-contractible wealth to pay for the job opportunity, it is always better for the employer to let him keep the money if he can commit, for example via a bond, not to spend it (i.e. if the relationship lasts longer than one period). Thus, for an employer to charge an entrance fee is always a bad idea.

4.1 Firm-specific human capital investments

Both bonds and entrance fees have the unappealing property of introducing moral hazard into the model. That is, the employee will have an incentive to claim that the employee shirks even if he does not, so that he can appropriate the bond or the entrance fee. Maybe that is the reason why we do not see bonds more frequently. However, a bond can appear in other shapes. Under reasonable assumptions one can interpret an employee's investment in firm-specific human capital as an implicit bond. The only feature of wealth that is of interest when using it for a bond is that the employee will be hurt by

losing it, and this is true for firm-specific human capital as well.

Below, it will be demonstrated how a firm-specific human capital investment can be used as a substitute for an entrance fee and at the same time solve the moral hazard problem in a self-enforcing static efficiency wage contract. The moral hazard problem disappears because if the employer fires the employee, he will lose the firm-specific human capital investment at the same time. The analogous reasoning applies for a multi-period model, but then it is for a bond instead of an entrance fee.

The timing is as follows. First, many risk-neutral employees who compete for a job opportunity can invest in firm-specific human capital at a cost $h \in \left[0, \bar{h}\right]$. The investment is sunk and gives an employee a private benefit of exactly h, if he is engaged by the firm; thus there are no *direct* gains from the investment. Then, a risk-neutral employer observes the education levels and employs a single employee. After the employee is hired there are only two possible outcomes of his private benefit from education: his chosen level of education if he remains hired and zero if he gets fired. It is easily seen that it is best for the employer to give the employee $w + h$ if he meets or exceeds the target and the minimum payment w_0 otherwise. Hence, now the incentive compatibility constraint becomes

$$p(w - w_0 + h) \geq C(\hat{e}) \tag{35}$$

and the expression for the actual effort which the employee will exert is

$$e(p, w) := C^{-1}((w - w_0 + h)p). \tag{36}$$

The two equations characterizing the solution, (19) and (22), now become

$$w^* = w_0 - h + pM'(p^*) \tag{37}$$

and

$$p^* \frac{B'(C^{-1}((p^*)^2 M(p^*)))}{C'(C^{-1}((p^*)^2 M(p^*)))} - 1 = 0. \tag{38}$$

To investigate the implications of the employee's firm-specific human capital investment, the employer's and the employee's utilities in equilibrium are differentiated with respect to the education level, h, chosen by the employee. The employer's utility in equilibrium is

$$U(p^*, w^*) = B(C^{-1}(p^*(w^* - w_0 + h))) - w^* - M(p^*) \tag{39}$$

and the employee's utility in equilibrium is

$$\begin{aligned} V(p^*, w^*) &= w^* - C(C^{-1}(p^*(w^* - w_0 + h))) \\ &= w^* - p^*(w^* - w_0 + h). \end{aligned} \tag{40}$$

Equation (38) reveals that the optimal monitoring accuracy, p^*, is independent of h. Knowing this, it is easy to differentiate equation (37) to get

$$\frac{dw^*}{dh} = -1. \tag{41}$$

This will be of help when differentiating (39) and (40) with respect to the education level chosen by the employee:

$$\frac{dU(p^*, w^*)}{dh} = \frac{B'(e^*)}{C'(e^*)} p^* \left(\frac{dw^*}{dh} + 1 \right) - \frac{dw^*}{dh} = 1 \tag{42}$$

and

$$\frac{dV(p^*, w^*)}{dh} = \frac{dw^*}{dh} - p^* \left(\frac{dw^*}{dh} + 1 \right) = -1. \tag{43}$$

It can now be seen that the firm-specific human capital investment indirectly redistributes wealth from the employee to the employer. The amount that is taken away from the employee and given to the employer is exactly the same as the former's private benefit from his education level.

Now, go back in time to when many employees compete for the job opportunity. The goal is to find a Nash equilibrium in education level choices. Firstly, any situation in which more than one employee chooses a positive level of education cannot be a Nash equilibrium. This is because all the employees who chose a positive level of education and were not hired would then have preferred to choose another education level (zero, or higher than the hired employee). Secondly, any situation in which the hired employee receives rents, i.e. any situation in which

$$V(p^*, w^*) = w^* - p^*(w^* - w_0 + h) > w_0 \tag{44}$$

or more simply, in which

$$h < (1 - p^*) p^* M'(p^*), \tag{45}$$

cannot be a Nash equilibrium, since then another employee would have wanted to undercut him by choosing a higher education level. And finally, any situation in which the hired employee receives negative rents, i.e. any situation in which

$$h > (1 - p^*) p^* M'(p^*), \tag{46}$$

cannot be a Nash equilibrium, since then the hired employee would have been better off by choosing a zero education level.

Thus, the only remaining candidate for a Nash equilibrium is that one employee chooses the education level

$$h = (1 - p^*)p^*M'(p^*), \tag{47}$$

and all the others choose the education level zero. And this is indeed a Nash equilibrium. Hence, the following proposition can be stated:

Proposition 11 *If employees can freely choose to invest in firm-specific human capital or if their potential education level, $\bar{h}$, is higher than $(1 - p^*)p^*M'(p^*)$, the hired employee will receive no rents, i.e. his individual rationality constraint will be binding.*

And if the upper bound of the education choice is binding for the employee, it can be stated that:

Proposition 12 *If the employee's potential education level, $\bar{h}$, is less than $(1 - p^*)p^*M'(p^*)$, the hired employee will receive rents, i.e. his individual rationality constraint will not be binding.*

5 Extensions and further research

5.1 Firm commitment, or explicit bonding contracts

A natural extension of the model is to allow the employer to commit himself to future control variables. So far, no restriction on the equilibrium wage path has been imposed on the model. In a setting where the employer cannot commit to future control variables, such a restriction is obviously not needed since equilibrium wages are always higher than the limited liability constraint. In the first example below, a minimum wage restriction that coincides with the employee's limited liability is imposed. Then it is discussed what happens in the absence of that restriction. Assume for the sake of simplicity that $T = 2$. Using the expression for the employee's ex post utility in the last period (26) to substitute for V^*_{t+1}, the incentive compatibility constraints (9) and (8) now become

$$e_1 = C^{-1}\left(p_1(w_1 - w_0 + \delta(1 - p_2)(w_2 - w_0)\right) \tag{48}$$

and

$$e_2 = C^{-1}\left(p_2(w_2 - w_0)\right). \tag{49}$$

The employer will use these expressions to substitute for e_1 and e_2 in his utility function. Hence he will face the following maximization problem:

$$\underset{p_1,p_2,w_1,w_2}{Max} \quad B(e_1) - w_1 - M(p_1) + \delta(B(e_2) - w_2 - M(p_2))$$

subject to the constraints $w_1, w_2 \geq w_0$ and $p_1, p_2 \in [0,1]$. Form the Lagrangian

$$\begin{aligned} L &= B(e_1) - w_1 - M(p_1) + \delta(B(e_2) - w_2 - M(p_2)) \\ &\quad +\lambda_1(w_1 - w_0) + \lambda_2(w_2 - w_0) \end{aligned}$$

and denote the solution to this problem $(p_1^*, p_2^*, w_1^*, w_2^*)$. Furthermore, let the corresponding amount of executed effort, $e_1(p_1^*, p_2^*, w_1^*, w_2^*)$ and $e_2(p_2^*, w_2^*)$ be denoted e_1^* and e_2^* respectively. The Kuhn-Tucker conditions are then:

$$\frac{B'(e_1^*)}{C'(e_1^*)}(w_1^* - w_0 + \delta(1 - p_2^*)(w_2^* - w_0)) - M'(p_1^*) = 0, \tag{50}$$

$$\frac{B'(e_1^*)}{C'(e_1^*)}p_1^* - 1 + \lambda_1 = 0, \tag{51}$$

$$\lambda_1(w_1^* - w_0) = 0, \tag{52}$$

$$\frac{B'(e_2^*)}{C'(e_2^*)}\delta(w_2^* - w_0) - \frac{B'(e_1^*)}{C'(e_1^*)}\delta p_1^*(w_2^* - w_0) - \delta M'(p_2^*) = 0, \tag{53}$$

$$\frac{B'(e_1^*)}{C'(e_1^*)}\delta p_1^*(1 - p_2^*) + \frac{B'(e_2^*)}{C'(e_2^*)}\delta p_2^* - \delta + \lambda_2 = 0, \tag{54}$$

and

$$\lambda_2(w_2^* - w_0) = 0. \tag{55}$$

It is easily seen that λ_2 must be zero, because otherwise the condition (55) tells us that w_2^* must be equal to w_0, which substituted into condition (53) together with the incentive constraint (49) implies zero effort in the second period. Then the employer would rather skip the last period. λ_1 on the other hand cannot be zero, because if it is, the conditions (51) and (54) require that $B'(e_2^*)/C'(e_2^*) = 1$, which again substituted into condition (53) together with the incentive constraint (49) implies zero effort in the second period. This is an example of the general result that is stated in the following proposition:

Proposition 13 *If the employer can commit himself to future control variables, the optimal wage path will be constituted by the minimum wage in all periods but the last period.*

The wage path is illustrated in Figure 2.

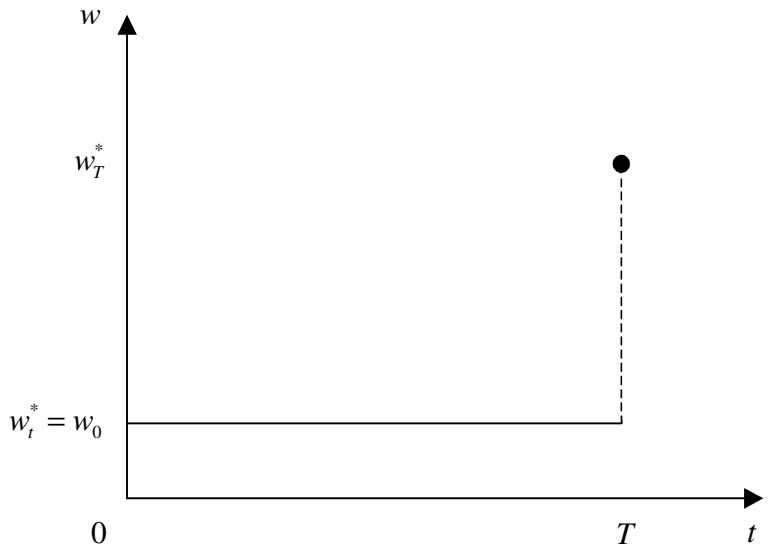

Figure 2

This result is not surprising. If the employee believes the employer's promises of future rewards, the existence of a wage premium in earlier periods than the last period is a conspicuous inefficiency. Without changing the incentives in one period and adding extra incentives in subsequent periods, a wage premium in that period can be replaced by the present value of this wage premium in the last period at no cost to the employer.

The minimum wage restriction is binding in the first period, and without it the employer would choose a lower wage in the first period together with a promise of a higher second period wage. In fact, he would minimize monitoring costs and approach the first best solution in the second period by an infinitely large fee in the first period together with an infinitely large second-period wage. In practice, even if there is no minimum wage restriction, the equilibrium wage may at least be restricted to be non-negative. But the intuition is straightforward: if the employer is able to make commitments he should save as much as possible of the payment for the last period.

5.2 Savings

If the limited liability constraint is affected by the employee's earlier wages, the resulting optimal wage path is affected as well. Say for example that the employer's wage is used entirely to add up liability. Then every dollar that is given to the employee in the first period makes the employer able to subtract one dollar from the employee's paycheck in all the remaining periods. But a

rational employee who foresees this reaction of the employer would of course spend all his money on consumption, leaving nothing to save. In a more general setting one could include the mechanism that employees compete for a job opportunity with commitments to save, thus gradually lowering their limited liability constraints. This may be done to the extent that all the hired employee's surplus is given up to the employer. It is noteworthy that a commitment to save can be made even if the employee cannot afford to buy a bond. The analogical reasoning applies for firm specific-human capital. This may be the reason why we so often see employers letting their employees educate themselves during working hours.

6 Concluding remarks

The unique optimal wage path derived in this paper shows similarities with one of the possible optimal wage paths described in Lazear (1981). There, the employee receives a wage lower than the value of his marginal product over his lifetime, but is compensated by a large lump sum in the last period to set present values of payment and marginal products equal. There is an important difference though; here the employee still receives some rents inherited from the efficiency wage mechanism, even if they are strongly diminished by the length of the relationship.

Another possibility that Lazear presents is the same as that first found in Becker and Stigler (1974). The employee posts a bond initially, is paid interest on it, and gets it back when the relationship is over. The result is a negative compensation in the first period, a constant compensation in intermediate periods lower than the higher compensation in the last period and in total, no rents to the employee. There are two obvious differences compared to the result in this paper: the first period compensation and the employee's rents. The reason for the first is only technical. In this paper, the idea of a bond is captured by the limited liability constraint. If the employee has initial wealth that could be used for a bond, it is simply assumed to lower the limited liability constraint, w_0. The employee doe not have to post the bond as long as the wealth that was meant for the bond is contractible. The reason for the second difference is more important. One of the goals with this paper has been to explain an upward sloping wage profile with a self enforcing contract, without introducing a reputation mechanism, a third part or the possibility to make commitments. If this is left out, the moral hazard problem of employer default is not a problem, as shown in for example Bhattacharya (1987), Carmichael (1983a, 1983b) and Malcomson (1984, 1986). And if the moral hazard problem is not present, it is not hard

to eliminate the employee's rents.

A third path proposed by Lazear is that the employee is paid less than the value of his marginal product initially, and receives a pension for some periods after the actual relationship is over. If one relaxes the assumption that the employer cannot commit to future control variables, and assumes that he can commit to pay a pension, then there is an infinite number of optimal contracts. The employer could for example distribute the last period's wage premium plus compensation for the resulting depreciation evenly over a number of following periods. The employer's gain from postponing some of the payment is exactly the same as the employee's loss, since they have identical discounting factors. But since the employer has to compensate for the employee's loss, both the employer and the employee will be indifferent between no pension and all kinds of such pension arrangements.

Lazear also finds that wage profiles which are flatter and more smoothly increasing than step functions suggested above can be optimal. He argues that in addition to the employee's incentive gains from a steep wage profile, there is a loss represented by a higher temptation for the employer to breach the contract. This is a trade-off that is excluded in the model analyzed in this paper. It simply assumes that the employer cannot fire the employee if he is not caught shirking. However, the employer is not tempted to breach the contract because he makes a positive profit in the last period. It is optimal to pay the higher wage in the last period, resulting from efficiency wages contrary to the delayed payment/bonding contract wage, not because it compensates the employee for previous periods, but because it yields the correct incentives in the last period!

Influential papers presenting efficiency wage models, such as Shapiro and Stiglitz (1984) and Bulow and Summers (1986), have ruled out the possibility of delayed payment/bonding contracts. This is mainly because they have not been satisfied with the treatment in earlier versions of the employer's moral hazard problem and the employee's wealth constraint. Maybe the most important contribution of this paper is that it builds a bridge between efficiency wages and delayed payment/bonding contracts; it demonstrates how a postponed payment can emerge in a self-enforcing contract, even if the employee has no initial wealth. Thus, in opposition to previous intuitions, there should be no antagonism between efficiency wages and delayed payment/bonding contracts. On the contrary, they are two compatible mechanisms which the profit maximizing firm should use in concord.[1]

[1]For other hybrid models of efficiency wages and delayed payment/bonding contracts, see Eaton and White (1982) and Dickens, Katz, Lang and Summers (1989).

References

Allgulin, Magnus (1999): Monitoring and Pay: General results, *SSE/EFI Working Paper Series in Economics and Finance* No. 340, Stockholm School of Economics.

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

Becker, Gary and George Stigler (1974): Law Enforcement, Malfeasance, and the Compensation of Enforcers, *Journal of Legal Studies* III, 1-18.

Bhattacharya, Sudipto (1987): Tournaments, Termination Schemes, and Forcing Contracts, Mimeo, School of Business, University of California, Berkeley.

Bull, Clive (1987): The Existence of Self-Enforcing Implicit Contracts, *Quarterly Journal of Economics* February, 147–159.

Bull, Clive (1985): Equilibrium Unemployment as a Worker Discipline Device: Comment, *American Economic Review* September, 890–891.

Bulow, Jeremy I. and Lawrence H. Summers (1986): A Theory of Dual Labor Markets with Applications to Industrial Policy, Discrimination, and Keynesian Unemployment, *Journal of Labor Economics* 4, 376–414.

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

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

Carmichael, H. Lorne (1983b): The Agent-Agents Problem: Payment by Relative Output, *Journal of Labor Economics* January, 37-50.

Carmichael, H. Lorne (1983a): Firm Specific Human Capital and Promotion Ladders, *Bell Journal* Spring, 241-258.

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

Dickens, W. T., L. F. Katz, K. Lang and L. H. Summers (1989): Employee Crime and the Monitoring Puzzle, *Journal of Labor Economics* vol. 7, no. 3, 331–347.

Eaton, C. and W. D. White (1982): Agent Compensation and the Limits of Bonding, *Economic Inquiry* XX, 330–343.

Lazear, Edward (1995): *Personnel Economics,* (Cambridge Mass.: MIT Press).

Lazear, Edward (1981): Agency, Earnings Profiles, Productivity and Hours Restrictions, *American Economic Review* LXXI, 606-620.

Malcomson, James (1986): Rank Order Contracts for a Principal with Many Agents, *Review of Economic Studies* October, 807-818.

Malcomson, James (1984): Work Incentives, Hierarchy, and Internal Labor Markets, *Journal of Political Economy* June, 486-507.

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