TRAINING, JOB SECURITY AND INCENTIVE WAGES

MARGARITA KATSIMI

CESIFO WORKING PAPER NO. 955
CATEGORY 4: LABOUR MARKETS
MAY 2003

TRAINING, JOB SECURITY AND INCENTIVE WAGES

Abstract

This paper considers the optimal level of firm-specific training by taking into account the positive effect of training on the expected duration of workers' current employment. In the framework of an efficiency wage model, a short expected job tenure represents a disamenity that reduces the penalty from shirking. As this disamenity increases, workers have an incentive to continue providing a positive level of effort only if they are compensated by a higher wage. We endogenize the employment separation rate by introducing firm-specific training. Firm-specific training creates a rent that is lost if the worker is separated from the firm. As a result, the firm will be more reluctant to fire its trained workforce in a recession. This implies that firm-specific training can decrease current wages as it implies a credible commitment to lower future labour turnover.

Margarita Katsimi
Dept of Int.and European Economic Studies
Athens University of Economics and Business
76, Patision Avenue
10434 Athens
Greece
Mkatsimi@aueb.gr

1 Introduction

As early developments of the human capital theory suggest, worker's uncertainty about future employment has a negative impact on investment in firm-specific training. According to Becker (1962, 1964) and Oi (1962), workers will be reluctant to bear the cost of this investment if there is a positive probability of being dismissed and not benefit from investment returns. In Booth and Chatterji (1989) workers will be induced to undertake specific training in sectors where there is a positive probability of being redundant if a redundancy payment is part of the firm's contract. This paper argues that the relationship between firm-specific training and the level of job security is two-fold:[1] Firstly, investment in firm-specific training creates a surplus in employment relationships. The firm can enjoy part or all of this rent as long as the worker remains employed. This implies that in a recession the optimal policy of the firm will be to retain the employment relationship even if the wage paid to the worker exceeds his marginal product in an alternative job. As a result, investment in firm-specific training becomes a credible commitment to lower future labour turnover. Secondly, a credible promise for high future employment will decrease the probability of being dismissed in a downturn and, in the framework of an efficiency wage model, it will lower wages. The ability of firm-specific training to provide this commitment mechanism will have a positive effect on the level of firm-specific training. Specifically, the firm will invest in firm specific training even if marginal cost exceeds expected marginal return since training 'buys' the firm a commitment technology. This implies that the marginal value of training exceeds expected marginal return.

The idea that job security will have a negative impact on efficiency wages has been analyzed by Katsimi (1995), Saint-Paul (1996) and Fella (2000).[2] All authors base their arguments on different versions of the Shapiro and Stiglitz (1984) model which shows that in an environment

---

[1] In this paper job security is defined as a negative function of the probability of being fired in a recession.

[2] Empirical evidence suggesting a wage premium based on the job's unemployment risk has been found by Abowd and Ashenfelter (1981), Adams (1985) and Li (1986). Research on this topic dates back to the pioneering work of Hall (1970, 1972).

where employers are unable to monitor workers' on-the-job effort costlessly, unemployment can induce workers to provide a positive level of effort. In this framework, a short expected job tenure represents a disamenity that reduces the penalty for shirking. As this disamenity increases, workers will have an incentive to continue providing a positive level of effort only if they are compensated by an increase in the wage rate. The optimal incentive wage will increase with the risk of being fired in a recession. In other words, the cost of labour decreases when the firm is expected to retain its workforce in a recession. This, however, requires either that the firm can commit to a low firing policy or that workers expect that low firing is the optimal future employment policy of the firm. In Saint-Paul (1996), firms' commitment to a future level of employment is represented by the fact that employment is set one period in advance. In Katsimi (1995) and Fella (2000), the imposition of redundancy payments by the government leads to a higher level of expected future employment in a recession.

This paper considers the case where the above commitment mechanism is endogenous. In the absence of an imposed turnover cost such as a redundancy payment, a profit-maximizing firm may have an incentive to voluntarily increase the cost of dismissing the existing workforce in a recession. In that context, the level of firm-specific training can be viewed as a 'commitment' to higher future employment in a recession. It is assumed that specific human capital is a team investment concerning all workforce and involves costs and returns only for the firm.[3] Forward looking workers perceive that a firm who has undertaken investment in firm-specific training will fire less in a recession. Since workers are uniformly trained at an optimal level, firing workers in a recession implies the loss of some of the return to training. If a recession occurs in the period when the investment return is realized, a reduction in employment implies the loss of the return from the human capital acquired by the dismissed workers. As a result, a positive level of training in the previous period will increase optimal employment in a recession. A higher level of expected employment in a downturn of the economy will decrease the probability of being fired

---

[3] For simplification we neglect the uncertainty stemming from quitting.

in a recession for each worker. In other words, firms' decision to undertake firm specific training improves the 'security' of workers' current employment and increases the penalty of shirking with a negative impact on incentive wages. This wage-reducing effect increases the value of human capital investment leading to a higher level of firm-specific training. *Ceteris paribus* the level of investment is higher the higher is the elasticity of the firing probability with respect to the level of training.

The rest of the paper is organized as follows: Section 2 presents the basic arguments of the paper in the framework a two-period model. In section 2.1, wage and employment determination is discussed. Firm's decision to invest in firm-specific human capital is described in section 1.3. Section 3 extends the basic model to an infinite horizon framework. The last section concludes.

2   A two-period model of employment and training

2.1   Wage and employment determination

We first present a two-period model that builds on Shapiro and Stiglitz (1984) shirking model by endogenizing workers' horizon on their job and allowing for investment in firm-specific training. We will first consider wage and employment determination in the absence of training. The economy consists of a continuum of firms indexed by j, where $j \in [0, 1]$. Each firm produces output by the following production function in each period:

$$Y_{j,t} = e g_{j,t}^i N_{j,t}^a \tag{1}$$

where Y is the level of output, N denotes the level of employment, e is the level of effort produced by each worker and and the subscript t denotes the time period, $t = 1, 2$. Output depends on a productivity shock that can be either high, g^H, or low, g^L, $g^H > g^L$. We will assume that g takes the high value in the first period. In the second period, g will take the low value with probability h whereas it will take the high value with probability $1 - h$.[4]

---

[4] This assumption is consistent with the two-state Markov process assumed for the infinite horizon case later on.

Firms set wages unilaterally in order to maximize profits. There is asymmetric information between employers and employees about the on-the-job effort of the latter: the monitoring technology is imperfect. Workers are risk neutral and their utility depends on the amount of goods they can consume with their wage, w, and on their job effort, e. The instantaneous utility, u, of the representative employed worker in each period is given by:

$$u_t = w_t - e \tag{2}$$

In the Shapiro and Stiglitz (1984) model, wages are paid before effort is observed and the penalty for shirking is the loss of the worker's job. However, in the framework of a finite horizon model this would imply no punishment for shirking in the last period. In the presence of a last period it is clear that firms should hold some of the workers pay for the end of the period in order to prevent workers from shirking. Thus, although in the infinite horizon version presented in a later section we will retain the timing of the Shapiro and Stiglitz (1984) model, in this section we will modify the timing along the lines of the simple life cycle incentive model presented by Carmichael (1989) in order to ensure enforceability. Specifically, we will assume that workers' pay consists of two parts: a legal minimum wage, w^m, that is paid to all workers independently from their level of effort and a bonus payment, B_t, that is paid after effort has been monitored only to non-shirkers.[5]

$$w_t = w^m + B_t \tag{3}$$

---

[5] Note that the presence of a legal minimum wage is crucial. In the absence of a legal minimum wage the cost minimizing wage paid by the firm will be the same in each period, workers enjoy no rents and the threat of being fired is ineffective [Carmichael (1989)].

Table 1 summarizes the timing of events.

Table 1

Period 1
* the prod. shock g_1 is realized, $g_1 = g^H$
* firms set wages, w_1 and employment, N_1
* workers choose effort, e and production takes place
* workers who chose $e = 0$ get fired and receive w^m with prob. p
* workers who chose $e > 0$, receive w^m and a bonus B_1
Period 2
* the prod. shock g_2 is realized
* firms set wages, w_1 and employment, N_1 :
-workers can lose their job with prob. hq
-unemployed can get a job with prob. φ
* workers choose effort, e and production takes place
* workers who chose $e = 0$ receive w^m with prob. p
* workers who chose $e > 0$ receive w^m and a bonus B_2

At the beginning of the first period, firms set employment and announce a bonus payment by taking into account the optimal behaviour of workers in the next stage. Then, each worker decides either to shirk and provide minimal effort ($e = 0$) or not to shirk in which case he provides some fixed positive level ($e = \bar{e}$). If a worker chooses to shirk, there is a probability p that his effort will be monitored in which case will not receive the bonus payment and he will be dismissed. In the second period, firms observe the productivity shock, g_2, and reset wages and the bonus payment while adjusting their workforce so that the marginal productivity condition for profit maximization is satisfied. Again, there is a probability p that the effort level of each worker will

be monitored, in which case if $e = 0$ the worker will only receive the minimum wage.

All employed workers may lose their job in the second period if the firm switches to a low state with probability q. Thus, uncertainty in the first period results from the unknown value of the productivity shock in the next period. We denote by q the probability that a worker will lose his job in a low state in period 2:

$$q = \frac{N_1^H - N_2^L}{N_1^H} \tag{4}$$

Clearly, the lower the expected level of employment in the low state, N_2^L, the higher the probability that a worker will be fired.[6]

Both workers and employers are forward-looking: workers decide on their level of effort by taking into account firms' current and future optimal behaviour and firms decide on the level of training, wages and employment by taking into account workers' current and future incentives. Thus, we solve our model by backward induction.

2.1.1 The last period

Firms set the value of the bonus payment so as to minimize labour cost subject to being able to prevent workers from shirking. Workers will provide the amount of effort that maximizes their utility. After the realization of the productivity shock in period 2, there is no uncertainty for workers who do not shirk. The value of not shirking in the last period is given by equation (2) when $e = \bar{e}$.

$$V_2^{NS} = w^m + B_2 - \bar{e} \tag{5}$$

Since shirking involves the risk of being caught and receiving only the minimum wage with probability p, the value of shirking in the second period is given by

---

[6] We assume that concerns about their reputation prevents firms from falsely claiming that cheating has occured in order to avoid the bonus payment. Although we do not model this incentives, young workers may provide effort at a higher wage if they suspect that they may be trated badly in the future.

$$EV_2^S = w^m + (1-p)B_2 \tag{6}$$

Each worker will choose a positive level of effort as long as the value of not shirking exceeds the value of shirking:

$$V_2^N \geq EV_2^S \tag{7}$$

Solving expression (7) as an equality gives the lower level of real wage at which workers will not have an incentive to shirk, w_2^E:

$$w_2^E = w^m + \frac{\bar{e}}{p} \tag{8}$$

Each profit maximizing firm will increase its workforce until the real wage equals the marginal product of labour :

$$w_2^E = a\bar{e}g_2^i N_2^{a-1} \tag{9}$$

Solving the profit maximization condition for N, gives the equilibrium level of employment:

$$N_2 = \left(\frac{pw^m + \bar{e}}{pg_2^i a}\right)^{\frac{1}{a-1}} \tag{10}$$

2.1.2 The first period

At the beginning of the first period, each worker selects an effort level in order to maximize his expected utility in the two periods by taking into account optimal policies in the next period. A worker who decides not to shirk in the first period will always face a probability of being fired in the next period in the presence of a negative productivity shock. Assuming a zero discount rate, the value of not shirking is given by

$$EV_1^N = w^m + B_1 - \bar{e} + (1-hq)\max[V_2^{NS}, V_2^S] + hqU_2 \tag{11}$$

where U_2 is the value of being unemployed in period 2, h is the probability of a recession, and q is the probability of being fired in a recession defined by equation (4).

At the beginning of the second period, workers may find themselves out of job either due to shirking in the first period or due to the firm's transition to a low state. The value from shirking in the first period can be written as:

$$EV_1^S = (1-p)\{w^m + B_1 + (1-hq)\max[V_2^{NS}, V_2^S] + hqU_2\} + p(w^m + U_2) \qquad (12)$$

The terms in the squiggled brackets on RHS of equation (12) states the expected utility of a shirker who does not get monitored in the first period and the remaining terms give his expected utility when he gets caught and fired in period 1.

Assuming no unemployment benefit, and denoting with φ the probability that a worker finds a job in the second period, the value of being unemployed is given by:

$$U_2 = \varphi \max\{EV_2^S, V_2^{NS}\} \qquad (13)$$

By using equations (11) and (12), we derive that the worker will provide a positive level of effort in the first period as long as

$$B_1 \geq \frac{\bar{e}}{p} - (1-hq)(1-\varphi)\max[U_2^{NS}, U_2^S] \qquad (14)$$

Substituting (14) into (13) gives the lower wage at which workers will provide a positive level of effort:

$$w_1^E = (w^m + \frac{\bar{e}}{p})(\varphi + hq(1-\varphi)) \qquad (15)$$

The wage that satisfies the NSC will depend positively on the probability of exiting unemployment since a high φ decreases the penalty from shirking. A better monitoring technology

reflected in a high p will increase the unemployment probability for each worker who decides to shirk, leading to a lower expected utility of shirking. Similarly, a decrease in the disutility of effort, $\bar{e}$, will imply a lower opportunity cost of not shirking, which will lower w_1. Finally, the efficiency wage will depend positively on the probability of being fired for cyclical reasons. As the probability of being fired in a low state increases, expected job tenure falls and so does workers' expected utility from being employed. Workers view a long job tenure as a form of compensation. Thus, if for any reason expected job tenure decreases, wages should increase for workers' expected utility to remain constant and vice versa. In the context of this model, a long job tenure implies a low value of q and h. One can see from equation (4) that the probability of being fired in a recession, q, depends negatively on the expected level of employment in the low state, N_2^L. In Figure 1, the non-shirking condition is satisfied at a wage rate equal to w^*. As expected job tenure increases, the expected utility from not shirking will now exceed the expected utility from shirking. This implies that the NSC is now satisfied for a lower level of wage. The NSC shifts to the left and the level of the efficiency wage decreases to w'.

Employment in the first period can be derived by the marginal productivity condition:

$$w_1^E = \bar{e} a g_1^H N_1^{a-1}$$

2.2 Firm-specific training

We will now consider the firms' decision to invest in specific training. We assume that in the first period the firm decides to invest on b units of firm-specific training for each worker. Firms bear the total cost of training , $c(b)N_1$. We assume that the workers' effort required by training is perfectly observable so that shirking during training is not an option for workers in the first period.[7]   Training is characterized by diseconomies of scale which implies that $c(b)$ is a twice

[7] Therefore, even if the cost of training represents an extra payment to workers it will not affect the incentive wage since all workers will receive this payment in the first period. Alternatively one can assume that training requires effort that is not perfectly observable. In that case, the cost would take the form of higher wages in the first period due to a higher e. However, this assumption would complicate the algebra without altering any of the results.

EVs-EVns
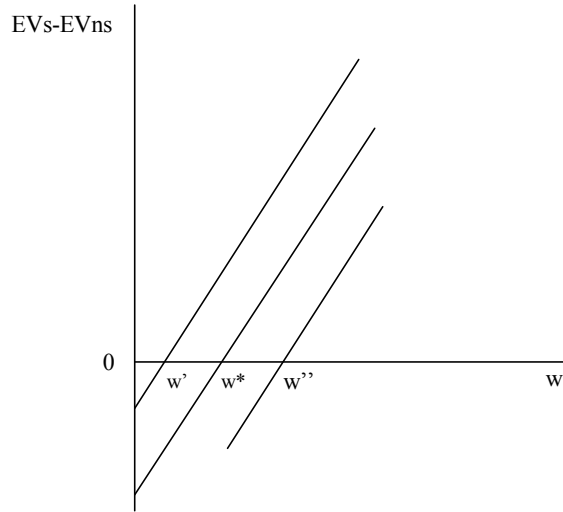
Figure 1: The non-shirking condition

differentiable, strictly convex function. Returns to training are realized in the second period. We assume that these returns by the firm are not threatened by the possibility of a premature quit. b units of training in the first period will increase productivity by $r(b)$ in the next period. $r(b)$ is a twice differentiable, strictly concave function that can represent an index of worker quality.[8] This implies that uncertainty over investment returns stems only by the uncertainty of the economic environment.[9] It is assumed that the expected return is a linear function of employment so that output in period 2 is defined as:

$$Y_2 = g^i (N_2^i)^a + r(b) \min\{N_1, N_2\} \tag{16}$$

We assume that training improves only firm-specific skills so that in period 2 the productivity

---

[8] According to Hart and Moutos (1995), this index may represent the speed at which the product is produced or the number of saleble units achieved at a given rate of production so that more investment leads to more saleable units.

[9] Hashimoto (1979, 1981) among others investigates the importance of uncertainty over investment returns to specific training when there is uncertainty over $r(b)$.

in firm k of a worker who worked in firm k in the previous period exceeds the productivity in firm k of a worker employed in firm l in the previous period by $r(b)$. The introduction of training makes the profit maximization decision of the firms dependent across the two periods. Expected profits in the two periods are given by

$$
\begin{aligned}
E(\Pi) &= \Pi_1 + E(\Pi_2) \\
&= Y_1 - w_1^E N_1 - c(b)N_1 + E[Y_2 - w_2^E N_2]
\end{aligned}
\tag{17}
$$

In period 2, there is no uncertainty and employment in the low state must satisfy the following FOC:

$$
g^L a N_2^{a-1} + r(b) = w_2
\tag{18}
$$

Solving equation (18) for N_2 gives the equilibrium level of employment in the low state in the second period.

$$
N_2^L = \left(\frac{pw^m + \bar{e} - r(b)p}{pga} \right)^{\frac{1}{a-1}}
\tag{19}
$$

From equation (19), it is easy to see that employment in a recession will depend positively on the level of b so that the introduction of training will lead to higher employment in period 2, $\frac{\vartheta N_2^L}{\vartheta b} > 0$. Firms will be more reluctant to fire workers after a negative productivity shock, since this is associated with losing part of the rent created by training. Note that in our analysis we assumed that monitoring takes place in per capita terms. If one were to assume monitoring per units of output, then one cannot exclude the possibility that the probability of being monitored in equation (20) would be a negative function of $r(b)$. However, the positive effect of training on N_2^L would still carry through if one assumes economies of scale in monitoring technology.[10] One

[10] For example, a camera can be used for monitoring the production of more than one unit of output. A proper assessment of the effect of productivity on the monitoring probability would, however, require endogenizing p. One

can see from equation (20) that an increase in N_2^L will have a negative impact on the efficiency wage in the first period.

In period 1, firms choose the level of employment and training that maximizes ex-ante profits, given the level of the real wage that satisfies the NSC. The first order condition with respect to employment from equation (17) is:

$$gaN_1^{a-1} + (1-h)r(b) = w_1 + \frac{\vartheta w_1}{\vartheta N_1}N_1 + c(b) \tag{20}$$

Solving for employment in the first period from equation (20), gives:

$$N_1 = (\frac{(pw^m + \bar{e})[\varphi + h(1-\varphi)] + p[c(b) - (1-h)r(b)]}{pagH})^{\frac{1}{a-1}} \tag{21}$$

The firm's choice variable for investment in specific training is b. After some rearrangements, the first order condition with respect to b can be written as:

$$
\begin{aligned}
F &= h\frac{\vartheta N_2^L}{\vartheta b}[r(b) + (w^m + \frac{\bar{e}}{p})(1-\varphi)] \\
&- \{c(b)_b N_1 - r(b)_b[hN_2^L + (1-h)N_1]\} = 0
\end{aligned}
\tag{22}
$$

The first part of equation (22) reflects the decrease in labour cost due to the reduction of the dismissal probability for cyclical reasons and the remaining part represents the difference between the marginal cost of training and the expected marginal return of investment in the second period. Solving equations (21) and (22) for N_1 and b, gives the optimal values of employment and investment in firm-specific training.

---

can assume that $p = f(m)$ where m denotes the level of monitoring and $f(m)' > 0$. Let us introduce a monitoring cost that is linear to the level of monitoring cm and solve for the optimal level of m. If the firm monitors the level of output, it would maximize the following profit function with respect to N_2 and m in the second period: $[g(N_2^i)^a + r(b)N_2](1-cm) - w(m)N_2$, where $\frac{\vartheta w}{\vartheta m} < 0$. Taking the two FOCs and combining them gives us the following condition for optimal m: $cgN_2^{a-1}[r(b),m] + cr(b) = \frac{-\vartheta w}{\vartheta m}$. After checking the effect of $r(b)$ on the optimal level of m by taking the total differential we can conclude that its sign and magnitude is ambiguous.

Proposition 1:

Assuming that b_0 is the level of training that equalizes marginal cost with expected marginal return to training, firms will choose a higher level of firm specific training, $b^* > b_0$, due to the positive impact of b on workers' expectations for future employment.

Proof: Assume that b_0 is the level of training which equalizes marginal cost with expected marginal return to investment. If one substitutes b_0 in equation (22), the second term of the equation will equal zero. Given that the remaining part of equation (22) is positive for positive values of b, $b = b_0$ does not satisfy equation (22). $F = 0$ will be satisfied for a higher level of b if $F_b < 0$. The derivative of equation (22) with respect to b is negative since $E(\Pi)_{bb} < 0$.[11]

Finally, we want to compare the level of b derived by equations (21) and (22) with the optimal level of training in the absence of uncertainty, $h = 0$.

Proposition 2:

Assume that $\bar{b}$ is the level of training that equalizes marginal cost with marginal return to training if the probability of shifting to a low state is zero. The optimal level of investment in firm-specific training , b^*, will exceed this level, $b^* > \bar{b}$, if the elasticity of the firing probability with respect to the level of training is significantly high.

Proof:

If $h = 0$, then equation (22) becomes $\bar{F} = r(b)_b - c(b)_b = 0$. Assume that $\bar{b}$ satisfies $\bar{F} = 0$. By substituting $\bar{b}$ in equation (21) and then in (22), we obtain:

$$F(\bar{b}) = w_{1b} - hr(\bar{b})_b[N_1(\bar{b}) - N_2^L(\bar{b})] + hr(\bar{b})\frac{\vartheta N_2^L}{\vartheta b} \tag{23}$$

Given that $F_b < 0$, $b^* > \bar{b}$ if $F(\bar{b}) > 0$.

---

[11] The stability condition $E(\Pi)_{N_1 N_1} < 0$ and $E(\Pi)_{N_1 N_1} E(\Pi)_{bb} - E(\Pi)_{N_1 b} E(\Pi)_{bN_1} > 0$ is satisfied for reasonable parameter values.

2.3 Firm-specific training as a commitment technology

Several shirking models including the original Shapiro and Stiglitz (1984) model show the negative effect of labour turnover on incentive wages. However, the usual assumption in the literature is that this turnover is exogenous to the firm's decision. Endogenizing the expected duration of employment is a natural step since firms can affect labour turnover in various ways. One obvious way is firm-specific training. As we showed above, training increases the probability of remaining employed in a low state, thereby reducing incentive wages. This mechanism has a positive impact on the optimal level of training. Furthermore, it is important to emphasize the credibility of training as a mechanism for reducing labour turnover. It is clear that if workers were not forward looking at the beginning of period 1, the firm would have an incentive to guarantee employment continuation in the low state ($q = 0$). This would imply lower wages from equation (15). However, such a promise would be time-inconsistent since the firm would have an incentive to deviate from that promise in the next period and set employment according to equation (10). Forward looking workers will choose their level of effort by taking into account the firm's optimal policy in the next period. By its decision to invest in b units of firm-specific training, the firm commits to the employment level defined by equation (19) in the low state. This commitment is credible since the firm bears the cost of training in the first period. Alternative ways of reducing labour turnover such as the imposition of firing costs or hiring costs are less credible since they impose a cost that the firm will bear in a future period. This implies that after workers make their effort choice, the firm will always have an incentive to try to reduce this cost. For example, since effort is a nonverifiable variable, there is always the possibility that the firm will falsely claim that workers cheated in order to avoid firing costs.

3 An infinite horizon model of employment and training

Next we wish to extend our model to an infinite horizon framework in order to see whether our main results depend crucially on the two-period framework of the previous analysis. The only

difference with the timing assumptions of the previous section is that wages are now paid at the beginning of each period after the realization of the productivity shock. This implies that the new setup is consistent with the efficiency wage model of Shapiro and Stiglitz.

3.1 Workers' behaviour

We assume that workers behave as in the shirking model of Saint-Paul $(1996)^{12}$. Fella (2000) makes similar assumptions in a continuous time model. Each worker will provide a positive level of effort in period t only if the expected present discounted income of being employed in period t+1, EV, exceeds the one of being unemployed in period t+1, EU, by a markup that equals the current level of effort divided by the monitoring probability p:

$$EV_{t+1} = EU_{t+1} + \frac{\bar{e}}{p} \tag{24}$$

If workers are risk neutral with a discount factor δ, the present discounted value of being employed in the two states can be written as:

$$V_t^H = w_t^H + \delta[(1-h)EV_{t+1}^H + h\frac{N_{t+1}^L}{N_t^H}EV_{t+1}^L + h(1 - \frac{N_{t+1}^L}{N_t^H})EU_{t+1})] \tag{25}$$

and

$$V_t^L = w_t^L + \delta[(1-h)EV_{t+1}^L + hEV_{t+1}^H] \tag{26}$$

Equations (25) and (26) say that the value of being employed in the current state equals the current wage plus the expected present discounted value of the next period. When the firm is in the high state, the expected present discounted value will depend on the probability that the worker will retain his position if the firm shifts to the low state, $\frac{N_{t+1}^L}{N_t^H}$. This probability is equivalent to probability q defined by equation (4). A worker who is employed in the low state will remain employed in the next period. Inserting (24) into (25) and (26) and after some manipulations we can solve for efficiency wages in the two states:

---

[12] We borrow the basic setup regarding workers' behaviour from chapter 7 of Saint-Paul (1996).

$$w_t^H \quad = \quad U_t - \delta EU_{t+1} + \frac{\bar{e}}{p}[1 - \delta + \delta h(1 - \frac{N_{t+1}^L}{N_t^H})] \tag{27}$$

$$w_t^L \quad = \quad U_t - \delta EU_{t+1} + \frac{\bar{e}}{p}(1 - \delta) \tag{28}$$

Assuming zero unemployment benefits, the value of being unemployed can be written recursively as:

$$U_t = \delta[(1 - \varphi)EU_{t+1} + \varphi(EU_{t+1} + \frac{\bar{e}}{p})] \tag{29}$$

where φ is the exit rate from unemployment. Firms take the exit rate from unemployment as given since, under our assumptions about g, in the steady state the exit rate is a function of average employment in each firm.[13]

Inserting (29) in (27) and (28) we get:

$$w_t^H \quad = \quad [1 - \delta(1 - \varphi) + \delta h(1 - \frac{N_{t+1}^L}{N_t^H})]\frac{\bar{e}}{p} \tag{30}$$

$$w_t^L \quad = \quad [1 - \delta(1 - \varphi)]\frac{\bar{e}}{p} \tag{31}$$

As expected, wages in both states will increase in the exit rate of the unemployment and in the required level of effort. A higher monitoring probability will induce workers not to shirk at a lower wage. Future employment in the low state will have a negative effect on efficiency wages as in the two-period model of the previous section.

3.2 Employment and training

Let us re-write the infinite horizon equivalent of the production function, equation (1), assuming non-linear returns to training:

$$Y_{j,t} = g_{j,t}^i b_{j,t-1}^{\beta}(N_{j,t}^i)^{1-\beta} \tag{32}$$

---

[13] This argument is demonstrated in Fella (2000) under the same assumptions about the stochastic environment.

where i takes a high or a low value. Equation (32) implies that output increases with training of the previous period, b_{t-1}.[14] Output depends on a productivity shock, g, that follows a two-state Markov process with symmetric transition probabilities:

$$P(g_{j,t+1} = g^H / g_{j,t} = g^L) = P(g_{j,t+1} = g^L / g_{j,t} = g^H) = h$$

and

$$P(g_{j,t+1} = g^H / g_{j,t} = g^H) = P(g_{j,t+1} = g^L / g_{j,t} = g^L) = 1 - h$$

For simplification we assume that firms never train their workforce in the low state, so that $b_t = 0$ in the low state. After the realization of the current period state, firms set the level of employment and wages and the level of firm specific training in the high state. In the low state, firms set only the level of employment and wages. The expected discounted value of the firm's profit at the beginning of period t in the high and low state are given respectively by:

$$V^H(b_{t-1}, g_t) = \max\{Y(g_t^H, b_{t-1}, N_t^H) - w_t^H(N_t^H, N_{t+1}^L)N_t^H - c(b_t)N_t^H +$$
$$\delta[(1-h)V^H(b_t, g_{t+1}^H) + hV^L(b_t, g_{t+1}^L)]\} \tag{33}$$

and

$$V^L(b_{t-1}, g_t) = \max\{Y(g_t^L, b_{t-1}, N_t^L) - w_t^L N_t^L +$$
$$\delta[(1-h)V^L(0, g_{t+1}^L) + hV^H(0, g_{t+1}^H)]\} \tag{34}$$

From equation (30), we can see that wages in the high state will depend on the probability of being fired if there is a switch to the low state. Therefore, future employment in the low state will affect current wages. In line with Obstfeld (1991), we define employment in a recession as a

---

[14] For simplification, we will assume that the training cost does not include payment to workers.

function of the state variables $N_t^L = f(b_{t-1}, g_t)$ which implies that future employment in the low state in (33) can be defined as a function of current training[15] :

$$N_{t+1}^L = f(b_t, g_{t+1}) \tag{35}$$

The first-order conditions for profit maximization in the high state can be written as:

$$(1 - \beta)g^H b_{t-1}^\beta (N_t^H)^{-\beta} = [1 - \delta(1 - \varphi) + \delta h]\frac{\bar{e}}{p} + c(b_t) \tag{36}$$

$$
\begin{aligned}
c(b_t)_b N_t^H &= \delta[(1 - h)\frac{\vartheta V^H(b_t, g_{t+1}^H)}{\vartheta b_t} + h\frac{\vartheta V^L(b_t, g_{t+1}^L)}{\vartheta b_t}] \\
&\quad - \frac{\vartheta w_t^H}{\vartheta N_{t+1}^L}\frac{\vartheta N_{t+1}^L}{\vartheta b_t}N_t^H
\end{aligned} \tag{37}
$$

$$\frac{\vartheta V^H(b_{t-1}, g_t^H)}{\vartheta b_{t-1}} = \beta g^H b_{t-1}^{\beta-1}(N_t^H)^{1-\beta} \tag{38}$$

Similarly, the first-order conditions in the low state are:

$$(1 - \beta)g^L b_{t-1}^\beta (N_t^L)^{-\beta} = [1 - \delta(1 - \varphi)]\frac{\bar{e}}{p} \tag{39}$$

$$\frac{\vartheta V^L(b_{t-1}, g_t^L)}{\vartheta b_{t-1}} = \beta g^L b_{t-1}^{\beta-1}(N_t^L)^{1-\beta} \tag{40}$$

By using (38) and (40) and the envelope theorem, we can rewrite equation (37) as:

$$
\begin{aligned}
&\delta[(1 - h)\beta g^H b_t^{\beta-1}(N_{t+1}^H)^{1-\beta} + h\beta g^L b_t^{\beta-1}(N_{t+1}^L)^{1-\beta}] + \frac{\bar{e}}{p}h\frac{\vartheta N_{t+1}^L}{\vartheta b_t} \\
&= c(b_t)_b N_t^H
\end{aligned} \tag{41}
$$

Employment in the two states is derived by solving (36) and (39) for N_t^H and N_t^L.

---

[15] This assumption is consistent with subgame perfection.

$$N_t^H = [\frac{(1-\beta)g^H}{[1-\delta(1-\varphi)+\delta h]\frac{\bar{e}}{p}+c(b_t)}]^{\frac{1}{\beta}}b_{t-1} \tag{42}$$

$$N_t^L = [\frac{(1-\beta)g^L}{[1-\delta(1-\varphi)]\frac{\bar{e}}{p}}]^{\frac{1}{\beta}}b_{t-1} \tag{43}$$

Equation (43) suggests that employment in the low state increases with training in the previous period and allows us to get a functional form for future employment in the low state defined by equation (35).

Proposition 3: *The marginal value of training is higher than its expected marginal return.*

Proof: The marginal value of training is given by the LHS of equation (41). Our proposition holds since as one can see from equations (35) and (43) $\frac{\vartheta N_{t+1}^L}{\vartheta b_t} > 0$.

Proposition 3 implies that an additional unit of training decreases the compensation workers require in order to provide a positive level of effort by improving their job security.

Finally, we substitute (42) and (43) in (41) which gives us the FOC for the optimal level of training in the current period:

$$c(b_t)_b X_1^{\frac{1}{\beta}} = [\beta\delta(1-h)g^H X_2^{\frac{1-\beta}{\beta}} + \beta\delta h g^L X_3^{\frac{\beta-1}{\beta}} + \frac{h\bar{e}X_3^{\frac{1}{\beta}}}{p}]b_{t-1}^{-1} \tag{44}$$

where $X_1(b_t) = \frac{(1-\beta)g^H}{[1-\delta(1-\varphi)+\delta h]\frac{\bar{e}}{p}+c(b_t)} > 0$, $X_2(b_{t+1}) = \frac{(1-\beta)g^H}{[1-\delta(1-\varphi)+\delta h]\frac{\bar{e}}{p}+c(b_{t+1})} > 0$ and $X_3 = \frac{(1-\beta)g^L}{[1-\delta(1-\varphi)]\frac{\bar{e}}{p}} > 0$.

Equation (44) is a second order non-linear difference equation. Since one can get solutions for the key economic variables from equations (42), (43) and (44), we do not specify the properties of the value function in the Bellman equations.[16]   In order to check for stability, we linearize (44) by taking a Taylor expansion around the steady state level of b.[17] Next, we define the

---

[16] Specifically, one should substitute the solution for the control variables back into the Bellman equations and specify the properties of the assumed value function so as to equate the RHS and the LHS.

[17] We assume that $c(b) = b^\gamma$, where $\gamma > 1$.The derived steady state level of training b_0 is $b_0 = [\frac{\beta\delta(1-\delta)\bar{e}}{p[\gamma(1-\beta)-\beta\delta]}]^{\frac{1}{\gamma}}$.

characteristic equation and solve for the two roots. The expression for the two roots appears to be very complicated and hence we can check for stability only by numerical simulation. We can conclude that for a wide range of reasonable parameter values we get one stable and one unstable root. Given that we have a backward looking solution, the condition for saddle path stability is satisfied for these values.[18]

4   Conclusions

The model presented above examines the decision to invest in firm-specific training in a two-state incentive wage model. It investigates a neglected positive aspect of job security provisions: the wage reduction associated with the lower unemployment risk. Several shirking models, including the original Shapiro and Stiglitz (1984) model, show the negative effect of the job separation rate on incentive wages. However, the usual assumption in the literature is that this rate is exogenous to the firm's decision. However, firms may affect labour turnover in various ways. We showed that firms can improve workers' expected job tenure by investing in firm specific training since firing a trained worker implies the loss of some rent. Forward looking workers expect that the probability of being fired in a recession will be lower the higher is the level of firm-specific training undertaken by the firm. This implies that training is a credible commitment to lower firing in a low state. In an efficiency wage framework, a longer expected job tenure implies a higher penalty for shirking. As a result, firms' investment in firm-specific training will have a negative effect on the efficiency wage. The ability of firm-specific training to provide a credible commitment to a lower job separation rate thereby reducing wage cost, will increase human capital investment above the level that equates marginal cost to expected marginal return.

---

[18] Results available uppon request.

References

Abowd, J. and O. Ashenfelter, 1981, Anticipated Unemployment, Temporary Layoffs, and Compensating Wage Differentials, in: S. Rosen, eds., Studies in Labor Markets,(University of Chicago Press, Chicago).

Adams, J.D., 1985, Permanent Differences in Unemployment and Permanent Wage Differentials, Quarterly Journal of Economics 100, 29-56.

Becker, G.S., 1962, Investment in Human Capital: A Theoretical Analysis, Journal of Political Economy 70, 9-49.

Becker, G.S., 1964, Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, (NBER, New York).

Booth, A.L. and M. Chatterji, 1989, Redundancy Payments and Firm-specific Training, Economica 56, 505-521.

Carmichael, H.L., 1989, Self-Enforcing Contracts, Shirking and Life Cycle Incentives, Journal of Economic Perspectives 3, 65-83

Fella, G., 2000, Efficiency Wage and Efficient Redundancy Pay, European Economic Review 44, 1473-1490.

Hall, R. E., 1970, Why is the Unemployment Rate so High at Full Employment, Brookings Papers on Economic Activity 1, 369-402.

Hall, R. E., 1972, Turnover in the Labour Force, Brookings Papers on Economic Activity 3, 709-765.

Hart, R.M. and T. Moutos, 1995, Human Capital, Employment and Bargaining, (Cambridge University Press, Cambridge).

Hashimoto, M., 1979, Bonus payments, on-the-job training and lifetime employment in Japan, Journal of Political Economy 87, 1086-1104.

Hashimoto, M., 1981, Firm-specific human capital as a shared investment, American Economic Review 71, 475-482.

Katsimi, M., 1995, Firing costs, job tenure and employment, Discussion Paper No. 451, University of Essex.

Li, H.L., 1986, Compensating differentials for cyclical and non-cyclical unemployment: The interaction between investors' and employees' risk aversion, Journal of Labor Economics 4, 277-300.

Obstfeld, M. 1991, A model of currency depreciation and the debt-inflation spiral, Journal of Economic Dynamics and Control 15, 151-177.

Oi, W., 1962, Labor as a quasi-fixed factor, Journal of Political Economy 70, 538-555.

Saint-Paul, G., 1996, Dual labor markets: A macroeconomic perspective, (MIT Press, Cambridge Massachusetts).

Shapiro, C. and J. Stiglitz, 1984, Equilibrium unemployment as a worker discipline device, American Economic Review 74, 433-444.

CESifo Working Paper Series

(for full list see www.cesifo.de)

_____