89-F-9

OPTIMAL LABOR CONTRACTS WITH NON-CONTRACTIBLE HUMAN CAPITAL

by

Yoshitsugu Kanemoto
University of Tokyo

and

W. Bentley Macleod
Queen's University

June 1989

OPTIMAL LABOR CONTRACTS WITH NON-CONTRACTIBLE HUMAN CAPITAL*

Yoshitsugu Kanemoto

Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

AND

W. Bentley MacLeod

Department of Economics, Queen's University, Kingston, Ontario, K7L 3N6, Canada; and Institut D'Analisi Economica, CSIC, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

*We would like to thank Kikawada Foundation and Social Science and Humanities Research Council in Canada for financial support. The second author is also grateful for the financial support provided by the Dirección General de Investigación Científica y Técnica of Spain during a visit to the Universitat Autònoma de Barcelona. We benefited from valuable comments from Hajime Miyazaki, Alistair Ulph, Masahiro Okuno, Takatoshi Ito, and other participants at the TCER-NBER-CEPR Conference on "Labor Relations and the Firm: Comparative Perspectives," Tokyo, January 7-8, 1989.

Abstract

The market for human capital is incomplete and therefore specific institutions must be designed to deal with this problem. An example of such an institution is lifetime employment (LTE), as observed in Japan. LTE coupled with an agreed-upon total wage bill will solve the moral hazard problem on the firm's side. On the worker's side, a tournament provides incentives for human capital investment. We show that the optimal contract will involve promotion of essentially all workers. Surprisingly, LTE can be more flexible than the policy of attaching wages to jobs that is often considered the "paradigm" American employment practice.

Introduction

The starting point for the current paper is the idea that the market for human capital is incomplete and therefore specific institutions must be designed to deal with this problem. An example of such an institution is the policy of attaching wages to jobs (WATJ) that is often considered the "paradigm" American employment practice.[1] We will show that the institution of lifetime employment (LTE), as observed in Japan, is another mechanism that will bring about efficient human capital investment. A possibly more surprising consequence of this analysis is that LTE is in a certain sense more flexible than WATJ.

As both Grout (1984) and Crawford (1988) have shown, in the absence of complete markets long term contracts are necessary to induce efficient relation-specific investment. However, firm-specific investment is often much too complicated and subtle to be contractible. Even if both the firm and the worker know whether or not the worker's investment is adequate, this does not necessarily mean that a third party can verify that. As Williamson (1985) has stressed, the lack of third party verifiability makes long term contracts very difficult to enforce[2].

Simply put, the problem is one of double sided moral hazard. In a competitive labor market, workers must be explicitly compensated for any investment into firm specific human capital. If a worker is paid up front, however, he/she will have an incentive to collect the payment, and not make the required investment. Given that it is impossible for third parties to observe and measure firm specific human capital, the worker cannot be punished for a breach of contract. On the other hand, if the worker makes the investment before payment, one has what Williamson describes as a hostage problem and the firm will have an incentive to renege on the original contract. Since the investment has no value on the market, anticipating this behavior, the worker will underinvest in human capital.

One of the important roles for the creation of a firm is to set up an institutional structure that will protect the worker against the hostage problem, and ensure that there is efficient human capital investment. A number of recent papers, notably Grossman and Hart (1986), and a survey by Holmström and Tirole (1989), have stressed the importance of incomplete contracting in understanding the nature of the firm. The firm will differ from spot market transactions in that it will be able to enforce efficiency enhancing contracts that are not feasible spot market equilibria.

In practice the firm carries out a large number of activities that supplement market transactions. In this paper we will concentrate our attention on only one aspect of the firm that is often associated with the paradigm Japanese firm, namely lifetime employment (LTE). Of course life time employment is not a characteristic of all Japanese workers, nor is it limited in practice to Japan alone. However since it is a characteristic that is often associated with Japan, we will suppose that LTE is the defining characteristic of Japanese firms. LTE is not a privilege that all workers enjoy, even in Japan. Depending on the state of demand, firms will also take on temporary workers. These may be married females, students, or farmers during the off season.

What is characteristic about the LTE institution is that at the time of employment it is usually clear whether one is a lifetime employee or a temporary employee. Workers given lifetime employment are expected to make significantly more contribution to the firm, and in particular are induced to invest in more firm-specific human capital. In this paper we solve for the optimal employment contract given the LTE institution. Our main contribution will be to show that in equilibrium the optimal contract involves treating workers that are ex ante the same, differently ex post. It is hoped that this result will provide some insight into the empirical observation that in Japan only a certain fraction of the work force actually enjoy the privilege of lifetime employment.

In fact a substantial number of Japanese workers face the standard employment-at-will doctrine. That is, the firm hires these workers only when demand conditions make it necessary. In addition to having less employment security, these workers on average receive lower wages. In our model this phenomenon is explained by supposing that the market for human capital is imperfect. We show that if the firm can commit itself to lifetime employment, and to an aggregate wage bill, then it will be possible for the firm to induce efficient human capital investment on the part of workers. If demand is stochastic, however, the firm will wish to adjust employment levels ex post. We assume that a worker's human capital investment will pay off only in the second period. Optimal capital accumulation then implies that in a period of high demand only a fraction of the workers are required to invest in human capital. These will be the permanent workers. Those workers not required to make the investment will be given a short term contract.

We assume that the labor market is competitive, and therefore ex ante all workers normally receive the same lifetime utility.[3] Life time employees, however, receive higher lifetime incomes to compensate them for the effort they exert in human capital accumulation. We assume that the default utilities are constant over time, and hence, in contrast to the implicit contract theory, there is no insurance role for the firm.

The reason that the institution of lifetime employment is able to simultaneously solve the hostage problem and provide incentives for human capital investment is based on the insight of Malcomson (1984).[4] Malcomson (1984) shows that the institution of WATJ can be viewed as a tournament in which workers compete for a limited number of good high paying jobs. The advantage of a tournament is that the hostage problem can be solved by the firm committing itself in advance to an aggregate wage bill for the permanent workers.

In contrast to a contract which specifies human capital investment of an individual worker, commitment on an aggregate wage bill is easy to enforce. A third party can easily verify whether or not the commitment is fulfilled. Even if a third party does not enforce the commitment, self-enforcement of the sort discussed in Bull (1987) is possible. This is particularly true in Japan where an enterprise union acts as a bargaining agent for workers and can enforce any wage package ex post. Note that concerning human capital investment of an individual worker, a union is one of the third parties which may have difficulty in verifying it.

A tournament requires that there be some form of employment guarantee. Otherwise, given that the labor market is assumed to be competitive, wages would be bid down to the competitive level ex post.

The existence of lifetime employment and an agreed-upon aggregate wage bill will solve the moral hazard problem on the firm's side. On the workers' side, efficient human capital investment is achieved by use of a tournament. This is done by dividing jobs into good and bad types, the good jobs having higher pay than the bad jobs. By adjusting the relative wages, and keeping the total wage bill fixed, competition for the good jobs will ensure that workers invest in the efficient level of human capital accumulation.

We show that the optimal contract will involve promotion of essentially all workers, with an insignificant number not receiving an wage increase. This is not only consistent with the evidence that workers in Japan faces a rising real wage profile, it also points to an advantage of the LTE institution over the WATJ system. In Malcomson (1984), the way that one can implement a tournament in a firm is to attach wages to jobs. For the wage policy to work as an effective incentive system, high paying jobs cannot disappear over the cycle. This implies that the number of promotions will be relatively rigid over time.

The LTE system is more flexible in this respect. Though the firm is committed to LTE at the time of hiring, it can adjust its hiring from period to period. In general, the optimal employment policy involves altering the number of permanent employees from period to period. This will imply that the US system of attaching wages to jobs will be relatively less flexible and therefore less efficient.[5]

After laying out the basic framework in section 1, the optimal contract is analyzed in section 2. Given the form of contract derived in section 2, the firm's hiring decision is analyzed in section 3. Section 4 contains our concluding comments.

1. The Model

Labor inputs in most modern corporations are complex amalgams of contractible, non-contractible, and imperfectly contractible elements. In order to model these elements in a simplest possible fashion, we assume that labor inputs can be divided into two types, contractible and non-contractible.[6]

The former type is observable by an outside party such as the court as well as the firm and the worker, and a contract which specifies these inputs can be enforced by a third party. Examples of this type would be work hours and the number of products (or intermediate products) produced. In Japan where legal enforcement is much less common than in the U.S., an outside party here should be interpreted as including the labor union. If the labor union can monitor labor inputs of an individual worker, it can prevent the firm from cheating by the threat of collective actions such as a strike.

Non-contractible inputs are not observable by an outside party (including the union) although both the firm and the worker can observe them (though possibly with a lag). An example would be efforts to learn correct skills to maintain the quality of the product. Because of the lack of third party verifiability, any reward for them must be self-enforcing.

We assume that non-contractible inputs have long-term effects although contractible ones do not. A typical example of such an input is a worker's effort to develop human capital. Another example is a worker's participation in improvements in production technology which is fairly common in Japanese firms. As will be seen later, without the long-term nature of non-contractible inputs there is no need for a *long-term* contract.

The profit of a firm in each period is $F(N,K,\theta)$, where N is the total amount of efficiency units of contractible labor provided by all the workers, K is the total amount of non-contractible human capital (or, the stock of know-hows that workers accumulated in the previous period), and finally θ is a random variable. The actual determination of N and K will be discussed below. The productivity of the firm is private information to the firm only. We assume that $F(N,K,\theta)$ is differentiable, concave, and strictly increasing in N and K. The firm maximizes the discounted profits, Π =

$E\{\sum_t F(N_t, K_t, \theta_t)\rho^t\}$, where ρ is the discount factor.

Our model is an overlapping generations model in which workers live two periods. The lifetime utility of a worker is given by $E\{u(c^1, \gamma^1) + \delta u(c^2, \gamma^2)\}$, where c^t is the consumption of the worker at age t, γ^t is the level of non-contractible effort that the worker of age t provides to develop human capital. We assume that γ^t is either 0 or 1.

The firm cannot directly observe γ, but it will know if a worker has acquired the right skills in the second period by observing his/her performance. Although both the firm and the worker know γ in the second producing γ must be self-enforcing. Workers in addition to supplying γ also supply effort that is contractible, e. We assume that e is exogenous and normalize it to one. For example, e is the work hours which are fixed legally.

The following assumptions are imposed on the preference structure of workers. First, $u(c, \gamma)$ is strictly increasing and strictly concave in c: $\frac{\partial u}{\partial c} > 0$, $\frac{\partial^2 u}{\partial c^2} < 0$. Second, non-contractible efforts γ are utility reducing: $u(c,0) > u(c,1)$. Third, the discount rate for the firm is lower than or equal to that of a worker: $\rho \geq \delta$. Because of this assumption, the worker will consume more in the first period than in the second period if no incentive problem exists.

The firm has a choice between two types of employment contracts. The first is a contract with employment at will: the firm can layoff workers as desired. The second is a two-period contract where the firm commits itself to keeping workers for two periods. However, the firm's commitment on wage payments is limited. The firm would want to make wages contingent on human capital investment, but commitment to such a contract is impossible because of the lack of third party verifiability. The firm can however commit itself to the aggregate wage bill, which opens up a possibility of using a tournament as an incentive scheme. Although the firm cannot layoff a worker with a two-period contract, workers are free to leave the firm after one period if they wish. Workers who have a two-period employment contract are called permanent workers, and those with a single-period contract are called temporary workers.

A worker has an access to the competitive spot market where he or she can receive the wage v. In the competitive market it is assumed that the worker will not produce any non-contractible effort γ. The amount of contractible effort in the competitive sector is also fixed at one. The utility of a worker in the competitive sector is then $u(v,0)$.

2. Labor Contracts

In this section, we first solve for cost minimizing wage contracts for both temporary and permanent workers. This yields the unit costs of temporary and permanent workers. We then characterize the profit maximizing employment policy in the next section. Throughout we will assume that the labor market is perfectly competitive, and that all workers will receive the same wage v on this market, regardless of the level of firm specific human capital investment.

Let us first examine the wages for workers that face an employment-at-will contract. At the beginning of each period these workers will be offered a take it or leave it contract. This will imply that in the last period their wage will be v, independent of the investment made in period one. Therefore in period one they will always choose an investment level equal to zero.[7] Hence, for there to be investment in human capital the worker must be guaranteed some share of the surplus in the following period. This will be possible with the institution of LTE.

Contracts for permanent workers are more complicated. In this case the firm offers a contract that will guarantee employment for two periods, combined with a well specified promotion rule. This commitment can be enforced through a legally binding contract that prohibits the firm from laying off a worker, because it is easy to verify whether or not a layoff has occurred. Legal enforcement is not, however, necessary. In Japan where enterprise unions are common, a self-enforcement mechanism as in Bull (1987) will work. If the firm lays off a worker, the union can punish the firm by going on a strike. Faced with this threat, the firm will not violate the contract.

Notice that a self-enforcement mechanism of this sort does not work for a contract directly specifying human capital investment. An individual worker cannot punish the firm for cheating because the firm can simply fire the worker and hire another worker. An enterprise union cannot punish the firm either because it is one of the third parties which are unable to verify the investment.

The promotion scheme is as follows. In period 1, the firm offers a wage w^1 which is the same for all permanent workers. In period 2, the firm offers a fraction λ of slots that pay $\bar{w}$ and $1-\lambda$ slots paying $\underline{w}$, where $\bar{w} \geq \underline{w}$. The second-period wages are paid after observing the results of training. In this promotion scheme, the firm makes commitment on the probability of promotion λ and wage levels of promoted and not promoted workers. Notice that this problem is different from a standard principal-agent model in that wages themselves cannot be conditional on investment levels. This is because these are not observable by a third party.

Formally all the firm can do is to choose which workers can be promoted and which cannot. Since the aggregate wage bill is fixed, this decision will not affect profits ex post. Therefore the firm will have an incentive to choose the promotion scheme that will result in workers making the efficient investment into human capital. The cost minimizing promotion scheme is the one that minimizes the unit labor costs, subject to all workers getting their default utility, and having the incentive to invest in human capital.

The program is as follows.

$$\min_{\{w^1,\bar{w},\underline{w},\lambda\in[0,1]\}} \{w^1 + \rho[\lambda\bar{w} + (1-\lambda)\underline{w}]\}$$

subject to

(IC) $$u(w^1,1) + \delta\{\lambda u(\bar{w},0)+(1-\lambda)u(\underline{w},0)\} \geq u(w^1,0) + \delta u(\underline{w},0),$$

(IR1) $$u(w^1,1) + \delta\{\lambda u(\bar{w},0)+(1-\lambda)u(\underline{w},0)\} \geq u(v,0) + \delta u(v,0),$$

(IR2) $$u(\underline{w},0) \geq u(v,0).$$

Condition (IC) says that the worker is better off investing in human capital (γ=1) and having a probability λ of being promoted to $\bar{w}$ than slacking off (γ=0) and having zero probability of promotion. Note that this condition assumes that all other permanent workers are making the investment. Condition (IR1) is the usual individual rationality constraint that the life-time utility of a permanent worker is higher than or equal to that of a temporary worker. Condition (IR2) is an individual rationality constraint in the second period: a worker who is not promoted is as well off as leaving the firm and joining the temporary pool.

The next proposition shows that λ=1 is a cost minimizing solution. This is a consequence of risk aversion. Because a worker is risk averse, the firm can reduce the total wage payment by increasing the probability of promotion, λ, and reducing the wage for the promoted job, $\bar{w}$, while keeping the expected utility constant.

Proposition 1. *A cost minimizing solution of λ is $\lambda = 1$. This is a unique solution when $\bar{w} > \underline{w}$.*

Proof:

The Lagrangian for the cost minimization problem is

$$\begin{aligned}\mathcal{L} = w^1 &+ \rho\{\lambda\bar{w} + (1-\lambda)\underline{w}\} \\ &- \alpha_1\{u(w^1,1) + \delta[\lambda u(\bar{w},0)+(1-\lambda)u(\underline{w},0)] - [u(w^1,0) + \delta u(\underline{w},0)]\} \\ &- \alpha_2\{u(w^1,1) + \delta[\lambda u(\bar{w},0)+(1-\lambda)u(\underline{w},0)] - (1+\delta)u(v,0)\} \\ &- \alpha_3\{u(\underline{w},0) - u(v,0)\}.\end{aligned}$$

The first order conditions for w^1, $\bar{w}$, and $\underline{w}$ are

$$1 - \alpha_1[u_c(w^1,1)-u_c(w^1,0)] - \alpha_2 u_c(w^1,1) = 0 \tag{2.1}$$

$$\rho - \delta(\alpha_1+\alpha_2)u_c(\bar{w},0) = 0 \tag{2.2}$$

$$(1-\lambda)[\rho - \delta(\alpha_1+\alpha_2)u_c(\underline{w},0)] + \alpha_1\delta u_c(\underline{w},0) - \alpha_3 u_c(\underline{w},0) = 0, \tag{2.3}$$

where the multipliers α_1, α_2, and α_3 are all nonnegative. The derivative of the Lagrangian with respect to λ is

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\lambda} &= \rho(\bar{w}-\underline{w}) - \delta(\alpha_1+\alpha_2)[u(\bar{w},0)-u(\underline{w},0)] \\ &= \delta(\alpha_1+\alpha_2)\{(\bar{w}-\underline{w})u_c(\bar{w},0) - [u(\bar{w},0)-u(\underline{w},0)]\} \quad \text{(from (2.2))} \\ &\leq 0 \quad \text{(from the concavity assumption).}\end{aligned}$$

Hence, a cost minimizing solution of λ is $\lambda=1$. This is a unique solution when $\bar{w}$ is strictly larger than $\underline{w}$. ***Q.E.D.***

Thus, we have a seniority system with a zero probability of not being promoted. Note however that, if a worker knows that everyone gets promoted, he or she does not have an incentive to provide effort. The solution must therefore be interpreted as a contract with a very small probability of not being promoted. For example, if the number of permanent workers is n, then n-1 workers will be promoted and one worker will not be promoted. Given that all other workers are making human capital investment, the probability of getting promoted is $\frac{n-1}{n}$ if a worker does the same, and the probability is zero (0) if he/she does not. The worker then has an incentive to invest in human capital even though the investment is not contractible.

The reason why the probability of getting promoted is (almost) one is very simple. Since workers are risk averse, they will always prefer a sure outcome to a lottery. Therefore all workers will receive the same wage in equilibrium in the second period, with the exception of a negligible number of workers who are not promoted.

With $\lambda=1$, constraints (IC) and (IR1) can be simplified as

(IC') $$u(w^1,1) + \delta u(\bar{w},0) \geq u(w^1,0) + \delta u(\underline{w},0),$$

(IR1') $$u(w^1,1) + \delta u(\bar{w},0) \geq u(v,0) + \delta u(v,0).$$

At the optimum, $w^1+\rho\bar{w}$ is minimized under constraints (IC'), (IR1'), and (IR2). Since $\underline{w}$ does not appear in the objective function, by setting $\underline{w} = v$ we can reduce constraints (IC') and (IR2) to

(IC") $$u(w^1,1) + \delta u(\bar{w},0) \geq u(w^1,0) + \delta u(v,0).$$

Thus, our problem is one of choosing w^1 and $\bar{w}$ in such a way to minimize $w^1+\rho\bar{w}$ subject to (IC") and (IR1').

Now, at least one of the two constraints (IC") and (IR') must be binding. Otherwise, one can reduce $\bar{w}$ without violating any of the constraints. Three cases are possible: (i) only constraint (IR') is binding, (ii) only constraint (IC") is binding, and (iii) both constraints are binding. Case (i) is obtained if the marginal utility of consumption is much higher at $\gamma=0$ than at $\gamma=1$, i.e., a worker enjoys consumption much more when he or she is not working hard. In such a case the second period wage $\bar{w}$ will be high enough to automatically satisfy the incentive constraint. In the opposite extreme where preference for the first period consumption is very strong, we obtain case (ii) where the incentive constraint cannot be satisfied if the life-time utility of a worker equals the market level. Case (iii) is obtained in an intermediate case.

The following proposition characterizes the cost minimizing contract for each of these cases.

Proposition 2. *The cost minimizing contract satisfies either one of the following three sets of conditions.*

$$\text{(i)} \quad \begin{cases} \dfrac{u_c(\bar{w},0)}{u_c(w^1,1)} = \dfrac{\rho}{\delta}, \\ u(w^1,1) + \delta u(\bar{w},0) = (1+\delta)u(v,0), \\ u(\bar{w},0) \geq u(v,0) + \frac{1}{\delta}[u(w^1,0) - u(w^1,1)]. \end{cases}$$

$$\text{(ii)} \quad \begin{cases} \dfrac{u_c(\bar{w},0)}{u_c(w^1,1)-u_c(w^1,0)} = \dfrac{\rho}{\delta}, \\ u(\bar{w},0) = u(v,0) + \frac{1}{\delta}[u(w^1,0) - u(w^1,1)], \\ u(w^1,1) + \delta u(\bar{w},0) \geq (1+\delta)u(v,0). \end{cases}$$

$$\text{(iii)} \quad \begin{cases} w^1 = \underline{w} = v \\ u(v,1) + \delta u(\bar{w},0) = (1+\delta)u(v,0) \\ \dfrac{u_c(\bar{w},0)}{u_c(v,1)} > \dfrac{\rho}{\delta} > \dfrac{u_c(\bar{w},0)}{u_c(v,1)-u_c(v,0)}. \end{cases}$$

Case (i) will not be obtained if $u_c(c,1) \geq u_c(c,0)$, and case (ii) will not be obtained when $u_c(c,1) \leq u_c(c,0)$.

Proof:

We have seen that the following three cases are possible: (i) only constraint (IR1') is binding, (ii) only constraint (IC") is binding, and (iii) both constraints are binding. Let us examine the three cases separately.

(i) If the only binding constraint is (IR1'), then the solution satisfies

$$\frac{u_c(\bar{w},0)}{u_c(w^1,1)} = \frac{\rho}{\delta} \geq 1,$$

$$u(\bar{w},0) \geq u(v,0) + \frac{1}{\delta}[u(w^1,0) - u(w^1,1)],$$

$$u(w^1,1) + \delta u(\bar{w},0) = (1+\delta)u(v,0).$$

From (IC"), we get $\bar{w} > v$, and the last two equations yield $w^1 \leq v$. Hence, we must have $w^1 \leq v < \bar{w}$ in this case.

We next show that this case will not be obtained if $u_c(c,1) \geq u_c(c,0)$. If this inequality holds, we have

$$\frac{u_c(\bar{w},0)}{u_c(w^1,0)} \geq \frac{u_c(\bar{w},0)}{u_c(w^1,1)} = \frac{\rho}{\delta} \geq 1,$$

which implies that $w^1 \geq \bar{w}$. Combining this with inequality $\bar{w} > v$ obtained from (IC'), we get $w^1 > v$. This contradicts the requirement that $w^1 \leq v$.

(ii) In this case, the optimal solution without the individual rationality constraint automatically satisfies the constraint. That is,

$$\frac{u_c(\bar{w},0)}{u_c(w^1,1)-u_c(w^1,0)} = \frac{\rho}{\delta},$$

$$u(\bar{w},0) = u(v,0) + \frac{1}{\delta}[u(w^1,0) - u(w^1,1)],$$

$$u(w^1,1) + \delta u(\bar{w},0) \geq (1+\delta)u(v,0).$$

From the last two conditions we have $w^1 \geq v$. The second equation implies $\bar{w}$

$> v$. If $u_c(c,1) \leq u_c(c,0)$, then this case will never occur, since the first equation would not hold.

(iii) When both constraints are binding, we have $w^1 = \underline{w} = v$, and $\bar{w}$ is uniquely determined by (IR') with equality. Then, we have $\bar{w} > v = w^1 = \underline{w}$. For the incentive constraint to be binding, we must have

$$\frac{u_c(\bar{w},0)}{u_c(v,1)} > \frac{\rho}{\delta} .$$

For the individual rationality constraint to be binding, inequality

$$\frac{u_c(\bar{w},0)}{u_c(v,1)-u_c(v,0)} < \frac{\rho}{\delta}$$

must be satisfied. *Q.E.D.*

It should be noted that a two-period contract is not necessary if non-contractible efforts are observed in the current period.[8] In such a case a promotion scheme can be implemented in a single period. That is, the firm can simply offer a bonus payment to a fixed fraction of workers at the end of the period. In our model, a two-period contract is necessary because the firm observes human capital investment only in the second period.

In cases (i) and (iii), the life-time utility of a permanent worker equals that of a temporary worker. A permanent worker has a higher life-time income, however, because he/she must be compensated for human capital investment. In case (ii), the life-time utility of a permanent worker is strictly higher than that of a temporary worker. When the discount factor is small, the incentive constraint cannot be satisfied unless the life-time utility exceeds the default level. This case requires rationing of permanent jobs

The first-best outcome is obtained in case (i) because the incentive constraint is not binding. In cases (ii) and (iii), however, the incentive constraint creates distortion in intertemporal wage profile, i.e., the wage profile is steeper than the first best.

In a special case where $u(c,\gamma)$ is additively separable, only case (iii) will be obtained because $u_c(c,1) = u_c(c,0)$. This yields the following corollary.

Corollary. *If $u(c,\gamma)$ is additively separable in c and γ, then the unique optimal solution is case (iii) in Proposition 2.*

Denote the minimized labor cost by $C(v) = w^1 + \rho\bar{w}$. It is not difficult to see that the minimized labor cost is higher than the labor cost of temporary workers, $(1+\rho)v$. We denote the labor cost differential by $d = C(v)-(1+\rho)v > 0$. The differential d is the minimum cost of inducing a worker to invest in human capital. Notice that, regardless of how the worker is used in the following period, if the firm wishes to have the worker make the human capital investment, then the cost C(v) is sunk as soon as the permanent worker is hired. This is very useful since it allows us to study the optimal contract separately from the issue of how many workers the firm wishes to make the human capital investment. If a firm plans to keep a worker for two periods, then the term d is exactly the marginal cost of human capital to the firm.

3. The Optimal Employment Policy

Let us now turn to the optimal employment policy of the firm. In each period the firm observes current productivity θ and decides how many new workers to hire, temporary and permanent. We avoid integer problems assuming that the number of workers is a continuous variable. Let M, n, and m denote the number of permanent workers from the previous period, the number of temporary workers, and the number of newly hired permanent workers, respectively.

We assume that non-contractible training of a permanent worker is embodied in the worker. This case can be contrasted with a case where trained workers improve upon production technology so that their contributions will remain in the firm even after they leave the firm. In the embodied training case, we can assume that the total amount of non-contractible human capital equals the number of permanent workers hired in the last period, i.e., $K_t = m_{t-1}$. Then, current profits gross of labor costs are $F(N,K,\theta) = F(M+n+m,M,\theta)$, which is concave in M, n, and m, since $F(N,K,\theta)$ is concave in N and K.[9]

The expected discounted profit of the firm is

$$E\{\sum_{t=1}^{\infty} \rho^{t-1}[F(m_{t-1}+n_t+m_t,m_{t-1},\theta_t)-vn_t-C(v)m_t]\}. \tag{3.1}$$

We assume that θ_t is i.i.d. That is, θ_t's are independent and have the same distribution $G(\theta_t)$. Then m_t and n_t are determined based only on θ_t and $M_t = m_{t-1}$. We can then take $s_t = (M_t,\theta_t)$ as state variables. The firm will choose $n(s_t)$ and $m(s_t)$ in each period.

Notice that since permanent employees are kept for life, their employment expense is a sunk cost for the firm that is formally equivalent to paying the value of their lifetime income in the first period. This is given by the term $C(v)m_t$. This is exactly what distinguishes lifetime employment from temporary employment. This does not however imply that ex post the firm cannot allocate the workers to jobs that do not use their human capital investments. All that LTE means is allocating workers to an unskilled job cannot involve a loss of utility. In Japan the firms will often lend excess workers out. Alternatively they can lay these workers off if they are given a sufficient severance pay. Both of these cases are implicitly included in this model as follows. The firm has already paid for the workers up front, therefore the only choice for the firm ex post is to compare employment on the job or in the alternative. Let $G(N,K,\theta)$ be the production function for work in the plant. Then the function F will include the possibility of hiring permanent workers out by defining it as follows:

$$F(N,K,\theta) = \max_{n\geq 0} \{G(N-n,K-n,\theta) + nv\}.$$

If G is concave, then so will be F. With this definition we are explicitly allowing the firm to optimally allocate permanent workers in the plant or at some alternative use.

The optimal policies will solve the following dynamic programming problem.

$$V(M_t,\theta_t) = \sup_{\{m_t,n_t\}} \{F(M_t+n_t+m_t,M_t,\theta_t)-vn_t-C(v)m_t + \rho\int V(m_t,\theta_{t+1})dG(\theta_{t+1})\}, \tag{3.2}$$

subject to the nonnegativity constraint,

$$n_t \geq 0, \tag{3.3}$$

$$m_t \geq 0. \tag{3.4}$$

and the initial condition,

$$m_0 = M_0, \tag{3.5}$$

where M_0 is a constant.

The following lemma shows that the value function is unique and is concave and continuous in M.

Lemma: *The value function $V(M,\theta)$ is unique. It is also concave and continuous in M.*

Proof:

Define the mapping,

$$T(V)(M,\theta) = \sup_{\{m\geq 0, n\geq 0\}} \{F(M+n+m,M,\theta)-vn-C(v)m + \rho\int V(m,\tilde{\theta})dG(\tilde{\theta})\}.$$

Then the value function is the fixed point of $T(V)$, i.e., $V(M,\theta) = T(V)(M,\theta)$.

Furthermore, under the norm $\|V\| = \sup_M \int |V(M,\theta)|dG(\theta)$ the operator $T(\cdot)$ is a contraction mapping in the space of functions continuous in M and measurable in θ. Hence the value function is the unique limit of the following algorithm.

$$V^0(M,\theta) = 0$$

$$V^{k+1}(M,\theta) = \sup_{\{m\geq 0, n\geq 0\}} \{F(M+n+m,M,\theta)-vn-C(v)m + \rho\int V^k(m,\tilde{\theta})dG(\tilde{\theta})\}.$$

Now, under our assumptions $F(M+n+m,M,\theta)-vn-C(v)m$ is concave and continuous in M, n, and m. It is then straightforward to check that $V^1(M,\theta)$ is concave and continuous in M. Applying the same argument recursively we can show that $V^k(M,\theta)$ is concave and continuous for any k. Since the limit of a concave function is concave, the value function $V(M,\theta) = \lim_{k\to\infty} V^k(M,\theta)$ is also concave. Furthermore under the norm we have defined the space of continuous functions in M is closed, and therefore $V(M.\theta)$ is also continuous in M.

Q.E.D.

From the results in Benveniste and Scheinkman (1979), the value function is differentiable. It is then straightforward to obtain the first order conditions for the dynamic programming problem. Define

$$\eta(M,\theta) = F_N - v, \tag{3.6}$$

$$\mu(M,\theta) = F_N + \rho\int V_M(m_t,\theta_{t+1})dG - C(v). \tag{3.7}$$

Then the first order conditions are

$$n_t\eta(M,\theta) = 0, \tag{3.8}$$

$$\eta(M,\theta) \leq 0, \tag{3.9}$$

$$m_t\mu(M,\theta) = 0, \tag{3.10}$$

$$\mu(M,\theta) \leq 0. \tag{3.11}$$

By the envelope property, the value function satisfies

$$V_M(m_{t-1},\theta_t) = F_N(N_t,K_t,\theta_t) + F_K(N_t,K_t,\theta_t) \tag{3.12}$$

$$V_\theta = F_\theta + \rho\int VdG_\theta. \tag{3.13}$$

Condition (3.9) shows that the value of marginal product of a temporary worker cannot exceed the labor cost of a temporary worker:

$$F_N \leq v.$$

From condition (3.8) the firm hires no temporary worker if the value of marginal product is strictly less than the labor cost, i.e., $F_N < v$.

Corresponding conditions for permanent workers are more complicated because the firm must keep a permanent worker for two periods. The value of marginal product of a permanent worker is F_N+F_K in the current period and V_M in the next period. Although the value of marginal product in the current period is known by the firm, the value in the next period, V_M, is uncertain. The expected value of their discounted sum is $F_N + \rho\int V_M(m_t,\theta_{t+1})dG$. Condition (3.11) says that this must be less than or equal to the labor cost of a permanent worker, $C(v)$. According to condition (3.10), the firm hires no permanent worker if the expected value is strictly less than the labor cost.

The following proposition shows that whenever the firm hires both temporary and permanent workers, the number of permanent workers hired is the same. Furthermore, if a period when the firm hires both types exists, then the firm always hires permanent workers whenever it hires temporary workers.

Proposition 3: *Suppose there exists a period when the firm hires both temporary and permanent workers. Then, if the firm hires a positive number of temporary workers, the firm always hires a constant number m^{**} of permanent workers, where m^{**} is positive and satisfies*

$$\rho\int V_M(m^{**},\theta)dG(\theta) = C(v) - v.$$

Proof:

In a period when the firm hires both temporary and permanent workers, we obtain $F_N = v$, and

$$\rho\int V_M(m_t,\theta)dG(\theta) = C(v) - v.$$

Given the concavity of the value function, the solution to this equation is unique. Hence, when the firm hires both temporary and permanent workers, the number of permanent workers hired is the same.

Next, we prove that whenever the firm hires a positive number of temporary workers, the firm always hires a positive number of permanent workers. Suppose the contrary. Then, since μ must be nonpositive and equals zero only at m^{**}, we must have $\mu < 0$. This implies that $m_t = 0$. Hence,

$$\mu = v + \rho\int V_M(0,\theta_{t+1})dG - C(v) < 0 .$$

However, concavity of V yields

$$\begin{aligned} & v + \rho\int V_M(0,\theta_{t+1})dG - C(v) \\ & \geq v + \rho\int V_M(m^{**},\theta_{t+1})dG - C(v) = 0 . \end{aligned}$$

This is a contradiction. *Q.E.D.*

A permanent worker is more productive than a temporary worker, but a permanent worker is more costly to the firm. If the productivity advantage is smaller than the extra cost of a permanent worker, hiring a permanent worker is not profitable. However, if the human capital embodied in the the workers is in some way a necessary input, then the firm will hire permanent workers every period. These results are summarized in the following proposition.

Proposition 4. *If $\rho F_K(N,K,\theta) < C(v) - (1+\rho)v$ for any N, K, and θ, then the firm hires no permanent worker. Suppose that $F_{NK} \geq 0$, then if $\rho\int F_K(N^*(\theta),0,\theta)dG(\theta) > C(v)$, then the firm hires a positive number of permanent workers every period. Here $N^*(\theta) = \underset{n\geq 0}{\mathrm{argmax}}\ \{F(n,0,\theta) - vn\}$. Suppose that $F_N(0,0,\theta) \geq v$ for every θ, then if $\rho\int F_K(N^*(\theta),0,\theta)dG(\theta) > C(v) - (1+\rho)v$, the firm will hire permanent workers in some states of nature.*

Proof:

Applying (3.9) and (3.12) to $V_M(m_t,\theta_{t+1})$ yields

$$V_M(m_t,\theta_{t+1}) \leq v + F_K.$$

Hence, we obtain

$$\begin{aligned}\mu(M,\theta) &= F_N + \rho\int V_M(m_t,\theta_{t+1})dG - C(v) \\ &\leq (1+\rho)v + \rho\int F_K dG - C(v) = \rho\int F_K dG - [C(v) - (1+\rho)v].\end{aligned}$$

This implies that if $\rho\int F_K dG < C(v) - (1+\rho)v$, then the firm hires no permanent worker. The first part of the proposition immediately follows from this result. Given that $F_{NK} \geq 0$, then the marginal product of a permanent worker in the following period is is at least as large as $\rho\int F(N^*(\theta),0,\theta)dG(\theta)$ at $m_t = 0$, from which we get the second result.

The final inequality follows by computing the marginal product of a permanent worker under the assumption that only temporary workers are hired each period. *Q.E.D.*

Notice that if human capital is sufficiently valuable, permanent workers will be hired every period. For example if one considers a Cobb-Douglas specification for F: $F(N,K,\theta) = \theta N^{\alpha}K^{1-\alpha}$, $\alpha \in (0,1)$, then permanent workers will be hired every period since the marginal product of capital is unbounded when $K = 0$. If no temporary workers are hired in a given period, then the number of permanent workers hired will in general depend on θ. In this case all workers will face a rising wage profile, and all workers will invest in human capital.

We may contrast this result to the WATJ institution. In this model the firm may hire a different number of permanent workers each period. All of them will be promoted if they invest in human capital. This is credible because at the time they are hired the firm commits itself to an expenditure of $m_tC(v)$ for two periods. As Malcomson (1984) has pointed out, the WATJ institution can achieve commitment by fixing the number of good jobs in the next period. This system may not be an optimal contract since the number of high paying jobs may be different from the number of workers that the firm wishes to have invest in human capital.

Secondly, this system may have less allocative flexibility than the LTE system. This is because in some periods the optimal allocation may involve a smaller number of high paying jobs. Since wages are attached to jobs, the firm cannot cut back on these jobs without adversely affecting the incentives in the system. In the LTE system the firm must pay the permanent workers their promised wage increase, but it may hire the workers out at the rate v.

4. Concluding Remarks

In this paper we introduced a simple overlapping generations model to study the effect of lifetime employment institution on the optimal employment policies. Lifetime employment allows contracts to be more efficient than they otherwise would be in a spot market environment. It allows the firm to commit itself to rewarding workers for investing in firm-specific human capital.

Because workers are homogeneous in our model, the optimal contract for LTE workers is a promotion scheme where (almost) all workers get promoted. Because promoted workers will be paid a higher wage than the market level, LTE workers face an upward sloping wage schedule. Furthermore, all of them make human capital investment in the first period.

Although a permanent worker is more productive than a temporary worker, the firm cannot lay him/her off. The firm therefore hires a temporary worker if demand is expected to fall in the next period. In our model where demand disturbances are i.i.d., the number of permanent workers hired is the same whenever the firm hires both temporary and permanent workers. Another consequence of the i.i.d. assumption is that, if a period when the firm hires both types exists, then the firm always hires permanent workers whenever it hires temporary workers. It has also been shown that if the human capital embodied in the the workers is in some way a necessary input, then the firm will hire permanent workers every period.

As discussed above, what the LTE system does in our model is to provide a mechanism for the firm to commit to a level of lifetime income for workers. This achieves higher human capital investment than otherwise. The optimal contract may not require all workers to invest in human capital. In our model such a possibility endogenously creates a distinction between permanent and temporary workers, as is observed in Japan. Since all workers are the same, both types of workers will normally receive the same lifetime utility, with the permanent workers receiving higher income to offset the cost of investing in human capital.

In practice workers will of course differ, and naturally firms would prefer better quality workers for LTE. This would create additional incentives for a distinction between the two types of workers, but it is not necessary for our theory. This consideration will imply, however, that testing of our model will require controlling for quality, as well as human capital levels. Preliminary work by Mincer and Higuchi (1988) does seem to suggest that much of the rising wage profile in Japan is due to human capital accumulation. What our model shows is that specific human capital can be a driving force for the LTE institution and a rising wage profile. This result is in contrast with the implicit contract model which would predict that risk averse workers should receive a flat wage over their lifetime.

REFERENCES

Benveniste, L. M. and Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics, *Econometrica* 47, 727-732.

Bull, C. (1987). The existence of self-enforcing implicit contracts, *Quarterly Journal of Economics* 102, 147-159.

Carmichael, L. H. (1983). Firm-specific human capital and promotion ladders, *Bell Journal of Economics* 14, 251-258.

Crawford, V. (1988). Long-term relationships governed by short-term contracts, *American Economic Review* 78, 485-499.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: a theory of vertical and lateral integration, *Journal of Political Economy* 94, 691-719.

Grout, P. (1984). Investment and wages in the absence of binding contracts: a Nash bargaining approach, *Econometrica* 52, 449-60.

Hart, O. D. and Holmström, B. (1987). The theory of contract, *in* "Advances in Economic Theory" (T. Bewley, Ed.), pp. 71-155, Cambridge University Press, London/New York.

Holmström, B. and Tirole, J. (1989). The theory of the firm, *in* "Handbook of Industrial Organization" (R. Schmalensee and R. Willig, Eds.) , North-Holland.

Krafcik, J. F. (1988). High performance manufacturing: an international study of auto assembly practice, mimeo, International Motor Vehicle Program, Cambridge, MA.

MacLeod, W. B. and Malcomson, J. M. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment, *Econometrica* 57, in print.

Malcomson, J. M. (1984). Work incentives, hierarchy, and internal labor markets, *Journal of Political Economy* 92, 486-507.

Malcomson, J. M. and Spinnewyn F. (1988). The multiperiod principal agent problem, *Review of Economic Studies* 55, 391-408.

Mincer, J. and Higuchi Y. (1988). Wage structures and labor turnover in the United States and Japan, *Journal of the Japanese and International Economies* 2, 97-133.

Williamson, O. E. (1985). "The Economic Institutions of Capitalism," Free Press, New York.

Footnotes

[1]See Williamson (1985) for a detailed discussion of labor institutions.

[2]See Hart and Holmström (1987) for a good survey of contract theory, and a discussion of the implications of incomplete contracting.

[3]As will be seen in section 2, an extreme case may exist where the incentive problem is so severe that human capital investment will be too low unless a permanent worker's life-time utility is higher than a temporary worker's. In such a case rationing of permanent jobs is necessary.

[4]Carmichael (1983) has also made a similar point.

[5]Of course many firms in the US do offer essentially lifetime employment contracts. Any differences between the US and Japan in the end must be a matter of degree and not absolutes. Many so called Japanese practices are found in US firms, and in car manufacturing at least many US plants are in fact more productive that the least productive Japanese plants. (see Krafcik (1988)).

[6]For a detailed discussion of the formal differences between contractible and non-contractible states see Hart and Holmström (1987) and MacLeod and Malcomson (1989).

[7]This result follows from our assumption that the labor market is competitive and that the firm has all the market power. Even if one assumed that the worker and the firm bargained ex post, one could still get the same result. This situation is formally studied in Grout (1984). In the absence of binding contracts, the wage contract is the outcome of a bilateral bargaining game. Since the investment into human capital is a sunk cost, in the bilateral bargaining game the worker and the firm will simply share

equally the productivity gain due to the investment. If receiving only a half of the surplus does not compensate for the investment, then the worker will not invest in human capital, even if employment is guaranteed. As the probability of reemployment falls, the incentive to carry out the investment goes to zero. Thus it is not strictly necessary to assume that the temporary worker has no market power to obtain the result that no human capital investment will occur when there is no employment guarantee in the second period.

[8]This observation was made by Malcomson and Spinnewyn (1988).

[9]Define $\tilde{F}(M,n,m,\theta) \equiv F(M+n+m,M,\theta)$ and denote $(N,K) = (M+n+m,M,\theta)$ by $(N,K) = G(m,n,M,\theta)$. Then $\tilde{F}(M,n,m,\theta) = F \circ G(m,n,M,\theta)$. Since F is concave in N and K, and G is linear in M, n, and m, it follows that $F \circ G$ is concave in M, n, and m.