THE FIRM AND WORKER LOYALTY

Jonathan Eaton

Working Papers in Economics E-88-30

The Hoover Institution
Stanford University
August 1988

Prepared for the "Symposium on Globalized Business Activities and the International Economy" held by the Research Institute of International Trade and Industry, Tokyo, Japan, June 23-24, 1988.

Hoover Institution
Working Paper No. E-88-30
August 1988

THE FIRM AND WORKER LOYALTY

Abstract

Empirical analysis has revealed two striking differences in U. S. and Japanese employment practices: Job tenure in Japan is longer, and wages rise faster with duration of employment. These characteristics may reflect compensation schemes that tie workers' incentives more closely to the long-term interests of the firm by deferring compensation until a time when it can be related more closely to the workers' previous effort. For deferred compensation to succeed in eliciting greater effort on behalf of the firm requires that the promise of later compensation for current effort be credible. If firms fulfil promises to pay deferred compensation in order to maintain their ability to defer compensation for workers hired subsequently then entry cannot occur to the point at which a marginal worker has zero net value. The constraint that the firm can offer only compensation schemes that will subsequently be in its interest to fulfil will in some circumstances imply that less effort can be elicited than would be the case if commitment to a scheme were perfect. This constraint will reduce effort more the higher the discount rate facing workers and firms. The lower cost of capital in Japan may therefore provide one explanation for the high levels of labor productivity achieved by Japanese firms.

Jonathan Eaton
National Fellow
Hoover Institution
(415) 723-0596
and
University of Virginia

I. Introduction

Japanese firms are renown for the high productivity of their domestic workers. Moreover, where they have expanded their activities abroad, worker productivity in Japanese-owned firms compares favorably with that in indigenously-owned or in other foreign-owned plants.[1]

Many observers attribute this success to the firm-employee relationship within Japanese-owned firms:

> Japanese workers' pride in their work and loyalty to their company are reflected in their capacity to produce goods that are not only competitive in price but reliable in quality.... The average Japanese laborer may accomplish no more than a loyal, hard-working American counterpart in a comparable factory, but loyalty to the company is typically higher and hard work more common. (Vogel, p. 131)

A critical component of worker loyalty to the firm is the typically longer-term nature of the relationship between workers and firms:

> They ... developed a seniority system of wage increases such that the newly-trained employees in whom the company invested so heavily would be motivated to remain.... The new philosophy incorporates many concepts from modern Western management.... [b]ut some basics of the pre-World War II Japanese system remain: long-term perspective, permanent employment, seniority, and company loyalty. (Vogel, pp. 133-134).

---

[1] See, for example, the discussion in Vogel (1979).

A perceived strength of the Japanese firm is its ability to harness its management and employees toward the pursuit of long-term goals rather than short-term profits:

> The Japanese firm is less interested in short-term profits and more concerned with the long run. Executives, ... when appropriate, boldly sacrifice profits for several years to build the groundwork for later success. (Vogel, p. 135)

Hashimoto and Raisian (1985) document two significant differences between employment practices in Japan and in the United States. The first is the Japanese system of "lifetime empoyment." The average tenure for a Japanese worker is much longer than for a U.S. worker: "By the time a typical worker reaches age 65, he would have had approximately five jobs in Japan and eleven jobs in the United States." (p. 725) "To summarize, long-term employment relationships are more prevalent in Japan than in the United States.... (p. 727)

A second difference is that the earnings profile of a Japanese working at a given firm is much steeper than that of a comparable U.S. worker. Hashimoto and Raisian find that "For all firm-size groups [small, medium, and large], growth rates between the peak-earnings year and the initial year are greater in Japan than in the United States.... More importantly, growth rates attributable to tenure are far greater in Japan than in the United States." (p. 732)

Hashimoto and Raisian point out that their findings on the average length of job tenure and the steepness of earnings profiles are consistent with at least three models of firm-employee relationships:

(i) Investment in firm-specific human capital provides one explanation for a positive relationship between job tenure and the steepness of the earnings profile. According to this interpretation, a worker's productivity at a particular job increases with his tenure at the firm as he accumulates firm-specific human capital. A positive earnings profile may emerge as the firm attempts to reduce the turnover of more experienced, and hence more productive, workers. In addition, firm-specific human capital may not be accumulated passively, but require effort on the part of the worker. To provide an incentive to undertake the effort necessary requires the firm to pay the worker a share of the return on his investment in human capital. A positively-sloped earnings profile would then reflect

the payment to a worker of his return on investment in human capital.[2]

(ii) A second hypothesis is that firms design compensation packages that result in steep earnings profiles to screen out potential low-productivity employees. According to this hypothesis, at the time that he seeks employment, a worker is better informed about his true productivity at the task in question than a potential employer. The employer can observe the worker's true productivity only after he has spent time with the firm. The firm can nevertheless ensure that it attracts only high-productivity workers by offering compensation that is less than that available in competing jobs during the period that the firm is ignorant of true productivity. It then rewards workers whom it finds to be highly productive with a level of compensation sufficient to ensure that the long-term reward of working for the firm exceeds that available elsewhere. Workers found to be unproductive are terminated or face compensation in line with their productivity.

---

[2]If the investment activity itself could be observed perfectly by the firm, it would not be necessary to delay compensation for the worker's effort in undertaking the investment until the return on the investment is realized. Compensating a worker for his effort in investing in firm-specific human capital only implies a positive earnings profile if the firm can observe the worker's investment only imperfectly. The firm may then choose to make increased compensation contingent upon the higher productivity actually achieved by the investment. Becker (1962), Hashimoto (1979), and Carmichael (1983) develop the theory of firm-specific human capital and its implications for job tenure and for earnings profiles.

Faced with such a compensation scheme, a worker who knows that his true productivity with the firm is low will seek a job elsewhere. Before the firm learns his true productivity he will earn less than he would elsewhere anyway, and once his true productivity is revealed he will fail to get a wage increase. But a worker who knows that he is highly productive at the task in question will be willing to forgo higher income during the trial period in anticipation of a high income once his true productivity is revealed to his employer.[3]

(iii) A third explanation appeals to the theory of agency. Even if workers are in principle identical, their productivity depends on their effort. The amount of effort that a worker exerts may only be observable to his employer after some time. A firm will consequently want to delay rewarding a worker for his effort until it knows how much effort was actually put in.[4]

In summary, the human capital, screening and agency hypotheses all provide explanations of a positive relationship between job tenure and earnings. They also suggest how higher worker productivity in Japanese firms may relate to the length of

---

[3]Spence (1974) and Salop and Salop (1976) explain a positively-sloped earnings profile on the basis of screening.

[4]Ross (1974) provides an early treatment of the design of incentive schemes to elicit effort on the part of an agent. Lazear (1979) models why, if its information on employee effort is delayed, a firm may not only defer compensation but also mandate a retirement age. Eaton and Rosen (1983) model and analyze the use of delayed compensation among executives of U.S. corporations.

tenure and the steepness of earnings profiles: Higher productivity in Japanese firms may be the consequence of: (i) more investment in firm-specific human capital; (ii) a more efficient matching of workers to tasks; (iii) greater effort on the part of workers towards achieving long-term goals of the firm. In each case the long-term relationship between workers and firms serves to unite the employee's objectives to those of the firm. A steeper earnings profile reflects the firm's desire to reward the employee for pursuing the firm's objectives.

All three explanations for long-term tenure and an upward-sloping earnings profile relate to the problem of asymmetric information between the firm and a worker. In the case of firm-specific human capital, the firm cannot observe the worker's investment activity itself, only its subsequent return. In screening, the firm cannot immediately observe the worker's productivity. If there is an agency problem, the firm cannot observe actual effort, and productivity is observed only with a lag. The first and third informational asymmetries constitute problems of moral hazard, the second one of adverse selection. By maintaining a long-term relationship with its employees and deferring compensation, the firm can mitigate the costs of these informational asymmetries. These can partially or totally overcome the problem that arises from the worker's inability to communicate contemporaneously his actual effort or his true productivity to his employer.

If, in fact, the success of Japanese firms in achieving high levels of worker productivity does derive from their superior ability to exploit the benefits of a long-term relationship and deferred compensation, the question then arises as to why Japanese firms have this superior ability. While there is now a substantial literature on how long-term tenure and a positively-sloped earnings profile can mitigate moral hazard and adverse selection problems among workers, less appears to be written on how these solutions actually introduce a problem of moral hazard on the part of the firm. Each of the explanations for a steep earnings profile requires that the firm commit to a compensation scheme that, once it comes time to execute, imposes a cost on the firm. According to the screening and agency hypothesis, the firm elicits efficient sorting or greater effort by deferring compensation from the period in which effort is actually put in to a subsequent period. In a competitive equilibrium the discounted value of the worker's effort will equal his wage. If the wage in the early period of employment falls short of productivity then in the later period of employment it must then exceed productivity.[5] The agency and delayed compensation explanations of a positive earnings profile therefore require firms to continue to employ workers even though the value of what they produce is less than they are paid.

---

[5]Indeed, this is the basis of Lazear's (1979) explanation of mandatory retirement.

The firm would raise current profits by firing these workers or reducing their pay to match their productivity. What keeps the firm from doing this, and why do new workers find the promise of future compensation for current effort believable?

Free entry is typically assumed to occur to a point at which the present discounted value of a firm is zero. A new firm that hires young workers and defers compensation earns a strictly positive cash flow initially. If entry occurs to the point at which the discounted present value of a new firm is zero then this initial positive cash flow must be offset by a negative one subsequently. The value of the firm must become negative at some point. A critical question is whether or not a firm with a negative value can be forced to continue operation and prevented from dissolving, thereby renegging on a promise of deferred compensation. If not, the equilibrium will be constrained by the requirement that any firm's value remain nonnegative.[6] This requirement may limit the firm's ability to elicit extraordinary effort or to screen efficiently.

This paper develops a model of the firm as a long-term

---

[6]In the case of the firm-specific human capital, since effort in the early period raises productivity later, deferring compensation does not mean that the earnings of an older worker exceed his marginal product. Hence deferring compensation in a competitive equilibrium does not imply that an ongoing firm has negative value. Nevertheless, if human capital is truly firm-specific, then the question remains as to why the firm is willing to pay a worker more than his current opportunity wage elsewhere.

employer of workers. Its purpose is to illustrate how the success of Japanese firms in eliciting effort and loyalty from their workers may derive from their superior ability to make credible promises of future compensation as a reward for current effort.

The analysis in Sections II, III and IV below characterizes the equilibrium of an economy in which compensation is deferred to overcome a problem of moral hazard: The firm can observe a worker's true effort only after he has been on the job for some time. A competitive equilibrium in which firms continue operation despite the expectation of ongoing losses is contrasted with one in which the value of a firm is constrained always to be nonnegative. In the absence of this constraint, deferred compensation can elicit the optimal amount of effort, that is, the amount of effort that equates the marginal product of an additional unit of effort to the firm to its marginal cost to the worker. If older workers are sufficiently productive, and the discount rate sufficiently low, then the nonnegativity constraint has no effect, and the appropriate contract can elicit optimal effort. Otherwise, the nonnegativity constraint results in a level of effort below the optimum. How far below depends, in particular, on the interest rate. As the interest rate facing workers and the firm rises, less effort is forthcoming.

Section V concludes with a brief discussion of what the model suggests about the implications of Japanese policy and the

overseas expansion of Japanese firms.

II. Production with Lagged Observation of Effort and Perfect Commitment to Wage Contracts

The value of a worker to his employer depends on the effort that he puts into his job. If the employer cannot observe effort while the worker is actually on the job, and if putting in effort is unpleasant for the worker, then a problem of moral hazard emerges. The employer has to provide an incentive for the worker to exert effort in his work.

If the employer can observe a worker's effort with a lag, however, then delaying compensation until information on effort is available, and making the amount of compensation contingent upon the effort observed, provides a means of eliciting greater effort from workers.

Consider an activity in which the value of a worker's activity in any period equals an amount e, which reflects his effort that period. Exerting effort imposes a disutility on the worker in amount v(e), where the function v(e) is twice differentiable and where $v'(e)>0$ and $v''(e)>0$. The employer can observe e only with a lag of one period. The minimal possible effort is $\underline{e}$ where $\underline{e}\geq 0$. Let $v(\underline{e})=0$.

Workers live two periods. In the first period they have the option of working in another activity which provides compensation

net of effort in amount u. If it is the case that $u>\underline{e}$ then the activity in which monitoring effort is problematic cannot operate unless workers can be encouraged to exert more than minimal effort. This is because the highest wage that an employer would be willing to pay for this activity, $\underline{e}$, would not be sufficiently attractive to lure them from the other activity. If, however, there exists a level of $e\geq\underline{e}$ such that $e-v(e)\geq u$ then <u>provided that the appropriate level of effort is elicited</u> this activity can be made attractive to workers and simultaneously leave the firm with a profit.

Say that workers try to maximize the expected discounted value of wages over two periods less the disutility of effort during those two periods. Note that in the second period of their work firms cannot elicit more than the minimal level of effort from any worker. The reason is that the worker's productivity cannot be observed until after his departure from the economy, which precludes compensating him on the basis of his observed effort that period.

Denote the discount factor that first-period workers apply to second period wages as β. To elicit an amount of effort $\tilde{e}$ requires both that the lifetime value of working in the activity in which monitoring is delayed is at least equal to that elsewhere. If compensation in the two periods of lifetime in this activity is w_1 and w_2, respectively, and the level of effort put forth in the first period is $\tilde{e}$, then this condition is that:

$$(1) \qquad w_1 + \beta w_2 - v(\tilde{e}) \geq u.$$

To elicit effort in the first period the firm denies second period compensation to any worker found to have exerted less effort than what was agreed upon. For this threat to succeed in eliciting effort in amount $\tilde{e}$ requires that:

$$(2) \qquad v(\tilde{e}) \leq \beta w_2.$$

I also assume that compensation is constrained to be nonnegative; i.e., that $w_1 \geq 0$ and $w_2 \geq 0$ and that second-period workers have no opportunities elsewhere.

Consider a firm that is normalized at a size such that it hires one new worker each period. In a stationary situation, the beginning-of-period value of an ongoing firm that has agreed to pay an old worker w_2' and can elicit effort $\tilde{e}$ from a young worker with a promise of lifetime wages w_1 and w_2 can be defined recursively as:

$$(3) \qquad q^O(w_2') = \max_{w_1, w_2, e} [\tilde{e} + \underline{e} - w_1 - w_2' + \beta q^O(w_2)],$$

where $\beta < 1$ is the firm's discount factor, assumed to be the same as a worker's. This maximization is subject to constraints (1)

and (2) above.

The firm's cost minimization will ensure that the first constraint binds. Substituting this constraint into expression (3) implies that the value of the ongoing firm is:

$$(4) \quad q^{O}(w_2') = \max_{w_2', e} [\tilde{e} + \underline{e} - u - v(\tilde{e})]/(1-\beta) - w_2'.$$

subject to the constraint (2).

The first-order condition for the optimal level of effort is that:

$$(5) \quad 1 - v'(\tilde{e}) = 0$$

for values of $\tilde{e} > \underline{e}$. This states that the optimal contract elicits effort at the level at which the marginal disutility of effort to a worker equals the marginal contribution of additional effort to output. I denote the level of effort that satisfies this condition as e*. Note, in particular, that e* is independent of the discount factor.

Any level of w_2 in the range $[v(e^*)/\beta, u/\beta]$ satisfies conditions for an optimum and the incentive-compatibility constraints on workers. As long as workers face the same discount rate as firms, the timing of compensation is irrelevant as long as the worker's incentive-compatibility condition (2) is

satisfied.

A new firm, that is, one that has no old workers, will have a value:

$$(6) \qquad q^N = \max_{w_1, w_2, e} \tilde{e} - w_1 + \beta q(w_2)],$$

again subject to constraints (1) and (2). The first-order condition (5) continues to determine the optimal level of e and cost minimization continues to impy that constraint (1) will bind. The value of a new firm is consequently:

$$(7) \qquad q^N = \max_{w_2}[e^* - u - v(e^*) + \beta w_2 + \beta q^O(w_2)]$$

$$= [e^* - u - v(e^*) + \beta \underline{e}]/(1-\beta).$$

Note that the value of a new firm, unlike that of an ongoing firm, is independent of w_2.

A crucial observation is that, as long as $w_2 > \underline{e}$, that is, as long as an old worker is paid more than his marginal product, then the value of a starting firm exceeds the value of an ongoing firm. The reason is that the first-period cash flow per worker of a new firm exceeds that of an ongoing firm by an amount $w_2 - \underline{e}$.

If new firms can credibly offer long-term compensation in which second-period compensation exceeds the worker's

contribution to output in that period, then the sector with costly monitoring will expand as long as the value of a new firm is positive.[7] Expansion will result in competition for workers in this activity to the point at which the return to labor in other sectors is bid up to the level:

$$u = e^* + \beta\underline{e} - v(e^*). \tag{8}$$

At this point $q^N = 0$, so that further entry will not occur.

But at this level of compensation the value of an ongoing firm is $\underline{e}-w_2$, which must be negative if $v(e^*)/\beta>\underline{e}$. In this case, since an ongoing firm can anticipate a perpetually negative cash flow, it has a negative value. The reason is that, in its initial period, delayed compensation allowed the firm to experience a strictly positive cash flow. In other words, the equilibrium involves a firm earning a positive cash flow of $\beta(w_2-\underline{e})$ in the first period of its existence and a negative cash flow of $(1-\beta)(\underline{e}-w_2)$ thereafter. When $w_2\geq v(e^*)/\beta$, the maximum value of an ongoing firm is $\underline{e}-v(e^*)/\beta$. Note, in particular, that cash flow is more negative the lower the discount factor β.

---

[7] Growth can literally take the form of entry of new firms or expansion of existing ones. The analysis here does not model the determinants of firm size. As long as firms are competitive, firm size does not affect the results.

III. Production with Lagged Observation of Effort and Imperfect Commitment to Wage Contracts

Since eliciting the optimal amount of effort may require, in competitive equilibrium, that the firm indefinitely run a negative cash flow, the ability of firms to commit to schemes of delayed compensation that are capable of eliciting the optimal amount of effort is questionable. In the equilibrium characterized above, a firm would have an incentive to enter into production its first period, earn e^*-w_1, and then dissolve, renegging on compensation promised in the second period. There is no incentive to continue operation since the firm has nothing to look forward to but a subsequent negative cash flow: Maintaining its ability to attract new young workers is not a sufficient incentive to repay old ones more than their marginal product.

In this case workers would not find the promise of future compensation sufficient to elicit e^* credible. They would anticipate that the firm would renege on its promise and refuse the wage contract described.

If firms can renege on delayed compensation by dissolving, then feasible compensation schemes will be subject to an incentive-compatibility constraint on firms as well as on workers. If dissolution absolves a firm from past wage commitments then the additional constraint takes the form that

$q^O(w_2') \geq 0$ in all periods.

In a stationary equilibrium $q^O(w_2')$ is independent of time and is the same for any firm after its initial period of existence. Hence if this condition is binding then $q^O(w_2) = 0$, which implies that cash flow is zero in all periods except, possibly, the initial one. Incorporating condition (5), cash flow π as a function of w_2 and $\tilde{e}$ of an ongoing firm is:

$$\pi = \tilde{e} - \underline{e} - u - v(\tilde{e}) + (\beta-1)w_2. \tag{9}$$

This magnitude is greater the lower w_2. That is, the less that the firm relies on delayed compensation the greater will be its cash flow after the initial period. The nonnegativity constraint on cash flow is consequently most easily satisfied when w_2 equals its lower bound of $v(\tilde{e})/\beta$. In this case cash flow becomes:

$$\pi = \tilde{e} + \underline{e} - u - v(\tilde{e})/\beta. \tag{10}$$

Given u, a firm can credibly offer only those wage contracts that elicit a level of effort that maintains a nonnegative cash flow. That is, the level of effort $\tilde{e}$ elicited must maintain a nonnegative value of π in expression (10) above. If the level of $\tilde{e}$ that satisfies the first-order condition (5) is consistent with a positive value of π in expression (10) then the incentive-compatibility constraint imposed on the firm is

consistent with a wage contract that elicits effort e*. Otherwise only a lower level of effort can be elicited. As the discount factor β falls, π becomes more negative. Hence, it becomes more difficult to obtain the optimal level of effort when the discount factor is low.

Since both the maximand (4) and expression (10) are concave functions of $\tilde{e}$, when the incentive-compatibility constraint (10) is binding, the firm will choose a wage contract that elicits the greatest possible effort consistent with a positive cash flow. This amount will be lower the lower is β. Hence the amount of effort actually elicited declines the more the future is discounted. This is in contrast to situations in which the incentive-compatibility constraint on firms is nonbinding, or in which firms can costlessly commit to deferred compensation, since e* is independent of β.

This result can be summarized in the following proposition:

Proposition 1: If eliciting optimal effort requires a firm to pay a senior worker more than his marginal product, and the firm is constrained to have nonnegative value in every period, then the level of effort that the firm can elicit from workers falls short of the unconstrained maximum. The amount by which it falls short will be greater the lower the discount factor in the economy.

Since the cash flow of an ongoing firm is zero, entry may no longer occur up to the point at which a new firm has zero value. Entry may only occur up to the point at which the firm's incentive-compatibility constraint is binding. The compensation provided by the other sector will then be:

$$u = \tilde{e} + \underline{e} - v(\tilde{e})/\beta. \tag{11}$$

This amount is less than what compensation would be in the absence of a binding incentive-compatibility constraint on the firm for two reasons. One is that the level of effort itself is less than e^*. The other is that the current-period value of deferred compensation determines the cash flow of an ongoing firm, while its (lower) discounted value determines the value of a new firm. The sector in which monitoring is delayed is consequently smaller relative to the other sector, and the overall level of compensation in the economy is lower.

IV. An Example

To make these ideas more concrete consider a specific example. Let the disutility of effort have the functional form:

$$v(e) = (e-\underline{e})^2/2 \qquad e \geq \underline{e}. \tag{12}$$
$$v(e) = 0 \qquad e < \underline{e}$$

where $\underline{e}<1$. In this example $e^* = 1+\underline{e}$. Eliciting this amount of effort requires that $w_2 \geq 1/2\beta$.

If the firm is able to write a wage contract that elicits this amount of effort then the level of compensation at which the value of a new firm is zero is given by:

$$u^N = 1/2 + (1+\beta)\underline{e}. \tag{13}$$

The value of an ongoing firm can be no greater than $(\underline{e}-1/2\beta)$, which is negative as long as $\underline{e}<1/2\beta$. In this case imposing the incentive-compatibility constraint on the firm that it always have positive value will modify the equilibrium.

Given u and $\tilde{e}$, the constraint that an ongoing firm earn nonnegative cash flow is that:

$$\tilde{e} + \underline{e} - u - (\tilde{e}-\underline{e})^2/2\beta \geq 0. \tag{14}$$

For $e \leq e^*$ firm profit is increasing in $\tilde{e}$. Hence if (14) constrains $\tilde{e}$ below e^* then a firm facing a given compensation requirement u will choose the maximum $\tilde{e}$ compatible with a nonnegative value of expression (14), which is:

$$\tilde{e} = \beta + \underline{e} + \sqrt{\beta+4\underline{e}-2u}\sqrt{\beta}. \tag{15}$$

which is defined only for:

$$u \leq (\beta + 4\underline{e})/2. \tag{16}$$

If u violates condition (16) then the firm's incentive-compatibility constraint precludes the activity requiring deferred compensation from occurring at all. No credible wage contract can elicit an amount of effort that prevents cash flow from turning negative for an ongoing firm.

Note that the level of effort that can be elicited depends positively on the discount factor β and negatively on the level of compensation in the other sector, u.

Whether or not entry will occur to the point at which new firms have zero value depends upon whether or not $\underline{e} \gtrless \beta/2$. If $\underline{e} \geq \beta/2$ then new firms can enter up to the point at which the value of a new firm is also zero. In this case condition (15) determines $\tilde{e}$. At this point the value of compensation in the competing sector is bid up to:

$$u^C = \underline{e} + \sqrt{2\beta\underline{e}}. \tag{17}$$

For $\underline{e} \geq \beta/2$ this level of compensation is consistent with condition (16). Hence, for $\underline{e} \in [\beta/2, 1/2\beta]$, entry will occur up to the point at which the value of a new firm is zero, just as in the unconstrained case. The level of effort that will be elicited is given by:

$$e^C = \beta + \underline{e} + \sqrt{\beta}(\sqrt{2\underline{e}} - \sqrt{\beta}). \tag{18}$$

Note that this level of effort and as well as the level of compensation increase with the discount factor β up to the point at which $\beta = 1/2\underline{e}$, where the unconstrained equilibrium is attainable. In this range both new and ongoing firms have zero cashflow. The value of deferred compensation corresponds exactly to the marginal product of a second-period worker, and first-period compensation equals the marginal product of a first-period worker.

Finally, for $\beta \in [0,\beta/2]$, a level of compensation given by u^C in expression (17) is no longer compatible with condition (16). There is no level of effort that can credibly be elicited consistent with this level of compensation. In this case the equilibrium is constrained by the condition that, whatever u is, the corresponding value of $\tilde{e}$ be feasible. Condition (16) consequently constrains the equilibrium. At this point firms can enter this activity only until the level of compensation attains:

$$u^E = (\beta + 4\underline{e})/2. \tag{19}$$

The implied amount of effort elicited in the first period is:

$$\tilde{e}^E = \beta + \underline{e}. \tag{20}$$

Note that once again both compensation and effort increase with β and $\underline{e}$.

A first-period worker produces $\tilde{e}^E$ but is paid only u^E. The firm earns an excess of $\beta/2-\underline{e}$ which offsets its loss from paying a second-period worker, who produces only $\underline{e}$, $\beta/2$.

Since the constraint that an ongoing firm have nonnegative cash flow is binding, the value of an ongoing firm is of course zero, as in the previous case discussed. A new firm, however, has value:

$$q^N = \beta/2 - \underline{e}, \tag{21}$$

which is positive. Entry does not occur up to the point at which the return on entry is driven to zero. The reason is that, once the required level of compensation is u^E, a new firm, in attempting to enter, would bid up the return to labor above the level consistent with the firm's incentive-compatibility constraint. It could not credibly promise workers that the compensation required to elicit the requisite amount of effort would be forthcoming.

Even though it is highly stylized and oversimplified, this example suggests the possibility of a quite rich variety of relationships bonding a worker and a firm, depending upon the appropriate discount factor and minimum worker productivity. If

minimum productivity and the discount factor are both very high then a firm can elicit the first-best amount of effort, the amount that equates its marginal productivity to its marginal disutitlity, since this does not require paying a worker more than his marginal product at any point. In this case the amount of effort and hence the productivity of a worker does not depend on the discount rate.

If, however, the discount factor and minimal effort are low then eliciting the optimal amount of effort requires the firm at some point to have a negative value. Even though the young workers in such a firm are paid less than their marginal product, older workers are paid in excess of their marginal product, with the net effect on cash flow negative.[8] If such a firm cannot be kept from dissolving then it cannot credibly offer a worker a compensation package that elicits the first-best amount of effort. A potential worker will anticipate that the firm will dissolve before he receives any deferred compensation. One possibility, if the minimal level of worker productivity is not too low, is that the firm will pay a worker his marginal productivity in each period. If the minimal level of worker productivity is very low, however, then in equilibrium the firm

---

[8]Introducing growth in the size of the firm does not affect this result as long as, at some point, the interest rate exceeds the growth rate of employment. Since this relationship between these rates is implied by the presence of any fixed asset in the economy, such as land, it is reasonable to impose as a constraint on the steady-state equilibrium of the economy.

will pay a worker more than his marginal productivity at the end of his relationship with the firm. In this situation, if an old firm has zero value then a new firm has positive value, but the constraint that a new firm's value remain positive after one period prevents firms from entering up to the point at which the value of a new firm is zero. Workers would not accept offers from new entrants in anticipation that, once facing higher employment costs, these firms would renege on their promised delayed compensation. In this type of equilibrium new firms will be observed to earn excess profit and old firms zero profit.

V. Implications for International Competition and Direct Foreign Investment

The model developed in Sections II, III, and IV focusses on one function of a firm, as an entity that can enter into long-term relationships that an individual, with finite life, cannot. The excess profit that an ongoing firm can earn by hiring young workers provides it an incentive to pay old workers more than their marginal product. Since its lifetime is infinite, or at least indefinite, workers perceive that it is in the firm's interest not to renege on deferred compensation. To do so would prevent it from earning a profit from hiring new workers. Because it can credibly promise to reward current effort with future compensation, it can elicit more effort than

would otherwise be forthcoming.

A prediction of the model is that productivity is higher in economies where the cost of capital is low. This is for two reasons. First, since workers discount the future less, deferring compensation is less costly. In addition, since firms discount the future less, the value of a new firm is more in line with that of an ongoing one.

One feature of the Japanese economy that may have made it a particularly fertile source of firms that can elicit high productivity is its high saving rate, relative to the level of investment. This has been a consequence of both a high level of private saving and government efforts to curtail budget deficits. The high level of saving, relative to investment, is reflected, among other things, in the large current account surpluses that have appeared since the liberalization of Japanese financial markets.[9] To the extent that international capital mobility remains imperfect, the cost of capital for Japanese firms is less than that elsewhere. The analysis above suggests why this will enhance their ability to achieve higher levels of worker-productivity.

Another feature of the economy is the implicit or explicit government financial backing of firms. According to Vogel (1979), at their initiation, at least, basic industries "were

---

[9]See, for example, the simulations discussed in Sachs and Boone (1988).

well-financed and ultimately backed by the government." (p. 133) Such backing also lowers the cost of capital for the firm.

To the extent that it is the firm's cost of capital that matters, even in economies in which workers discount the future more heavily, Japanese firms may be at an advantage, but the productivity levels that can be attained may not match those in Japan.

In summary, the ability of Japanese firms to exploit long-term relationships with their employees seems to have been a critical feature of their success in achieving high labor productivity. This paper has developed a very simple model to illustrate the role of the firm as an institution that can elicit the loyalty of workers that is manifested in greater productivity. It suggests that firms will be most successful at eliciting effort when their workers, and their own, rates of discount are low.

REFERENCES

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

Carmichael, Lorne (1983), "Firm Specific Human Capital and Promotion Ladders," Bell Journal of Economics, 14: 251-258.

Eaton, Jonathan and Harvey S. Rosen (1983), "Agency, Delayed Compensation, and the Structure of Executive Remuneration," Journal of Finance, 38: 1489-1505.

Hashimoto, Masanori (1979), "Bonus Payments, On-the-Job Training and Lifetime Employment in Japan," Journal of Political Economy, 87: 1086-1104.

Hashimoto, Masanori and John Raisian, (1985), "Employment Tenure and Earnings Profiles in Japan and the United States," American Economic Review, 75: 721-735.

Lazear, Edward P. (1979), "Why is There Mandatory Retirement?" Journal of Political Economy, 87: 1261-1284.

Ross, Stephen A. (1974), "On the Economic Theory of Agency and the Principle of Similarity," in M. F. Balch, D. L. McFadden, and S. Y. Yu, editors, Essays on Economic Behavior and Uncertainty. Amsterdam: North-Holland.

Sachs, Jeffrey and Peter Boone (1988), "Japanese Structural Adjustment and the Balance of Payments." Unpublished, Harvard University.

Salop, Joanne and Steven Salop (1976), "Self-Selection and Turnover in the Labor Market," Quarterly Journal of Economics, 90: 619-627.

Spence, A Michael (1974), Market Signaling. Cambridge, Massachusetts: Harvard University Press.

Vogel, Ezra F. (1979), Japan as Number One. Cambridge, Massachusetts: Harvard University Press.