

A Behavioral Economic Approach to Performance-based Wage Systems

Kohei Daido

School of Economics, Kwansei Gakuin University

1. Introduction

With the burst of the bubble economy and prolonged recession that followed in Japan, the way in which various features of the Japanese-style management—features that had been praised as supporting Japan's economic growth—were appraised drastically changed; subsequently, a need arose to reform those features in response to changes in the business environment in which firms were placed. In this context, a review of seniority wage systems and other reward systems began, and many firms started to introduce performance-based wage systems in which the amount of a worker's remuneration was determined based on the worker's performance. Although the operation of these systems seems to be simple, in practice they have not functioned very well. Furthermore, recently problems with these systems have often been pointed out.¹

The objective of this paper is to analyze performance-based wage systems from the perspective of contract theory. In addition, we focus on the effects that incentives have on the psychological aspects of agents. This is based on principal-agent models that incorporate the fruits of behavioral economics, which has made remarkable progress in recent years. The viewpoints of behavioral economics are important because introduction of the performance-based wage system not only acts as an incentive for workers, but it may also have a psychological effect on workers. This effect is considered as one of the reasons for the gap between the intended effects of performance-based wage systems of firms and the actual effects.

For example, because wages are linked to performance in the performance-based wage system, a wage gap will be generated among workers in the same workplace. As a result, the wage gap may have an effect on the motivation of workers, and the performance-based wage system may work in a different direction from what the firm first intended. There is also the issue of the difficulty of assessing performance as a basis of determining remuneration. It is not difficult to imagine that the same performance is evaluated differently

¹ See, for example, Jo (2004) and Takahashi (2004).

depending on who the assessor and the assessed are. In this case, it is possible that the gap in evaluation may have an effect on workers' motivation. Therefore, introduction of a performance-based wage system not only functions as a direct incentive for workers, but it may also have a psychological effect on workers; this generates a gap between the intended effect of introducing the system and the actual outcomes.

For these reasons, a more meaningful analysis of performance-based wage systems is possible by considering not only the direct effects of remuneration as an incentive, but also its psychological effects on agents. Therefore, in this paper we use a number of simple models and discuss the differences that emerge when considering the psychological effects of such incentives, in comparison with results obtained from standard models.²

The contents of this paper are as follows. In Section 2, we consider the situation where an agent's utility depends not only on the agent's own remuneration, but also on the remuneration of others. For this purpose, we explain the term, "social preferences," as defined by Fehr and Schmidt (1999). We then analyze a moral hazard model that incorporates social preferences, examine what effects the consideration of social preferences have on optimum reward systems, and study how the degree of dependence of remuneration on performance is determined based on the degree of social preferences and the difficulty of a project. If it can be understood that the widening of the wage gap means the expansion of performance-based pay, we would then be able to point out from the results in which situations performance-based pay is favorable.

In Section 3, we use a simple model to explain the "crowding-out effects" of motivation by incentives, as proposed by Frey (1997) and others. The crowding-out effects are a psychological concept based on a theory concerning the relation between external monetary rewards and intrinsic motivation, and they are also used by Takahashi (2004) as grounds for criticizing performance-based pay. In this section, therefore, we define the crowding-out effects and confirm, by presenting Frey (1997)'s model, that when these crowding-out

² For this reason, we do away with details in the models described in the following sections, at the risk of sacrificing rigorousness to a certain degree, in order to provide plain descriptions as much as possible. In each model, however, we list references of more detailed analyses, to which readers interested in more rigorous discussions can refer.

effects exist, they give rise to the possibility that performance-based incentives may be unfavorable. Lastly, in Section 4, we consider the psychological utility associated with the issue of delegation of responsibilities, and compare fixed wage contracts and pay-for-performance contracts. In reality, it may be difficult to judge whether performance-based reward systems are generally good or bad. Empirical studies in the field of labor economics, however, have shown a number of necessary conditions for performance-based pay to function effectively. Delegation of responsibilities is sometimes mentioned as one of them. It has been indicated that when a performance-based wage system is accompanied by the delegation of responsibilities and when assessment is done fairly (in the respect that when an agent succeeds in enhancing performance because he has been delegated responsibilities, the enhanced performance is reflected in the agent's remuneration), the performance-based wage system is more favorable than fixed wage contracts.

2. Performance-based Pay and the Wage Gap

Performance-based wage systems may differ widely in detail depending on the firms introducing those systems. Okunishi (2001), however, gave the following three points as principle features common to all performance-based wage systems:

- (i) Results, more than processes, are emphasized as a wage-determining factor.
- (ii) Short-term results, more than long-term results, are emphasized.
- (iii) They widen the wage gap.

As Okunishi (2001) points out in relation to the third point above that widening of the wage gap may be interpreted as an expansion of performance-based pay, it is important that wages for good and bad performances are widely differentiated in a performance-based wage system. This is so that the differences act on workers as incentives and performance-based wage systems function effectively. Therefore, the question in introducing a performance-based wage system is whether a reward system as an incentive mechanism that widens the wage gap is actually favorable for firms.

This question, however, is based on the assumption that remuneration only has a monetary effect on workers. It should be noted that wage gaps created in a performance-based wage system may have other effects on a worker's utility. In other words, it is important to consider that a worker is interested not only

in the amount of wages he receives, but also in the amount of his wages relative to those of another worker (a colleague in the same workplace, for example). Fehr and Schmidt (1999, 2003) conceived a utility function that reflects not only pecuniary benefits considered in standard economic models, but also the amount of others' gains; economic subjects with such a utility function are called economic subjects with social preferences, and their effect is analyzed.

In this section, we analyze a moral hazard model with limited liability constraints in which types of social preferences are incorporated into the utility function of agents. Through this analysis, we see how reward systems as an incentive mechanism will be when social preferences are taken into account. As a result, the optimum reward system is determined based on such parameters as the probability of a project's success and degree of social preferences; therefore, we consider the question of in which situations there should or should not be a wage gap, in other words, in which situations performance-based wage systems will be more favorable.

We first consider a standard model.³ We think of a case where there is a risk-neutral principal and agent. The agent carries out a project, and wage is paid depending on the results. The agent carries out the project at an effort level e , and the output x is generated as a result. Here, we suppose that the agent's effort level is either high or low, and the outcome is either a success (x_s) or failure (x_f). The cost of effort borne by the agent is $d > 0$ when the effort level is high and 0 when the effort level is low. The outputs are $x_s = x > 0$, $x_f = 0$. Obviously, it is considered that the higher the effort level, the higher the probability of success. More specifically, when the agent's effort level is high, the probability of success is p_1 , and when it is low, the probability of success is p_0 , where $0 < p_0 < p_1 < 1$. The principal cannot monitor the agent's effort level, and therefore the wage is dependent on the outcome. As the simplest form of performance-based wage, wage w_s is paid when a project is successful, and wage w_f is paid when a project fails. For the sake of simplicity, it is assumed that $w_s = w \geq 0$, $w_f = 0$.

In this case, let us analyze what wage w the principal will offer to the agent. The agent's utility is defined as wage minus the cost of effort. On the other

³ For more details on contract theory, see Itoh (2003). The model in this section is based, in particular, on the model of limited liability constraints in chapter 5.1.

hand, the principal's utility is the project's output minus the wage. Assume that x is now sufficiently large, and the principal hopes that the agent's effort level is high. In this case, the principal needs to choose from w that meets the conditions shown below and that will minimize her expected payment.⁴

$$p_1w - d \geq p_0w, \quad (\text{IC})$$

$$p_1w - d \geq 0, \quad (\text{IR})$$

$$w \geq 0 \quad (\text{LL})$$

The left side of (IC) is the agent's expected utility when his effort level is high, and the right side is the expected utility when the effort level is low. Therefore, (IC) is a condition for the agent to choose a high effort level over a low effort level (incentive compatibility constraint). (IR) is the condition for the agent to agree to this contract when the reservation utility the agent gains by utilizing an external opportunity rather than accept the contract is zero (individual rationality constraint). Lastly, (LL) is the limited liability constraint where a contract in which the agent's wage becomes negative cannot be enforced.

The analysis of this case is easy. If we sort out (IC), we obtain the condition $w \geq d/\Delta p$, where $\Delta p \equiv p_1 - p_0$. Considering that the principal hopes to minimize w as much as possible and that $w = d/\Delta p$ always satisfies (IR), the optimum wage in this case is $w^* = d/\Delta p (> 0)$.

We next discuss a model in which the social preferences of agents are incorporated in the above model. To consider the comparison of wages between agents, assume now that there are two agents. To make our analysis simple, the two agents are homogenous; there is no correlation between their productions, and the two are independently engaged in their projects and are individually paid wages in accordance with their outputs.⁵

Here, we formulate social preferences in accordance with Fehr and Schmidt (1999). In doing so we consider, as an item in the agent's utility function in addition to wage and the cost of effort, the utility S gained when comparing one's own wage (w) with another's (colleague's) wage (\bar{w}), as shown below.⁶

⁴ In accordance with the custom in this field, the principal is hereafter called "she" and the agent, "he."

⁵ Itoh (2004) analyzes a case where wage is determined in accordance with each agent's output and what features the optimum contracts would have depending on social preferences.

⁶ For the purpose of this paper, we call the utility that is considered additionally when a psychological factor such as S is included the "psychological utility."

$$S \equiv -\alpha \max\{\bar{w} - w, 0\} - \alpha \lambda \{w - \bar{w}, 0\}. \quad (1)$$

$\alpha > 0$ is a coefficient showing the degree of social preferences. In the first term, when the colleague's wage is higher than the agent's own wage, the agent will be subject to psychological disutility expressed as the difference in the wages multiplied by α . There may be a number of cases depending on the value of λ . If $\lambda > 0$, the agent is averse to any wage gap between him and the colleague. This is called the "inequity aversion." If $1 > \lambda$, the agent is averse to any wage gap regardless of whether his wage is higher or lower than the colleague's; however, the disutility the agent feels when his wage is higher than the colleague's is smaller than the disutility he feels when his wage is lower than the colleague's. This is called the "loss aversion." If $0 > \lambda > -1$, it can be interpreted as indicating that the agent feels psychological (positive) utility when his wage is higher than the colleague's. We assume below that $-1 < \lambda < 1$.

Let us obtain the optimum wage in this case. The incentive compatibility constraint of the agent in this case is as follows:

$$\begin{aligned} p_1 w - d - \alpha w(1 - p_1)p_1 - \alpha \lambda w p_1(1 - p_1) \\ \geq p_0 w - \alpha w(1 - p_0)p_1 - \alpha \lambda w p_0(1 - p_1). \end{aligned} \quad (\text{ICS})$$

The third term on the left side of (ICS) indicates the psychological disutility the agent is subjected to when the agent fails at the probability of $1 - p_1$ and receives no wage and the colleague succeeds at the probability of p_1 and receives wage w . Similarly, the fourth term indicates the psychological utility that arises when the agent succeeds at the probability of p_1 and receives wage w and the colleague fails at the probability of $1 - p_1$ and receives no wage.⁷ The same applies to the right side of (ICS), but in this case the agent's probability for success is p_0 , because his effort level is low.

On the other hand, the agent's individual rationality constraint in this case is as follows:

$$p_1 w - d - \alpha w(1 - p_1)p_1 - \alpha \lambda w p_1(1 - p_1) \geq 0, \quad (\text{IRS})$$

(ICS) and (IRS) may be rewritten as follows:

⁷ Note that the model is based on the assumption that the colleague has selected a high effort level. In cases where social preferences are considered, the condition of $w_f = 0$ needs to be checked, but it is omitted here.

$$w \geq \frac{d}{\Delta_p} \cdot \frac{1}{1+\alpha\{p_1-\lambda(1-p_1)\}}, \quad (\text{ICS}')$$

$$w \geq \frac{d}{p_1} \cdot \frac{1}{1-\alpha(1-p_1)(1+\lambda)}. \quad (\text{IRS}')$$

The question here is the relation between the right sides of (ICS') and (IRS'). When α is smaller than a certain value, the right side of (ICS') is greater than the right side of (IRS').⁸ Here, as in the case where social preferences are not considered, we focus on the case where the incentive compatibility constraint is effective. In this case, the optimum wage, which is determined where (ICS') takes a sign of equality, is as follows:

$$w^{**} = \frac{d}{\Delta_p} \cdot \frac{1}{1+\alpha\{p_1-\lambda(1-p_1)\}}$$

When w^{**} and w^* are compared, the relation of their magnitude depends on the values of p_1 and λ . If $\lambda < p_1/(1-p_1)$, then $w^{**} < w^*$. In other words, if we were to consider social preferences, the performance-based wage should be lowered instead to weaken the incentive. In particular, this relation always holds when $\lambda < 0$ or $1/2 < p_1$. When $\lambda < 0$, in other words, when the agent is averse to his own wage being smaller than that of his colleague's and rather prefers that his own wage be higher, it is better to lower the performance-based wage. This is because the preference of this type of agent to succeed regardless of the results of his colleague directs the agent to choose a high effort level. Moreover, if the principal were to consider this point, the agent would select a high effort level even when the wage provided as an incentive is low. When $1/2 < p_1$, in other words, when there is a high probability of a project's success as long as an effort is made, there is an increased possibility that if the agent were to choose a low effort level he alone would fail and receive no wage even though his colleague would succeed and receive a wage. In this case, when $\lambda < 0$, the agent will obviously be averse to his own wage being smaller than his colleague's, and even when $0 < \lambda$, the agent will be averse to any wage gap between himself and his colleague. Therefore, even if a wage provided as an incentive were low, the agent would select a high effort level. Consequently, in the case where $\lambda < p_1/(1-p_1)$, it is preferable that the wage gap between those with a high output and those with a less than high output be narrowed, and that

⁸ To be precise, $\frac{1}{(1-p_1)(1+\lambda)+l-1} \geq \alpha$ where $l \equiv p_1/p_0 > 1$.

the tendency towards performance-based wages should be decreased.

On the other hand, if $p_1 / (1 - p_1) < \lambda$, then $w^{**} > w^*$. In other words, if we were to consider social preferences, a wage that is strongly dependent on performance is favorable. For this condition to hold, it must at least be that $p_1 < 1/2$. This corresponds to a situation where even at a high effort level, the probability of a project's success is low. Similarly, for this case to hold, it must obviously be that $0 < \lambda$ (inequity aversion). In this case, there are two opposing incentives; namely, an incentive for the agent to make an effort so as to avoid a situation where a colleague receives a wage but the agent does not, and another incentive for the agent not to make an effort so as to avoid a situation where the agent alone receives a wage. However, the more the agent is averse to being the only one to receive a wage, the stronger the latter incentive will become. Moreover, since the probability of the project's success is very low to begin with, there is a high probability that the colleague would fail and receive no wage. Also, as long as $0 < \lambda$ the agent can lessen the psychological disutility by failing in the project. To give an incentive for this type of an agent to select a high effort level, the wage must be further increased when the project is successful. These results suggest that when there is a preference for uniformity among workers, the tendency towards performance-based wages should be strengthened and the wage gap widened in order to steer a difficult project towards success.

3. Motivation Crowding-out

Recently we often hear of discussions about motivation crowding-out as one of the criticisms of performance-based wage systems. These discussions are based on findings from psychological research by Deci (1975) and others which state that external wages undermine the intrinsic motivation of workers and thus cause their performance to decline.⁹ For example, when a performance-based wage is suddenly introduced and paid to a worker who finds joy in the very act of performing the work he is assigned, the worker feels as if he is working for money, and as a result his job satisfaction and motivation declines.¹⁰

This point has been confirmed through a number of economic experiments.

⁹ As far as I know, Frey (1997) was the first to use the term "crowding-out" in this context.

¹⁰ See Frey (1997), and Frey and Jegen (2001) for other discussions on motivation crowding-out.

For example, Fehr and Gächter (2002) and Irlenbusch and Sliwka (2005) obtained results from experiments showing that when performance-based wage systems are introduced, the effort levels of agents decline and their efficiency is inhibited in comparison with fixed wage systems. Gneezy and Rustichini (2000) confirm that there is no monotonous relation between monetary incentives for performance and their outcomes. In other words, if performance-based wages have already been introduced, the performance of workers would be enhanced the more the incentives are increased. However, if performance-based wages are introduced and paid to workers who were not paid such wages before, their performance would rather decline. Like Gneezy and Rustichini (2000), Pokorny (2004) obtained results which indicate that incentives and performance are in a non-monotonous relation. These results also showed that when incentives are introduced and then gradually increased, performance at first begins to rise until it reaches a certain point after which it starts to decline, forming a shape of an inverted U.

Frey (1997) calls this characteristic of wages as the “hidden cost of reward.” In this section, we examine this point using his model. An agent’s utility may be expressed as $U = B - C$, where B is the benefit and C is the cost. If both B and C are dependent on output x and incentive w provided by the principal, they can be expressed as $B(x, w)$ and $C(x, w)$, respectively. With respect to the output, B is a concave function and C is a convex function.¹¹ In this case, the agent’s optimum output x^* is determined at a point that fulfills $B_x = C_x$. Moreover, when x^* is dependent on the incentive provided by the principal, we can express its effect dx^*/dw , by partial differentiation of $B_x = C_x$ by w , as follows:

$$\frac{dx^*}{dw} = \frac{B_{xw} - C_{xw}}{C_{xx} - B_{xx}}. \quad (2)$$

The denominator on the right side of (2) is positive. On the other hand, in terms of the numerator, C_{xw} expresses the price effect, and B_{xw} expresses the crowding effect. In other words, here the crowding effect is the effect of the incentive on the marginal benefit. According to Frey (1997), there is a crowding-in effect when $B_{xw} > 0$ and a crowding-out effect when $B_{xw} < 0$. If we

¹¹ In other words, we assume $B_x > 0$, $B_{xx} < 0$, $C_x > 0$, $C_{xx} > 0$. B_i is B ’s partial derivative regarding i , and B_{ij} is B ’s second order partial derivative regarding j , where $i, j = x, w$. The same applies to C .

assume that $C_{xw} = 0$ for the sake of simplification, then it would be $dx^*/dw < 0$ in the case of crowding out. This result indicates that by increasing the wage as an incentive, the agent's optimum output level rather declines. Moreover, Frey (2004) explains that when $dx^*/dw < 0$, the optimum wage should be lowered. This shows that when wages provided as an incentive have a negative effect on the agent's marginal benefit, there is a probability that by increasing the incentive the agent's output may decline. So if this is taken into consideration it can be interpreted that it is desirable for the principal to lessen the incentive. This points to the possibility of an increased incentive from performance-based wages to degrade performance by undermining intrinsic motivation, and can be interpreted as supporting criticisms made against performance-based wage systems from this viewpoint. Obviously, the opposite is true when $B_{xw} > 0$. It can be understood from the result that when wages provided as an incentive are fulfilling their expected function of motivating the agent to action, the incentive works to improve the agent's output.¹²

Like Frey (1997), Grepperud and Pederson (2006) analyze the crowding-out effect by using a linear contract in the principal-agent model. Here, a linear contract is defined as a wage contract expressed as a sum of a fixed wage and a performance-dependent piece rate wage. For example, a linear contract can be expressed as $w = \gamma + \beta x$, where γ is the fixed wage and β is the measure to which the wage is dependent on the output (incentive intensity). In many contract theory models that were based on the linear contract, the focus of analysis was on the trade-off between incentive and risk. In other words, the argument was that when the process is not evaluated (or cannot be evaluated) and the resultant performance-based wages are considered as having a risk, performance-dependent contracts are not favorable for those agents who are averse to such a risk.¹³ In contrast, Grepperud and Pederson (2006) see their model as analyzing not the trade-off between incentive and risk, but rather the trade-off between incentive and motivation. Details about the model are

¹² On criticisms against performance-based pay from the viewpoint of the crowding-out effect, see Takahashi (2004). On problems related to those criticisms, see Nakamura (2006).

¹³ There are recent studies that doubt the negative correlation between incentive and risk. For example, Prendergast (2002) analyzes cases where there is a positive correlation between the two. Ishiguro (2005) briefly summarizes the theoretical models of these two cases.

omitted in this paper as the essential points are the same as Frey's (1997). Grepperud and Pederson (2006) consider a case of crowding out as one where the agent's cost function is dependent not only on his chosen effort level, but also on the incentive β set by the principal, and as β increases the cost as well as the marginal cost related to the effort level increase. They show that in this case the response of the optimum effort level to the incentive becomes weaker and the optimum incentive also becomes weaker, in comparison with a standard model.

On the other hand, the above results are strongly dependent on the direct assumptions of crowding out.¹⁴ In contrast, Sliwka (2003) considers a model in which the motivation crowding-out effect is analyzed endogenously. In this model, two types of agents are assumed. One is the type found in the standard principal-agent model, and the other is the "reliable agents" who are not opportunistic and always choose the effort level instructed by the principal. Obviously, the principal will only need to offer fixed wages to the reliable agents. In this case, if the fraction of reliable agents is high the optimum incentive would be weaker than in a standard case. Also, when the cost for assessing the output is large, fixed wages are more favorable than performance-based wages. Moreover, Sliwka (2003) discusses a case in which an agent's gain is influenced not only by the gain of other agents, but also by beliefs about their reliability. In this case, it is indicated that a strong incentive will function as a signal that other agents are not reliable types and as a result the agent will lower his effort level.

4. Delegation of Responsibilities and Assessment of Outputs

We understand from the above discussions that in cases where psychological factors are considered, favorable conditions arise when incentives different from those in a standard case are provided. The optimum incentive, however, varies depending on the workings of psychological factors and on the extent of their effect. In this respect, theoretically it may not be too surprising that there are many different views as to the estimation of performance-based wage systems. On the other hand, empirical research has indicated a number of preconditions for the effective functioning of performance-based wage systems.

¹⁴ On this point, Grepperud and Pederson (2006) concede in the footnote that their model on intrinsic motivation and crowding out is elementary and further research is needed.

For example, Ohtake and Karato (2003) mention the conditions that are common to white-collar and blue-collar workers alike, namely, (1) work sharing and clarification of roles, (2) granting responsibility for work, and (3) opportunities for development of competencies.¹⁵ As regards to white-collar workers, they also point out that granting greater discretion in work is an important factor.

In this section, we therefore focus particularly on the delegation of responsibilities and introduce Daido (2006), who analyzes types of contracts by considering the psychological utility that arises from the possibility of delegation. In this case, we consider a situation where a risk-neutral agent will carry out a project, and a principal will provide either fixed wage contracts (FWC) or pay-for-performance contracts (PPC). In the case of PPC, the principal will decide whether or not to delegate responsibilities. Here, we consider that responsibilities are not delegated when the principal gives instructions on how to execute the project, and that responsibilities are delegated when the agent has to find his own way of executing the project. In this case, it is probable that even if the agent's effort level is the same, the project's outcome will differ depending on whether or not responsibilities are delegated, in other words, depending on how the project is executed. If we consider this point, in addition to normal expected utility there is also the possibility that utility based on psychological factors may arise in the agent. Here, we note the possibility that psychological utility may arise respectively in cases where responsibilities are not delegated and in cases where responsibilities are delegated.

We first consider the case where responsibilities are not delegated. In this case, the psychological utility of the agent originates in resigning himself to a low level of outcome because responsibilities are not delegated. On the other hand, if he had been delegated responsibilities he would have found an efficient way of executing the project that would have realized a high level of outcome, and because of that he would have received a high wage. In other words, the gap between the wage commensurate with the outcome that would have been achieved had he been delegated responsibilities and the wage given to him when responsibilities are not delegated causes psychological utility (or disutility).¹⁶

¹⁵ Prior to this, Genda, Kambayashi and Shinozaki (2001) made a similar remark.

¹⁶ There is the probability that even when the agent is delegated responsibilities, he may not be able to find an effective way to execute the project; however, such a probability is excluded here. In other words, in the sense that the agent we deal with here has a better way of executing the project, he is a relatively competent agent.

Secondly, we think of psychological utility when responsibilities are delegated. In this case, there is the possibility that the agent will find a more efficient way of executing the project and attain a high level of outcome; however, the question is how the principal will assess the increased outcomes. If the principal fails to fully recognize the outcomes that have been enhanced because responsibilities were delegated, the agent may feel psychological disutility as a result of the underestimation. We compare each case of PPC taking into account the delegation and non-delegation of responsibilities and psychological utility accompanying that to FWC, and assess performance-based wages through the model analysis.

We now briefly describe the model. Assume there is a risk-neutral principal and agent, and the agent will engage in a project. The project's output x is dependent on the agent's effort level $e \geq 0$, and we assume that $x=e$. The output is verifiable. The principal can choose either a fixed wage contract (FWC) or a pay-for-performance contract (PPC). If the principal chooses PPC, she will then choose either to give instructions on how to execute the work (non-delegation of responsibilities) or leave to the agent the task of finding a way to execute it (delegation of responsibilities). Each type of contract is considered as follows. With regard to FWC, we assume that the agent will choose the effort level that the principal instructs him with and at fixed wage w .¹⁷ With regard to PPC, we assume the linear contract $w(x) = \gamma + \beta x$.

Regarding the delegation of responsibilities, we consider the following. Let μe express the project's outcome when the agent is allowed to choose how to execute it. When $\mu > 1$ (or $\mu < 1$), delegating responsibilities to the agent will yield higher (or lower) output. We look at the case of $\mu > 1$. In this case, the abovementioned psychological utility associated with the possibility of delegation of responsibilities is considered as follows. We first consider a case where responsibilities are not delegated. The source of psychological utility in such a case is the gap between the wage based on outcomes that would have been realized had responsibilities been delegated and the wage based on

¹⁷ For such a contract to be possible, we need to think of the possibility of monitoring the agent's effort level and the possibility of monitoring itself. However, we ignore such considerations here, and assume the agent is a "reliable agent" described by Sliwka (2003) above.

outcomes that are realized without the delegation of responsibilities. When $\mu > 1$, if the former is larger than the latter then the agent will have a negative psychological utility. The agent's expected utility, when the psychological disutility arising from the non-delegation of responsibilities is considered, is as follows:

$$U = \gamma + \beta e - c(e) - \alpha(\mu - 1)\beta e,$$

where $c(e)$ is the cost of the agent's effort. The last term expresses the psychological disutility resulting from the gap between the wage dependent on outcomes that would have been achieved had responsibilities been delegated and the actual wage.¹⁸

Next we formulate psychological utility in the case where responsibilities are delegated. In this case, output $x = \mu e$ is achieved. This, however, is the output gained based on a method that the agent chose in executing the project, and how the output is evaluated will depend on the principal. Therefore, the measure of the principal's evaluation of this output is expressed as θ (≥ 1). When $\mu > \theta$, the principal underestimates the agent's output, and the agent is likely to be dissatisfied with the fact that even though he attained a high level of output it was not fully recognized. On the other hand, when $\mu < \theta$, the principal overestimates the output. We consider only the case of $\mu > \theta$ below. It can be interpreted here that the larger the value of θ , the more likely that the principal is making a fairer assessment. In this case, the agent's expected utility is expressed as follows:

$$U = \gamma + \beta\theta e - c(e) - \alpha(\mu - \theta)\beta e.$$

The principal's expected profit is obtained by subtracting the expected value of the agent's wage from the expected value of the outcome. In other words, it is $e - w$ in FWC, $e - (\gamma + \beta e)$ in non-delegated PPC, and $\mu e - (\gamma + \beta\theta e)$ in delegated PPC.

Let us now look at the results obtained from the above model. When we compare the case of FWC where the agent's effort level can be monitored and such an effort level can be forced on the agent, and the case of PPC where psychological utility is not considered, the agent's optimum effort level and the

¹⁸ Here, we basically follow the formulation of (1), and $\alpha > 0$ expresses the degree of psychological utility. Moreover, if we follow (1), obviously we could consider the case of $\mu < 1$, but it will be omitted here.

principal's expected profit will correspond as long as the agent is risk-neutral. In other words, in such an environment the principal's expected profit will be the same whether she chooses FWC or PPC.¹⁹

Let us now compare FWC and non-delegated PPC. First, the optimum effort level is higher in FWC. In other words, when we consider psychological utility, introducing PPC as an incentive lowers the effort level. This can be understood as a motivation crowding-out effect in the sense that the incentive will have a negative effect on the effort level.²⁰ The principal's expected profit will also be higher in FWC. Thus, even though non-delegated PPC and FWC are of the same value when psychological utility is not considered, FWC becomes more favorable for the principal than non-delegated PPC when psychological utility is taken into account. This shows that introduction of a performance-based wage system that is not accompanied by the delegation of responsibilities may bring about an unfavorable outcome for the principal.

Lastly, we compare delegated PPC with the last two cases. In this case, the principal will have to make more payments to the agent when her assessment of the output is fairer (when θ is higher), but, at the same time the agent's psychological disutility can be reduced; as a result, it will induce the agent to work at a high effort level. Moreover, the higher θ becomes the greater the latter's effect; therefore the principal's expected profit will be an increasing function of θ . As a result, as the value of θ becomes larger delegated PPC will become more favorable for the principal than non-delegated PPC or FWC. We can understand from these results that a performance-based wage system will function as a favorable system for firms that introduce it when it is accompanied by the delegation of responsibilities and when the outputs are assessed fairly.

5. Conclusion

In this paper, we analyzed principal-agent models in which psychological

¹⁹ As already mentioned, we ignore the cost of monitoring the effort level and output. Pendergast (2002) assumes that the cost of monitoring output is larger than the cost of monitoring effort levels. On the other hand, Barth, et.al. (2006) consider the cost of monitoring effort levels in FWC. If we consider these points, we can probably say that it is not very clear which monitoring cost is the larger of the two.

²⁰ It should be noted that the crowding-out effect here means that the effort level declines only in comparison between FWC and PPC. It does not mean that when the incentive intensity rises in PPC, the effort level declines.

aspects were incorporated into the agents' utility, and examined the effects psychological utility has on incentives through comparison with results obtained from standard models. From these results, we discussed how the value of performance-based wage systems could be interpreted from the point of view of behavioral economics. Using simple models, we analyzed whether a wage gap should be widened or not when agents have social preferences, in accordance with the degree of those preferences and with the difficulty of a project. We also analyzed that high wages do not necessarily raise the level of the performance of agents and that if performance-based wage systems were to be introduced, they need to be accompanied by delegation of responsibilities.

Against the backdrop of the rapid development of behavioral economics, there are other interesting models related to the topics dealt with in this paper. For example, in relation to the model of social preferences in Section 2, Itoh (2004), as mentioned in the footnote, analyzes a case where the wages of two agents are determined based on each other's outputs and shows what the optimum contracts would be in light of the nature of social preferences. Neilson and Stowe (2004) make a similar analysis in a case where a linear contract is proposed to an agent with social preferences. As a reference related to the crowding-out of intrinsic motivation in Section 3, Bénabou and Tirole (2003) show that when the principal personally has information about the difficulty of a project or about the agent's ability, the wage the principal offers to the agent will function as a signal on the agent's ability, and a high wage may decrease the agent's motivation.

As we have seen above, much research has been done based on findings in the field of psychology to build new models that will further expand past economic theories and allow researchers to deal with a wider range of topics. A somewhat more unified model that can more comprehensively explain these research results may become necessary in the future. It is hoped that such a new model would give meaningful new interpretations of the analysis of performance-based wage systems, which was examined in this paper, and of other important topics in labor economics.

References

- Barth, Erling, Bernt Bratsberg, Torbjörn Hægeland, and Oddbjørn Raaum. 2006. Who pays for performance? IZA Discussion Paper no. 2142.
- Bénabou, Roland, and Jean Tirole. 2003. Intrinsic and extrinsic motivation. *Review of*

- Economic Studies* 70:489-520.
- Daido, Kohei. 2006. Incentives, delegation, and performance evaluation. Mimeo.
- Fehr, Ernst, and Klaus M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, no.3:817-68.
- . 2003. Theories of fairness and reciprocity: Evidence and economic applications. In *Advances in economics and econometrics: Theory and applications, Eighth World Congress, Volume 1*, ed. Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, 208-57. NY: Cambridge University Press.
- Fehr, Ernst, and Simon Gächter. 2002. Do incentive contracts crowd out voluntary cooperation? *Institute for Empirical Research in Economics*. University of Zurich, Working Paper no. 34.
- Frey, Bruno S. 1997. *Not just for the money: An economic theory of personal motivation*. Cheltenham, UK and Brookfield, VT: Edward Elgar Publishing.
- Frey, Bruno S., and Reto Jegen. 2001. Motivation crowding theory. *Journal of Economic Surveys* 15, no.5:589-611.
- Genda, Yuji, Ryo Kambayashi, and Takehisa Shinozaki. 2001. Seikashugi to noryoku kaihatsu: Kekka toshiten rodo iyoku [Performance-based pay and development of competencies: Incentives as a result]. *Soshiki Kagaku* (Organizational Science) 34, no.3:18-31.
- Gneezy, Uri, and Aldo Rustichini. 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics* 115, no.3:791-810.
- Grepperud, Sverre, and Pål Andreas Pederson. 2006. Crowding effects and work ethics. *LABOUR* 20, no.1:125-38.
- Irlenbusch, Bernd, and Dirk Sliwka. 2005. Incentives, decision frames, and motivation crowding out - An experimental investigation. IZA Discussion Paper no. 1758.
- Ishiguro, Shingo. 2005. Shokumu hatsumei no keizai bunseki: Keiyaku rironteki sekkin [Economic analysis of employee's invention: An approach from contract theory]. *Nippon Rodo Kenkyu Zasshi* (The Japanese Journal of Labour Studies), no.541:24-33.
- Itoh, Hideshi. 2003. *Keiyaku no keizai riron* [Economic theory of contracts]. Tokyo: Yuhikaku.
- . 2004. Moral hazard and other-regarding preferences. *Japanese Economic Review* 55, no.1:18-45.
- Jo, Shigeyuki. 2004. *Uchigawa kara mita Fujitsu "Seikashugi" no hokai* [Insider's view of the collapse of Fujitsu's "performance-based pay system"]. Tokyo: Kobunsha.

- Nakamura, Keisuke. 2006. *Seikashugi no shinjitsu* [The truth about performance-based pay]. Tokyo: Toyo Keizai Shinposha.
- Neilson, William S. and Jill Stowe. 2004. Piece-rate contracts for other-regarding workers. Mimeo.
- Ohtake, Fumio, and Koji Karato. 2003. Seikashugi chingin seido to rodo iyoku [Performance-based wage systems and incentives]. *Keizai Kenkyu* (The Economic Review) 54, no.3:1-20.
- Okunishi, Yoshio. 2001. “Seikashugi” chingin donyu no joken [Conditions for introducing “Performance-based” wage]. *Soshiki Kagaku* (Organizational Science) 34, no.3:6-17.
- Pokorny, Kathrin. 2004. “Pay-but don’t pay too much: An experimental study on the impact of incentives. Mimeo.
- Prendergast, Canice. 2002. The tenuous tradeoff between risk and incentives. *Journal of Political Economy* 110, no.5:1071-102.
- Sliwka, Dirk. 2003. On the hidden costs of incentive schemes. IZA Discussion Paper no. 844.
- Takahashi, Nobuo. 2004. *Kyomo no seikashugi: Nippon gata nenkosei fukkatsu no susume* [Falseness of performance-based pay: Proposal on the revival of the Japanese-style seniority system]. Tokyo: Nikkei Business Publications.