International Journarl of Mangement Science
Vol 4, No 2, November 1998

RESPONSE TIME, INCENTIVE SYSTEM, AND LONG-TERM RELATIONSHIP*

Hosun Rhim
College of Business and Economics, Hanyang Univeristy
Seoul, 133-791, Korea

(Received August 1998)

ABSTRACT

This paper presents an incentive system to reduce response time from a supplier. The incentive system is expressed as a contract between an assembler and a supplier who have a long-term relationship. We produce the optimal payment scheme and expected total cost, when the assembler is farsighted. We show that the farsighted assembler obtains higher effort level from the supplier than the myopic assembler. We also show that the expected total cost of the farsighted assembler is smaller in the long run, although it is initially higher than that of the myopic assembler.

1. INTRODUCTION

In this paper, we present an incentive system of a farsighted assembler who wants to reduce response time from a myopic supplier under a long-term relationship. The response time is defined as the time from which order is placed to the time when the part is arrived. We assume that response time is the only performance measure for the incentive system. The assembler is assumed to play a role of a supply chain leader, as it offers an incentive contract to the supplier. Although various types of incentives are being used in practice, we focus only on payment schemes. By this, the long-term relationship between the assembler and the supplier is addressed.

In the early 1980s, some leading companies introduced a new dimension of competition: time-based competition. Time-based competition is a new philosophy that seeks closeness to the customer by delivering high-valued products and

* The research of the author was supported by research grants(1998) of Hanyang University

services in the least amount of time [4]. The goal of the time-based competitor is to eliminate non-value-adding time[2] continuously so that the system response time of the whole supply chain is reduced, thereby bringing the customer closer.

In time-based competition, response time is not a given parameter anymore, but it is treated as a variable of strategic importance. However, in most mathematical models of production problems, response time has been regarded as a given system parameter. Because of the growing strategic value of response time, researchers have started to examine the relationships between the response time and other factors such as the batch size, set-up time and incentives.

The work of Karmarkar [7] is one of the first papers that address this issue. In the paper, he presents the relationship between response time and the batch size, using a single-stage queueing model. Bitran and Tirupati [3] refine Karmarkar's model by adding a simple job release mechanism based on the continuous review policy of (R,Q) type. Kim and Tang [9] provide another extension of Karmarkar's model, in which they consider the Kanban system as a job release mechanism.

Not much research effort has been dedicated to the issue of incentive schemes regarding the response time, while physical factors such as batch size have been considered. To our knowledge, Tang [12] first addressed this issue in the context of a production problem. He considers the case in which the supplier offers a contract that includes his order response time as a variable. However, our model substantially differs from Tang's in the following ways. First, the contract-offering subject is different. In his model, the supplier offers the contract, while, in our model, the assembler does. Second, a set of alternative contracts is evaluated in Tang's model. We, however, design an optimal contract incentive itself. Third, Tang's model is still a single-person, single-period decision making model, while our model is a two-person, two-period model based on game theory. Fourth, Tang's model is based on a specific job-release mechanism that corresponds to the linear control rule, while our model is general enough to be applied to any system configuration.

Our model draws on the works of Holmström [6] and Grossman and Hart [5]. Their modeling approach has been developed by several researchers and extensively applied to various areas including accounting, finance, industrial organization and marketing (see Ackere [1]). However, this modeling approach has not been applied to the issue of response time reduction between supplier and assembler.

We construct a two-period model to explain a long-term relationship of time-

1) According to Stalk [11], in a traditional manufacturing system, value is added to products for only 0.05% to 2.5% of the time that they are in the factory.

based competitors with their suppliers, as continuous improvement of time-based competitors cannot be achieved without a long-term relationship. The first and second periods represent the present and the future, respectively. The assembler of the two period model is farsighted in the sense that she considers the effect of current contract in the future. However, we suppose that suppliers behave myopically. This is not unusual especially when time-based competitors transplant their system into the other countries (see Kenney and Florida [8]). To observe the effects of the long-term relationship to the incentive contract, we compare the contract of the farsighted assembler with that of the myopic assembler who changes its supplier in each period.

Although the assembler can hire more than one supplier for each part, we assume that only one supplier is hired and that the supplier already has the ability to reduce response time. This assumption is realistic considering the following two aspects. First, if the benefit to the assembler from the response time reduction is independent of the other suppliers, the result can be directly applied to the multiple supplier case. Second, intensified competitive pressures have forced assemblers to rationalize their supplier structure and to reduce the number of suppliers. Womack et al.[13] observed that the mass producers were trying to cut the number of suppliers to each assembly plant to between 350 and 500 and had largely reached this goal. Similarly, we assume that a proper supplier was selected through this rationalization process.

The remainder of the paper is structured as follows. In section 2, we build a two-period model. In section 3, the farsighted assembler is compared with the myopic one. Section 4 concludes.

2. PROBLEM FORMULATION

In this section, we build a two-period model to address the long-term relationship between an assembler and a supplier. Each period consists of three stages of actions; in the first stage, the assembler presents the supplier with a contract that describes the payment scheme; in the second stage, the supplier performs at some effort level; in the final stage, the assembler compensates the supplier according to the realized response time. In the first period, the assembler offers an incentive scheme considering the future (the second period), and updates the incentive scheme after observing the supplier's first period performance (response time).

We assume that the assembler (or the principal) observes the supplier's (or the agent's) response time, but cannot infer the supplier's effort level. Response time is a random variable whose distribution depends on the effort level. Supplier's effort is defined as any activity that reduces response time (for example, investment in new equipment, R&D expenditures, organization changes, and so on). The assembler demands as much effort from the supplier as possible, since this can result in reduced response time. However, the supplier needs more compensation since more effort incurs more disutility to the supplier. While supplier's effort level affects the welfare of both parties, the assembler has a function of prescribing payment rules; that is, before the supplier chooses an action, the assembler sets a rule that specifies the compensation to the supplier as a function of supplier's response time. This phenomenon is quite common in practice, since the firm size of the suppliers is typically much smaller than that of the assembler who employs thousands of workers.

Notation and technical assumptions of the model are as follows:

l_t : response time of the supplier at period t, $t = 1,2$;

s_t : payment to the supplier or equivalently income of the supplier at period t, $t = 1,2$;

$c(l_t)$: total cost of the assembler, which is a function of response time l_t;

e_t : effort level of the supplier at period t, $t = 1,2$;

$U(s_t)$: utility of the supplier from income s_t;

$V(e_t)$: disutility of the supplier from effort e_t;

m : reservation level or minimum utility level of the supplier, $m>0$.

a) *In the first period, response time l_1 has only two values: αL and L where α is given on $(0,1)$. In the second period, response time l_2 has values of αl_1 and l_1.* We assume that within a short period it has only two representative values: improved and unimproved response time. (See figure 1.) We assume that the response time, once reduced, does not increase in the next period; i.e. once a process for response time reduction is installed in a system, any realized improvements are fully absorbed into the system so that the response time does not increase, even when no effort is exerted.

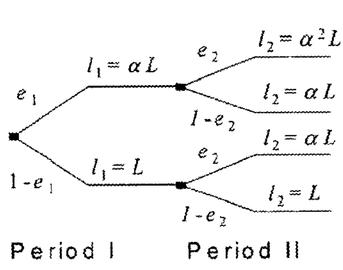

Figure 1. State-space of the two-period model

b) *We assume that* $P\{l_t = \alpha \cdot l_{t-1} \mid \text{given } l_{t-1}, \text{effort level} = e_t\} = e_t$ *and* $P\{l_t = l_{t-1} \mid \text{given } l_{t-1}, \text{effort level} = e_t\} = 1 - e_t$, *where* $l_0 = L$, $e_t \in [0,1]$ and $t = 1,2$. Response time is influenced by the effort devoted by the supplier and other uncertain factors. The probability of improved response time increases as effort level increases, while that of unimproved response time decreases.

c) *The order quantity is fixed and supplier's income* s_t *is a function of response time* l_t*: i.e.* $s_t = s_t(l_t)$. Since order quantity is fixed, it does not affect the supplier's income. Income s_t cannot be a function of effort e_t since the assembler cannot observe the effort.

d) *Supplier's utility function for income* s_t *and effort level* e_t *is additively separable and hence can be expressed as* $U(s_t)-V(e_t)$, *where U is utility from income* s_t *and V is disutility from effort* e_t. The additively separable utility function is assumed for mathematical tractability. Supplier's effort is defined to be every activity that reduces the response time. Effort has a value to the assembler since it increases the likelihood of a favorable outcome. However, it is a disutility to the supplier.

e) *Supplier is risk averse and* $U(s_t) = 2s_t^{1/2}$; *that is,* $U(s_t) \geq 0$, $U'(s_t) > 0$, *and* $U''(s_t) < 0$. Supplier's income s_t is a random variable since s_t is a function of response time l_t. If the supplier were risk neutral, this problem would have a trivial solution: i.e., the supplier would bear all the risks. For simplicity, we use a power utility function whose relative risk aversion is constant for all s_t. Various styles of utility functions and resulting compensation schemes are investigated by Basu et al. [2].

f) *The marginal disutility for effort increases with effort so that we assume the following quadratic function:* $V(e_t) = \gamma e_t^2$, $\gamma > 0$. It implies that the resources become more valuable as invested amount increases. Again, we assume quadratic function for mathematical tractability.

g) *We assume a single supplier with a single part.* This is an assumption to model reduced number of suppliers in the leading companies. However, the model developed under this assumption can be applied to mutiple-supplier or multiple-part problems, if total cost incurred by each supplier or product is independent of each other.

h) *The cost of a part,* $c(l_t)$ *is a linear function of response time* l_t*: i.e.* $c(l_t) = c \cdot l_t$. *The total cost incurred from the part is the sum of assembly cost and payment to the supplier: i.e.* $c \cdot l_t + s_t(l_t)$. The linearity of the assembly cost is assumed for simplicity. Since we assume a single supplier with a single part, the total cost is the sum of assembly cost and payment for the part.

i) *The assembler's only objective is to minimize its expected total cost.* Minimizing expected total cost is equivalent to assuming that the firm is risk neutral. Reduced response time may result in price increase and in turn profit increase for the assembler; however we assume that these effects are already considered when estimating the cost coefficient c.

j) *The effort response function and utility function of the supplier are common knowledge; that is,* the supplier does not possess any private information other than the actual effort level.

k) *The assembler knows the minimum level of expected utility of the supplier and must guarantee at least the minimum level. The minimum level m is given exogenously.* If the assembler does not guarantee the minimum, the supplier will find other businesses and the supplier relation cannot be maintained. We need to note that the supplier's second-period minimum utility level does not increase as its first period response time decreases. Requiring a higher utility level based on the performance is not common in a partnership. At Toyota, for example, suppliers are not expected to commit themselves to delivering at unrealistically low prices from the beginning but must be prepared instead to lower their price continually over the life of the model [13].

Given the assumptions and problem description, we next formulate the problem. Following the standard procedure of backward induction, we start from the second period and then move on to the first period.

2.1 Second Period Model

Second period is used as a future time point. In this period, the assembler chooses the payment scheme $s^2(l^2)$ to minimize its expected cost. From assumption (b) and (h), the expected total cost of the assembler in the second period, ETC_2, can be expressed as:

$$ETC_2 \equiv \sum_{l_2 \in \{\alpha l_1, l_1\}} \left[c \cdot l_2 + s_2(l_2)\right] P(l_2 \mid l_1, e_2) \tag{1}$$

Since the assembler cannot observe the actual effort level of the supplier, the payment scheme is not a function of effort level, but of response time; however, an assembler can predict the effort level for any payment scheme by taking into account the following two conditions.

The first condition is often called *participation constraint* since it guarantees a utility at least equal to what the supplier could achieve in other activities such as supplying to other assemblers. Since the supplier's utility function is the difference between the utility from income $s_2(l_2)$ and disutility from effort e_2, the assembler's payment must satisfy the following condition:

$$\sum_{l_2 \in \{\alpha l_1, l_1\}} \left\{U\left[s_2(l_2)\right] - V(e_2)\right\} P(1_2 | l_1, e_2) \geq m \tag{2}$$

The second condition is called *incentive compatibility constraint.* It takes into account the behavior of a rational supplier who maximizes her utility function. If the assembler offers payment scheme $s_2(l_2)$ considering this condition, the supplier cannot improve her utility by deviating from the assembler's expectation. This condition is expressed as follows:

$$e_2^* = \arg\max_{e_2 \in [0,1]} \sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P(l_2 \mid l_1, e_2) - V(e_2) \tag{3}$$

Let $f(e_2) = \sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P(l_2|l_1, e_2) - V(e_2)$. Since $V(e_2)$ is convex and other terms are linear in e_2, $f(e_2)$ is concave so that there exists a unique optimum. The effort level e_2^* which satisfies the first-order condition of $f(e_2)$ maximizes the equation. Thus the assembler's problem can be rewritten as follows:

$$\min_{s_2(l_2), e_2} \sum_{l_2} [c \cdot l_2 + s_2(l_2)]P(l_2|l_1, e_2) \tag{4}$$

$$\text{subject to: } \sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P(l_2|l_1, e_2) - m - V(e_2) \geq 0 \tag{5}$$

$$\sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P_{e_2}(l_2|l_1, e_2) - V'(e_2) = 0 \tag{6}$$

where. $P_{e_2}(l_2|l_1, e_2) = \partial P(l_2|l_1, e_2)/\partial e_2$.

To produce optimal $s_2(l_2)$, we use the Kuhn-Tucker method [10]. Let λ_2 and μ_2 be the Lagrangean multipliers corresponding to Eq. (5) and (6), respectively. Then Lagrangean Lf_2 is:

$$\begin{aligned} Min\ Lf_2 = {} & \sum_{l_2} [c \cdot l_2 + s_2(l_2)]P(l_2|l_1, e_2) - \lambda_2 \Big[\sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P(l_2|l_1, e_2) - m - V(e_2)\Big] \\ & - \mu_2 \Big[\sum_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P_{e_2}(l_2|l_1, e_2) - V'(e_2)\Big] \end{aligned} \tag{7}$$

At the optimum, the first derivative of Lf_2 with respect to $s_2(l_2)$ for given l_2 is equal to zero, which yields the following lemma. (Proofs of lemmas, propositions, and corollaries are provided in the appendix.)

Lemma 1. For every response time l_2,

$$1/U'(s_2(l_2)) = \lambda_2 + \mu_2\, P_{e_2}(l_2|l_1, e_2)/P(l_2|l_1, e_2) \text{ for } e_2 \in [0,1],$$

where. $P_{e_2}(l_2|l_1, e_2) = \partial P(l_2|l_1, e_2)/\partial e_2$.

Lemma 1 and assumption (e) lead to derive the following payment scheme and utility of the supplier:

$$s_2(l_2) = \left(\lambda_2 + \mu_2 \frac{P_{e_2}(l_2|l_1,e_2)}{P(l_2|l_1,e_2)} \right)^2, \text{ and} \tag{8}$$

$$U(s_2(l_2)) = 2\left(\lambda_2 + \mu_2 \frac{P_{e_2}(l_2|l_1,e_2)}{P(l_2|l_1,e_2)} \right) \tag{9}$$

Lagrangean multiplier λ_2 may have any nonnegative value. However, we show that λ_2 always has a positive value and forces the participation constraint tight. This result allows us to obtain the expressions for λ_2 and μ_2, and therefore the payoff scheme $s_2(l_2)$. Also, we show that $\mu_2 > 0$.

Lemma 2. i) $\lambda_2 > 0$ and $\lambda_2 = (m + V(e_2))/2$

ii) $$s_2(l_2) = \begin{cases} \left[\dfrac{m + V(e_2)}{2} + \dfrac{V'(e_2)(1-e_2)}{2} \right]^2 & \text{if } l_2 = \alpha l_1 \\ \left[\dfrac{m + V(e_2)}{2} - \dfrac{V'(e_2)e_2}{2} \right]^2 & \text{if } l_2 = l_1 \end{cases} \tag{10}$$

iii) $\mu_2 > 0$

In this lemma, the payment scheme $s_2(l_2)$ consists of two parts: *base-salary* (λ_2^2) and *commission* $(s_2(l_2) - \lambda_2^2)$. Base-salary sets a payment level, and commission adjusts it based on response time. Positive base-salary requests that the assembler pay the supplier the minimum amount. Positive commission indicates that the optimal payment scheme deviates from the Pareto-optimal risk sharing. If the assembler could observe the supplier's effort level, she could use a forcing contract. Under this contract *(first-best solution)*, the supplier selects a proper action without the incentive-compatibility constraint. However, the payment scheme, $s_2(l_2)$ trades off some of the risk-sharing benefits coming from provision of incentives *(second-best solution)*.

With the results obtained up until now, we express the objective function of the model with given parameters and effort level e_2. The expected total cost of Eq. (4) consists of two parts: expected assembly cost and expected payment:

Expected assembly cost

$$= (c\alpha l_1)e_2 + (cl_1)(1-e_2) = cl_1[1-(1-\alpha)e_2], (e_2 \in [0,1]) \tag{11}$$

Expected payment

$$= s(\alpha l_1)e_2 + s(l_1)(1-e_2) = \frac{[m + V(e_2)]^2}{4} + \frac{V'(e_2)^2 e_2(1-e_2)}{4}, (e_2 \in [0,1]) \tag{12}$$

Eq. (11) and (12) yield the following expected total cost.

$$ETC_2(e_2|l_1) = cl_1[1-(1-\alpha)e_2] + \frac{[m+V(e_2)]^2}{4} + \frac{V'(e_2)^2 e_2(1-e_2)}{4}, (e_2 \in [0,1]) \quad (13)$$

The assembler sets the optimal effort level which minimizes expected total cost $ETC_2(e_2|l_1)$. Since Eq. (13) holds on the compact set of $e \in [0,1]$, there exists a minimum total cost. Since $V(e)$ is a quadratic function, we can show that $ETC_2(e_2|l_1)$ has a certain shape as in the following lemma.

Lemma 3. $ETC_2(e_2|l_1)$ is convex or convex-concave on $[0,1]$.

Lemma 3 implies that global optimum is obtained at boundary points or the left-most point on (0,1), where the first order condition holds. Let $e_2^0 = \min\{e_2 \mid \partial ETC_2 / \partial e_2 = 0$ and $e_2 \in (0,1)\}$. Then, optimal effort level $e_2^* = \operatorname{argmin}_{e_2 \in \{0, e_2^0, 1\}} \{ETC_2(e_2|l_1)\}$. If it is not unique, larger e_2 is selected because it is more desirable without any additional cost.

3.2 First-period model

In the first period, a farsighted assembler minimizes expected total costs, considering the second period costs. Thus, the objective function is expressed as follows:

$$\sum_{l_1 \in \{\alpha L, L\}} [cl_1 + s_1(l_1)] P(l_1|e_1) + \delta \sum_{l_1 \in \{\alpha L, L\}} \sum_{l_2 \in \{\alpha l, l\}} [cl_2 + s_2(l_2)] P(l_2|1_1, e_2) P(l_1|e_1) .$$

The second-period expected total costs conditional on l_1 are already obtained in the previous section. The expectation of these amount is discounted by the discount factor $\delta \in [0,1]$. If $\delta = 0$, the model is collapsed to a single period model of a myopic assembler. Since the supplier is assumed to be myopic, constraints are not different from those of the second period model. The formulation of the first period problem is:

$$\underset{s_1(l_1), e_1}{Min} \sum_{l_1 \in \{\alpha L, L\}} [cl_1 + s_1(l_1)] P(l_1|e_1) + \delta \sum_{l_1 \in \{\alpha L, L\}} \sum_{l_2} [cl_2 + s_2(l_2)] P(l_2|1_1, e_2) P(l_1|e_1) \quad (14)$$

Subject to: $$\sum_{l_1 \in \{\alpha L, L\}} U[s_1(l_1)] P(l_1|e_1) - m - V(e_1) \geq 0 \quad (15)$$

$$\sum_{l_1 \in \{\alpha L, L\}} U[s_1(l_1)] P_{l_1}(l_1|e_1) - V'(e_1) = 0 \quad (16)$$

To produce an optimal payment scheme $s_1(l_1)$, we follow the same procedure used in the second period. The Lagrangean Lf_1 of Eq. (14)-(16) is:

$$Min \;\; Lf_1 = \sum_{l_1} [cl_1 + s_1(l_1)] P(l_1|e_1) + \delta \sum_{l_1} \sum_{l_2} \left[cl_2 + s_2^*(l_2)\right] P(l_2|l_1, e_2^*) P(l_1|e_1)$$

$$- \lambda_1 \left\{ \sum_{l_1} U \left[s_1(l_1)\right] P(l_1|e_1) - m - V(e_1) \right\}$$

$$- \mu_1 \left\{ \sum_{l_1} U \left[s_1(l_1)\right] P_{e_1}(l_1|e_1) - V'(e_1) \right\}$$

Optimal $s_1(l_1)$ for a given l_1 satisfies the following three conditions:

i) $1/U'(s_1(l_1)) = \lambda_1 + \mu_1 \, P_{e_1}(l_1|e_1)/P\,(l_1|e_1)$ (17)

ii) $\sum_{l_1} U[s_1(l_1)] P(l_1|e_1) - m - V(e_1) = 0$ (18)

iii) $\sum_{l_1} U[s_1(l_1)] P_{e_1}(l_1|e_1) - V'(e_1) = 0$ (19)

Eq. (17)-(19) are equal to the conditions of the second-period model if e_1 is replaced by e_2. Thus the payment scheme $s_1(l_1)$ has the same form as $s_2(l_2)$:

$$s_1(l_1) = \begin{cases} \left[\dfrac{m + V(e_1)}{2} + \dfrac{V'(e_1)(1-e_1)}{2} \right]^2 & \text{if } l_1 = \alpha L \\ \left[\dfrac{m + V(e_1)}{2} - \dfrac{V'(e_1)e_1}{2} \right]^2 & \text{if } l_1 = L \end{cases}$$

Then, the 2-period expected total cost of the farsighted assembler $ETC_f\left(e_1, e_2^*\right)$ is expressed as a function of e_1:

$$ETC_f(e_1, e_2^*) = cL[1 - (1-\alpha)e_1] + \frac{[m + V(e_1)]^2}{4} + \frac{V'(e_1)^2 e_1 (1-e_1)}{4}$$
$$+ \delta\left[ETC_2(e_2^*|l_1 = \alpha L)e_1 + ETC_2(e_2^*|l_1 = L)(1-e_1)\right] \quad (20)$$

To obtain an optimal effort level e_1^*, we again need to find out the shape of Eq. (20). Since the terms related with the second-period are canceled out in the second derivative of $ETC_f(e_1, e_2^*)$, the result of Lemma 3 also holds for $ETC_f(e_1, e_2^*)$.

Lemma 4. $ETC_f(e_1, e_2^*)$ is a convex or convex-concave function. Let $e_1^0 = \min\{e_1 | \partial ETC_f(e_1, e_2^*)/\partial e_1 = 0$ and $e_1 \in (0,1)\}$. Then, $e_1^* = \text{argmin}_{e_1 \in \{0, e_1^0, 1\}} \left\{ETC_f(e_1, e_2^*)\right.$. The same tie-breaking rule as the second-period model holds.

3. MYOPIC VS. FARSIGHTED ASSEMBLERS

In this section, we investigate the effects of long-term relationship between the assembler and the supplier, using the model developed in the previous section. The long-term relationship is maintained by farsighted assemblers. We compare a farsighted assembler with a myopic one. In the traditional relationship of the myopic assembler, the supplier is replaced every period, and each period starts from the unimproved state again. Comparison is focused on the differences in optimal effort levels, costs, and payment schemes.

Additional notation is as follows:

j : index for assembler types; farsighted assembler f and myopic assembler s,

e_{tj} : effort level of assembler type $j \in \{f, s\}$ in period $t \in \{1, 2\}$,

$ETC_t(e_{tj})$: expected total cost in a single period $t \in \{1, 2\}$ for given e_{tj},

$$ETC_t(e_{tj}) = \sum_{l_{t-1}} ETC_t(e_{tj}|l_{t-1})P(l_{t-1}|e_{t-1}), \text{ where } P(L \mid e_0) = 1,$$

$ETC_j(e_{1j}, e_{2j})$: expected total cost of assembler type $j \in \{f, s\}$ for given (e_{1j}, e_{2j}),

$$ETC_j(e_{1j}, e_{2j}) = ETC_1(e_{1j}) + \delta\, ETC_2(e_{2j}),$$

First, we compare the optimal effort levels. The farsighted assembler anticipates cost reduction from shorter response time in the second period. In the following lemma, we show that shorter response time incurs less cost in the second period.

Lemma 5. $ETC_2\left(e_2^*(l_1)|l_1 = \alpha L\right) \le ETC_2\left(e_2^*(l_1)|l_1 = L\right)$.

To reduce the costs in the second period, the farsighted assembler requires more effort from the supplier than the myopic assembler as stated in Proposition 1.

Proposition 1. $e_{1f}^* \ge e_{1s}^*$ where e_{1j}^* is optimal effort level for type j assembler.

Different effort levels result in different cost savings. The myopic assembler minimizes her cost period by period, while the farsighted assembler does for the whole periods. Thus, an immediate corollary follows.

Corollary 1 : i) $ETC_1(e_{1f}^*) \ge ETC_1(e_{1s}^*)$.

ii) $ETC_1(e_{1f}^*) + \delta\, ETC_2(e_{2f}^*) \le ETC_1(e_{sf}^*) + \delta\, ETC_2(e_{sf}^*)$.

This corollary maintains that in the initial stage of improvement, the long-term relation based on time-based competition costs more to the assembler than the traditional relationship. The reduced cost by improvement, however, dominates the initial cost in the long run. So the assembler implementing long-term relationship should be more patient and willing to invest more resources in the initial stage.

Finally, we compare the optimal payment scheme of the farsighted assembler with that of the myopic assembler. As we noticed, the optimal payment schemes have an identical functional form. The only difference is in the optimal effort level. Thus, we investigate the effect of the effort levels on the payment scheme. To facilitate the explanation, we define $GAP(e_1)$, the difference of the commission. That is, $GAP(e_1) \equiv s_1(\alpha L) - s_1(L)$.

Proposition 2.

i) If $m \le \gamma$, $dGAP(e_1)/de_1 > 0$.

ii) Suppose $m \le \gamma$. Then, $dGAP(e_1)/de_1 > 0$, if $e_1 \in \left[0, \frac{\gamma + \sqrt{\gamma^2 + 3\gamma m}}{3\gamma}\right)$;

$dGAP(e_1)/de_1 \le 0$, otherwise.

iii) $\partial \lambda_1{}^2/\partial e_1 > 0$.

This proposition examines two components of the payment scheme: base-salary and commission. Part (i) maintains that the assembler increases *GAP* as he expects higher effort level from the supplier. However, this is true only when increasing the effort level is not so costly to the supplier. If increasing the effort level is relatively expensive as in Part (ii), the assembler starts to decrease *GAP* after a certain point. It is because the increase of the base-salary of Part (iii) cannot make up the disutility of the supplier any more after the point.

5. CONCLUSIONS

In this paper, we investigate a payment scheme of a farsighted assembler to reduce order response time from a myopic supplier. The major findings are as follows: first, the farsighted assembler requires higher effort level than the myopic one. Higher effort level in the present increases the possibility of response time reduction, which in turn reduces the costs in the future;

Second, the cost of the farsighted assembler in the initial stage of the improvement is higher than that of the myopic one. In the long run, however, the total discounted cost is smaller when the assembler is farsighted. This implies that improvement can not be accomplished without investment by the assembler in the initial stage;

Third, the incentive scheme of the farsighted assembler is different from that of the myopic one, although both incentive schemes consist of base salary and commission. The compensation gap that is tied to the response time is larger when the assembler is farsighted. However, this is true only if the disutility from the effort is not substantial to the supplier. Also, the base-salary of the incentive scheme needs to be increased with the increase of the effort level. The increased risk from the incentive scheme should be compensated by the base-salary adjustments.

APPENDIX

Proof of Lemma 1: The optimal payment scheme $s_2^*(l_2)$ satisfies $\partial L_2/\partial s_2 = 0$ for a given l_2.

i) Case 1: $e_2 \neq 0, 1$. When $l_2 = \alpha l_1$,

$$\frac{\partial L_2}{\partial s_2} = \frac{\partial}{\partial s_2}\left[(c\alpha l_1 + s(\alpha l_1))e_2 - \lambda_2 U(s_2)e_2 - \mu_2 U(s_2)\right] = 0, \text{ or}$$

$$\frac{1}{U'(s_2)} = \lambda_2 + \mu_2 \frac{1}{e_2}, \text{ or} \tag{21}$$

$$\frac{1}{U'(s_2)} = \lambda_2 + \mu_2 \frac{P_{e_2}(\alpha l_1|e_2)}{P(\alpha l_1|e_2)} \quad (e_2 \neq 0) \tag{22}$$

When $l_2 = l_1$, by letting $\partial L_2/\partial s_2 = 0$,

$$\frac{1}{U'(s_2)} = \lambda_2 + \mu_2 \frac{(-1)}{(1-e_2)} = \lambda_2 + \mu_2 \frac{P_{e_2}(l_1|e_2)}{P(l_1|e_2)} \quad (e_2 \neq 1) \tag{23}$$

ii) Case 2: $e_2 = 0$ or 1. If $e_2 = 1$, then $l_2 = \alpha l_1$ and Eq. (22) holds. Thus,

$$\frac{1}{U'(s_2(\alpha l_1))} = \lambda_1 + \mu_1 \tag{24}$$

If $e_2 = 0$, then $l_2 = l_1$ and Eq. (23) holds, and

$$\frac{1}{U'(s_2(l_1))} = \lambda_1 - \mu_1 \tag{25}$$

Since these are special cases of Eq. (22) and (23), this lemma holds for all $e_2 = [0,1]$.

Proof of Lemma 2 : i) Case 1: $e_2 \neq 0$, 1. Suppose $\lambda_2 = 0$. Then,

$$\frac{1}{U'(s_2(l_1))} = \begin{cases} \mu_2 e_2^{-1} & \text{if } l_2 = \alpha l_1 \\ -\mu_2(1-e_2)^{-1} & \text{if } l_2 = l_1 \end{cases} \tag{26}$$

Since $U'(s_2(l_2)) > 0$ from assumption (e), left-hand side of Eq. (26) is always positive; however, the right-hand side can have a negative value, which is a contradiction.

Case 2: $e_2 = 0$ or 1. Same contradiction is deduced from the supposition $\lambda_2 = 0$.

ii) $(\lambda_2 > 0)$ forces a tight participation constraint.

$$\sum\nolimits_{l_2 \in \{\alpha l_1, l_1\}} U[s_2(l_2)]P(l_2|l_1, e_2) - V(e_2) = m \tag{27}$$

This result also enables us to derive λ_2 with Eq. (9) as follows:

$$\lambda_2 = \{m + V(e_2)\}/2. \tag{28}$$

In order to obtain μ_2, we first consider the case that $e_2 \neq 0,1$. From Eq. (6) and (9),

$$\mu_2 = \frac{V'(e_2)}{2[e_2^{-1} + (1-e_2)^{-1}]} \quad (e_2 \neq 0,1) \tag{29}$$

Thus, if $e \neq 0,1$,

$$s_2^*(l_2) = \begin{cases} \left[\frac{m+V(e_2)}{2} + \frac{V'(e_2)(1-e_2)}{2}\right]^2 & \text{if } l_2 = \alpha l_1 \\ \left[\frac{m+V(e_2)}{2} - \frac{V'(e_2)e_2}{2}\right]^2 & \text{if } l_2 = l_1 \end{cases} \tag{30}$$

However Eq. (30) holds for two boundary points $e_2 = 0$, 1. Suppose $e_2 = 1$. Then, from Eq. (27),

$$2(\lambda_2 + \mu_2) = m + V(1). \tag{31}$$

Suppose $e_2 = 0$. Then, from the same equation,

$$2(\lambda_2 - \mu_2) = m + V(0) = m\,. \tag{32}$$

From Eq. (31) and (32),

$$\lambda_2 = [2m + V(1)]/4, \text{ and} \tag{33}$$

$$\mu_2 = V(1)/4 \tag{34}$$

Then, if $e \neq 0,1$,

$$s_2(l_2) = \begin{cases} [m + V(1)]^2/4 & \text{if } e_2 = 1 \\ m^2/4 & \text{if } e_2 = 0 \end{cases} \tag{35}$$

Eq. (35) is a special case of Eq. (30).

iii) Case 1: $e_2 \neq 0$, 1: From Eq (29), $\mu_2 > 0$.

Case 2: $e_2 = 0$ or 1: From Eq. (34), $\mu_2 > 0$.

Proof of Lemma 3: Since $V(e_2) = \gamma e_2^2$,

$$ETC_2(e_2) = cl_1[1-(1-\alpha)e_2] + \frac{[m+\gamma e_2^2]^2}{4} + (\gamma e_2)^2(e_2 - e_2^2) \tag{36}$$

$$\partial ETC_2(e_2)/\partial e_2 = -cl_1(1-\alpha) + m\gamma e_2 + 3\gamma e_2 - 3\gamma^2 e_2^3 \tag{37}$$

$$\partial^2 ETC_2(e_2)/\partial e_2^2 = m\gamma + 6\gamma^2 e_2 - 9\gamma^2 e_2^2 \tag{38}$$

Let $\partial^2 ETC_2(e_2)/\partial e_2^2 = 0$. Then, the solutions e_{2-}., e_{2+} are:

$$e_{2-} = \frac{\gamma - \sqrt{\gamma^2 + \gamma m}}{3\gamma} < 0, \quad e_{2+} = \frac{\gamma + \sqrt{\gamma^2 + \gamma m}}{3\gamma} < 0,$$

Since e_{2+} can have any positive value, $ETC_2(e_2)$ is convex or convex-concave on [0,1]

Proof of Lemma 5: We can prove this lemma by showing $dTEC_2(e_2 \mid l_1)/dl_1 \geq 0$ since $\alpha L \leq L$.

$$\frac{dETC_2(e_2|l_1)}{dl_1} + \frac{\partial ETC_2}{\partial l_1} + \frac{\partial ETC_2}{\partial e_2} \cdot \frac{\partial e_2}{\partial l_1}$$

$\partial ETC_2/\partial e_2 = 0$ for optimal e_2. From Eq. (13), $\frac{dETC_2}{dl_1} = c[1-(1-\alpha)l_2] \geq 0$ □

Proof of Proposition 1: The proposition can be proved by showing $de_1^*/d\delta \geq 0$, since δ for the farsighted assembler is greater than zero.

i) Case 1: $0 \leq e_1^* < 1$. Holding other parameters except δ fixed, define $w(e_1, \delta) = dETCf/de_1$. Then,

$$w(e_1,\delta) = -cL(1-\alpha) + \frac{[m+V(e_1)]V'(e_1)}{2} + \frac{2V''(e_1)e_1(1-e_1) + (V'(e_1))^2(1-2e_1)}{4} + \delta\left[ETC_2(e_2^* \mid l_1 = \alpha L) - ETC_2(e_2^* \mid l_1 = L)\right] \tag{39}$$

From the cost minimizing condition,. $w(e_1,\delta)$ Thus, $\left(\frac{\partial w}{\partial e_1}\right)de_1 + \left(\frac{\partial w}{\partial \delta}\right)d\delta = 0$, or

$$\frac{de_1}{d\delta} = -\left(\frac{\partial w}{\partial \delta}\right)\Big/\left(\frac{\partial w}{\partial e_1}\right) \tag{40}$$

From Eq. (A.16) and Lemma 6,

$$\frac{\partial w}{\partial \delta} = ETC_2(e_2^* | l_1 = \alpha L) - ETC_2(e_2^* | l_1 = L) \leq 0 \tag{41}$$

From Lemma 5, $ETCf$ is convex. On $[0,e_1^*],(0 \le e_1^* < 1)$. Thus,

$$\frac{\partial w}{\partial e_1} \ge 0 \tag{42}$$

From Eq. (41)-(42), $de_1/d\delta \ge 0$.

ii) Case 2: $e_1^* = 1$.

Let $e_1^{\#} = \min\{e_1 \mid \partial ETCf(e_1)/\partial e_1 = 0 \text{ and } e_1 \ge 0\}$. If $e_1^{\#} \ge 1$, case 1 shows that e_1^* remains at one when δ increases. Suppose that $e_1^{\#} < 1$. In this case, we need to show that $ETCf\big|_{e_1=e_1^{\#}}$ increases faster than $ETCf\big|_{e_1=1}$, as δ increases. That is, we show that $\left.\frac{dETCf}{d\delta}\right|_{e_1=1} < \left.\frac{dETCf}{d\delta}\right|_{e_1=e_1^{\#}}$.

$$\left.\frac{dETCf}{d\delta}\right|_{e_1=1} = ETC_2\left(e_2^*\middle|l_1 = \alpha L\right) \tag{43}$$

$$\left.\frac{dETCf}{d\delta}\right|_{e_1=e_1^{\#}} = \left(\frac{\partial ETC_2}{\partial\delta}\right)d\delta + \left(\frac{\partial ETCf}{\partial e_1}\right)de_1 \tag{44}$$

Since $\partial ETCf / \partial e_1 = 0$, Eq. (44) is

$$\left.\frac{dETCf}{d\delta}\right|_{e_1=e_1^{\#}} = ETC_2(e_2^*\big|l_1 = \alpha L)e_1^{\#} + ETC_2(e_2^*\big|l_1 = L)(1-e_1^{\#}) \tag{45}$$

From Lemma 5, since $0 \le e_1^{\#} < 1$,

$$\left.\frac{dETCf}{d\delta}\right|_{e_1=1} < \left.\frac{dETCf}{d\delta}\right|_{e_1=e_1^{\#}} \tag{46}$$ □

Proof of Proposition 2: i) $GAP(e_1) = \frac{1}{2}\left[mV' + VV' + (V')^2/2 - (V')^2 e_1\right]$

$$\partial GAP/\partial e_1 = V''\left(m + V + V'(1-2e_1)\right)/2 = \gamma(m + 2\gamma e_1 - 3e_1^2)$$

Let. $f(e_1) = m + 2\gamma e_1 - 3e_1^2$ .Since $f(0) > 0$ and $f(1)=m-\gamma$, $\partial GAP(e_1)/\partial e_1 > 0$, if $m > r$.

ii) Let $f(e_1) = 0$. Then, the solutions e_{1-}, e_{1+} are:

$$e_{1-} = \frac{\gamma - \sqrt{\gamma^2 + 3\gamma m}}{3\gamma};$$

$$e_{1+} = \frac{\gamma + \sqrt{\gamma^2 + 3\gamma m}}{3\gamma}.$$

Thus, $dGAP(e_1)/de_1 > 0$, if $e_1 \in \left[0, (\gamma + \sqrt{\gamma^2 + 3\gamma m})/3\gamma\right)$

iii) Obvious from Eq. (28). □

REFERENCES

[1] Ackere, A. V., "The Principal / Agent Paradigm: Its Relevance to Various Functional Fields", *European Journal of Operational Research,* 70 (1993), 83-103.

[2] Basu, A. K., R. Lal, V. Srinivasan, and R. Staelin, "Sales Force Compensation Plans: An Agency Theoretic Perspective," Marketing Science, 4/4 (1985), 267-291.

[3] Bitran, G. R. and D. Tirupati, "Lot Sizing Under (Q,R) Policy in a Capacity Constrained Manufacturing Facility," *Robotics & Computer-Integrated Manufacturing,* 1/3 (1984), 327-337.

[4] Blackburn, J.D. and L. N. V. Wassenhove, "From IE to JIT to Time-Based Competition", *INSEAD,* 93/46/TM/SM (1993).

[5] Grossman, S. J. and O. D. Hart, "An Analysis of the Principal-Agent Problem", *Econometrica,* 51/ January (1983), 7-45.

[6] Holmström, B., "Moral Hazard and Observability, " *Bell Journal of Economics,* 10/Spring (1979), 74-91.

[7] Karmarkar, U. S., "Lotsizes, Lead Times and In-process Inventories," *Management Science,* 33/3 (1987), 409-418.

[8] Kenney, M. and R. Florida, *Beyond Mass Production, The Japanese System and its Transfer to the U. S.,* Oxford University Press, Oxford, (1993)

[9] Kim, I. and C. S. Tang, "Lead Time and Response Time in a Kanban Control System," *European Journal of Operational Research,* 101/3 (1997), 474-485.

[10] Kuhn, H.W., and A.W. Tucker: "Nonlinear Programming," in Jerzy Neyman (ed.), *Proceedings of the Second Berkeley Symposium,* University of California Press, Berkeley, (1951), 481-492.

[11] Stalk, G., Jr., "Time-The Next Source of Competitive Advantage", *HBR,* Jul.-Aug. (1988), 41-51.

[12] Tang, C. S., "Supply Contracts with Special Incentives", *Working Paper, UCLA* (1991).

[13] Womack, J. P., D. T. Jones, and D. Roos, *The Machine that Changed the World,* Harper Perennial 1990.