

# *Trainers' Dilemma of Choosing between Training and Promotion\**

Kyota Eguchi\*\*

The University of Tsukuba  
Institute of Policy and Planning Sciences

First version: October 2001

This version: April 2002

---

\* The earlier version of this paper was presented at the 2001 Annual Meeting of the Japanese Economic Association at Hitotsubashi University. I am very grateful to Takashi Shimizu, Michihiro Kandori, Hitoshi Matsushima, Hideshi Itoh, Yoshiyasu Ono, Fumio Ohtake, Yukiko Abe, and seminar participants at the University of Tokyo, Kyoto University, and ISER in Osaka University. This research is partially supported by Grant-in-Aid for Scientific and Research of the Japanese Ministry of Education.

JEL Classification Numbers: J24, J31

\*\* Correspondence: Kyota Eguchi, The University of Tsukuba, Institute of Policy and Planning Sciences, 1-1-1 Tennoudai, Tsukuba, Ibaragi 305-8573, Japan. Phone & Fax: +81-298-53-5375.

e-mail: [eguchi@sk.tsukuba.ac.jp](mailto:eguchi@sk.tsukuba.ac.jp)

# *Trainers' Dilemma of Choosing between Training and Promotion*

## **Abstract**

It is observed that old employees play a significant role in training young employees especially in Japanese firms. Why do old employees provide this training for young employees who are their rivals in the promotion competition? The more training that trainers devote to trainees, the less likely will trainers be promoted. We show that training generalists rather than training specialists can soften the trainers' dilemma of choosing between training and promotion. As opposed to the specialist scheme, the amount of training provided by a trainer does not decrease that trainer's promotion probability in the generalist scheme. Hence, even if the productivity of the generalists is lower than that of the specialists, the generalist scheme can implement a higher level of training and improve the firms' profit and social welfare.

Key Words: Generalists, Specialists, Trainers' Incentives, and Promotion

JEL Classification Numbers: J24, J31.

## **1. Introduction**

Numerous studies have been based on the human capital theory and show the importance of human investment and skill accumulation. Indeed, workers often learn skills through on-the-job training or schooling (off-the-job training), and trainers play a significant role in training courses. School trainers might be professional instructors and not employees, though it is often the case that trainers are co-workers of the trainees. In particular, in the case of on-the-job training, trainers and trainees are likely to belong to the same firm, which can cause the trainers and trainees to be rivals of each other in the promotion competition. Thus, we have a problem: What are the incentives for trainers to instruct trainees and thus help their rivals at their own expense? In other words, how do firms encourage trainers to provide training for trainees?

An answer to this question may be the separation of old workers as trainers and young workers as trainees in the promotion competition. If firms can commit to treat generations of workers separately in terms of promotion, then the old workers, as trainers, will not worry about losing to the young workers in the promotion competition and would be willing to devote time and skills to training. Very often, however, firms can not commit to that separation and are willing to promote good young workers with high productivity rather than old less productive workers. Actually, the seniority rule is rarely established in Japan or with white collar workers in the U.S. It is observed that young talented employees are promoted faster than older ones with lower ability when the age difference between older and younger is small. So firms are unlikely to promote the old workers with lower productivity. The more time and energy trainers devote to trainees, the less likely those trainers will get the promotion prize. Thus, firms must offer a high payment as incentive for the contribution of old workers as trainers.

We show that training generalists can soften the trainers' dilemma of choosing between training and promotion. We call the multiskilled persons generalists. They can work at various workshops and are contrary to specialists who can work only at a particular workshop. The productivity of the generalists is less than that of the specialists which is one of the merits of specialization. We consider the simple case of a firm with two different jobs and the choice of two kinds of skill schemes. One is the specialist scheme: workers with the specialized skill (specialists) can complete only one job. The other is the generalist scheme: workers with multiskills (generalists) can complete both jobs.

In the skill specialization scheme, the trainer's training level directly influences his own possibility of winning the promotion competition. If a trainer is shirking on

training, the trainee's productivity is likely to be low and thus the trainer can face high probability of promotion instead of getting low training reward. Under the generalist scheme, training provided by a trainer for a trainee affects the promotion probability of not only himself, but also the other trainers; if one trainer is shirking on training, the amount of training provided by the other trainer decreases the shirking trainer's probability of promotion. In the specialist scheme, this effect disappears since the probability of promotion the shirking trainer faces is independent of other trainers' behavior. The amount of training provided by a trainer does not decrease that trainer's promotion probability in the generalist scheme as it does in the specialist scheme. Hence, the aspects of the specialist scheme that discourage trainers are not likely to appear in the generalist scheme. If the firm selects the production scheme of the generalists rather than that of the specialists, the firm can lower payment for the contribution of trainers and still encourage trainers to provide training for trainees under the conflict between training and promotions. When the productivity of the generalists is sufficiently large, firms can increase profits. Trainers' training vs. promotion dilemma enlarges the effectiveness of the generalist scheme even if the merits of specialization and division of labor exist. The generalist scheme is more competitive than the specialist one for the trainers' promotion.

Indeed, in Japanese firms, as Koike (1977) and Aoki (1988) indicate, workers learn a wide range of skills through job rotation and on-the-job training. Old employees play a significant role in instructing young ones on the job. Training generalists and old employees as trainers are often observed in Japanese firms. Ohkusa and Ohtake (1997) have found that information sharing enhances the effect of profit sharing on productivity in Japanese firms. Information sharing is very relevant to the generalist scheme. Also, Ichniowski, Shaw, and Prennushi (1997) have found that innovative changes in human resource management such as team work, flexible job assignments, employment security, or job rotation also improve the productivity of the finishing line in the U.S. steel industry. Osterman (1994) and Ichniowski, Kochan, Levine, Olson, and Strauss (1996) point out this trend of innovative changes in human resource management in the U.S. These studies imply that the generalist scheme is not limited to the Japanese firms. Our conclusion that multiskill accumulation enhances trainers' incentives supports these studies.

The effects of generalists have been considered from the viewpoint of sudden stochastic shocks. Aoki (1986) analyzes horizontal or vertical firm structures and shows that the firms with horizontal information processing can improve profit under moderate stochastic shocks. Although Aoki (1986) pays much attention to information sharing

among employees, information sharing is very relevant to generalists. Itoh (1987) shows that training generalists improves firms' profits under moderate stochastic shock and is dominated by training specialists under drastic shocks or very stable states. Koike (1977) (1991) mentions that employees learn a wide range of skills through job rotation and on-the-job training in Japanese firms. When a worker is absent from the workshop, other workers can complete his task on his behalf. These studies are based on the viewpoint of training generalists under uncertainty. Moreover, Carmichael and MacLeod (1993) indicate that multiskilled workers will cooperate with labor-saving technological change in cases where singly skilled workers will not. Though we also are concerned with multiskilled vs. specialized workers, our analysis is different from these studies; we consider the effects of training generalists from the viewpoint of trainers' incentives.

In this paper we analyze the dilemma of trainers in choosing between training incentives and promotion. This paper is organized as follows. In section 2, we explain the basic model and consider the skill specialization case. Section 3 investigates the effects of multiskill accumulation. After an extensional case is described in chapter 4, conclusion and discussion are drawn in section 5.

## **2. The Basic Model**

We consider a simple model where a firm has two management jobs, A and B. These jobs are different and essential to the firm's activities. Workers with high productivity are promoted to these management jobs. It is necessary for young workers to receive the training provided by old workers in order to complete the management jobs A and B. Old workers as trainers provide some amount of training and increase the productivity of young workers as trainees.

There are two kinds of production schemes: specialist and generalist. Specialists have necessary skills specific to either job A or B, whereas the generalist has skills that can be applied to both jobs in the incumbent firm. We assume specialized merits appear: the productivity of workers with specialized skills is greater than that of workers with general skill. Both the specialized and general skills are specific to the incumbent firm and irrelevant to the other firms.

First, we consider the specialist scheme: a trainer with only one specialized skill instructs the trainee in this skill (figure 1). We call the specialized skill that is specific to the management position A (B), the skill A (B). The trainer with skill A

provides some amount of training of skill A for the trainee. Since we assume that skills A and B are symmetric, it is sufficient to consider only the skill A case. The same results that we show later can be applied to the skill B case.

When a trainee receives training of the skill A, the trainee's productivity  $T$  at the management position A is followed by

$$T = t + \varepsilon,$$

where the amount of training devoted to the trainee is denoted as  $t$  and the trainee's potential ability as  $\varepsilon \in E$ . His productivity at the management position B is zero since he knows nothing about the skill B. The training level  $t$  devoted to a trainee is nonnegative. Although the distribution function of the trainee's potential ability  $\Phi(\varepsilon)$  is known, the trainee's potential ability  $\varepsilon$  is unknown to everyone during training, but is then observable after training. The mean of  $\varepsilon$  is zero, and the density function  $\phi(\varepsilon)$  is differentiable. The amount of training provided by the trainer is verifiable, and thus the firm can offer a wage scheme contingent on the amount of training:  $w(t)$ . As we indicate later, trainers' utility function is linear with respect to wage, and thereby any risk problem does not appear. Hence, even if the firm can observe only a signal with some noises on the training level provided by a trainer but cannot observe the precise training level, our results are not influenced at all.

A trainer's productivity is given by  $\bar{T}$ . After the trainee's productivity  $T$  is revealed, either the trainer or the trainee with the skill A (B) is promoted to the management position A (B) in the firm. If  $T \leq \bar{T}$ , the trainer wins in the promotion competition. However, if  $T > \bar{T}$ , he loses. Hence, when the trainer provides the amount of training  $t$ , he is promoted with the probability  $\Phi(\bar{T} - t)$ . Since both jobs A and B are different and essential to the firm, promotion probability of the trainer and trainee with the skill A is independent of the results of the training of the other skill B. The winner receives the promotion payment  $v$ .

There is a constraint on the promotion payment the firm faces:  $v \geq \bar{v} (> 0)$ . The constraint of promotion payment  $\bar{v}$  is exogeneously given. This constraint means a labor market pressure the incumbent firm faces. Promotion provides additional signals on promoted employees' ability for outside firms. Outside firms know that promoted workers are likely to have higher abilities, and thus outside firms might take the promoted workers from the incumbent firm. If there is labor market pressure like this, the incumbent firm must offer a positive level of promotion payment  $\bar{v}$  to discourage the promoted employees from quitting the firm. The firm must offer the promotion

payment  $\bar{v}$  at least. The incumbent firm knows its employees' ability, but outside firms cannot observe it in this situation, and thus promotion is likely to provide an effective signal on promoted workers' ability. Employees with high ability tend to be promoted and outside firms can observe who are promoted.<sup>1</sup>

The trainer's expected utility is given by

$$U(t) \equiv w(t) - c(t) + \Phi(\bar{T} - t)v,$$

where  $c(t)$  is a training cost function satisfied as follows:

$$c' > 0, c'' > 0, c(0) = 0, \text{ and } c'(0) = 0. \quad \dots(1)$$

Timing of players' decisions is given as follows:

- 1) The firm specifies wage schemes contingent on the amount of training to trainers and a promotion payment  $v$ .
- 2) Trainers provide some amount of training for trainees.
- 3) Productivity of trainees is revealed. Workers with higher productivity are promoted and receive the promotion prize.

We consider the trainer's incentive problem. As we show later, there is an optimal training level  $t^\#$  for the firm, and the firm is willing to pay nothing if  $t \neq t^\#$ . Thus, the wage profile is offered:

$$w(t) = \begin{cases} w & \text{if } t = t^\# \\ 0 & \text{otherwise.} \end{cases}$$

The wage scheme implies that the firm cannot take something pecuniary from workers, so the firm chooses to pay nothing as the severest punishment. The firm cannot punish shirking trainers punitively. This setting is consistent with the real world since labor law often prohibits a firm from receiving something as a kind of bond from workers. Note

---

<sup>1</sup> Waldman (1984) and Ricart I Costa (1988) analyze that the effect of extracting workers by other outside firms distorts promotion speed and the promoted range of employees with different abilities under asymmetric information on employees' abilities between the incumbent firm and other outside ones. Gibbons and Katz (1991) consider which type of workers is likely to be fired in the similar situation.

that the firm cannot commit to the dismissal of shirking trainers. When productivities of trainers and trainees are revealed, the firm is willing to promote workers with higher productivity. Even if a trainer is shirking on the training, he can be promoted when his productivity exceeds his trainee's productivity. It is *ex post* optimal for the firm to promote workers with higher productivity. Trainers take the firm's *ex post* optimal action into account, and thus they realize that dismissal of shirking trainers as a punishment is not enforceable. Thus, the firm faces the trade off relationship between high wage payment and trainers' incentives of training although no risk problem appears.

When the trainer faces the wage scheme, he will provide either the amount of training  $t=t^\#$  or zero. If he does not choose the training requirement  $t^\#$ , he receives no wage. The zero training level minimizes his training cost and maximizes the probability of promotion, and hence zero training leads to the highest utility among  $t \neq t^\#$ . Therefore, using (1), the trainer's incentive problem is as follows:

$$U(t^\#) = w - c(t^\#) + \Phi(\bar{T} - t^\#)v \geq \Phi(\bar{T})v = U(0). \quad \dots(2)$$

The incentive compatibility is binding on the equilibrium, so that the wage level  $w$  is determined.

For simplicity, we consider only output of management sections.<sup>2</sup> The training course provides no output. Since either the trainer or the trainee is promoted to the management position A, the firm's expected output for position A is given by

$$\begin{aligned} y &\equiv \int \max\{\bar{T}, T\} \phi(\varepsilon) d\varepsilon \\ &= \Phi(\bar{T} - t) \bar{T} + \int_{\bar{T}-t}^{\bar{T}} (t + \varepsilon) \phi(\varepsilon) d\varepsilon. \end{aligned}$$

Thus, the firm's profit is  $\Pi = Y - 2w - 2v$ , where  $Y=2y$ , since the management positions A and B are symmetric and the trainees' wage is zero. Substituting (2) into this equation,

---

<sup>2</sup> The following case is considered: total output of workers with skill A is given by  $y = \alpha y_1 + \beta y_2$ , where  $y_1$  is the promoted worker's productivity and  $y_2$  is the non-promoted worker's productivity, and  $\alpha$  and  $\beta$  are constant coefficients and mean the importance of management and non-management positions. Under  $\alpha > \beta$ , the firm is willing to promote the more productive worker. In this paper,  $\alpha = 1$  and  $\beta = 0$  are assumed for simplicity. This assumption is not crucial to our results.

the firm's expected profit under  $t = t^\#$  is given by

$$\Pi = 2 \left[ \Phi(\bar{T} - t^\#) \bar{T} + \int_{\bar{T} - t^\#}^{t^\# + \varepsilon} (t^\# + \varepsilon) \phi(\varepsilon) d\varepsilon - c(t^\#) - \left[ \Phi(\bar{T}) - \Phi(\bar{T} - t^\#) \right] (v - \bar{v}) \right] \dots(3)$$

subject to  $v \geq \bar{v}$ . Since it holds that  $-\left[ \Phi(\bar{T}) - \Phi(\bar{T} - t^\#) \right] (v - \bar{v}) < 0$ , the firm is willing to offer the minimum promotion payment:  $v = \bar{v}$ . The optimal training level  $t^\#$  is given by

$$\frac{\partial \Pi}{\partial t^\#} = 2 \left[ 1 - \Phi(\bar{T} - t^\#) - c' - \phi(\bar{T} - t^\#) \bar{v} \right] = 0. \dots(4)$$

This equation (4) as the first order condition holds if  $1 - \Phi(\bar{T}) > \phi(\bar{T}) \bar{v}$ . Consider the existence of a solution for equation (4), denoting the function  $F(t)$  as follows:

$$F(t) \equiv 1 - \Phi(\bar{T} - t) - c' - \phi(\bar{T} - t) \bar{v}. \dots(5)$$

We assume throughout this paper as follows:

***Assumption 1***

$$F(0) = 1 - \Phi(\bar{T}) - \phi(\bar{T}) \bar{v} > 0$$

From (1),  $F(t) \rightarrow -\infty$  as  $t \rightarrow +\infty$ . Since  $\phi(\varepsilon)$  is differentiable,  $F(t)$  is differentiable. Therefore, a training level satisfying (4) exists under assumption 1.

Trainers provide the training requirement  $t^\#$ . A high training level decreases the probability of promotion for the trainer. The firm must offer a higher wage when the trainer has a conflict between training and promotion than when there is no conflict. Without the conflict, the first best training level is enforced by the first order condition:  $1 - \Phi(\bar{T} - t) - c' = 0$ . Hence, the training level implemented in this setting is lower than when no conflict exists.

**3. The Generalist Scheme**

In this section we consider the generalist scheme. A trainee receives the training provided by a trainer with the general skill. There are two training pairs

( $k \in \{i, j\}$ ) of a trainer and a trainee (see figure 2). We refer to the trainer (trainee) of the training pair  $k$  as trainer  $k$  (trainee  $k$ ). Let us denote the total amount of training provided by trainer  $k$  for trainee  $k$  as  $t_k$ . Trainers and trainees as generalists can work toward both management positions A and B. Trainers' productivity in the management positions A and B is  $\delta\bar{T}$  ( $0 < \delta < 1$ ). This implies that the generalists are dominated by specialists from the viewpoint of productivity (the merits of specialization and division of labor), though the generalists can complete both management jobs A and B. The trainee  $k$ 's productivity in the management positions A and B is given by

$$\delta T_k = \delta(t_k + \varepsilon_k), \quad (k \in \{i, j\}).$$

The general skill leads to symmetric performance in the management jobs A and B. Each trainer's contribution to the training is observable. Each trainee's potential abilities  $\varepsilon_k$  ( $k \in \{i, j\}$ ) is unknown during training, but is observable after training. Furthermore,  $\varepsilon_k$  is independently and identically distributed. The firm offers a wage scheme contingent on the trainer's amount of training. As it has been shown in the specialist scheme, there is the optimal training level of a trainer for the firm  $t_k^*$ , and thus the wage scheme is offered as follows:

$$w_k(t_k) = \begin{cases} w_k & \text{if } t_k = t_k^* \\ 0 & \text{otherwise} \end{cases} \quad (k \in \{i, j\}).$$

Next, we consider the probability of promotion of the trainer  $k$ . Contrary to the skill specialization case, the promotion probability of trainers is influenced by two trainees' performances. If both trainees' productivity is less than the trainers',  $T_i \leq \bar{T}$  and  $T_j \leq \bar{T}$ , both trainers win the promotion competition. In this case, the firm's total output is  $2\delta\bar{T}$ . The state occurs with the probability  $\Phi(\bar{T} - t_i)\Phi(\bar{T} - t_j)$ . If trainee  $k$ 's productivity exceeds the trainers' productivity  $\delta\bar{T}$ , but the other trainee  $l$ 's does not, then trainee  $k$  and one of the trainers are promoted. This situation occurs and a trainer is promoted with the probability  $\frac{1}{2} \prod_{k \neq l \in \{i, j\}} [1 - \Phi(\bar{T} - t_k)] \Phi(\bar{T} - t_l)$ . The output is  $\delta(\bar{T} + T_k)$ . Finally, if both trainees' productivity is greater than the trainers', neither of the trainers is promoted. The firm's output is  $\delta(T_i + T_j)$ . This case occurs with the probability  $\prod_{k \in \{i, j\}} [1 - \Phi(\bar{T} - t_k)]$ . We denote the expected promotion payment of trainer  $k$ , given the amount of training provided by the other trainer  $l$ , as  $V^k(t_k | t_l)$ :

$$\begin{aligned}
V^k(t_k|t_l) &\equiv v \left[ \Phi(\bar{T} - t_k) \Phi(\bar{T} - t_l) \right. \\
&\quad \left. + \frac{1}{2} \left[ (1 - \Phi(\bar{T} - t_k)) \Phi(\bar{T} - t_l) \right. \right. \\
&\quad \left. \left. + \Phi(\bar{T} - t_k) (1 - \Phi(\bar{T} - t_l)) \right] \right] \quad (k \neq l \in \{i, j\}) \\
&= \frac{v}{2} \left[ \Phi(\bar{T} - t_k) + \Phi(\bar{T} - t_l) \right].
\end{aligned}$$

The expected utility of trainer  $k$ , given the training level of the other trainer  $l$ , is given by

$$U^k(t_k|t_l) \equiv w_k(t_k) - c(t_k) + V^k(t_k|t_l) \quad (k \neq l \in \{i, j\}).$$

In the same manner as the skill specialization case, the zero training level leads to the highest expected utility among  $t_k \neq t_k^*$ :

$$U^k(0|t_l) = V^k(0|t_l) = \frac{v}{2} \left[ \Phi(\bar{T} - t_l) + \Phi(\bar{T}) \right] \quad (k \neq l \in \{i, j\}).$$

Since workers with the general skill can work at both management positions A and B, you might think that the firm can choose the following punishment for the shirking trainer when one trainee and one trainer is promoted. If a trainer deviates from the training requirement, the firm might promote the other trainer. If the firm can punish the shirking trainer like this, the firm can give more training incentives to the trainers and lower the wage of trainers. Hence, the generalist scheme can dominate the specialist scheme. However, in this paper, we consider the case of the firm that cannot use that kind of punishment for the tie-breaking of trainers. Even if the firm cannot use that kind of punishment for the tie-breaking of trainers, we show that the generalist scheme can dominate the specialist one.

If the following case is considered, our setting is reasonable. This is the case in which the trainers' productivity involves very little noise. If trainers' productivity is vibrated very little near  $\bar{T}$ , the firm is willing to promote the trainer with higher productivity *ex post*. Hence, the firm cannot commit to not promoting the shirking trainer while promoting the other trainer when a trainee and a trainer are promoted. Trainers take the firm's *ex post* optimal behavior into account, and thus the punishment scheme is not enforceable. If the noise is very small, we can ignore the existence of the

noise when measuring the trainers' productivity *ex ante* as in this model.

Moreover, you may think that the following incentive device is useful: the firm offers a bonus to a trainer  $k$  when his trainee  $k$  is promoted. The existence of bonus seems to enhance trainers' incentive of training. However, as we show in next chapter, the bonus scheme does not play any significant role. The firm is always unwilling to offer the bonus.

To make trainers choose  $t_k = t_k^*$  ( $k \in \{i, j\}$ ), the firm must satisfy the incentive compatibility:  $U^k(t_k^*|t_l^*) \geq U^k(0|t_l^*)$ . Since the firm is willing to decrease the wage level, the incentive compatibility is binding on the equilibrium. The trainers provide the required training amount  $t_k^*$ , and thus the following equations hold on the equilibrium:

$$w_k = c(t_k^*) - V^k(t_k^*|t_l^*) + V^k(0|t_l^*) \quad (k \neq l \in \{i, j\}). \quad \dots(6)$$

You may think there would be an equilibrium such that both trainers deviate from the required training level even if the incentive compatible contract is offered. However, Nash implementation is equivalent to dominant strategy equilibrium implementation in our model. Hence, we can show that trainers always provide the required training level regardless of the other trainers' behavior when the firm offers the incentive compatible contract.

**Proposition 1**

When the firm offers an incentive compatible contract on training in the generalist scheme, the optimal training requirement for the firm is enforceable as a unique Nash equilibrium.

***Proof***

When incentive compatible contracts are offered:

$$w_k - c(t_k^*) + \frac{v}{2} \left[ \Phi(\bar{T} - t_l^*) + \Phi(\bar{T} - t_k^*) \right] \geq \frac{v}{2} \left[ \Phi(\bar{T} - t_l^*) + \Phi(\bar{T}) \right], \quad (k \neq l \in \{i, j\})$$

it is obtained that  $w_k - c(t_k^*) + \frac{v}{2} \left[ \Phi(\bar{T} - t_k^*) - \Phi(\bar{T}) \right] \geq 0$ . This means that one trainer's action does not depend on the other trainer's action. The trainer  $k$  has an incentive for providing the required level of training regardless of the other trainer's action. Hence,

trainers are willing to provide the training requirement for trainees. ■

This proposition indicates that there is no equilibrium in which trainers deviate from the required training level  $t_k^*$  ( $k \in \{i, j\}$ ) when the firm offers an incentive compatible contract to trainers. If the firm offers the incentive compatible contract, the firm can always maximize its profit.

The firm's expected output is

$$\begin{aligned}
Y^G \equiv & \Phi(\bar{T} - t_i)\Phi(\bar{T} - t_j)2\delta\bar{T} \\
& + (1 - \Phi(\bar{T} - t_i))\Phi(\bar{T} - t_j)\delta(\bar{T} + E(T_i|T_i > \bar{T})) \\
& + \Phi(\bar{T} - t_i)(1 - \Phi(\bar{T} - t_j))\delta(\bar{T} + E(T_j|T_j > \bar{T})) \\
& + (1 - \Phi(\bar{T} - t_i))(1 - \Phi(\bar{T} - t_j))\delta(E(T_i|T_i > \bar{T}) + E(T_j|T_j > \bar{T}))
\end{aligned} \tag{7}$$

where  $E(T_k|T_k > \bar{T}) = t_k + \frac{\int_{\bar{T}-t_k}^{\infty} \varepsilon_k \phi(\varepsilon_k) d\varepsilon_k}{1 - \Phi(\bar{T} - t_k)}$  ( $k \in \{i, j\}$ ). The firm's profit under the generalist scheme is given by

$$\Pi^G = Y^G - w_i - w_j - 2v.$$

Substituting (6) and (7) into the above equation, the optimal training requirements are given by

$$\frac{\partial \Pi^G}{\partial t_k} = \delta \left[ 1 - \Phi(\bar{T} - t_k) \right] - c' - \frac{v}{2} \phi(\bar{T} - t_k) = 0 \quad (k \in \{i, j\}). \tag{8}$$

The determinant of training requirement  $t_k$  of the trainer  $k$  is independent of the other trainer's action. For any distribution of  $\varepsilon_k$  ( $k \in \{i, j\}$ ), from the symmetricity, it is obtained that  $t_i^* = t_j^*$ . We define the amount of training per trainer on the equilibrium as follows:

$$t^* \equiv t_i^* = t_j^*. \tag{9}$$

Hence, equations (7) and (8) are replaced by

$$Y^G = 2\delta \left[ \Phi(\bar{T} - t^*)\bar{T} + (1 - \Phi(\bar{T} - t^*))t^* + \int_{\bar{T}-t^*} \varepsilon \phi(\varepsilon) d\varepsilon \right], \quad \dots(7)'$$

and

$$\delta \left[ 1 - \Phi(\bar{T} - t^*) \right] - c' - \frac{v}{2} \phi(\bar{T} - t^*) = 0. \quad \dots(8)'$$

Using (9), it is clear that  $V^k(t^*|t^*) = v\Phi(\bar{T} - t^*)$  and  $V^k(0|t^*) = \frac{v}{2} \left[ \Phi(\bar{T} - t^*) + \Phi(\bar{T}) \right]$ .

Thus, the firm's profit is given by

$$\begin{aligned} \Pi^G = 2 \left[ \delta \left\{ \Phi(\bar{T} - t^*)\bar{T} + (1 - \Phi(\bar{T} - t^*))t^* + \int_{\bar{T}-t^*} \varepsilon \phi(\varepsilon) d\varepsilon \right\} \right. \\ \left. - c(t^*) - \frac{v}{2} \left[ \Phi(\bar{T}) - \Phi(\bar{T} - t^*) \right] - v \right] \quad \dots(10) \end{aligned}$$

As similar to the specialist scheme, the firm is willing to offer  $v = \bar{v}$ .

In the same manner as the specialist scheme, we define the function as follows:

$$F^G(t) \equiv \delta \left[ 1 - \Phi(\bar{T} - t) \right] - c' - \frac{\bar{v}}{2} \phi(\bar{T} - t). \quad \dots(11)$$

Equation (8) holds if  $\delta$  is sufficiently large. From assumption 1, it is obvious that  $F^G(0) > F(0) > 0$  under  $\delta = 1$ . Hence, as we show as follows, there is the critical point  $\delta = \underline{\delta}$  such that a positive training level is implemented under the generalist scheme.

**Lemma 1**

There is a productivity of generalists  $\delta = \underline{\delta}$  such that a positive training level is enforced under  $\delta > \underline{\delta}$ .

***Proof***

See Appendix.

Under  $\delta < \underline{\delta}$ , it holds that  $F^G(t) < 0$  for any  $t$ , and thus a positive training amount is not implemented. If  $\delta > \underline{\delta}$ , a positive training level  $\tilde{t} = \arg \max F^G(t)$  exists and  $F^G(\tilde{t}) > 0$  holds. From the continuity of  $F^G(t)$  with respect to  $t$ , positive training amount is realized in this case. Under  $\delta = \underline{\delta}$ , a positive training amount may be implemented. However, if the zero training level can only yield the maximum value of  $F^G(t)$ , a positive training amount is not implemented. Therefore, a positive training level

is always enforced under  $\delta > \underline{\delta}$ .

Now, we assume the density function  $\phi$  as follows:

***Assumption 2***

The density function  $\phi$  satisfies that  $\forall \varepsilon \in E \quad \phi(\varepsilon) > 0$ .

This assumption is very natural and not strong. We compare the firm's profit in the generalist scheme with that in the specialist one.

***Lemma 2***

It holds that  $\Pi^G(t^*) > \Pi(t^\#)$  and  $t^* > t^\#$  under  $\delta = 1$ .

***Proof***

First, we show  $\Pi^G(t^*) > \Pi(t^\#)$  under  $\delta = 1$ . Clearly, it holds under  $\delta = 1$  and assumption 2 that  $F^G(t) > F(t)$  for any  $t$ . From (3) and (10), it holds under  $\delta = 1$  that  $\Pi^G(t) > \Pi(t)$  for any  $t$ . Hence,  $F(t^\#) = 0 < F^G(t^\#)$  and  $\Pi^G(t^\#) > \Pi(t^\#)$ . Obviously,  $\Pi^G(t^*) \geq \Pi^G(t^\#) > \Pi(t^\#)$  because  $t^*$  maximizes the firm's profit in the generalist scheme.

Next, we define the maximum of  $t^\#$  as  $t_{Max}^\#$ . Since  $\Pi^G(t_{Max}^\#) > \Pi(t_{Max}^\#)$  and  $F^G(t_{Max}^\#) > F(t_{Max}^\#) = 0$ , there is a training level  $t'$ :

$$t' > t_{Max}^\#,$$

where  $F^G(t') = 0$  and  $\Pi^G(t') > \Pi(t_{Max}^\#)$  hold. See figure 3. This result is brought by  $\lim_{t \rightarrow +\infty} F^G(t) = -\infty$  and the continuity of  $F^G(t)$  with respect to  $t$ . This implies that under the generalist scheme,  $t'$  maximizes the firm's profit locally near  $t_{Max}^\#$  under  $\delta = 1$ .

Next, we will show that a global maximum point exists under  $t > t_{Max}^\#$ . In other words, we show that the training level  $t < t_{Max}^\#$  does not maximize the firm's profit globally. We consider a training level  $t''$ :  $t'' < t_{Max}^\#$ . Since it holds under  $\delta = 1$  that  $\Pi^G(t) > \Pi(t)$  for any  $t$ , we get

$$\Pi^G(t'') = \Pi^G(t_{Max}^\#) - \int_{t''}^{t_{Max}^\#} F^G(t) dt > \Pi(t_{Max}^\#) - \int_{t''}^{t_{Max}^\#} F(t) dt = \Pi(t'').$$

Since it is clear under  $\delta = 1$  that  $F^G(t) > F(t)$  for any  $t$ , it holds that  $\int_{t''}^{t_{Max}^\#} F^G(t) dt > \int_{t''}^{t_{Max}^\#} F(t) dt$ .  $\Pi(t_{Max}^\#)$  is the global maximum for the firm's profit in the

specialist scheme, and thus it is obtained that  $\int_{t^n}^{t_{Max}^\#} F(t)dt \geq 0$ . Hence,  $\int_{t^n}^{t_{Max}^\#} F^G(t)dt > \int_{t^n}^{t_{Max}^\#} F(t)dt \geq 0$  holds under  $t'' < t_{Max}^\#$ , then it is obtained that  $\Pi^G(t'') < \Pi^G(t_{Max}^\#)$ . Therefore, all global maximum points exist under  $t > t_{Max}^\#$  when  $\delta = 1$ . Since  $t_{Max}^\#$  is the maximum of  $t^\#$ , any solution  $t^*$  is more than  $t^\#$ . ■

In the case of  $\delta = 1$  the merits of specialization and division of labor disappear. The increase of promotion probability caused by trainers' shirking actions is smaller in the generalist scheme than in the specialist scheme, and thus the firm can lower the wage of trainers in the generalist scheme. The effect of decreasing wage increases the firm's profit and training level.

Next, we consider the case of  $\delta < 1$ . Although the merits of specialization and division of labor are sacrificed under the generalist scheme in this case, this scheme enhances training. This is a significant merit. When the training merit exceeds the demerit of less productivity, the firm's profit in the generalist scheme is greater than under the specialist scheme. As we show later, since we can show that  $\Pi^G(t)$  is continuous with respect to  $\delta$ , the generalist scheme improves the firm's profit when  $\delta$  is sufficiently large.

### **Lemma 3**

$\Pi^G(t^*(\delta))$  is continuous with respect to  $\delta$ .

### ***Proof***

See Appendix.

As lemma 3 indicates,  $\Pi^G(t^*(\delta))$  is continuous with respect to  $\delta$ . However,  $t^*(\delta)$  is not always continuous with respect to  $\delta$ . The function  $F^G(t)$  is not always monotonously decreasing with respect to  $t$  although  $F^G(t) \rightarrow -\infty$  as  $t \rightarrow +\infty$ . The case like figure 4 can happen. When it holds that  $-\int_{t_1}^{t_2} F^G(t)dt = \int_{t_2}^{t_3} F^G(t)dt$  in figure 4, two optimal solutions exist:  $t^*(\delta) = t_1$  or  $t_3$ . In this case, if  $\delta$  is slightly changed, optimal solution can jump from  $t_1$  to  $t_3$  or *vice versa*. The relationship between  $t^*$  and  $\delta$  can be drawn like figure 5.

Moreover, we can show easily that the firm's profit increases with respect to  $\delta$ .

**Lemma 4**

It holds that  $\Pi^G(t^*(\delta_1)) > \Pi^G(t^*(\delta_2))$  if  $\delta_1 > \delta_2$  and that  $t^*(\delta_1) > t^*(\delta_2)$  if  $\delta_1 > \delta_2 > \underline{\delta}$ .

***Proof***

See Appendix.

Using lemma 2, 3, and 4, the curve of  $\Pi^G(t^*(\delta))$  is represented in figure 6. Note that under  $\delta < \underline{\delta}$ ,  $t^*(\delta) = 0$  and thus  $\Pi^G(0) = 2\delta \left\{ \Phi(\bar{T})\bar{T} + \int_{\bar{T}} \varepsilon \phi(\varepsilon) d\varepsilon \right\} - \bar{v}$ . The curve of  $\Pi^G(0)$  is linear with respect to  $\delta$  under  $\delta < \underline{\delta}$ . Hence, as figure 6 implies, we can easily show that the firm's profit implemented in the generalist scheme is greater than it should be in the specialist one when  $\delta$  is sufficiently large. Moreover, it is shown that the training level is higher in the generalist rather than the specialist scheme.

**Proposition 2**

It holds that  $\Pi^G(t^*) > \Pi(t^\#)$  and  $t^* > t^\#$  if  $\delta \in (0, 1)$  is sufficiently large.

***Proof***

See Appendix.

Although  $\Pi^G(t^*(\delta), \delta)$  is continuous with respect to  $\delta$ ,  $t^*(\delta)$  is not always continuous with respect to  $\delta$ . However,  $t^*(\delta)$  is an increasing function with respect to  $\delta$ . It is sufficient to show  $\lim_{\delta \rightarrow 1} t^*(\delta) > t^\#$  for proposition 2. From lemma 2,  $t^*(\delta) > t^\#$  holds under  $\delta = 1$ . If  $\lim_{\delta \rightarrow 1} t^*(\delta) = t^*(1)$  holds, it is clear that  $\lim_{\delta \rightarrow 1} t^*(\delta) = t^*(1) > t^\#$ . When  $t^*(\delta)$  is continuous with respect to  $\delta$  at the point of  $\delta = 1$ , proposition 2 holds (see figure 7-(A)).

However, it can be that  $t^*(\delta)$  is not continuous with respect to  $\delta$  at the point of  $\delta = 1$ . We consider the case. If  $\lim_{\delta \rightarrow 1} t^*(\delta) \leq t^\#$  holds, as figure 7-(B) implies, proposition 2 does not hold. We can show that this case does not occur. Hence, even if  $\lim_{\delta \rightarrow 1} t^*(\delta) < t^*(1)$  holds, it is always obtained that  $t^\# < \lim_{\delta \rightarrow 1} t^*(\delta) < t^*(1)$ . See figure 7-(C).

Therefore, proposition 2 holds.

In the specialist scheme, there are two promotion competitions segmented by the management positions A and B. The trainer with skill A competes only with his

trainee for management job A. On the other hand, there is only one large promotion competition in the generalist scheme. A trainer competes with not only his trainee but also the other trainer and trainee for two management positions A and B. Note that there are two workers per management position in both schemes. Under the generalist scheme, however, the amount of training provided by a trainer decreases the promotion probability of the other trainer. Hence, the benefit to the trainer (i.e. the chance of promotion) who deviates from the training requirement  $t^*$  is less in the generalist scheme than in the specialist scheme since the other trainer provides the training requirement on the equilibrium. The existence of the second trainer decreases the benefit which first trainer would otherwise receive from deviating, so the firm can lower the wage level, implement more training, and increase profit.

We consider social welfare. Social welfare in the specialist and generalist schemes are denoted as  $W$  and  $W^G$ . Since wage and promotion prize are just transfers from the firm to workers, these disappear from the viewpoint of social welfare:

$$W = 2 \left[ \delta \left( \Phi(\bar{T} - t^\#) \bar{T} + (1 - \Phi(\bar{T} - t^\#)) t^\# + \int_{\bar{T} - t^\#}^{\infty} \varepsilon \phi(\varepsilon) d\varepsilon - c(t^\#) \right) \right], \quad \dots(12)$$

and

$$W^G = 2 \left[ \delta \left\{ \Phi(\bar{T} - t^*) \bar{T} + (1 - \Phi(\bar{T} - t^*)) t^* + \int_{\bar{T} - t^*}^{\infty} \varepsilon \phi(\varepsilon) d\varepsilon \right\} - c(t^*) \right]. \quad \dots(13)$$

We can show that social welfare in the generalist scheme exceeds that in the specialist one if  $\delta$  is sufficiently large.

**Proposition 3**

$W^G > W$  holds if  $\delta \in (0, 1)$  is sufficiently large.

***Proof***

See Appendix.

This proposition means that social welfare in the generalist scheme exceeds that in the specialist one when the productivity of the general skill is sufficiently large. That the generalist scheme implements more training is essential to improving social welfare. The generalist scheme solves the trainers' dilemma of training vs. promotion, and always encourages trainers to provide training for trainees if social welfare is improved. The conflict between training and promotion leads to less training than the first best

amount of training without the conflict. Since the generalist scheme softens the conflict and encourages trainers to provide training, social welfare can be improved.

#### **4. Extention**

Trainers can get wage and promotion payment in this model. You may think that the firm offers the payment scheme dependent on promotion of their trainees. In the generalist scheme, trainers face the competition which trainees are promoted. If trainee  $k$  is promoted, trainer  $k$  can get a payment. However, payment schemes like this contingent on promotion of trainees do not play any significant role.

First, we consider this scheme in the specialist scheme. This payment is denoted as  $z$ . Since a trainer gets  $z$  as a kind of insurance when he is not promoted, his expected utility is given by

$$U(t) \equiv w(t) - c(t) + \Phi(\bar{T} - t)\bar{v} + \lceil \lceil - \Phi(\bar{T} - t) \rceil z .$$

There are two cases: shirking trainers get  $z$  or not. If a shirking trainer gets  $z$ , his utility is as follows:

$$U(0) = \Phi(\bar{T})\bar{v} + \lceil \lceil - \Phi(\bar{T}) \rceil z .$$

In this case, using incentive compatibility:  $U(t) \geq U(0)$ ,

$$w(t) = c(t) + \lceil \lceil \Phi(\bar{T}) - \Phi(\bar{T} - t) \rceil (\bar{v} - z) . \quad \dots(14)$$

The expected profit of the firm is given by

$$\begin{aligned} \Pi &= 2 \left[ y - c(t) - \lceil \lceil \Phi(\bar{T}) - \Phi(\bar{T} - t) \rceil (\bar{v} - z) - \bar{v} - \lceil 1 - \Phi(\bar{T} - t) \rceil z \right] \\ &= 2 \left[ y - c(t) - \lceil \lceil \Phi(\bar{T}) - \Phi(\bar{T} - t) \rceil \bar{v} - \bar{v} - \lceil 1 - \Phi(\bar{T}) \rceil z \right] . \end{aligned}$$

Clearly, the firm is willing to offer  $z = 0$ . As (14) indicates, the payment contingent on promotion of trainees  $z$  can reduce wage level by  $\lceil \lceil \Phi(\bar{T}) - \Phi(\bar{T} - t) \rceil z$ . However, since the firm pays this payment  $\lceil \lceil - \Phi(\bar{T} - t) \rceil z$ , payment cost increases for  $\lceil \lceil - \Phi(\bar{T}) \rceil z$  by providing the payment  $z$  after all. Thus,  $z = 0$  is optimal for the firm.

Next, we consider the case such that the shirking trainer cannot get  $z$ . In this case, from  $U(0) = \Phi(\bar{T})\bar{v}$ , it similarly holds that

$$w(t) = c(t) + \left[ \Phi(\bar{T}) - \Phi(\bar{T} - t) \right] \bar{v} - \left[ 1 - \Phi(\bar{T} - t) \right] z.$$

After all, the payment  $z$  just decreases wage level, and thus  $z$  cannot affect the firm's profit and trainers' action. The firm's profit is the same as (3). Hence, the payment contingent on promotion of trainees is not effective at all.

Whether shirking trainers can get the payment contingent on promotion of trainers or not, the firm has no incentives of offering this payment scheme and then this payment scheme disappears. Similarly, the above statement can be applied to the case of generalist scheme.

## **5. Conclusion and Discussion**

We have shown that the generalist scheme can increase training and profit. The generalist scheme decreases the trainers' pressure in the promotion competition, and thus a lower payment is sufficient to induce trainers to provide training for rival trainees. Even if the merits of specialization and division of labor exist, the generalist scheme can improve training, and then profit can increase.

In this model, limited liability is crucial to our results. Since there is no risk problem, no limited liability leads to the first best allocation. The firm offers wage contingent on the first best training level and punishes trainers punitively when their training does not lead to the training requirement. The firm can absorb the entire welfare from trainers. Thus, trainers' dilemma disappears and the specialist scheme always dominates the generalist one because of productivity merit. However, labor law protects workers from firms' punitive treatments and limited liability of workers is observed in the real world. The setting of limited liability and no risk problem (risk neutral agents) leads to make models simple and significant.

We have considered the case where the firm cannot make a promise that old workers as trainers and young workers as trainees are treated separately. If the firm can commit to separate treatment on promotion, then the trainers' dilemma does not appear and the first best level of training can be implemented. In this case, trainers do not compete with trainees but with the other trainers. However, our setting is crucial in the real world. It is observed that young workers with high productivity often overtake old

workers in the promotion ladder. Itoh (1994) mentions that to treat different generations of employees separately in the promotion competition solves the trainers' dilemma. We partially agree with his statement because Japanese firms do treat employees with a large age difference separately. Also, there is the strict seniority rule for blue collar workers in the U.S. However, if there is a small difference in age, younger employees with high productivity can be promoted faster than older ones with low productivity. Employees who are about five years older can play a significant role in the training of younger employees. Hence, the situation described in this paper is very important to the understanding of promotion and training in the real world. Strictly separate treatment according to workers' age and tenure rarely exists except for bureaucrats. Hence, our setting is relevant to the real world.

In this paper, we consider a simple case: a worker is promoted to each of the management positions A and B. It is not crucial to our results that only one employee is promoted to each of the management positions. Even if  $N$  workers are promoted to each position, the generalist scheme can lead to more profit than the specialist one. Although a trainer for the management position A has  $N-1$  trainers and  $N$  trainees as rivals in the specialist scheme, a trainer competes with  $2N-1$  trainers and  $2N$  trainees in the generalist scheme. Hence, the generalist scheme can lead to more profit. We have considered the case of the firm that cannot punish the shirking trainer under a tie-breaking of trainers: the firm cannot punish the shirking trainer and promote the other one when only one of the trainers is promoted. This setting implies that our results are independent of the number of workers. Also, the result depends on the number of different jobs. If there is only one kind of job in the firm, the effect of generalist scheme does not exist. An increase in the variety of jobs strengthens the effectiveness of the generalist scheme.

As Koike (1977) and Aoki (1988) mentioned, in Japanese firms, instructing generalists, team work, and job rotation play significant roles. Recently, innovations in human resource management such as team work, flexible job assignments, and job rotation improve profit and workers' motivation in the U.S. (Ichniowski, Shaw, and Prennushi (1997)). We have shown a new view to support the above studies on how a generalist scheme affects the trainers' dilemma of choosing between training and promotion.

## **Appendix**

**Proof of lemma 1**

The function  $F^G(t)$  is continuous with respect to  $t$ . Since  $F^G(0)$  is a real value and it holds that  $F^G(t) \rightarrow -\infty$  as  $t \rightarrow +\infty$ , the maximum value of  $F^G(t)$  always exists under  $t \geq 0$ . Clearly,  $F^G(t)$  is continuous with respect to  $\delta$ . From assumption 1, it holds that  $F^G(0) > F(0) > 0$  under  $\delta = 1$ . Thus, there is a level of  $\delta = \underline{\delta}$  such that

$$\max_t \left\{ F^G(t) = \underline{\delta} \left[ 1 - \Phi(\bar{T} - t) \right] \left[ -c' - \frac{\bar{v}}{2} \phi(\bar{T} - t) \right] \right\} = 0.$$

Under  $\delta < \underline{\delta}$ , the maximum of  $F^G(t)$  is negative, and hence a positive training level is not enforced. On the other hand, since the maximum of  $F^G(t)$  is positive under  $\delta > \underline{\delta}$ , a positive training level is realized as the equilibrium because of continuity of  $F^G(t)$  with respect to  $t$ . Under  $\delta = \underline{\delta}$ , only the zero training level can lead to the maximum value of  $F^G(t)$ . Therefore, a positive training level is always implemented under  $\delta > \underline{\delta}$ . ■

**Proof of lemma 3**

We consider the function:

$$\begin{aligned} \Pi^G(t, \delta) \equiv & 2 \left[ \delta \left\{ \Phi(\bar{T} - t) \bar{T} + (1 - \Phi(\bar{T} - t)) t + \int_{\bar{T}-t}^{\bar{T}} \varepsilon \phi(\varepsilon) d\varepsilon \right\} \right. \\ & \left. - c(t) - \frac{\bar{v}}{2} \left[ \Phi(\bar{T}) - \Phi(\bar{T} - t) \right] - \bar{v} \right] \end{aligned}$$

The function  $\Pi^G(t, \delta)$  is continuous with respect to  $t$  given any level of  $\delta$ . Since  $\lim_{t \rightarrow +\infty} \Pi^G(t, \delta) = -\infty$  and  $\Pi^G(0, \delta) = 2 \left[ \delta \left\{ \Phi(\bar{T}) \bar{T} + \int_{\bar{T}}^{\bar{T}} \varepsilon \phi(\varepsilon) d\varepsilon \right\} - \bar{v} \right]$  is a constant real value from (1), there is the training level:  $t^*(\delta) = \arg \max_t \Pi^G(t, \delta)$ . It is sufficient to show for any  $\delta_1$

$$\lim_{\delta \rightarrow \delta_1} \Pi^G(t^*(\delta), \delta) = \Pi^G(t^*(\delta_1), \delta_1). \quad \dots(A1)$$

(A1) implies that the function  $\Pi^G(t^*(\delta), \delta)$  is continuous with respect to  $\delta$  at the point of  $\delta = \delta_1$ . If  $\lim_{\delta \rightarrow \delta_1} t^*(\delta) = t^*(\delta_1)$  holds, (A1) is clearly obtained.

Next, we consider the case of  $\lim_{\delta \rightarrow \delta_1} t^*(\delta) = t' \neq t^*(\delta_1)$ : optimal training level is not continuous at the point of  $\delta_1$ . In this case,  $\lim_{\delta \rightarrow \delta_1} \Pi^G(t^*(\delta), \delta) = \Pi^G(t', \delta_1)$  and  $t' \neq t^*(\delta_1)$ . From the definition of  $t'$  and continuity of  $F^G(t, \delta)$ ,  $\lim_{\delta \rightarrow \delta_1} F^G(t, \delta) = F^G(t', \delta_1) = 0$  holds. This implies that  $t'$  is a solution of  $F^G(t, \delta) = 0$  under  $\delta = \delta_1$ . Hence, it holds that  $\lim_{\delta \rightarrow \delta_1} \Pi^G(t^*(\delta), \delta) = \Pi^G(t^*(\delta_1), \delta_1)$ . Therefore, it is obtained that  $\lim_{\delta \rightarrow \delta_1} \Pi^G(t^*(\delta), \delta) = \Pi^G(t^*(\delta_1), \delta_1)$  even if  $\lim_{\delta \rightarrow \delta_1} t^*(\delta) = t' \neq t^*(\delta_1)$  holds. In other words,  $\Pi^G(t^*(\delta))$  is continuous with respect to  $\delta$ . ■

#### Proof of lemma 4

Clearly, the following inequality holds:

$$\Pi^G(t^*(\delta_1), \delta_1) \geq \Pi^G(t^*(\delta_2), \delta_1) > \Pi^G(t^*(\delta_2), \delta_2) \quad \text{if } \delta_1 > \delta_2.$$

The first inequality holds because  $t^*(\delta_1)$  realizes the maximum value of the profit under  $\delta = \delta_1$  and the second one is clear.

Next, we will show  $t^*(\delta_1) > t^*(\delta_2)$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . For simple notation,  $t_1 = t^*(\delta_1)$  and  $t_2 = t^*(\delta_2)$ . Suppose that  $t_1 = t^*(\delta_1) < t^*(\delta_2) = t_2$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . Clearly,  $\Pi^G(t, \delta_1) > \Pi^G(t, \delta_2)$  holds for any  $t$ . Hence, the following inequality is obtained:

$$\begin{aligned} \Pi^G(t_1, \delta_1) &= \Pi^G(t_2, \delta_1) - \int_{t_1}^{t_2} F^G(t, \delta_1) dt \\ &> \Pi^G(t_2, \delta_2) - \int_{t_1}^{t_2} F^G(t, \delta_2) dt = \Pi^G(t_1, \delta_2), \end{aligned} \quad \dots(\text{A2})$$

where  $F^G(t, \delta) \equiv \delta \left[ 1 - \Phi(\bar{T} - t) \right] - c' - \frac{\bar{v}}{2} \phi(\bar{T} - t)$ . Since  $F^G(t, \delta_1) > F^G(t, \delta_2)$  holds for any  $t$ , and  $t_2$  maximizes the function  $\Pi^G(t, \delta_2)$  under  $\delta = \delta_2$ , it is obtained that  $\int_{t_1}^{t_2} F^G(t, \delta_1) dt > \int_{t_1}^{t_2} F^G(t, \delta_2) dt \geq 0$ . From this inequality and (A2),  $\Pi^G(t_1, \delta_1) < \Pi^G(t_2, \delta_1)$ . This contradicts  $t_1 = t^*(\delta_1) \equiv \arg \max \Pi^G(t, \delta_1)$ . The result is caused by the assumption of  $t_1 = t^*(\delta_1) < t^*(\delta_2) = t_2$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . Hence, we get

$$t_1 = t^*(\delta_1) \geq t^*(\delta_2) = t_2 \quad \text{under } \delta_1 > \delta_2 > \underline{\delta}. \quad \dots(\text{A3})$$

Finally, we show that (A3) holds strictly:  $t_1 = t^*(\delta_1) > t^*(\delta_2) = t_2$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . Suppose that  $t^*(\delta_1) = t^*(\delta_2) = t^*$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . This indicates  $F^G(t^*, \delta_1) = F^G(t^*, \delta_2) = 0$ . However, from  $\delta_1 > \delta_2 > \underline{\delta}$ ,

$$\begin{aligned} F^G(t^*, \delta_1) &= \delta_1 \left[ 1 - \Phi(\bar{T} - t^*) \right] - c' - \frac{\bar{v}}{2} \phi(\bar{T} - t^*) \\ &> \delta_2 \left[ 1 - \Phi(\bar{T} - t^*) \right] - c' - \frac{\bar{v}}{2} \phi(\bar{T} - t^*) = F^G(t^*, \delta_2) \end{aligned}$$

This is a contradiction. Hence,  $t^*(\delta_1) \neq t^*(\delta_2)$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . From this result and (A3), it holds that  $t^*(\delta_1) > t^*(\delta_2)$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . ■

### **Proof of proposition 2**

From lemma 2, 3, and 4, it is clear that  $\Pi^G(t^*) > \Pi(t^\#)$  holds if  $\delta \in (0, 1)$  is sufficiently large (see figure 6).

Next, we consider the implemented training level. There are two cases to consider:  $t^*(\delta)$  is continuous with respect to  $\delta$  at  $\delta = 1$  or not.

[1] If  $t^*(\delta)$  is continuous with respect to  $\delta$  at  $\delta = 1$ ,  $\lim_{\delta \rightarrow 1} t^*(\delta) = t^*(1)$  holds. From lemma 2,  $\lim_{\delta \rightarrow 1} t^*(\delta) = t^*(1) > t^\#$ . This case can be drawn in figure 7-(A). Hence, when  $\delta$  is sufficiently large,  $t^*(\delta) > t^\#$  holds.

[2] We consider the other case:  $t^*(\delta)$  is not continuous at  $\delta = 1$ . We will show  $\lim_{\delta \rightarrow 1} t^*(\delta) > t_{Max}^\#$ . Now, for convenience, we denote as  $\lim_{\delta \rightarrow 1} t^*(\delta) = t'$ . From the definition of  $t'$  and the continuity of  $F^G(t, \delta)$  with respect to  $\delta$ , it holds that  $\lim_{\delta \rightarrow 1} F^G(t, \delta) = F^G(t', \delta_1) = 0$ . This implies that  $t'$  is a solution under  $\delta = 1$  and  $t'$  maximizes  $\Pi^G(t', 1)$ . From lemma 2,  $t' > t_{Max}^\#$  holds under  $\delta = 1$ . Therefore,  $\lim_{\delta \rightarrow 1} t^*(\delta) = t' > t_{Max}^\#$ . We have proved that  $t^* > t^\#$  when  $\delta$  is sufficiently large. ■

### **Proof of proposition 3**

First, we consider the case of  $\delta = 1$ . In this case,  $W(t) = W^G(t)$  holds under any  $t$ . From (3), (10), (12), and (13),

$$W(t^\#) = \Pi(t^\#) + 2\bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^\#)] + 2\bar{v}$$

and

$$W^G(t^*) = \Pi^G(t^*) + \bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^*)] + 2\bar{v}. \quad \dots(\text{A4})$$

Using (A4), it is obtained that

$$W(t^\#) = W^G(t^\#) = \Pi^G(t^\#) + \bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^\#)] + 2\bar{v}.$$

Since  $t^\# < t^*$  under  $\delta = 1$  from lemma 2, it is clear that  $\bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^\#)] < \bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^*)]$ . Using  $\Pi^G(t^\#) < \Pi^G(t^*)$ , it is obvious that

$$W(t^\#) = W^G(t^\#) < W^G(t^*) \quad \text{under } \delta = 1. \quad \dots(\text{A5})$$

Lemma 4 indicates that  $\Pi^G(t^*(\delta))$  and  $t^*(\delta)$  are increasing functions with respect to  $\delta$  under  $\delta > \underline{\delta}$ . Hence, it is obvious that  $\Pi^G(t^*(\delta_1)) > \Pi^G(t^*(\delta_2))$  and  $\bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^*(\delta_1))] > \bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t^*(\delta_2))]$  under  $\delta_1 > \delta_2 > \underline{\delta}$ . Therefore,

$$W^G(t^*(\delta_1)) > W^G(t^*(\delta_2)) \quad \text{under } \delta_1 > \delta_2 > \underline{\delta}. \quad \dots(\text{A6})$$

Welfare in the generalist scheme  $W^G$  increases with respect to the optimal training requirement  $t^*(\delta)$  which increases with respect to  $\delta$ . From proposition 2, it holds that  $t^*(\delta) > t^\#$  when  $\delta \in (0, 1)$  is sufficiently large. This implies that the training requirement in the general scheme can be more than that in the specialist scheme. In the proof of proposition 2, it has shown that  $\lim_{\delta \rightarrow 1} t^*(\delta) = t' > t^\#$ . Then, it holds that

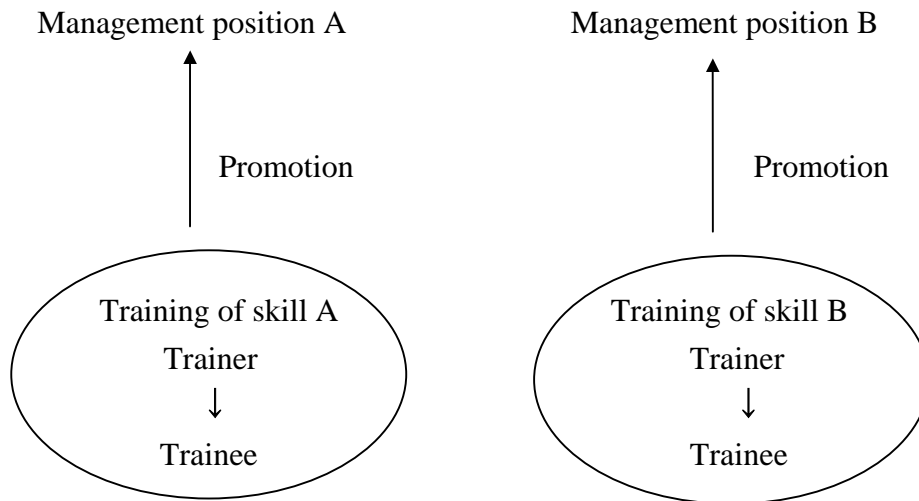
$$\lim_{\delta \rightarrow 1} W^G(t^*(\delta)) = W^G(t') = \Pi^G(t') + \bar{v}[\Phi(\bar{T}) - \Phi(\bar{T} - t')] + 2\bar{v}.$$

From  $t' > t^\#$ , it is obtained that  $\lim_{\delta \rightarrow 1} W^G(t^*(\delta)) = W^G(t') > W(t^\#)$ . Hence,  $W^G > W$  holds if  $\delta \in (0, 1)$  is sufficiently large. ■

## References

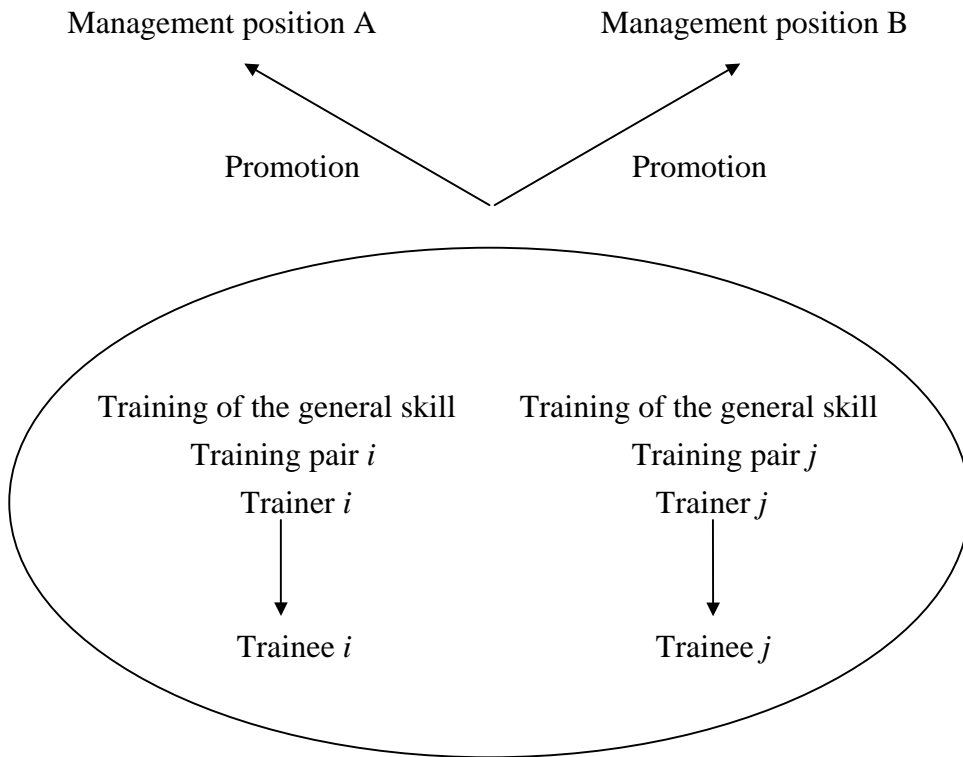
- Aoki, Masahiko (1986)** "Horizontal vs. Vertical Information Structure of the Firm",  
*American Economic Review*, 76, 971-983.
- **(1988)** "Information, Incentives, and Bargaining in the Japanese  
Economy", (Cambridge, Cambridge: *Cambridge University Press*, 1988).
- Carmichael, Lorne, and Bentley MacLeod (1993)** "Multiskilling, Technical Change  
and the Japanese Firm", *Economic Journal* , 103, 142-160.
- Gibbons, Robert, and Lawrence F. Katz (1991)** "Layoffs and Lemons",  
*Journal of Labor Economics*, 9, 351-380.
- Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi (1997)** "The Effects of  
Human Resource Management Practices on Productivity: A Study of  
Steel Finishing Lines", *American Economic Review*, 87, 291-313.
- , **Thomas Kochan, David Levine, Craig Olson, and Geroqe Strauss (1996)**  
"What Works at work: Overview and Assessment",  
*Industrial Relations*, 35, 299-333.
- Itoh Hideshi (1987)** "Information Processing Capacities of the Firm",  
*Journal of the Japanese and International Economies*, 1, 299-326.
- **(1994)** "Japanese Human Resource Management from the Viewpoint of  
Incentive Theory", in Aoki, M. and R. Dore, eds., "The Japanese Firm:  
The Sources of Competitive Strength", Oxford: Oxford University Press.
- Koike, Kazuo (1977)** "Participation of Unions in Workshops: Comparative Analysis on  
Industrial Relations in Japan and the U.S.",  
("Syokuba no Rodo-Kumiai to Sanka: Roshi-Kankei no Nichibei Hikaku"  
in Japanese), (Tokyo: Toyo-Keizai Shinpo, 1977).
- **(1991)** "Economics of Job" ("Shigoto no Keizaigaku" in Japanese),  
(Tokyo: *Toyo-Keizai Shinpo*, 1991)
- Ohkusa, Yasushi, and Fumio Ohtake (1997)**, "The Productivity Effects of  
Information Sharing, Profit Sharing, and ESOPs",  
*Journal of the Japanese and International Economies*, 11, 385-402.
- Osterman, Paul (1994)** "How Common is Workplace Transformation and Who  
Adopts It?", *Industrial and Labor Relation Review*, 47, 173-188.
- Ricart I Costa, Joan E., (1988)** "Managerial Task Assignment and Promotion",  
*Econometrica*, 56, 449-466.
- Waldman, Michael, (1984)** "Job Assignment, Signaling, and Efficiency",  
*Rand Journal of Economics*, 15, 255-267.

## The specialist scheme

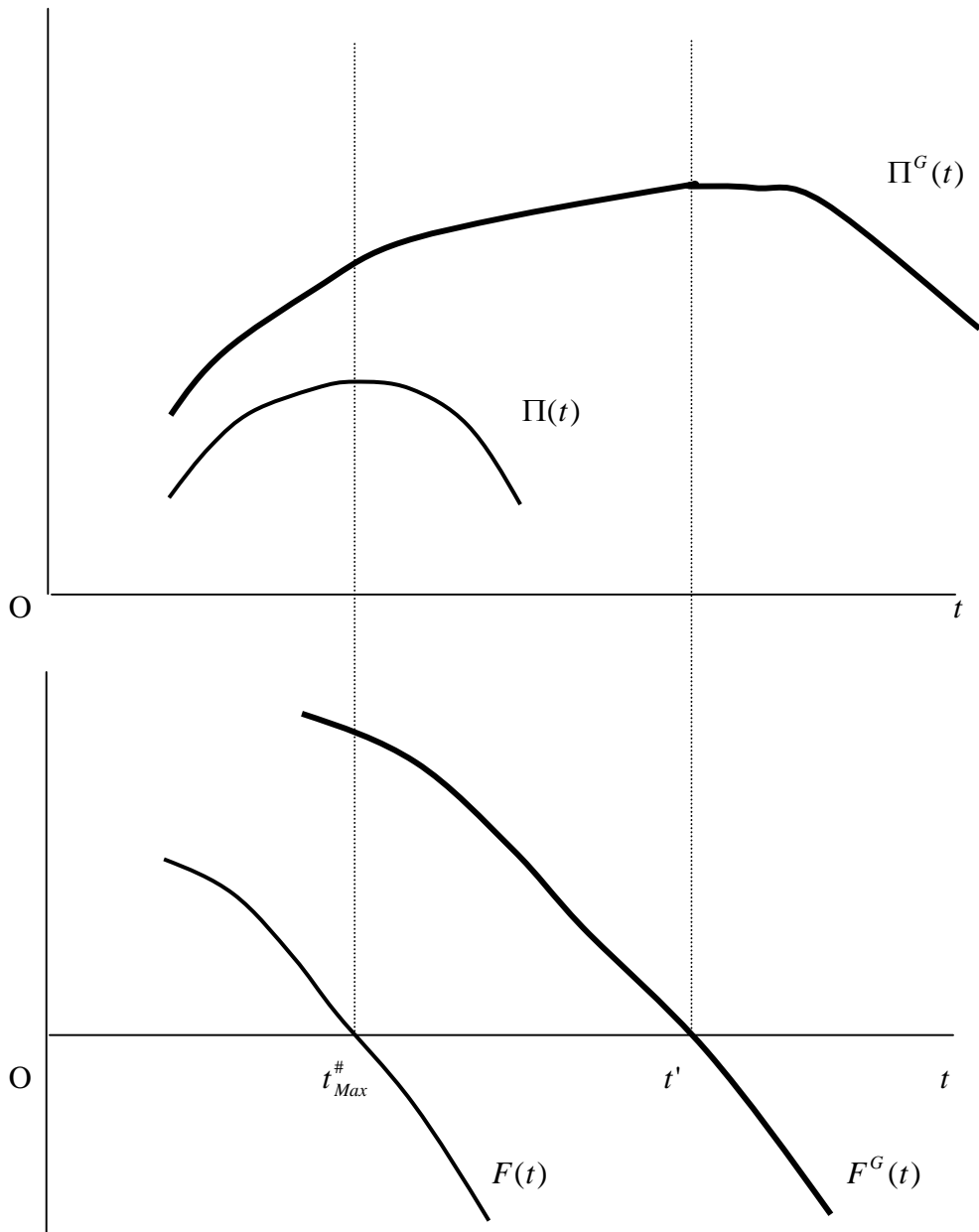


**Figure 1**

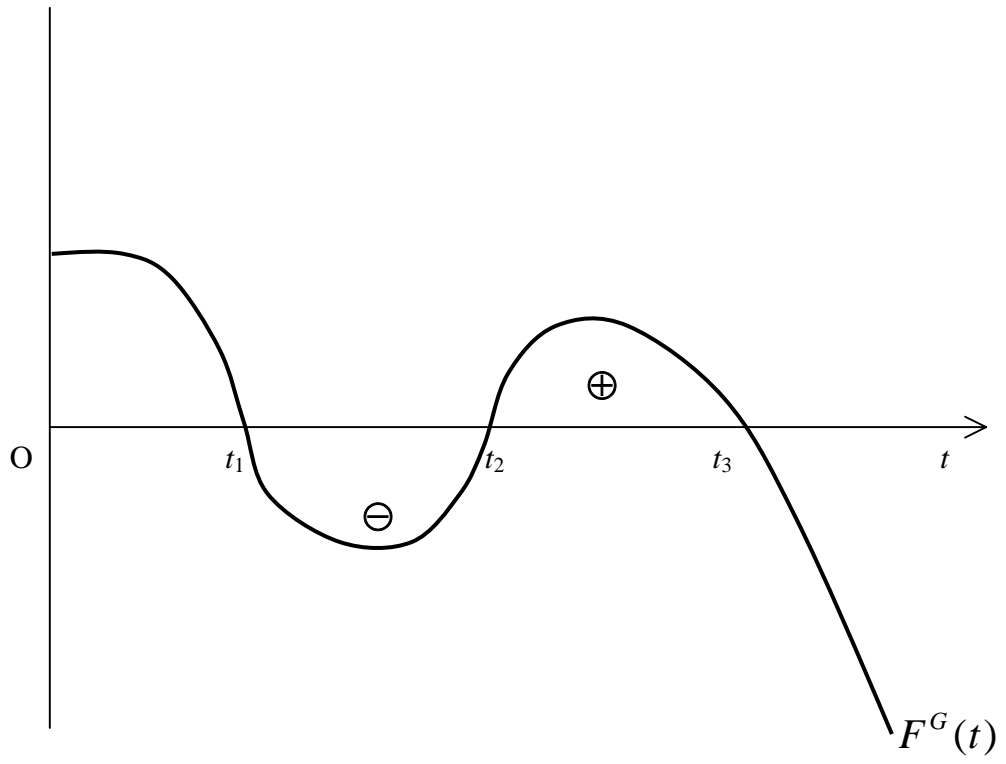
## The generalist scheme



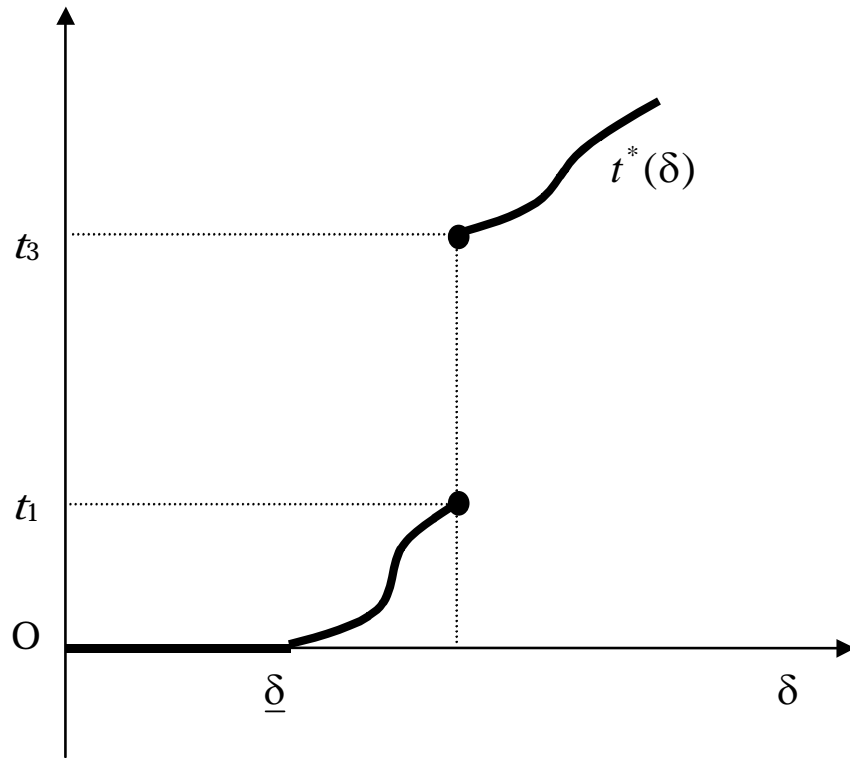
**Figure 2**



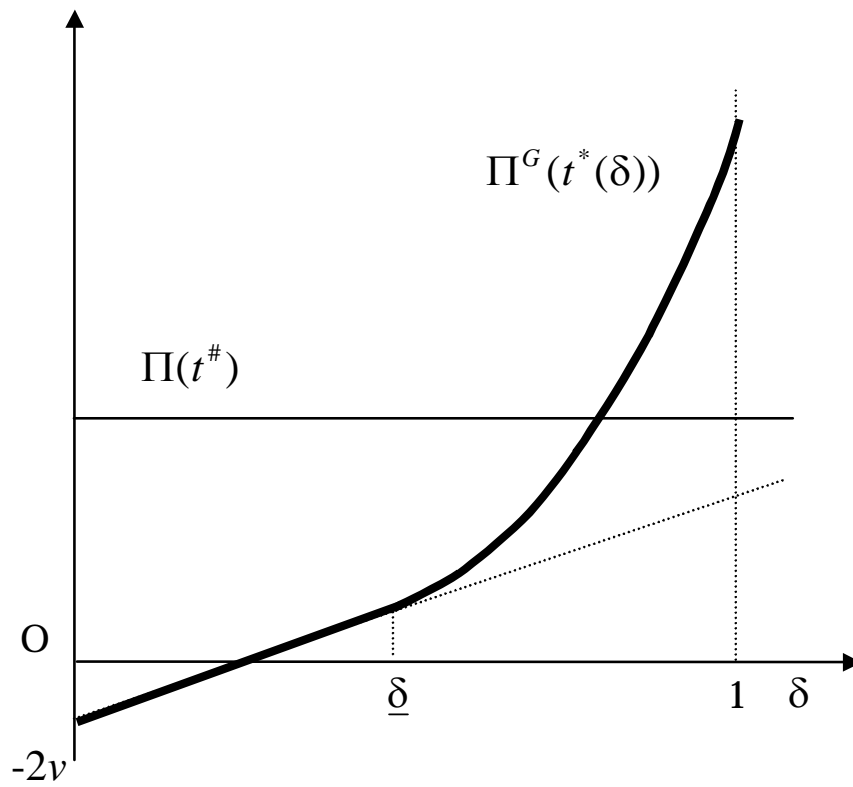
**Figure 3**



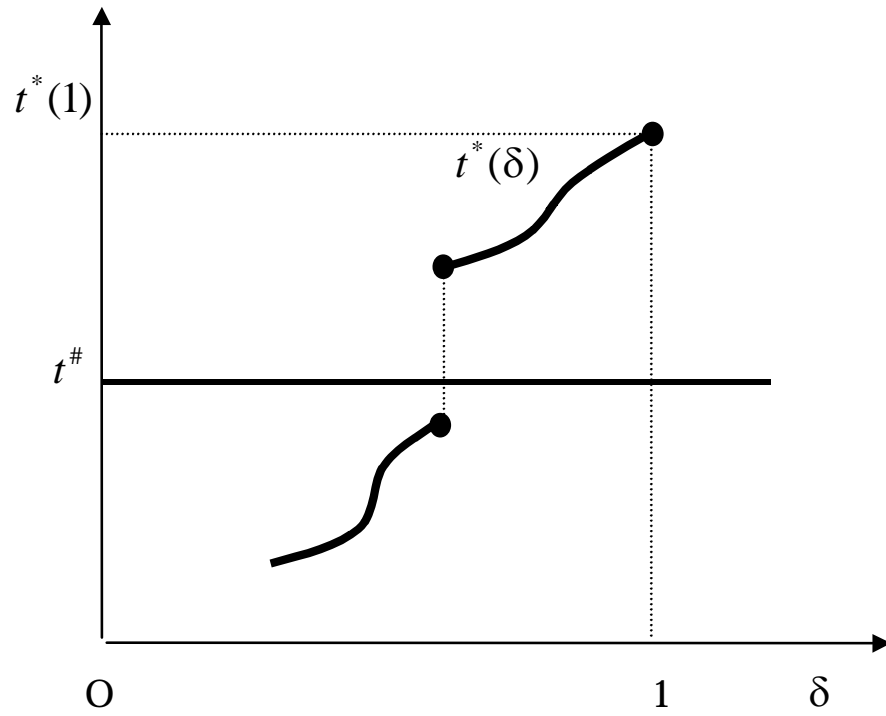
**Figure 4**



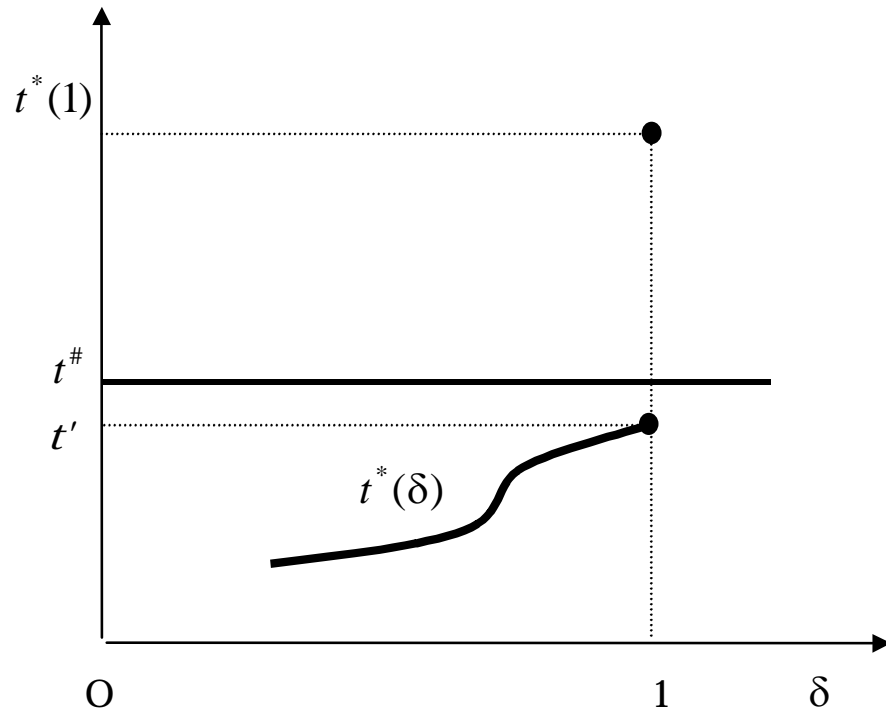
**Figure 5**



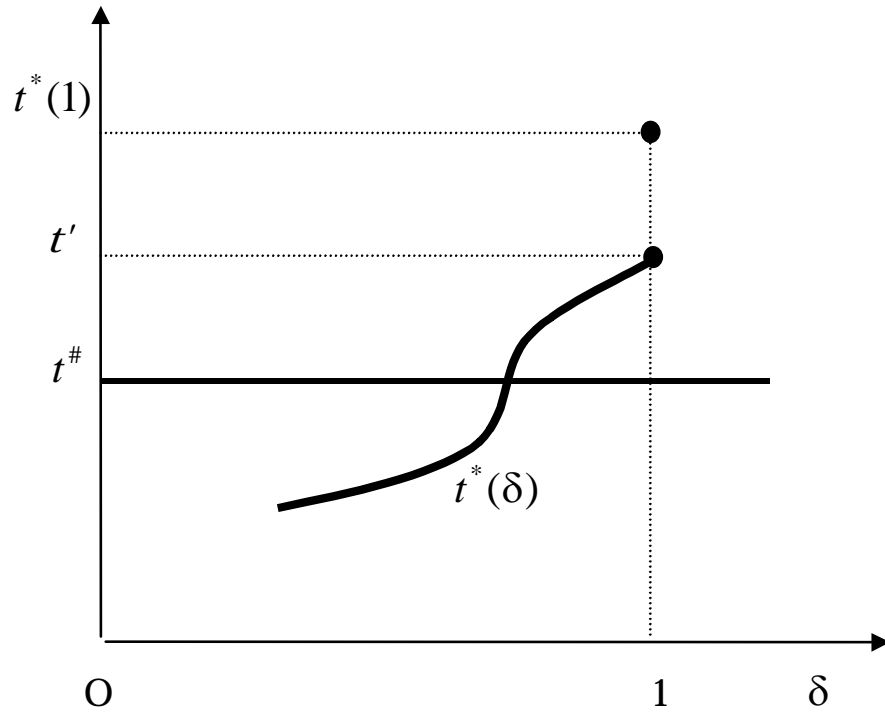
**Figure 6**



**Figure 7-(A)**



**Figure 7-(B)**



**Figure 7-(C)**