GCOE Discussion Paper Series

Global COE Program
Human Behavior and Socioeconomic Dynamics

Discussion Paper No. 252

Dynamically Sabotage-Proof Tournaments

Junichiro Ishida

March 2012

GCOE Secretariat
Graduate School of Economics
OSAKA UNIVERSITY
1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

Dynamically Sabotage-Proof Tournaments*

Junichiro Ishida

ISER, Osaka University

March 26, 2012

Abstract

This paper explores the consequences of sabotage for the design of incentive contracts. The possibility of sabotage gives rise to a dynamic concern, similar to the Ratchet effect, which distorts the agents' incentives. We first show that the mere possibility of sabotage may make it impossible to implement the first-best effort, and then offer two distinct incentive schemes, fast track and late selection, to circumvent this problem. The present model offers a mechanism through which these two schemes arise in a unified framework.

JEL Classification Codes: J41; M12.
Key Words: Sabotage; Tournament; Fast track; Late selection.

*Correspondence to: Junichiro Ishida, Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan. E-mail address: jishida@iser.osaka-u.ac.jp

1 Introduction

In this paper, we study the design of incentive contracts when agents may resort to destructive sabotage activities in a two-period tournament model with strategic information revelation. We consider two agents with different innate ability who choose between three alternatives: exerting productive effort, sabotage effort or no effort at all. To illustrate the nature of the problem, we start the analysis by showing that sabotage tends to be an effective strategy for low-ability agents, especially when they are faced with high-ability opponents: that is, when the perceived difference in innate ability is sufficiently large, it becomes the preferred option for the less able agent to sabotage the opponent to fill this gap. This fact gives rise to a serious dynamic implication when agents gain some information about each other over time, because a high-ability agent essentially runs a risk of becoming the target of sabotage by signaling his high ability in early stages.

This dynamic concern gives rise to a new type of inefficiency that has not been considered in the literature, yielding critical implications for the design of optimal incentive schemes. We first show that any tournament that can implement the first-best effort (always inducing productive effort) is not sabotage-proof under fairly plausible conditions: in other words, there exists no contract that can implement the first-best effort in both periods when sabotage is a viable option. The logic behind this result is fairly simple. When the first-best effort is implemented in the first period, any difference in the productivity must be attributed to the difference in innate ability. There then inevitably arises a situation where the perceived *ex post* difference in innate ability is so large that it is optimal for the less able agent to resort to sabotage activities in the second period. This result thus indicates that although the costs arising from sabotage activities are well recognized in the static setting, the possibility of sabotage invites more serious problems in the dynamic setting than previously recognized.

Given this impossibility result, the main purpose of the paper is to explore ways in which to mitigate sabotage in search of the second best, with a particular focus on how much weight to place on the ranking in each period. The main issue here is how to devise a tournament that can mitigate the incentive for low-ability agents to exert sabotage effort in the second period. In general, there are two distinct ways to achieve this goal. We show that a tournament can be made sabotage-proof by shifting the weight in either direction, i.e., either towards the ranking in the first period or towards that in the second. While these two schemes can equally prevent sabotage from actually taking place on the equilibrium path, each comes at a cost with different implications.

First, consider a tournament which places more weight on the ranking in the first period:

that is, high-powered incentives are provided in the first period, followed by low-powered incentives in the second. The key aspect of this scheme is the pay compression to make sabotage less effective in the second period. Since less is at stake in the second period, this reduces the incentive for low-ability agents to exert sabotage effort in that period. There is a cost associated with this scheme, however, because with low-powered incentives, it also fails to induce desirable productive effort from low-ability agents. For expositional purposes, we refer to this as the fast-track scheme since this scheme rewards the first-period winner more heavily. An illuminating property of this type of incentive scheme is that it always induces productive effort from high-ability agents at a potential cost that (discouraged) low-ability agents exert no effort in the second period. Note that since effort is positively related to ability, i.e., more able agents exert more effort, effective inputs from each agent are highly diversified under the fast-track scheme. The fast-track scheme then implies a less steep wage-tenure profile and a more dispersed earnings distribution among employees.

Besides the fast-track scheme, there is another, more novel, way to mitigate sabotage in this setting. Now consider a tournament which places more weight on the ranking in the second period: that is, low-powered incentives are provided in the first period, complemented with high-powered incentives in the second. Since more is at stake in the second period, high-ability agents have an incentive to conceal their private information in order not to get too much ahead of others. For expositional purposes, we refer to this as the late-selection scheme since this scheme rewards the second-period winner more heavily. An illuminating property of this type of incentive scheme is that it always induces productive effort from low-ability agents at a potential cost that high-ability agents exert no effort in the first period. It is perhaps worth emphasizing that this (that is, only low-ability agents exert effort) can never happen in the static setting under the assumption that the marginal return to effort is larger for high-ability agents. Note that effort is negatively related to ability, which makes effective inputs from each agent fairly homogenized under this scheme. The late-selection scheme then implies a steeper wage-tenure profile and a more compressed distribution of earnings among employees.

The present model thus provides a potential explanation for both fast track and late selection in a unified framework, from a previously unexplored perspective which rests on the possibility of sabotage. At the same time, the results obtained here also yield several testable implications and predictions for differences in promotion patterns, wage dynamics and dispersion across countries as well as across industries. One factor that proves to be critical, among some others, is the nature of production technologies. In general, the fast-track scheme becomes the preferred choice when the production process exhibits strong complementarities between ability and effort and hence values diversity, rather than homogeneity, in inputs. The late-selection scheme is, on the other hand, more likely to be efficient when the production process values homogeneity in inputs.

Besides these implications, the paper also raises a theoretical issue of some interest, identifying a dynamic inefficiency that arises from the mere possibility of sabotage. It is intuitively clear that sabotage generally entails a significant loss of efficiency. An obvious cost of sabotage is that it substitutes for more productive alternatives. An occurrence of sabotage also leads to another inefficiency, as pointed out by Chen (2003), that the principal may fail to select the most deserving agent in the presence of sabotage activities. Note that both of these costs are direct consequences of sabotage since they arise only when sabotage actually takes place on the equilibrium path. On the contrary, we argue that the possibility of sabotage gives rise to another type of dynamic inefficiency that distorts the agents' behavior in order to manipulate information. The dynamic inefficiency suggested here is less visible or hidden in that this type of inefficiency arises even when sabotage does not actually take place on the equilibrium path.

The paper is related to an extensive literature that examines the optimal timing of promotion or compensation with different approaches and focuses.[1] Just to name a few, one of the most influential approaches to explain delayed compensations is the incentive approach by Lazear (1979), who posits that compensations should be delayed to maintain career incentives. The learning approach emphasizes that some attributes of a worker are not immediately observable, and promotions are necessarily delayed as the evaluator needs to collect information about the worker.[2] The signaling approach, most notably by Waldman (1984), states that promotions in early stages tend to be inefficiently few because a promotion signals the worker's ability, which in turn raises the retention wage.[3] Prendergast (1992) argues that Japanese firms are able to delay promotion since (i) important decisions are delegated to middle managers, and/or (ii) workers are limited in their mobility in the labor market. Morita (2004) provides an alternative explanation for a steeper wage-tenure profile, often observed in Japan, by attributing it to the relative importance of managerial ability. As an explanation for fast track, on the other hand, Meyer (1982) constructs a two-period model and shows that it is optimal to bias the second-period tournament in favor of the first-period winner.

[1] For more extensive surveys, see Gibbons and Waldman (1999a) and Prendergast (1999).

[2] Related to this, Ishida (2006) shows that seniority becomes even more crucial for promotion decisions when the principal (evaluator) possesses more accurate information about the worker's innate attributes than the worker himself.

[3] Owan (2004) applies this approach to explain the US-Japanese differences by complementing it with firm-specific human capital. Similarly, Ishida (2004) also applies this approach by complementing it with signaling through schooling. Bernhardt (1995) extends this signaling approach to account for fast-track promotions as well as other stylized facts of internal labor markets.

Gibbons and Waldman (1999b, 2006) show that workers who receive large wage increases early in their stay at one level of the hierarchy are promoted quickly to the next level because workers who receive large wage increases are likely to be those with high ability.

On the theoretical side, the paper is also related to works that focus on negative aspects of relative compensation schemes. A seminal paper on sabotage in a tournament is Lazear (1989) who suggests that the pay compression can be effective in mitigating sabotage activities. This logic plays an important role in the fast-track scheme. The paper is most closely related to Chen (2003) who shows in a static tournament that able members are likely to be subject to sabotage attacks, illustrating an inefficiency that the most able member might not have the best chance of being promoted. The focus of the present paper differs from Chen (2003) as it is placed on dynamic interactions between the agents and the ways in which to mitigate sabotage activities.[4] Chen (2005) also argues that external recruitment can be utilized as an effective tool against sabotage activities.[5] The paper is also related to a literature that deals with collusion under relative performance evaluation. Along this line, Ishiguro (2004) shows that the principal can prevent collusion by offering asymmetric contracts, even though agents are symmetric with respect to productive abilities. This paper is similar in spirit as it seeks for sabotage-proof contracts, instead of collusion-proof, where agents have no incentive to exert sabotage effort.

2 The model

2.1 Environment

Consider a two-period model in which a principal (female) hires two agents (male), denoted by $i \in \{1, 2\}$, to produce output. Each agent differs in his innate ability $\eta_i \in \{L, H\}$ which eventually amounts to a difference in the propensity for sabotage. We say that agent i possesses high (low) ability if $\eta_i = H$ ($\eta_i = L$). The prior probability of the ability type is denoted as $\text{prob}\{\eta_i = H\} = \theta_i \in [0, 1]$ where we allow for $\theta_1 \neq \theta_2$. Without loss of generality, we assume $\theta_1 \geq \theta_2$. The ability type is the agent's private information while the prior probabilities are common knowledge. Define $\mu_{i,t} \equiv \text{prob}\{\eta_{-i} = H \mid \Omega_{i,t}\}$, $i \neq -i$ as agent i's belief about the other agent's ability type at the beginning of period t, based on his information set $\Omega_{i,t}$. By construction, $\mu_{i,1} = \theta_{-i}$, $i = 1, 2$.

[4] Chen (2003) discusses several schemes to mitigate sabotage activities, although not in a formal analysis.

[5] In this paper, we thus seek for ways to prevent sabotage in situations where external recruitment cannot be used effectively, possibly for reasons suggested in Chan (1996) or Waldman (2003).

2.2 Production

In each period t, each agent must decide his effort levels $a_{i,t} = (e_{i,t}, d_{i,t}) \in \{0,1\}^2$, where $e_{i,t}$ denotes the level of productive effort while $d_{i,t}$ denotes the level of (destructive) sabotage effort. A departure from the conventional setup is the additional option of sabotage. The cost of effort is denoted as $c(e_{i,t}, d_{i,t})$ where

$$c(0,0) = 0, \ c(1,0) = c, \ c(0,1) = (1+\lambda)c, \ c(1,1) = C.$$

We assume that $\lambda > 0$, i.e., sabotage effort is more costly than productive effort.[6] We also assume that C is so large that it is never optimal to choose $e_{i,t} = d_{i,t} = 1$ (exerting both productive and sabotage effort simultaneously), e.g., because there is simply not enough time to do both.

The individual productivity of each agent depends on his ability and effort choice as well as the other agent's effort choice. The individual productivity is denoted as $y_{i,t} = y^{\eta_i}(e_{i,t}, d_{-i,t})$ where

$$y^H(1,0) = h, \ y^H(0,0) = y^H(1,1) = m, \ y^H(0,1) = l,$$
$$y^L(1,0) = m, \ y^L(1,1) = y^L(0,0) = l, \ y^L(0,1) = 0.$$

We assume that $h - m > m - l > l > 0$, i.e., productive effort is complementary to ability (the marginal value of productive effort is larger for the high-ability type). This specification implies that the role of sabotage effort is to negate the other agent's productive effort, which gives comparative advantage in sabotage to the low-ability type.[7]

2.3 Information

The crux of the model is its information structure. In this model, we consider a situation where the agents have access to more precise information regarding their productivities than the principal.[8] More precisely, each agent i can observe the opponent's productivity $y_{-i,t}$ in each period (although not the ability type nor the effort choice). The principal, on the other

[6] The cost of sabotage is interpreted broadly here, including various psychological costs (of engaging in anti-social behavior) as well as potential losses of reputation and more explicit punishments when the agent gets caught. The cost of sabotage also reflects how closely the agents interact with each other: the cost tends to be small when they work closely with each other.

[7] The fact that $y^H(0,0) = y^L(1,0)$ is also important for the subsequent analysis as it indicates that the high-ability type can pretend to be of the low-ability type by exerting no effort. Although it might appear somewhat restrictive to assume this, it simply implies that the high-ability type can always control his effort level to keep pace with others. This assumption is a simple and tractable way to capture this aspect.

[8] This is based on the premise that colleagues, who interact on the daily basis, typically have more information about each other than their boss. Note, however, that this assumption is not essential: virtually the same argument should hold even when the agents' information is noisy and imperfect although the analysis would be exceedingly complicated.

hand, can only observe each agent's output, which is an imperfect signal of the productivity. Let $\hat{y}_{i,t}$ denote the output which is given by

$$\hat{y}_{i,t} = y_{i,t} + \varepsilon_{i,t}.$$

The noise term $\varepsilon_{i,t}$ is drawn i.i.d. from some well-defined distribution with strictly positive density on its support and has zero mean.

Given the true productivities $(y_{1,t}, y_{2,t})$, the probability that agent 1 outperforms agent 2, as observed by the principal, is given by

$$\text{prob}\{y_{1,t} - y_{2,t} \geq \varepsilon_{2,t} - \varepsilon_{1,t}\} \equiv G(y_{1,t} - y_{2,t}).$$

Since each agent's output is subject to a shock drawn independently from the same distribution, G has the following properties: (i) it is strictly increasing in its argument; (ii) $G(0) = 0.5$; (iii) $\lim_{x\to-\infty} G(x) = 0$ and $\lim_{x\to\infty} G(x) = 1$; and (iv) the corresponding density g is single-peaked and symmetric around zero, i.e., $\text{argmax}_x g(x) = 0$ and $g(x) = g(-x)$ for all x. These properties further imply that $G(x) + G(-x) = 1$ and $G(x) - G(0) > G(x + x_0) - G(x_0)$ for any $x > 0$ and $x_0 > 0$, which we repeatedly use in the subsequent analysis.

2.4 Contracts

We place several restrictions on the class of contracts that the principal can offer. First, the principal faces a liquidity constraint so that the wages paid to the agents must be nonnegative.[9] Second, the relative ranking is the only contractible measure of performances, and hence the principal must rely on a rank-order tournament to motivate the agents.[10] This actually implies two more restrictions: (i) the principal cannot condition the second-period wages on the first-period outcome; (ii) she is not allowed to offer asymmetric contracts contingent on each agent's identity.[11] Since the principal must rely on relative performance

[9] In a typical case, the analysis would be much simpler without the liquidity constraint, because the principal can extract all the rents from the agents. This is not the case in a model with heterogeneous agents: in such a case, as in our model, the participation constraints differ across the agents, some of which do not bind at the optimum.

[10] When agents are asymmetric with respect to the productivity, it is often optimal for the principal to set up a tournament with a handicap. See Lazear and Rosen (1984). Meyer (1992) also shows that the possibility of biased tournaments have important dynamic implications. Sice our main concern, i.e., strategic information revelation due to sabotage concerns, is independent of this issue, we assume that it is not feasible to bias the tournament one way or the other in order to isolate this effect.

[11] One reason for this is that the relative ranking is not verifiable to a third party. In such a case, if a contract is not symmetric where one agent is offered a low prize while the other is offered a high prize, the principal is *ex post* always tempted to choose the one with the low prize to save the wage cost: note that at this point, the effort levels are already chosen, so that the principal has no strong incentives to honestly choose the "right" winner.

measures to motivate the agents, sabotage evidently becomes a serious issue for all parties involved in the transactions.

A contract that covers both periods can then be written as $W \equiv (w_1, z_1, w_2, z_2)$ where w_t and z_t are the wages for the winner and the loser, respectively, in period t. We assume that the principal retains full bargaining power to propose a contract (a take-it-or-leave-it offer). Then, under the liquidity constraint, the loser always receives the minimum, which is zero, because the incentive effects are determined solely by the wage spread $w_t - z_t$; see Appendix A for more details. Consequently, any optimal contract can generically be written as (w_1, w_2).

2.5 Preferences

Both the principal and the agents are assumed to be risk-neutral and maximize the sum of the expected payoffs with no time discounting. Each agent maximizes the expected wage minus the cost of effort over the two periods. In period 2, each agent simply solves a static problem by maximizing the expected instantaneous payoff conditional on the belief about the opponent's type. The effort choice in period 1, on the other hand, may change the opponent's belief, and each agent must take into account this dynamic effect.

The principal maximizes the expected profit π_t minus the wage costs over the two periods. The gross profit for the principal, excluding the wage costs, is given by

$$\pi_t = y_{1,t} + y_{2,t} - D(d_{1,t} + d_{2,t}).$$

Here, we assume that sabotage effort, once exerted, may generate negative externalities within the firm and lowers the firm's profit by $D \geq 0$.[12] Taking the ability distribution (θ_1, θ_2) as given, the principal designs a contract (a tournament) so as to maximize

$$\max_{(w_1, z_1, w_2, z_2)} \quad \sum_{t=1}^{2} \Big(E(\pi_t \mid \theta_1, \theta_2) - w_t - z_t \Big),$$

subject to the agents' IC constraints which are discussed in more detail in Appendix A.

3 Optimal contracts

The possibility of sabotage in a dynamic setting, along with information asymmetry, potentially yields a plethora of equilibria, including many uninteresting ones for the purpose of

[12] There are several conceivable channels through which the occurrence of sabotage adversely affects the firm's performances: for instance, a firm which allows sabotage to take place with positive probability may face a difficulty in attracting high-ability agents, and that could easily lower the firm's productivity in the long run.

the current analysis. For this reason, we make three assumptions on the parameter space to focus our attention to more relevant cases.

Assumption 1: λ is neither too large nor too low. More precisely,

$$\frac{G(h-l)-G(m-l)}{G(m-l)-G(0)} > 1+\lambda \geq \frac{\theta_1(G(h-l)-G(m-l))+(1-\theta_1)(G(m-l)-G(0))}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))}.$$

Assumption 2: c is sufficiently small, so that it is always optimal to induce productive effort whenever it is feasible to do so.

Assumption 3: D is sufficiently large, so that it is never optimal to let sabotage take place on the equilibrium path.

Some remarks are in order. First, the possibility of sabotage has no bite when the cost of sabotage is either too large or too small: if the cost is too large, the option of sabotage actually plays no role; if the cost is too small, it is not possible to induce productive effort by any incentive scheme.[13] Assumption 1 rules out these apparently uninteresting cases. Assumption 2 is fairly standard, meaning that productive effort is sufficiently valuable for the principal. As we will see, however, this does not imply that the optimal contract always implements $e_{i,t}=1$ because there are cases where it is simply not feasible to do so. Finally, Assumption 3 states that the external negative effect of sabotage is prohibitively large, which allows us to focus on the class of sabotage-proof contracts.[14]

3.1 Optimal contracts when sabotage is not a viable option: a benchmark

We first consider a benchmark case where sabotage effort is not a viable option in order to single out the impact of potential sabotage activities. Given that the cost of effort is sufficiently small (Assumption 2), the optimal contract is the one that implements $e_{i,1} = e_{i,2} = 1$ for any possible contingency with the minimum cost. We first establish the following result (all the proofs are relegated to Appendix B).

Proposition 1 *Suppose that sabotage effort is not a viable option. The optimal contract is then given by*

$$w_1 = w_1^* \equiv \frac{c}{\theta_1[G(h-l)-G(h-m)]+(1-\theta_1)[G(m-l)-G(0)]}, \; w_2 = w_2^* \equiv \frac{c}{G(h-l)-G(h-m)},$$

which implements $e_{i,1} = e_{i,2} = 1$ on any equilibrium path.

[13] To be more precise, when the cost is smaller than the lowerbound (but large than one), there exists no contract that can induce productive effort from the low-ability type for a given prior. The optimal contract in this case is the one that induces productive effort from the high-ability type and no effort from the low-ability type.

[14] We say that a contract is sabotage-proof if sabotage never takes place on the equilibrium path. This assumption plays only an auxiliary role because sabotage effort naturally entails an inefficient use of resources, and the optimal contract is indeed sabotage-proof in many cases even without this assumption.

3.2 Optimal contracts with sabotage activities

We now introduce sabotage effort into the model and see how the mere possibility of sabotage activities alters the optimal structure of incentives. The additional option of sabotage imposes an additional constraint which immensely influences the agents' behavior.

Proposition 2 *Suppose that it is optimal for the high-ability type to exert productive effort in period 2. Then, there exists some threshold* $\bar{\mu} \in (0,1)$ *such that the low-ability type never exerts productive effort in period 2 for* $\mu_{i,2} > \bar{\mu}$ *if*

$$\frac{G(h-l) - G(m-l)}{G(m-l) - G(0)} > 1 + \lambda.$$

The proposition 2 implies that, under Assumption 1, there exists no contract that can implement $e_{i,1} = e_{i,2} = 1$ with probability one. This is because if $e_{i,1} = 1$ for $i = 1, 2$, the agents can perfectly identify the opponent's ability type, and hence $\mu_{i,2} \in \{0, 1\}$. The situation described in Proposition 2 is then bound to arise whenever the low-ability type is matched against the high-ability type. In this situation, as we have already seen, it is not possible to induce productive effort from both types, indicating that the principal is unable to implement the first-best effort. The main problem here is the possibility of sabotage activities that might take place in period 2. We now consider two distinct incentive schemes that can suppress this possibility.

One possible scheme is to provide low-powered incentives in period 2 and directly suppress sabotage activities, as suggested by Lazear (1989): fast track is used as as a way to protect the high-ability type from sabotage attacks from colleagues by designating the winner early on. The well-known cost associated with this scheme is that with low-powered incentives, the principal may also fail to induce desirable productive effort. An illuminating property of this type of incentive scheme is that it always induces productive effort from the high-ability type. We refer to this incentive scheme as the fast-track scheme since more weight is placed on early performances.[15]

The fast-track scheme: This scheme offers high-powered incentives in period 1, followed by low-powered incentives in period 2. Sabotage activities are circumvented by leaving smaller rents for the low-ability type to exploit in period 2.

In this dynamic setting, there is another way to circumvent sabotage activities in period 2. Suppose that the principal provides low-powered incentives in period 1, complemented

[15] Under this scheme, one's lifetime income is influenced more by early performances, implying that winners (well-paid agents) are designated early on. Throughout the analysis, we use the term "fast track" strictly in this sense.

with high-powered incentives in period 2. Since more is at stake in period 2, the high-ability type would rather choose to exert no effort in order to conceal his ability type. If $\bar{\mu} > \theta_1$, this can prevent sabotage activities because the productivity reveals no relevant information and hence $\mu_{i,2} = \theta_{-i}$ with probability one. An illuminating property of this type of incentive scheme is that it always induces productive effort from the low-ability type, in contract to the fast-track scheme. We refer to this incentive scheme as the late-selection scheme since more weight is placed on late performances.

The late-selection scheme: This scheme offers low-powered incentives in period 1, complemented with high-powered incentives in period 2. Sabotage activities are circumvented by letting the high-ability type keep pace with the low-ability type.

One can show that the optimal contract must belong to either class under Assumptions 2 and 3. To see this, note that the ability type is not identifiable only when the late-selection scheme is implemented; for any other cases, the ability type is perfectly identified, leading to the same outcome in period 2.[16] Provided that effort is sufficiently valuable, it is then optimal for the principal to implement the fast-track scheme (which induces productive effort from both types in period 1), if not the late-selection scheme (which induces productive effort only from the low-ability type in period 1).

It is intuitively clear that revealing no information would be of no help if θ_1 is very large to begin with: in this case, the dispersion in the perceived ability is too large, from the low-ability type's viewpoint, and sabotage would occur anyway. In the current setup, therefore, the late-selection scheme works only if θ_1 is sufficiently small. If the late-selection scheme cannot be made sabotage-proof, the fast-track scheme, which can always be made sabotage-proof under Assumption 1, must emerge as the optimal choice.

Proposition 3 *If*

$$\theta_1(G(h-m) - G(m-l)) \geq (1-\theta_1)(G(m-l) - G(0)), \tag{1}$$

the late-selection scheme cannot be made sabotage-proof. The fast-track scheme is optimal where the optimal contract is given by

$$w_1 = w_1^{\mathrm{FT}} \equiv \frac{c}{\theta_1(G(h-l) - G(h-m)) + (1-\theta_1)(G(m-l) - G(0))}, \ w_2 = w_2^{\mathrm{FT}} \equiv \frac{c}{G(h-l) - G(m-l)}.$$

Under the optimal fast-track scheme, (i) the high-ability type exerts productive effort in both periods; (ii) the low-ability type exerts productive effort in period 1 and no effort in period 2.

[16] To be more precise, we need some (weak) restrictions on the beliefs off the equilibrium path. See Appendix B for more detail.

When θ_1 is sufficiently small, however, a carefully designed late-selection scheme can prevent sabotage from taking place on the equilibrium path. In this case, both schemes can be sabotage-proof, and which scheme emerges as optimal depends on several exogenous parameters of the model. We now state the following which is the main result of this paper.

Proposition 4 *Suppose that (1) does not hold, so that both schemes can be made sabotage-proof. Moreover, define $\bar{\theta} \equiv (\theta_1 + \theta_2)/2$. Then, if*

$$\Delta(\bar{\theta}) \equiv \bar{\theta}h + (1-\bar{\theta})l - m \geq \frac{(G(m-l) - G(0)) - (G(h-l) - G(m-l))}{2(G(h-l) - G(m-l))(G(m-l) - G(0))}c, \tag{2}$$

the fast-track scheme is optimal where the optimal contract is given by $w_1 = w_1^{\text{FT}}$ and $w_2 = w_2^{\text{FT}}$. Otherwise, the late-selection scheme is optimal where the optimal contract is given by

$$w_1 = w_1^{\text{LS}} \equiv \frac{c}{G(m-l) - G(0)},\ w_2 = w_2^{\text{LS}} \equiv \frac{c}{\theta_1(G(h-l) - G(h-m)) + (1-\theta_1)(G(m-l) - G(0))}.$$

Under the optimal late-selection scheme, (i) the high-ability type exerts no effort in period 1 and productive effort in period 2; (ii) the low-ability type exerts productive effort in both periods.

Under the fast-track scheme, since the ability type is perfectly identifiable in period 2, the principal must compress the pay structure in order to mitigate potential sabotage activities. This obviously comes at a cost: with low-powered incentives, it also fails to induce desirable productive effort from the low-ability type as well. Compared with the case without sabotage activities, the fast-track scheme provides weaker incentives in period 2, i.e., $w_2^* > w_2^{\text{FT}}$, leading to a less steep wage-tenure profile. Under the late-selection scheme, on the other hand, the principal starts with weaker incentives ($w_1^* > w_1^{\text{LS}}$) and raises the stake later on ($w_2^{\text{LS}} > w_1^{\text{LS}}$), leading to a steeper wage-tenure wage profile.

Moreover, the optimal form of contract also has some impact on the wage dispersion across workers. Under the fast-track scheme, effort is positively related to ability, and the distribution of effective inputs is more diverse; under the late-selection scheme, effort is negatively related to ability, and the distribution of effective inputs is more homogenized. As a consequence, the earnings distribution tends to be more compressed under the late-selection scheme than under the fast-track scheme. To see this, we define the wage dispersion by the difference in the expected lifetime income between the two ability types. Letting $WD^j(\Theta)$, $j = FT, LS$, denote the wage dispersion under each respective scheme as a function of $\Theta \equiv (\theta_1, \theta_2, h, m, l, c)$, we obtain

$$\begin{aligned} WD^{\text{FT}}(\Theta) &\equiv G(h-m)w_1^{\text{FT}} + G(h-l)w_2^{\text{FT}} - (1-G(h-m))w_1^{\text{FT}} - (1-G(h-l))w_2^{\text{FT}} \\ &= (2G(h-m)-1)w_1^{\text{FT}} + (2G(h-l)-1)w_2^{\text{FT}}, \end{aligned}$$

$$WD^{\mathrm{LS}}(\Theta) \equiv G(0)w_1^{\mathrm{LS}} + G(h-m)w_2^{\mathrm{LS}} - (1-G(0))w_1^{\mathrm{LS}} - (1-G(h-m))w_2^{\mathrm{LS}}$$
$$= (2G(h-m)-1)w_2^{\mathrm{LS}}.$$

Note that $WD^{\mathrm{FT}}(\Theta) - WD^{\mathrm{LS}}(\Theta) = (2G(h-l)-1)w_2^{\mathrm{FT}} > 0$ for any Θ, i.e., if an economy switched from the late-selection scheme to the fast-track scheme, it always ends up with a larger wage dispersion.[17]

4 Implications: fast track versus late selection

When (1) fails to hold, the optimal form of contract is determined by such exogenous parameters as $(\theta_1, \theta_2, h, m, l, c)$ as well as the noise distribution G. Condition (2) in Proposition 5 is especially crucial, yielding several predictions and implications for promotion patterns, earnings dynamics and earnings dispersion. We in particular raise three key determining factors – the diversity of the population, knowledge intensity of the underlying production process and the accuracy of performance evaluation – and discuss them in turn.

4.1 Diversity vs homogeneity

A critical determinant is the difference between $h-m$ (the marginal output of productive effort for the high-ability type) and $m-l$ (the marginal output for the low-ability type), which captures the degree of complementarity between ability and effort in the production process. When the degree of complementarity is large, the production process tends to value diversity: the marginal value of effort depends heavily on the ability type, so that the high-ability type's effort is much more critical than the low-ability type's. As $h-m$ approaches $m-l$, on the other hand, the difference in innate ability gradually loses its significance and the production process tends to value homogeneity: the marginal value does not depend strongly on the ability type, so that the low-ability type's effort gains its importance in a relative sense.

To see when condition (2) holds, it is perhaps clear to focus on the signs of $\Delta(\bar{\theta})$ and $\Gamma \equiv (G(m-l) - G(0)) - (G(h-l) - G(m-l))$. Fix h and l arbitrarily and let m move between l and $(h+l)/2$. It is then easy to see that $\Delta(\bar{\theta})$ is strictly decreasing in m while Γ is strictly increasing. Moreover, as $m \to l$, $\Delta(\bar{\theta}) > 0$ and $\Gamma < 0$ so that the fast-track scheme emerges as the optimal option. At the other end, as $m \to (h+l)/2$, $\Delta(\bar{\theta}) < 0$ if $\bar{\theta} < 0.5$ and

[17] This thought experiment thus indicates that there is a force in the fast-track scheme to enlarge the wage dispersion. Obviously, what we really need to compare is the wage dispersion when the fast-track scheme is optimal with that when the late-selection is optimal. To be more precise, consider two distinct economies, A and B, each of which is endowed with a different set of parameters $\Theta^j = (\theta_1^j, \theta_2^j, h^j, m^j, l^j, c^j)$, $j = A, B$. Suppose further that the fast-track scheme is optimal in economy A while the late-selection scheme is optimal in economy B. What we need to show is then $WD^{\mathrm{FT}}(\Theta^{\mathrm{A}}) > WD^{\mathrm{LS}}(\Theta^{\mathrm{B}})$. Although there are some exceptional cases, the fact that $WD^{\mathrm{FT}}(\Theta) > WD^{\mathrm{LS}}(\Theta)$ for any given Θ implies that it holds for a wide range of parameters.

$\Gamma > 0$ so that the late-selection scheme is preferred. In general, the fast-track (late-selection) scheme is more effective when the underlying technology values diversity (homogeneity) in inputs. An important question to ask at this juncture is then what this difference between $h - m$ and $m - l$ embodies. Here, we provide two possible interpretations of this result in order to draw more practical implications of the present model.

Production technologies: The first interpretation is that the degree of complementarity captures the nature of the underlying production technologies. For instance, the production process tends to value diversity when the productivity depends more crucially on the best idea or the luckiest draw and hence disproportionately reflects the input of a few highly talented individuals. This feature is more common in industries such as software, fashion and entertainment which value new ideas and creativity (the so-called "superstar" industry), and/or in firms where important decisions are made by a selected few individuals (centralized organizations).[18] Under these circumstances, it is evidently costly to force highly productive workers to keep pace with others because the overall productivity hinges crucially on their inputs. The fast-track scheme is effective in this sense since it effectively protects those productive workers from sabotage attacks, thereby allowing them to be relatively free of sabotage concerns in early stages.

The late-selection scheme, on the other hand, tends to be the better option when the underlying technology values homogeneity in inputs. This feature is more common in industries such as automobile and consumer electronics which require careful and precise implementation of tasks throughout the entire production process, and/or in firms where more important decisions are delegated to middle managers and shop-floor workers (decentralized organizations).[19] Under these circumstances, it is certainly costly to have a group of discouraged workers since their presence could significantly pull down the overall productivity. The late-selection scheme is effective in this sense since it minimizes a chance that any given worker is totally discouraged, thereby keeping workers' morals relatively high throughout their career.

Human capital distribution: A large difference between $h - m$ and $m - l$ also means that the marginal value of effort is much higher for the high-ability type than for the low-ability type. Another, more straightforward, interpretation is then that it reflects the distribution of talent or human capital: the difference is more likely to be large when the population is more diverse, from top to bottom, in terms of the market productivity. The analysis then reveals that the fast-track scheme works better with a more diverse group of workers.

[18] Rosen (1981) raises entertainment as a typical example of a superstar industry. Also, Grossman and Maggi (2000) raise software and fashion as industries where the outstanding performance of one or a few persons is critical.

[19] Kremer (1993) argues the importance of implementing all the tasks equally well in manufacturing.

4.2 Knowledge intensity

For any given (h, m, l), an increase in $\bar{\theta}$ raises $\Delta(\bar{\theta})$ and thus favors the fast-track scheme. This suggests that the fast-track scheme works better in firms which predominantly comprise of highly qualified workers and professionals. Firms in knowledge-intensive industries or those that require professional skills supposedly fall into this category. When $\bar{\theta}$ is relatively high to begin with, it is more costly to let productive workers keep pace with others, and the fast-track scheme emerges as the better alternative to deal with potential sabotage.

4.3 Evaluation accuracy

The last factor is the shape of the noise distribution G, which reflects the accuracy of performance evaluation. The distribution degenerates towards the mean as the variance of the noise terms $\varepsilon_{i,t}$ converges to zero. This means that an improvement in the accuracy of evaluation tends to raise Γ, thereby favoring the late-selection scheme. When performance evaluation is noisier and less accurate, on the other hand, the noise distribution has fat tails and Γ eventually becomes negative, thus favoring the fast-track scheme.

The intuition behind this result can be seen by comparing the first-period wage under the late-selection scheme and the second-period wage under the fast-track scheme.[20]. In the first period under the late-selection scheme, the optimal contract must induce productive effort from the low-ability type when the high-ability type exerts no effort. Since the high-ability type exerting no effort is equivalent to the low-ability type exerting productive effort, this is effectively a tournament between two homogeneous agents from the low-ability type's viewpoint. On equal footing, an improvement in the accuracy of evaluation raises the marginal value of effort (the marginal increase in the winning probability) and hence reduces the agency cost. This draws clear contrast to the second period under the fast-track scheme, where the optimal contract induces productive effort only from the high-ability type. Here, the situation is a tournament with heterogeneous agents: an improvement in the accuracy of evaluation lowers the marginal value of effort and hence raises the agency cost, because the high-ability type is more likely to win the competition anyway without costly effort.

5 Conclusion

This paper constructs a two-period model of a tournament to illustrate dynamic inefficiencies that arise from the possibility that agents may engage in sabotage activities. We first show that when sabotage is a viable option, it is impossible to implement the first-best effort under

[20]Since $w_2^{\text{LS}} = w_1^{\text{FT}}$, the difference in the agency cost is given by $w_1^{\text{LS}} - w_2^{\text{FT}}$

fairly plausible circumstances. Given this result, we then show that the wage compression needs to occur only in either the first or second period, depending on the difference between the high-ability and low-ability types. The fast-track scheme which rewards the first-period winner more heavily is optimal when the production process values diversity in inputs; the late-selection scheme which rewards the second-period winner more heavily is optimal when the low-ability type it values homogeneity. We argue that this result provides a mechanism through which both fast track and late selection arise in a unified framework from a previously unexplored angle.

As a final note, the model abstracts away from many potentially critical aspects since the analysis is already highly complicated with an additional option of sabotage and dynamic information revelation. For this reason, there are several avenues to extend the current analysis. For instance, it might be interesting to extend the model to a multi-period setting to explore more thoroughly the dynamic aspect of incentive provision. In this case, it is expected that the speed of learning matters in that the fast-track scheme becomes more valuable as the agents gain information about each other more slowly: in fact, if the speed of learning is so slow that the agents learn nothing about each other until they retire, there is no point in adopting the late-selection scheme. Conversely, this implies that the late-selection scheme tends to work better in firms where workers interact closely with each other.[21]

Another possibility is to relax the assumption that the principal can commit to a long-term contract (no renegotiation at the interim stage). This assumption effectively rules out the effects that arise from learning between the principal and the agents, as the principal is unable to utilize any information acquired subsequently.[22] Although we believe that this commitment assumption reasonably captures reality,[23] there is much more to see when the principal cannot commit to the second-period contract from the outset.

Finally, and perhaps more importantly, we also exclude outside labor markets altogether from consideration. A particularly intriguing case arises when a promotion (or the ranking) is observable to outside parties and hence works as a signal. Along this line, MacLeod and Malcomson (1988) show that workers may exert extra effort to secure a promotion when it is observed by the outside market. The question is then what would happen with the option of sabotage, which could pose a number of interesting scenarios. First, with a similar logic,

[21]This might partially explain the popularity of the late-selection scheme among Japanese firms which tend to place more emphasis on team work.

[22]Theoretical models along this line include Harris and Holmstrom (1982) and Waldman (1984) among many others. Chiappori et al. (1999) provide a general empirical framework and evidence that learning, along with downward rigidity, is an important characteristic of wage formation.

[23]In many cases, wages are tied to ranks and do not vary substantially with individual traits once ranks are fixed.

one can conjecture that the presence of the outside market could intensify the incentive for low-ability agents to exert sabotage effort. In addition, if sabotage is expected to take place on the equilibrium path, a promotion loses its value as a signal (of the promoted agent's innate ability), which may diminish the incentive to exert productive effort.[24] On the other hand, with available outside options, high-ability agents may deviate by exerting productive effort and then leave the firm. This could make the late-selection scheme less enforceable, indicating that the availability of outside options is an important factor behind it. In any event, the addition of outside markets is a realistic extension which could yield substantial implications, and it is an important aspect to be considered in future.

Appendix A: the IC constraints

By assumption, we rule out the possibility that $e_{i,t} = d_{i,t} = 1$. This implies that each agent has three alternatives (productive effort, sabotage effort and no effort), and three constraints are hence sufficient to characterize the preferences among them. Since we only look at sabotage-proof contracts, the opponent's type-contingent equilibrium effort choice in each period can be written as $(e^H_{-i,t}, e^L_{-i,t})$. Let $P^\eta_{e,d}(\mu_{i,t}; e^H_{-i,t}, e^L_{-i,t})$ denote the probability that agent i outperforms the opponent, conditional on the ability type η and the effort choice (e, d), taking the opponent's effort profile and the belief as given. To save notation, we sometimes write $P^\eta_{e,d}(\mu)$ or simply $P^\eta_{e,d}$ wherever it is not confusing. Given this definition, we can show that

$$P^\eta_{e,d}(\mu; e^H_{-i,t}, e^L_{-i,t}) = \mu G(y^\eta(e, 0) - y^H(e^H_{-i,t}, d)) + (1 - \mu)(G(y^\eta(e, 0) - y^L(e^L_{-i,t}, d)).$$

Let $u^\eta_t(e, d)$ denote the expected payoff in period $t + 1$, conditional on the ability type and the current effort choice: note that since period 2 is the last period, $u^\eta_2(e, d) = 0$ for any η and (e, d). Given this, the IC constraint in period t is obtained as follows. First, productive effort is preferred to no effort if

$$P^\eta_{1,0} w_t + (1 - P^\eta_{1,0}) z_t - c + u^\eta_t(1, 0) \geq P^\eta_{0,0} w_t + (1 - P^\eta_{0,0}) z_t + u^\eta_t(0, 0). \tag{A.1}$$

Second, productive effort is preferred to sabotage effort if

$$P^\eta_{1,0} w_t + (1 - P^a_\eta) z_t - c + u^\eta_t(1, 0) > P^\eta_{0,1} w_t + (1 - P^\eta_{0,1}) z_t - (1 + \lambda) c + u^\eta_t(0, 1). \tag{A.2}$$

[24] At the same time, as the signaling value of promotion diminishes, the firm may need to pay less to retain the promoted worker.

Finally, no effort is preferred to sabotage effort if

$$P^{\eta}_{0,0}w_t + (1 - P^{\eta}_{0,0})z_t + u^{\eta}_t(0, 0) > P^{\eta}_{0,1}w_t + (1 - P^{\eta}_{0,1})z_t - (1 + \lambda)c + u^{\eta}_t(0, 1). \quad \text{(A.3)}$$

Notice that the IC constraints depend only on the wage spread $w_t - z_t$, and hence we can conclude that $z_t = 0$.

We need to construct a tournament that induces either productive effort or no effort. Although the agents face dynamic incentives in period 1 and static incentives in period 2, static incentives are sufficient for most cases we consider,[25] and hence they repeatedly appear throughout the proofs. To make them as readable and comprehensible as possible, we will stick with the following notations to represent the static IC constraints.

Taking the agent's own ability η, the belief μ, and the opponent's type-contingent effort choice as given,

(i) Productive effort is preferred to no effort if $w_t \geq w^{\eta}_{\text{PN}}(\mu; e^H_{-i,t}, e^L_{-i,t})$;

(ii) Productive effort is preferred to sabotage effort if $w^{\eta}_{\text{PS}}(\mu; e^H_{-i,t}, e^L_{-i,t}) > w_t$;

(iii) No effort is preferred to sabotage effort if $w^{\eta}_{\text{NS}}(\mu; e^H_{-i,t}, e^L_{-i,t}) > w_t$.

To induce productive effort, therefore, it must be that

$$w^{\eta}_{\text{PS}}(\mu; e^H_{-i,t}, e^L_{-i,t}) > w_t \geq w^{\eta}_{\text{PN}}(\mu; e^H_{-i,t}, e^L_{-i,t}).$$

Similarly, to induce no effort, it must be that

$$\min\{w^{\eta}_{\text{PN}}(\mu; e^H_{-i,t}, e^L_{-i,t}), w^{\eta}_{\text{NS}}(\mu; e^H_{-i,t}, e^L_{-i,t})\} > w_t.$$

These notations will appear repeatedly throughout this Appendix.

Appendix B: the proofs

Proof of Proposition 1: We first examine the optimal second-period contract. To induce productive effort from both types, we need to have $w_2 \geq w^H_{PN}$ and $w_2 \geq w^L_{PN}$. Note that $w^L_{PN} > w^H_{PN}$ for any $\mu_{i,2}$ and $(e^H_{-i,2}, e^L_{-i,2})$ so that $w_2 \geq w^H_{PN}$ holds if $w_2 \geq w^L_{PN}$. To induce productive effort with probability one, therefore, we need to focus on w^L_{PN}.

[25] If the effort choice in period 1 does not affect the expected utility in period 2, the IC constraints in period 1 virtually become static.

If both types exert productive effort in period 1, then $\mu_{i,2} \in \{0,1\}$. To induce productive effort for any possible contingency, therefore, a second-period contract must satisfy both $w_2 \geq w^L_{PN}(1;1,1)$ and $w_2 \geq w^L_{PN}(0;1,1)$. Note that $w^L_{PN}(\theta;1,1)$ is increasing in θ. This implies that the second-period optimal contract is given by

$$w_2 = w^L_{PN}(1;1,1) = \frac{c}{G(h-l) - G(h-m)}.$$

We now shift attention to the first-period problem. To induce productive effort from both types, the contract must satisfy (A.1) which can be written as

$$w_1 \geq \frac{c + u_1^{\eta_i}(0,0) - u_1^{\eta_i}(1,0)}{P^{\eta_i}_{1,0} - P^{\eta_i}_{0,0}}.$$

The first-period constraint thus hinges critically on $u_1^{\eta_i}(0,0) - u_1^{\eta_i}(1,0)$. The key question here is whether an agent can gain from unilaterally deviating from the equilibrium strategy to manipulate the opponent's belief. To this end, throughout the analysis, we place two (fairly natural) restrictions on the way the agents form their beliefs off the equilibrium path.

Restriction A: If $y_{-i,1} = l$, $\mu_{i,2} = 0$;

Restriction B: If $y_{-i,1} = h$, $\mu_{i,2} = 1$.

Under these restrictions, it is easy to see that the low-ability type cannot manipulate the opponent's belief because $\mu_{-i,2} = 0$ regardless of the effort choice in period 1 as long as $e^H_{-i,1} = 1$. This implies that $u_1^{\eta_i}(0,0) = u_1^{\eta_i}(1,0)$, and the problem is totally identical to that in period 2. The low-ability type exerts productive effort if $w_2 \geq w^L_{PN}(\theta_{-i};1,1)$. Since $w^L_{PN}(\theta;1,1)$ is increasing in θ, a candidate contract is given by $w_1 = w^L_{PN}(\theta_1;1,1)$.

We now verify whether this candidate contract can induce productive effort from the high-ability type. Suppose that the high-ability type unilaterally deviates and chooses $e_{i,1} = 0$. This signals to the opponent that the agent is of the low-ability type. This does not influence the opponent's behavior, however, because the candidate contract always induces productive effort for any realized belief. Again, the problem is reduced to the static one where the high-ability type exerts productive effort if $w_1 \geq w^H_{PN}(\theta_1;1,1)$. Since $w^L_{PN}(\theta_1;1,1) > w^H_{PN}(\theta_1;1,1)$, the candidate contract can indeed induce productive effort from the high-ability type, and the optimal contract in period 1 is hence given by

$$w_1 = w^L_{PN}(\theta_1; 1, 1) = \frac{c}{\theta_1(G(h-l)-G(h-m)) + (1-\theta_1)(G(m-l)-G(0))}.$$

Q.E.D.

Proof of Proposition 2: To induce productive effort, a contract must satisfy $w^\eta_{PS}(\mu) > w_2 \geq w^\eta_{PN}(\mu)$. This implies that there exists no contract that can induce productive effort from the low-ability type if $w^L_{PN}(\mu) > w^L_{PS}(\mu)$ or, equivalently,

$$\frac{P^L_{0,1}(\mu; e^H_{-i,2}, e^L_{-i,2}) - P^L_{1,0}(\mu; e^H_{-i,2}, e^L_{-i,2})}{P^L_{1,0}(\mu; e^H_{-i,2}, e^L_{-i,2}) - P^L_{0,0}(\mu; e^H_{-i,2}, e^L_{-i,2})} > \lambda. \tag{A.4}$$

We now show that there exists a threshold $\bar{\mu} \in (0, 1)$ such that (A.4) holds for $\mu > \bar{\mu}$ when $e^H_{-i,2} = 1$. To this end, it suffices to show that

$$\frac{P^L_{0,1}(1; 1, e^L_{-i,2}) - P^L_{1,0}(1; 1, e^L_{-i,2})}{P^L_{1,0}(1; 1, e^L_{-i,2}) - P^L_{0,0}(1; 1, e^L_{-i,2})} > \lambda,$$

which can be written as

$$\frac{G(h-m) - G(m-l)}{G(h-l) - G(h-m)} > \lambda \Leftrightarrow \frac{G(h-l) - G(m-l)}{G(h-l) - G(h-m)} > 1 + \lambda. \tag{A.5}$$

This holds by Assumption 1 because $G(m-l) - G(0) > G(h-l) - G(h-m)$.

Q.E.D.

Proof of Proposition 3: We first establish the following result which we will repeatedly use throughout the proof.

Lemma 1 *Productive effort always dominates sabotage effort for the high-ability type (in the static setting).*

PROOF: For the high-ability type, productive effort is preferred to $\lambda c > (P^H_{0,1} - P^H_{1,0})w_2$ where

$$P^H_{0,1} = \mu_{i,2} G(m - y^H(e^H_{-i,2}, 1)) + (1-\mu_{i,2}) G(m - y^L(e^L_{-i,2}, 1)),$$

$$P^H_{1,0} = \mu_{i,2} G(h - y^H(e^H_{-i,2}, 0)) + (1-\mu_{i,2}) G(h - y^L(e^L_{-i,2}, 0)).$$

This condition always holds for any $w_2 \geq 0$ if $P^H_{0,1} < P^H_{1,0}$. Sufficient conditions for this are $G(h - y^H(e^H_{-i,2}, 0)) > G(m - y^H(e^H_{-i,2}, 1))$ and $G(h - y^L(e^L_{-i,2}, 0)) > G(m - y^L(e^L_{-i,2}, 1))$, which hold for any $(e^H_{-i,2}, e^L_{-i,2})$.

Q.E.D.

This result means that we do not need to worry about the possibility of sabotage for the high-ability type, which substantially simplifies the analysis. Given this result, we derive the optimal fast-track and late-selection contracts, when they exist, in turn. The obtained results are used for both Propositions 3 and 4.

1. The optimal fast-track contract

Sketch of the proof: We seek for the optimal contract within a class of contracts that induce productive effort from both types in period 1 (the fast-track scheme). At this point, however, it is not clear what could happen in period 2 under the fast-track scheme. We first show that under Assumption 1, there exists no second-period contract which induces productive effort from the low-ability type (Step 1-1). This implies that the best the principal can do in period 2 is to induce productive effort from the high-ability type for any realized belief. There is indeed such a second-period contract which we obtain in Step 1-2. Finally, in Step 1-3, we construct a first-period contract which induces productive effort from both of the types.

Step 1-1: Showing that under Assumption 1, there exists no second-period contract which induces productive effort from the low-ability type.

Under the fast-track scheme, each agent knows the opponent's ability type with precision in period 2. This means that there are two cases we need to consider, depending on the opponent's ability type.

We first consider the case where the opponent turns out to be of the high-ability type ($\mu_{i,2} = 1$). Suppose, on the contrary, that the low-ability type exerts productive effort in this contingency. If this is the case, it follows from Proposition 2 that the high-ability opponent must exert no effort or, more precisely,

$$w_{PN}^{H}(0; e_{-i,2}^{H}, 1) = \frac{c}{G(h-m) - G(0)} > w_2. \tag{A.6}$$

However, given that the high-ability opponent exerts no effort, it is also necessary to have

$$w_2 \geq w_{PN}^{L}(1; 0, e_{-i,2}^{L}) = \frac{c}{G(m-l) - G(0)}. \tag{A.7}$$

This is a contradiction since (A.6) and (A.7) do not hold simultaneously. We can thus conclude that the low-ability type does not exert productive effort when the opponent happens to be of the high-ability type.

Now suppose that the opponent turns out to be of the low-ability type ($\mu_{i,2} = 0$). In this case, we need to have $w_2 \geq w^L_{PN}(0; 1, 1)$ if the low-ability type is to exert productive effort. For a contract to be sabotage-proof for any possible contingency, it is also necessary that

$$w^L_{NS}(1; 1, e^L_{-i,2}) = \frac{(1+\lambda)c}{P^L_{0,1}(1; 1, e^L_{-i,2}) - P^L_{0,0}(1; 1, e^L_{-i,2})} > w_2.$$

These conditions cannot hold simultaneously, however, since

$$w^L_{PN}(0; e^H_{-i,2}, 1) = \frac{c}{G(m-l) - G(0)} > w^L_{NS}(1; 1, e^L_{-i,2}) = \frac{(1+\lambda)c}{G(h-l) - G(m-l)},$$

by Assumption 1.

Step 1-2: Constructing a second-period contract which always induces productive effort from the high-ability type with the lowest cost.

We now know that the best the principal can do in period 2 is to induce productive effort from the high-ability type for both $\mu_{i,2} = 0$ and $\mu_{i,2} = 1$. To induce productive effort in any contingency, a contract must satisfy $w_2 \geq \max\{w^H_{PN}(1; 1, 0), w^H_{PN}(0; 1, 0)\}$ where

$$w^H_{PN}(1; 1, 0) = \frac{c}{G(h-m) - G(0)}, \; w^H_{PN}(0; 1, 0) = \frac{c}{G(h-l) - G(m-l)}.$$

Since $w^H_{PN}(0; 1, 0) > w^H_{PN}(1; 1, 0)$ by assumption, a candidate contract is given by $w_2 = w^H_{PN}(0; 1, 0)$.

For this to be optimal, we also need to show that it is sabotage-proof. The condition for this is given by $\min\{w^L_{NS}(1; 1, 0), w^L_{NS}(0; 1, 0)\} > w^H_{PN}(0; 1, 0)$. Since $\lim_{\mu \to 0} w^L_{NS}(\mu; 1, 0) = \infty$, it suffices to show that

$$w^L_{NS}(1; 1, 0) = \frac{(1+\lambda)c}{G(h-l) - G(m-l)} > \frac{c}{G(h-l) - G(m-l)}, \tag{A.8}$$

which holds as long as $\lambda > 0$.

Step 1-3: Constructing a first-period contract which induces productive effort from both types with the lowest cost.

We first consider the low-ability type's problem. In this case, it is important to note that the low-ability type cannot influence the opponent's belief when he is expected to exert productive effort.

This means that the first-period problem is identical to the second-period problem, and the same constraints apply for this case. To induce productive effort from the low-ability type, it is necessary that, given some prior belief, $w^L_{PS}(\theta_{-i};1,1) > w_1 \geq w^L_{PN}(\theta_{-i};1,1)$. There exists some w_1 that satisfies this if $w^L_{PS}(\theta_{-i};1,1) > w^L_{PN}(\theta_{-i};1,1)$. This condition can be written as

$$\frac{\lambda c}{\theta_{-i}(G(h-m)-G(m-l))} > \frac{c}{\theta_{-i}(G(h-l)-G(h-m))+(1-\theta_{-i})(G(m-l)-G(0))}.$$

For this to hold for any (θ_1,θ_2), we need to have

$$\lambda > \frac{\theta_1(G(h-m)-G(m-l))}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))},$$

which holds by Assumption 1.

We now consider the high-ability type's problem. As above, under the fast-track scheme, any deviation does not change the opponent's behavior: the high-ability type always exerts productive effort while the low-ability type always exerts no effort, regardless of the belief. This implies that $u^H_{1,0} = u^H_{0,0} = u^H_{0,1}$, and there is thus no point in manipulating the opponent's belief by deviating from the equilibrium strategy. The problem is again static where the optimal contract must satisfy $w^H_{PS}(\theta_{-i};1,1) > w_1 \geq w^H_{PN}(\theta_{-i};1,1)$. Since $w^L_{PN}(\theta_{-i};1,1) > w^H_{PN}(\theta_{-i};1,1)$ and $w^L_{PN}(\mu;1,1)$ is increasing in μ, we need to look at the case where $\theta_{-i} = \theta_1$.

In sum, the candidate contract under the fast-track scheme is given by

$$w_1 = w_1^{\mathrm{FT}} \equiv \frac{c}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))},\ w_2 = w_2^{\mathrm{FT}} \equiv \frac{c}{G(h-l)-G(m-l)},$$

which is sabotage-proof under Assumption 1. Under the maintained assumptions, this is the optimal contract within the class of contracts which induce productive effort from both types in period 1.

2. The optimal late-selection contract

Sketch of the proof: We now seek for the optimal contract within the class of contracts that induce productive effort from the low-ability type and no effort from the high-ability type in period 1. In this case, it is fairly straightforward to see what could happen in period 2 because no information is revealed in period 1. Under Assumption 2, the optimal second-period contract is the one that induces productive effort from both types with the lowest cost, which we obtain in Step 2-1. Given this, we then examine the optimal first-period contract. In Step 2-2, we construct a candidate first-period contract which induces productive effort from the low-ability type. Finally, in Step 2-3, we make sure that the high-ability type exerts no effort under this candidate contract.

Step 2-1: Constructing a second-period contract which induces productive effort from both types with the lowest cost.

Under the late-selection scheme, the ability type is not identifiable in period 2, i.e., $\mu_{i,2} = \theta_{-i}$, on the equilibrium path. The problem is thus exactly the same as the first-period problem under the fast-track contract. As we have seen, the optimal contract in this situation is

$$w_2 = w^L_{PN}(\theta_1; 1, 1) = \frac{c}{\theta_1(G(h-l) - G(h-m)) + (1-\theta_1)(G(m-l) - G(0))}. \tag{A.9}$$

Step 2-2: Constructing a first-period contract which induces productive effort from the low-ability type with the lowest cost.

The dynamic incentive matters for the first-period problem, and we need to compute the expected second-period payoff $u^\eta_1(e, d)$. To this end, we place an additional restriction, on top of Restrictions A and B used in the proof of Proposition 1, on the way the agents form their beliefs off the equilibrium path.

Restriction C: If $d_{i,1} = 1$, $\mu_{i,2} = \theta_{-i}$.

This restriction may need more clarification. Although we assume that the opponent's effort level is not directly observable, each agent i knows his own ability η_i and effort $e_{i,1}$ and observes his productivity $y_{i,1}$. The level of sabotage effort, chosen by the opponent, is thus precisely inferred from this available information. In this paper, we take a rather conservative stance that this type of deviation adds no information to the prior.

Under this restriction, we now drive a first-period contract which induces productive effort from the low-ability type with the lowest cost. With the possibility of sabotage, there are two ways to deviate from the equilibrium. Suppose first that the low-ability type deviates and chooses $d_{i,1} = 1$. Since this type of deviation would not affect the opponent's belief by assumption, one can see that $u^L_1(1, 0) = u^L_1(0, 1)$, which reduces the problem to the static one. Now suppose that $e_{i,1} = 0$, which signals his true ability type to the opponent. Given the optimal contract in period 2, however, this would not change the opponent's behavior because $w^L_{PN}(\theta_1; 1, 1) > w^L_{PN}(0; 1, 1) > w^H_{PN}(0; 1, 1)$.

This means that $u_1^L(1,0) = u_1^L(0,0)$, which again reduces the problem to the static one. To induce productive effort from the low-ability type, it is necessary that

$$w_{PS}^L(\theta_1; 0, 1) > w_1 \geq w_{PN}^L(\theta_1; 0, 1) = \frac{c}{G(m-l) - G(0)}. \tag{A.10}$$

Note that $w_{PS}^L(\theta_1; 0, 1)$ goes to infinity so that sabotage effort is never chosen. The candidate first-period contract is the one that satisfies (A.10) with the lowest cost, i.e., $w_1 = w_{PN}^L(\theta_1; 0, 1)$.

Step 2-3: Verifying that the high-ability type has no incentive to deviate under the candidate contract obtained thus far.

The argument thus far yields a candidate contract $(w_1 = w_{PN}^L(\theta_1; 0, 1), w_2 = w_{PN}^L(\theta_1; 1, 1))$. For this to be the optimal late-selection scheme, one needs to verify that the high-ability type has no incentive to deviate, by exerting either sabotage effort or productive effort. This last step is the most tedious one to make.

While there are two ways for the high-ability type to deviate from the equilibrium path, a deviation by exerting sabotage effort does not impose much of a problem since $u_1^H(1,0) = u_1^H(0,1)$ for the same reason as above. If the high-ability type deviates and chooses $e_{i,1} = 1$, on the other hand, this signals to the opponent his true type. This may lead the opponent to exert sabotage effort in period 2 if he happens to be of the low-ability type. If this is the case, a deviation from the equilibrium path matters, and we need to explicitly compute $u_1^H(0,0)$ and $u_1^H(1,0)$.

First, it is straightforward to obtain the expected payoff on the equilibrium path:

$$u_{0,0}^H = (\theta_{-i} G(0) + (1 - \theta_{-i}) G(h - l)) w_2 - c.$$

It is, on the other hand, more complicated to see what happens off the equilibrium path. To this end, we need to establish the following result.

Lemma 2 *Suppose that the high-ability type deviates and chooses $e_{i,1} = 1$, and moreover that the second-period contract is given by (A.9). Then, in period 2, (i) the low-ability type exerts sabotage effort; (ii) the high-ability type exerts productive effort.*

Proof: The low-ability type exerts sabotage effort if $w_{PN}^L(\theta_1; 1, 1) \geq \max\{w_{PS}^L(1; 1, 0), w_{NS}^L(1; 1, 0)\}$.

Note that $w^L_{NS}(1;1,0) > w^L_{PS}(1;1,0)$ because

$$\frac{(1+\lambda)c}{G(h-l)-G(m-l)} > \frac{\lambda c}{G(h-m)-G(m-l)} \Leftrightarrow \frac{G(h-m)-G(m-l)}{G(h-l)-G(h-m)} > \lambda,$$

which holds by Assumption 1. The condition is hence satisfied if

$$\frac{c}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))} \geq \frac{(1+\lambda)c}{G(h-l)-G(m-l)},$$

which holds by Assumption 1.

Given that $w_2 = w^L_{PN}(\theta_1;1,1)$, the high-ability type exerts productive effort if

$$w^L_{PN}(\theta_1;1,1) \geq \frac{c}{\theta_{-i}(G(h-m)-G(0))+(1-\theta_{-i})(G(m-l)-G(0))}.$$

This can be written as

$$1 \geq \frac{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))}{\theta_{-i}(G(h-m)-G(0))+(1-\theta_{-i})(G(m-l)-G(0))},$$

which holds for any θ_{-i}.

Q.E.D.

This implies that the expected payoff when the agent deviates by choosing $e_{i,1}=1$ is given by $u^H_1(1,0) = (\theta_{-i}G(0)+(1-\theta_{-i})G(m-l))w_2 - c$. Then, the high-ability type chooses no effort if

$$\min\{\frac{(1+\lambda)c}{P^H_{0,1}(\theta_{-i};0,1)-P^H_{0,0}(\theta_{-i};0,1)}, \frac{c+(1-\theta_{-i})(G(h-l)-G(m-l))w_2}{P^H_{1,0}(\theta_{-i};0,1)-P^H_{0,0}(\theta_{-i};0,1)}\} > w_1.$$

The candidate contract $(w_1 = w^L_{PN}(\theta_1;1,1), w_2 = w^L_{PN}(\theta_1;0,1))$ is optimal if

$$\min\{\frac{(1+\lambda)c}{G(m-l)-G(0)}, \frac{c+(1-\theta_{-\mathrm{i}})(G(h-l)-G(m-l))w^L_{PN}(\theta_1;1,1)}{G(h-m)-G(0)}\} > \frac{c}{G(m-l)-G(0)}.$$

Since this condition is harder to satisfy when θ_{-i} is larger, it suffices to show that

$$\frac{c+(1-\theta_1)(G(h-l)-G(m-l))w^L_{PN}(\theta_1;1,1)}{G(h-m)-G(0)} > \frac{c}{G(m-l)-G(0)}.$$

Substituting $w^L_{PN}(\theta_1;1,1)$, we obtain

$$\frac{(1-\theta_1)(G(h-l)-G(m-l))}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))} > \frac{G(h-m)-G(m-l)}{G(m-l)-G(0)},$$

which is further simplified to

$$(1-\theta_1)(G(m-l)-G(0)) > \theta_1(G(h-m)-G(m-l)). \tag{A.11}$$

When this condition fails to hold, there exists no contract that induces no effort from the high-ability type, and sabotage occurs on the equilibrium path with positive probability.

In sum, the optimal late-selection contract is given by

$$w_1 = w_1^{\mathrm{LS}} \equiv \frac{c}{G(m-l)-G(0)},\ w_2 = w_2^{\mathrm{LS}} \equiv \frac{c}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))}.$$

Under the maintained assumptions and (A.11), this is the optimal contract within the class of contracts that induce productive effort only from the low-ability type in period 1.

3. The optimality of the first-track scheme.

Given this preceding argument, it is straightforward to prove the proposition. If (A.11) does not hold, then the fast-track scheme is the only possibility to prevent sabotage. Then, the optimal fast-track contract $(w_1^{\mathrm{FT}}, w_2^{\mathrm{FT}})$ must emerge as the optimal choice under the restriction that the optimal contract must be sabotage-proof.

Q.E.D.

Proof of Proposition 4: Suppose that (A.11) holds, so that both schemes can be made sabotage-proof. In this case, we need to compare the two contracts. Let π^j, $j = FT, LS$, denote the expected output under each respective scheme, which is given by

$$\pi^{\mathrm{FT}} = (\theta_1+\theta_2)h + (2-\theta_1-\theta_2)m - w_1^{\mathrm{FT}} + (\theta_1+\theta_2)h + (2-\theta_1-\theta_2)l - w_2^{\mathrm{FT}},$$

$$\pi^{\mathrm{LS}} = 2m - w_1^{\mathrm{LS}} + (\theta_1+\theta_2)h + (2-\theta_1-\theta_2)m - w_2^{\mathrm{LS}}.$$

Letting $\bar{\theta} \equiv (\theta_1+\theta_2)/2$, the principal adopts the fast-track scheme over the late-selection scheme if

$$\bar{\theta}h + (1-\bar{\theta})l - \frac{w_1^{\mathrm{FT}}+w_2^{\mathrm{FT}}}{2} \geq m - \frac{w_1^{\mathrm{LS}}+w_2^{\mathrm{LS}}}{2}.$$

Substituting the optimal wages, we obtain (2).

Q.E.D.

References

Acemoglu, Daron and Pischke, Jörn-Steffen, Structure of Wages and Investment in General Training, *Journal of Political Economy*, 107, 539-72.

Aoki, Masahiko, 1986, Horizontal vs. Vertical Information Structure of the Firm, *American Economic Review*, 76, 971-83.

Ariga, Kenn; Brunello, Giorgio and Ohkusa, Yasushi, 2000, *Internal Labour Markets in Japan*, Cambridge University Press.

Bernhardt, Dan, 1995, Strategic Promotion and Compensation, *Review of Economic Studies*, 62, 315-39.

Chan, William, 1996, External Recruitment versus Internal Promotion, *Journal of Labor Economics*, 14, 555-70.

Chiappori, Pierre-Andre; Salanie, Bernard and Valentin, Julie, 1999, Early Starters and Late Beginners, *Journal of Political Economy*, 107, 731-60.

Chen, Kong-Pin, 2003, Sabotage in Promotion Tournaments, *Journal of Law, Economics, and Organization*, 19, 119-40.

Chen, Kong-Pin, 2005, External Recruitment as an Incentive Device, *Journal of Labor Economics*, 23, 259-77.

Gibbons, Robert and Waldman, Michael, 1999a, Careers in Organizations: Theory and Evidence, in *Handbook of Labor Economics*, vol. 3, eds., Ashenfelter, Orley C. and Card, David, Amsterdam: North Holland.

Gibbons, Robert and Waldman, Michael, 1999b, A Theory of Wage and Promotion Dynamics inside Firms, *Quarterly Journal of Economics*, 114, 1321-58.

Grossman, Gene M. and Maggi, Giovanni, 2000, Diversity and Trade, *American Economic Review*, 90, 1255-75.

Ishida, Junichiro, 2004, Signaling and Strategicallly Delayed Promotion, *Labour Economics*, 11, 687-700.

Ishida, Junichiro, 2006, Optimal Promotion Policies with the Looking-Glass Effect, *Journal of Labor Economics*, 24, 857-77.

Ishiguro, Shingo, 2004, Collusion and Discrimination in Organizations, *Journal of Economic Theory*, 116, 357-69.

Kremer, Michael, 1993, The O-Ring Theory of Economic Development, *Quarterly Journal of Economics*, 108, 551-75.

Lazear, Edward P., 1979, Why is There Mandatory Retirement?, *Journal of Political Economy*, 87, 1261-84.

Lazear, Edward P., 1989, Pay Equality and Industrial Politics, *Journal of Political Economy*, 97, 561-80.

Lazear, Edward P. and Rosen, Sherwin, 1981, Rank-Order Tournaments as Optimum Labor Contracts, *Journal of Political Economy*, 89, 841-64.

MacLeod, W. Bentley and Malcomson, James M., 1988, Reputation and Hierarchy in Dynamic Models of Employment, *Journal of Political Economy*, 96, 832-54.

Meyer, Margaret A., 1992, Biased Contests and Moral Hazard: Implications for Career Profiles, *Annales d'Economie et de Statistique*, 25/26, 165-87.

Morita, Hodaka, 2004, Firm Survival and Specific Human Capital, mimeo, University of New South Wales.

Morita, Hodaka, 2005, Multi-Skilling, Delegation and Continuous Process Improvement: A Comparative Analysis of US-Japanese Work Organizations, *Economica*, 72, 69-93.

Osawa, Naoto; Kamiyama, Kazushige; Nakamura, Koji; Noguchi, Tomohiro and Eiji Maeda, 2002, An Examination of Structural Changes in Employment and Wages in Japan, Bank of Japan Research Papers.

Owan, Hideo, 2004, Promotion, Turnover, Earnings, and Firm-Sponsored Training, *Journal of Labor Economics*, 22, 955-78.

Prendergast, Canice, 1992, Career Development and Specific Human Capital Collection, *Journal of the Japanese and International Economies*, 6, 207-27.

Prendergast, Canice, 1999, The Provision of Incentives in Firms, *Journal of Economic Literature*, 37, 7-63.

Rosen, Sherwin, 1981, The Economics of Superstars, *American Economic Review*, 71, 845-58.

Spilerman, Seymour and Ishida, Hiroshi, 1996, Stratification and Attainment in a Large Japanese Firm, in *Generating Social Stratification: Toward a New Generation of Research*, ed., Kerckhoff, Alan, Westview Press.

Waldman, Michael, 1984, Job Assignments, Signalling and Efficiency, *Rand Journal of Economics*, 15, 255-67.

Waldman, Michael, 2003, Ex Ante versus Ex Post Optimal Promotion Rules: the Case of Internal Promotion, *Economic Inquiry*, 41, 27-41.

Appendix C: the case where Assumption 1 fails to hold (not for publication)

In what follows, we consider the case where Assumption 1 fails to hold while maintaining the other assumptions, namely Assumptions 2 and 3. To this end, define

$$\Lambda_h \equiv \frac{G(h-l)-G(m-l)}{G(m-l)-G(0)},\ \Lambda_l \equiv \frac{\theta_1(G(h-l)-G(m-l))+(1-\theta_1)(G(m-l)-G(0))}{\theta_1(G(h-l)-G(h-m))+(1-\theta_1)(G(m-l)-G(0))}.$$

Assumption 1 states that $\Lambda_h > 1+\lambda \geq \Lambda_l$ where we maintain $\Lambda_h > \Lambda_l$.

Suppose first that the cost of sabotage is so large that

$$1+\lambda \geq \Lambda_h.$$

In this case, the optimal contract in the benchmark case can induce productive effort from the low-ability type because we have $w^L_{PS}(1) \geq w^L_{PN}(1)$ which can be written as

$$w^L_{PS}(1;1,e^L_{-i,2}) = \frac{\lambda c}{G(h-m)-G(m-l)} \geq w^L_{PN}(1;1,e^{\mathrm{L}}_{-i,2}) = \frac{c}{G(h-l)-G(h-m)}.$$

Under Assumption 2, therefore, the optimal contract in this case is the one shown in Proposition 2 (the benchmark contract).

Now suppose that the cost of sabotage is so small that

$$\Lambda_l > 1+\lambda > 1. \tag{A.12}$$

In this case, there exists no contract that can induce productive effort from the low-ability type for a given prior, even if he gains no additional information. The best one can hope for in this contingency is then to induce productive effort from the high-ability type and no effort from the low-ability type in both periods under Assumption 3, if such a contract exists. It is also important to note that since it is not possible to induce productive effort from the low-ability type, the ability type is perfectly identifiable in period 2 under any optimal contract so that $\mu_{i,2} \in \{0,1\}$.

We start with the second-period problem. Given that the high-ability type always exerts productive effort, to induce no effort from the low-ability type, the second-period contract must satisfy

$$\frac{c}{\mu_{i,2}(G(h-l)-G(h-m))+(1-\mu_{i,2})(G(m-l)-G(0))} > w_2.$$

$$\frac{(1+\lambda)c}{\mu_{i,2}(G(h-l)-G(m-l))+(1-\mu_{i,2})(G(l)-G(0))} > w_2.$$

for $\mu_{-i,2} = 0, 1$. In other words, to always induce no effort from the low-ability type, we must have

$$\min\{\frac{c}{G(m-l)-G(0)}, \frac{(1+\lambda)c}{G(h-l)-G(m-l)}, \frac{(1+\lambda)c}{G(l)-G(0)}\} > w_2.$$

To induce productive effort from the high-ability type, the contract must satisfy

$$w_2 \geq \frac{c}{\mu_{i,2}(G(h-m)-G(0))+(1-\mu_{i,2})(G(h-l)-G(m-l))}.$$

The high-ability type always exerts productive effort if

$$w_2 = w_2^{**} \equiv \frac{c}{G(h-l)-G(m-l)}.$$

Under this contract, the low-ability type always exerts no effort if

$$\min\{\frac{c}{G(m-l)-G(0)}, \frac{(1+\lambda)c}{G(h-l)-G(m-l)}, \frac{(1+\lambda)c}{G(l)-G(0)}\} > \frac{c}{G(h-l)-G(m-l)},$$

which can be written as

$$G(h-l)-G(m-l) > G(m-l)-G(0).$$

This condition holds when $\Lambda_h > \Lambda_l > 1+\lambda$, so that the optimal second-period contract is given by $w_2 = w_2^{**}$.

We now turn to the first-period problem. Since each agent's behavior is independent of the opponent's ability type (the high-ability type always exerts productive effort while the low-ability type always exerts no effort), the problem is reduced to a static one. Given this, we again seek for a contract which always induces productive effort from the high-ability type and no effort from the low-ability type. A candidate contract is then obtained as

$$w_1 = w_1^{**} \equiv \frac{c}{\theta_2(G(h-m)-G(0))+(1-\theta_2)(G(h-l)-G(m-l))}.$$

Under this contract, the low-ability type always exerts no effort if

$$\frac{c}{\mu_{i,1}(G(h-l)-G(h-m))+(1-\mu_{i,1})(G(m-l)-G(0))} > w_1^{**}.$$

$$\frac{(1+\lambda)c}{\mu_{i,1}(G(h-l)-G(m-l))+(1-\mu_{i,1})(G(l)-G(0))} > w_1^{**}.$$

As can be seen from above, these conditions obviously hold for any $\mu_{i,1}$.

In sum, the optima contract when $\Lambda_l > 1 + \lambda$ is

$$w_1 = w_1^{**},\ w_2 = w_2^{**},$$

which always induces productive effort from the high-ability type and no effort from the low-ability type. One can view this as an extreme form of the fast-track scheme.

Appendix D: a sufficient condition for no equilibrium sabotage (not for publication)

Throughout the analysis, we presuppose that it is never optimal to let the agents choose sabotage effort on the equilibrium path (Assumption 3). As we will see below, however, the possibility of equilibrium sabotage can easily be ruled out under some additional conditions, even without imposing Assumption 3. Here, we derive a sufficient condition for the optimality of sabotage-proof contracts when $D = 0$ (no external cost of sabotage incurred by the principal).

Suppose that the principal designs a contract which allows sabotage to take place in period 2. In this case, under the assumption that c is sufficiently small, it is evidently optimal to implement $(e_1^H = 1, e_1^L = 1)$ in period 1. Given this, the best she can do in period 2 is that: (i) a high-ability agent always exerts effort; (ii) a low-ability agent exerts effort when matched with a low-ability agent and sabotage effort when matched with a high-ability agent. This effort profile can be implemented by offering

$$w_2 = w^L(1; 1, 0) = \frac{(1+\lambda)c}{G(h-l) - G(m-l)}.$$

See (A.8) for detail, implying that the principal must offer at least more than the optimal fast-track contract, i.e., $w_2 > w_2^{\mathrm{FT}}$.

Since it costs more to achieve this, it must result in higher output than the optimal (sabotage-proof) fast-track contract if it is to be optimal. In other words, a sufficient condition for no equilibrium sabotage is

$$\theta_1 h + (1-\theta_1)l + \theta_2 h + (1-\theta_2)l \geq \theta_1(2\theta_2 h + (1-\theta_2)(m+l)) + (1-\theta_1)(\theta_2(m+l) + 2(1-\theta_2)m).$$

The left-hand side is the expected total output in period 2 under the contract that allows sabotage to take place, while the right-hand side is the output under the optimal fast-track contract. This condition can be written as

$$(\theta_1 + \theta_2 - 2\theta_1\theta_2)(h - m) \geq 2(1 - \theta_1)(1 - \theta_2)(m - l),$$

To see this condition more clearly, suppose that $\theta_1 = \theta_2 = \theta$. The condition is then reduced to

$$\theta(h - m) \geq (1 - \theta)(m - l),$$

which fails to hold for any θ if $h - m$ is sufficiently larger than $m - l$. Alternatively, for any (h, m, l) satisfying $h - m > m - l$, the condition fails to hold if θ is at least large than one half.