Completing a Project on Time: When to Study and When to Wing It

Yonggyun Kim*

Duke University

yonggyun.kim@duke.edu

July 9, 2019

Abstract

I study a dynamic agency problem such that the agent has multi-technologies to complete a task: the *winging it* technology requires one breakthrough with a low arrival rate, whereas the *studying* technology requires two breakthroughs with high arrival rates. The intermediate breakthrough of the studying technology is assumed to be observed by both the principal and the agent. To derive the incentive scheme, the principal compares an immediate payoff (to induce winging it) and a deadline extension (to induce studying) to maximize her expected payoff. When the winging it and studying technologies are equally efficient, the recommended action of the agent in the optimal contract would be one of three types: (i) winging it always; (ii) studying always; (iii) switching once from studying to winging it. The form of the optimal contract depends on the payoff of the project and the effectiveness of the studying technology.

JEL Classification: J41, L14, D86

Keywords: Dynamic Agency, Skill Improvement, Speed of Completion, Multi-technology, Poisson Arrival

*I am grateful to Attila Ambrus, Arjada Bardhi, Fei Li, Pino Lopomo, Curtis Taylor and Huseyin Yildirim for comments and suggestions.

A woodsman was once asked, "What would you do if you had just five minutes to chop down a tree?" He answered, "I would spend the first two and a half minutes sharpening my axe." (Jaccard, 1956, p.12)[1]

1 Introduction

A traditional interpretation of *skill* in economics is productivity, which means that the level of skill determines how much cost it would take to produce a certain number of products or how many products one can produce with a constrained amount of resources.[2] Nonetheless, in practice, another key nature of skill is *speed*–how fast one can finish a given task. Many firms and workers can invest in skills to expedite the completion of some tasks. For example, automobile companies invest in 3D printing and collaborative robots to speed up the manufacturing processes (Koenig, 2019). A woodman, who wants to chop down a tree, may spend some time in sharpening the ax. A research assistant (RA), who needs to complete an empirical project, may learn advanced skills such as STATA or R to facilitate the project.

The obvious benefit of investing in skills related to speed is that once advanced skills are acquired, the task can be completed faster than before. On the contrary, since an agent's effort is constrained, there are opportunity costs of investing in such skills: the agent is forgoing chances to complete the task with a current mediocre skill. Automobile companies can divert resources from investment to production. The woodman can try to chop down the tree with the dull ax rather than sharpening the ax. The RA has an option to complete the task with a basic skill such as Excel. Therefore, workers or companies often need to choose between investing in an advanced skill (studying) and completing tasks with a current basic skill (winging it).

This paper studies a dynamic principal agent model with multi-technologies that reflect the economic situations described above. A principal delegates a project to be completed to an agent. The project requires an ultimate breakthrough and running the project incurs a flow cost to the principal. The agent initially has a basic skill and can choose to complete the project with the basic skill but the breakthrough arrives with a low rate. The agent can also invest in an advanced skill, and then the agent can acquire the skill with an arrival rate, which is higher than the arrival rate of the breakthrough with the basic skill. Once the advanced skill is obtained, the breakthrough arrives with a rate higher than the basic skill's arrival rate. The agent is also able to divert efforts for a private benefit. Then, an

[1] I thank Huseyin Yildirim for bringing this quote to my attention.

[2] A classic paper that holds this viewpoint is Spence (1973), which assumes that the agent's skill determines the marginal cost.

action of the agent with the basic skill can be summarized as a choice among the *winging it* technology (completing the project with the basic skill), the *studying* technology (investing in the advanced skill), and the *leisure* (shirking for a private benefit). I assume that no further skill can be acquired once the advanced skill is obtained. Hence, the agent with the advanced skill chooses between the working technology (completing the project with the advanced skill) and the leisure. The agent's choice of technologies is hidden to the principal, but the agent and the principal can observe the skill improvement and the completion of the project. Both the principal and the agent are risk neutral, do not discount the future, and the agent is protected by limited liability. I study how the principal should optimally design the contract and how recommended actions change over time.

I start by analyzing the first best case, i.e., the case such that the agent's action is also observable to the principal. In this case, there are two candidates for the optimal recommended action schedule: the winging it only schedule and the studying only schedule, of which recommended actions are to choose the specified technology until the ultimate breakthrough or the skill improvement is attained. It is shown that one of these two schedules would be the optimal schedule and the principal chooses the schedule with the shorter expected duration. This is because the principal wants to minimize the expected flow cost, which is proportional to the expected duration. Hence, it is natural to define the notion of efficiency for the technologies based on the expected duration of the above schedules: one technology is more efficient than the other if and only if the expected duration of the schedule corresponding to the technology is shorter than the other.

The main analysis of this paper is to solve the case such that the agent's action is unobservable to the principal, the agent's skill improvement is verifiable, and both technologies are equally efficient, i.e., recommending each technology is indifferent to the principal in the first best. In this case, the hidden action assumption drives the principal to employ a finite deadline to induce the agent not to shirk. At each instance of time, the recommended action is determined by comparing the principal's expected payoffs from each technology. Unlike to the first best case, the principal would no longer be indifferent between recommending each technology. This is because the finite deadline distorts the probability of ultimate breakthrough and the expected duration for each technology in a different way. If the principal wants to induce the agent to wing it, the agent needs to be compensated by the immediate payment upon the success of the project. On the contrary, if the principal wants to induce the agent to study, since the skill improvement does not give an immediate benefit to the principal, the principal extends the deadline upon the skill improvement, i.e., gives more time to the agent to complete the project with the advanced skill.

By comparing the expected payoffs from each incentive scheme, I show that incentive compatibility constraints always bind and derive that the optimal contract would take one of the following forms:

1. **Winging it always**: The agent is recommended to wing it until the deadline and immediately paid upon the ultimate breakthrough. If the agent does not make the breakthrough until the deadline, the contract is terminated.

2. **Studying always**: The agent is recommended to study until the deadline. If the agent improves the skill, the deadline is extended and the agent is recommended to complete the project with the advanced skill. If the agent does not make the skill improvement until the deadline, the contract is terminated.

3. **Switching once from studying to winging it**: The agent is recommended to study until the intermediate deadline and if the agent improves the skill before the intermediate deadline, the deadline is extended and the agent is recommended to complete the project with the advanced skill. If the agent does not make the skill improvement until the intermediate deadline, the agent is recommended to wing it until the deadline and immediately paid upon the ultimate breakthrough. If the agent does not make the breakthrough until the deadline, the contract is terminated.

Moreover, the form of the optimal contract is determined by the payoff of the project and the effectiveness of the studying technology, which is defined as the ratio of the ultimate breakthrough arrival rate with the advanced skill to the skill improvement arrival rate.

When the technologies are not equally efficient, the optimal contract may take a form other than the aforementioned forms. I present numerical examples such that there are two switches of recommended actions in the optimal contract or incentive compatibility constraints do not bind. From these examples, we can see how the efficiency of the technologies affects the optimal contract.

In the rest of this section, I discuss the related literature. In Section 2, I introduce the model. The first best case is analyzed in Section 3. In Section 4, I characterize the optimal contract when the agent's action is not observable to the principal and both technologies are equally efficient. The numerical examples for unequally efficient technologies case are presented in Section 5, then I conclude and list possible extensions of this model in Section 6.

1.1 Related Literature

The dynamic moral hazard literature has been enriched by recent developments of continuous time methods in contract theory. One strand of the literature utilizes Poisson processes to set up problems (Biais et al., 2010; Green and Taylor, 2016a; Bonatti and Hörner, 2017; Varas, 2017; Sun and Tian, 2017), and I follow this approach since this paper focuses on the completion of the project.

The most closely related study is the one by Green and Taylor (2016a), who study a model in which multiple breakthroughs are needed to complete a project and an agent needs to exert an effort (unobservable to the principal) to achieve breakthroughs. If the agent were assumed to complete the project only with the advanced skill, this paper would be identical to their model: the skill improvement serves as the first breakthrough and the completion of the project serves as the second breakthrough.[3] However, the option to complete the project with the basic skill, which is not considered by Green and Taylor (2016a), allows the agent to face a technology allocation problem between completing the project with the basic skill and improving the skill.

The technology allocation problem is naturally related to the multitasking problem in the sense that the agent has multiple options to pursue. In the seminal paper, Holmstrom and Milgrom (1991) considers an economic situation in which a production worker faces multiple tasks such as producing outputs and maintaining quality in a static environment.[4] Several subsequent multitasking problems are also explored in dynamic setups (Manso, 2011; Capponi and Frei, 2015; Varas, 2017; Szydlowski, forthcoming). A common assumption on these previous studies is that each task has a different payoff structure.[5] For example, Manso (2011) studies a two-armed bandit problem in a simple agency model with two periods. The main assumption is that if the agent chooses to explore (or chooses the risky arm), the payoff is stochastic and if the agent chooses to exploit (or chooses the safe arm), the payoff is constant. In contrast, the two technologies in this paper are same in the payoff structure. The difference of these technologies is 'how' the ultimate breakthrough is made–by the basic skill or by the advanced skill.

2 Model

A principal (she) hires an agent (he) to complete a project. The project is conducted in continuous time and can be potentially operated over an infinite horizon: $t \in [0, \infty)$. The

[3] To be precise, it is identical to the tangible first breakthrough case of the working paper version of the paper (Green and Taylor, 2016b). In the published version of the paper, they only consider the case such that the principal cannot observe the first breakthrough.

[4] Dewatripont et al. (2000) and Laux (2001) also study multitasking problems in static environments.

[5] The only paper that does not have this assumption is Varas (2017). He considers a dynamic model with a Poisson process in which the agent chooses between a good project and a bad project. These projects look identical to the principal and yield the same payoff, but differ in the rate of failure.

project requires an ultimate breakthrough and I denote it as the success or use the term "the project succeeds." When the project succeeds, the principal realizes a payoff $\Pi > 0$ and the game ends. While the project is running, the principal incurs an operating cost of $c > 0$ per unit time. The principal is assumed to have an infinite amount of resources to fund the project while the agent is protected by the limited liability. The principal and the agent are both risk neutral and patient, i.e., do not discount the future.

The distinctive feature of this model is that the arrival rate of the breakthrough depends on the agent's skill. I assume that there are two levels of skills: a basic skill and an advanced skill. At the beginning of the game, the agent is only equipped with a basic skill. The agent may acquire an advanced skill by investing in the skill. Denote the agent with the basic skill as the *low* type and the agent with the advanced skill as the *high* type. Assume that the skill improvement is publicly observable, thus, the principal knows the agent's type.

At each instance of time t, the agent allocates 1 unit of effort to the completion of the main task (a_t), the skill improvement (b_t), and the leisure (l_t): $a_t + b_t + l_t = 1$ and $a_t,\ b_t,\ l_t \geq 0$. The allocation of efforts is unobservable to the principal. Let λ be a triple $(\lambda_L, \lambda_S, \lambda_H)$ such that $c < \lambda_L \Pi$, $\lambda_L < \lambda_S$, and $\lambda_L < \lambda_H$. The arrival rate for the main breakthrough is $\lambda_L a_t$ for the low type agent and $\lambda_H a_t$ for the high type agent. The high type agent always assigns $b_t = 0$ because there is no room for improving the skill. The skill improvement for the low type agent arrives with the rate $\lambda_S b_t$ and the low type agent becomes the high type agent when the skill is improved.[6] The agent (regardless of the type) receives ϕl_t as a private flow benefit and ϕ is assumed to be $0 < \phi < c$.

Then, the low type agent's action can be considered as a choice among three technologies: (i) the winging it technology (a_t) with which the success requires one breakthrough with a low arrival rate (λ_L); (ii) the studying technology (b_t) with which the success requires two breakthrough with high arrival rates ($\lambda_S,\ \lambda_H$); (iii) the leisure (l_t) with which the success never arrives. For the high type agent, an action can be understood as a choice between the working technology (a_t) and the leisure.

3 Observable Action and Skill Improvement

In this section, I assume that the agent's allocation of efforts and type are observable to the principal and characterize the first best contract.

[6]Since the skill improvement is assumed to be arrived with a Poisson arrival rate $\lambda_S b_t$, the current effort level is the sole factor of the skill improvement. If one wants to consider that a past effort affects the chance of the current skill improvement, the arrival rate would need to be proportional to the cumulative effort level.

I consider two benchmark effort schedules and will show that one of the schedules is the first-best schedule. Let τ_s be the (random) date that the skill is improved and τ_m be the (random) date that the main task is completed. Then, the two schedules are defined as follows:

Winging It Only Schedule

$$a_t = 1 \quad \text{for} \quad t \leq \tau_m,$$

Studying Only Schedule

$$b_t = 1 \quad \text{for} \quad t \leq \tau_s \quad \text{and} \quad a_t = 1 \quad \text{for} \quad \tau_s < t \leq \tau_m.$$

In other words, the winging it only schedule means that the agent tries to complete the task only with the basic skill, i.e., only chooses the winging it arm, whereas the studying only schedule means that the agent tries to attain the skill first and then completes the main task with the advanced skill, i.e., only chooses the studying technology until the skill is improved. Note that the low type agent does not switch technologies over time for both schedules.

The probability distribution function of τ_m for the winging it only schedule is given by $\lambda_L e^{-\lambda_L \tau_m}$. On the other hand, for the studying only schedule, the probability distribution of τ_m conditional on skill improvement at τ_s is $\lambda_L e^{-\lambda_L(\tau_m - \tau_s)}$ for $\tau_m > \tau_s$ and 0 for $\tau_m \leq \tau_s$. The marginal probability distribution of τ_s for the studying only schedule is given by $\lambda_S e^{-\lambda_S \tau_s}$. Then, the expected profits for both schedules are given as follows:

Winging It Only Schedule:
$$E_w^* \equiv \int_0^\infty \left(\Pi - \int_0^{\tau_m} c\ dt \right) \lambda_L e^{-\lambda_L \tau_m} d\tau_m = \Pi - \frac{c}{\lambda_L}$$

Studying Only Schedule:
$$E_s^* \equiv \int_0^\infty \int_{\tau_s}^\infty \left(\Pi - \int_0^{\tau_m} c dt \right) \lambda_H e^{-\lambda_H(\tau_m - \tau_s)} d\tau_m \ \lambda_S e^{-\lambda_S \tau_s} d\tau_s = \Pi - \frac{c}{\lambda_S} - \frac{c}{\lambda_H}$$

By comparing the expected profits, we can easily see that the winging it only schedule is indifferent to the studying only schedule if and only if

$$\frac{1}{\lambda_L} = \frac{1}{\lambda_S} + \frac{1}{\lambda_H}. \tag{3.1}$$

Denote that the winging it technology is more efficient if the left hand side of (3.1) is greater than the right hand side, the studying technology is more efficient if the right hand side is greater than the left hand side, and both technologies are equally efficient if the equality holds. Note that $1/\lambda_L$ is the expected duration of the winging it only schedule and $1/\lambda_S + 1/\lambda_H$ is the expected duration of the studying only schedule. Therefore, the efficiency relation gives clear intuition for comparing two schedules: the shorter the expected duration is, the greater the expected payoff is. Furthermore, the following proposition shows that the one of two benchmark schedules is the first-best schedule.

Proposition 3.1. *Suppose that the agent's allocation of efforts and skill improvement are observable to the principal. If the winging it technology is more efficient ($1/\lambda_L < 1/\lambda_S + 1/\lambda_H$), the winging it only schedule is the first best schedule, i.e., it gives the highest expected profit to the principal. If the studying technology is more efficient ($1/\lambda_L > 1/\lambda_S + 1/\lambda_H$), the studying only schedule it the first best schedule.*

4 Optimal Contracts for Equally Efficient Technologies

In this section, I derive optimal contracts for the case where both technologies are equally efficient, the skill improvement is observable to the principal and the agent, but the agent's action is unobservable to the principal.

4.1 Contract

At the beginning of the game, the principal offers a contract to the agent and fully commits to all contractual terms. If the agent rejects the offer, the principal and the agent receive zero. Recall that the agent is low type at the time a contract is proposed, and the arrivals of the success and the skill improvement are observed by both players. Note that if the agent has not made the success or the skill improvement, the calendar time would be the only relevant variable that summarizes the public history. A contract is denoted by $\Gamma^L \equiv \{a^L, b^L, R^L, \Gamma^H, T^L\}$, where each variable is defined as follows at the calendar time t:

- $a_t^L \in [0, 1]$: the recommended effort to the completion of the main task conditional on no success and no skill improvement;
- $b_t^L \in [0, 1]$: the recommended effort to the skill improvement conditional on no success and no skill improvement;

- $R_t^L \geq 0$: the monetary payment from the principal to the agent for success conditional on no skill improvement;[7]

- $\Gamma_t^H \equiv \{a^{H,t}, R^{H,t}, T^{H,t}\}$: an updated contract for skill improvement at time t conditional on no success;
  - $a_s^{H,t} \in [0,1]$: the recommended effort to the completion of the main task at time $s \geq t$ conditional on skill improvement at time t and no success;
  - $R_s^{H,t} \geq 0$: a monetary payment from the principal to the agent for success at time $s \geq t$ conditional on skill improvement at time t and no success;
  - $T^{H,t} \geq t$: the date at which the project is terminated conditional on skill improvement at time t and no success;

- $T^L \geq 0$: the date at which the project is terminated conditional on no success and no skill improvement.

Action processes $a^{H,t}$ and (a^L, b^L) induce probability distributions $\mathbb{P}^{a^{H,t}}$ over a date of success τ_m and $\mathbb{P}^{a^L,b^L}$ over a pair of dates of the skill improvement and the success (τ_m, τ_s). Let $\mathbb{E}^{a^{H,t}}$ and $\mathbb{E}^{a^L,b^L}$ denote the corresponding expectation operators. If the agent is high type and adheres to the recommended action of Γ_t^H, the principal's expected utility at time t is given by

$$P_t^H(\Gamma_t^H) = \mathbb{E}^{a^{H,t}}\left[\left(\Pi - R_{\tau_m}^{H,t}\right) \cdot \mathbf{1}_{\{t \leq \tau_m \leq T^{H,t}\}} - \int_t^{T^{H,t} \wedge \tau_m} c\ ds\right],$$

where the first term in the expectation is the net profit from the success and the second term is the (cumulative) operating cost. The agent's expected utility is given by

$$U_t^H(\Gamma_t^H) = \mathbb{E}^{a^{H,t}}\left[R_{\tau_m}^{H,t} \cdot \mathbf{1}_{\{t \leq \tau_m \leq T^{H,t}\}} + \int_t^{T^{H,t} \wedge \tau_m} \phi(1 - a_s^{H,t}) ds\right],$$

where the first term is the payoff from the success and the second term is the benefit from the leisure.

If the agent is low type and adheres to the recommended actions of Γ^L, the principal's

[7] Since both the principal and the agent are risk neutral and do not discount the future, without loss of generality, all monetary payment to the agent can be backloaded (see, e.g., Ray (2002)).

(ex ante) expected utility is given by

$$P_0^L(\Gamma^L) = \mathbb{E}^{a^L,b^L}\left[\left(\Pi - R_{\tau_m}^L\right)\cdot \mathbf{1}_{\{\tau_m \le \tau_s \wedge T^L\}} + P_{\tau_s}^H(\Gamma_{\tau_s}^H)\cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T^L\}} - \int_0^{T^L\wedge\tau_m\wedge\tau_s} c\ dt\right],$$

where the first term is the net profit from the success, the second term is the expected payoff from the skill improvement at time s, and the last term is the (cumulative) operating cost. The agent's expected utility is given by

$$U_0^L(\Gamma^L) = \mathbb{E}^{a^L,b^L}\left[R_{\tau_m}^L\cdot \mathbf{1}_{\{\tau_m \le T^L\}} + U_{\tau_s}^H(\Gamma_{\tau_s}^H)\cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T^L\}} + \int_0^{T^L\wedge\tau_m\wedge\tau_s} \phi(1 - a_t^L - b_t^L)\ dt\right],$$

where the firms term is the payoff from the success, the second term is the expected payoff from the skill improvement at time s, and the last term is the benefit from the leisure.

Definition 4.1. A contract $\Gamma^L = \left\{a^L, b^L, R^L, \Gamma^H, T^L\right\}$ is incentive compatible if

1. for all $t \le T^L$, the recommended action profile $a^{H,t}$ maximizes the high type agent's expected utility, i.e.,

   $$U_t^H(\Gamma_t^H) \ge \mathbb{E}^{\tilde{a}}\left[R_{\tau_m}^{H,t}\cdot \mathbf{1}_{\{\tau_m \le T^{H,t}\}} + \int_t^{T^{H,t}\wedge\tau_m} \phi(1-\tilde{a}_s)ds\right]$$

   for any action process $\tilde{a} \in \{\{a_s\}_{t\le s\le T^{H,t}}\ :\ a_s \in [0,1]\}$;

2. the recommended action profile (a^L, b^L) maximizes the low type agent's expected utility, i.e.,

   $$U_0^L(\Gamma^L) \ge \mathbb{E}^{\tilde{a},\tilde{b}}\left[R_{\tau_m}^L\cdot \mathbf{1}_{\{\tau_m \le T^L\}} + U_{\tau_s}^H(\Gamma_{\tau_s}^H)\cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T^L\}} + \int_0^{T^L\wedge\tau_m\wedge\tau_s} \phi(1 - \tilde{a}_t - \tilde{b}_t)\ dt\right]$$

   for any action process $(\tilde{a},\tilde{b}) \in \{\{a_t, b_t\}_{0\le t\le T^L}\ :\ (a_t, b_t, a_t + b_t) \in [0,1]^3\}$.

The objective of the principal is to find a contract Γ^L that maximizes her ex ante expected utility $P_0^L(\Gamma^L)$ subject to the incentive compatibility constraint and the individual rationality constraint, i.e., $U_0^L(\Gamma^L) \ge 0$. Denote such contract as an *optimal contract*.

4.2 The Principal's Problems

To derive the optimal contract, I consider the agent's promised utility for each type as a state variable and write a contract recursively.[8] Denote u_H $(u_L) \in \mathbb{R}_+$ as a promised utility for the high (low) type agent.

4.2.1 When the agent is high type

The Hamilton-Jacobi-Bellman (HJB) equation for the high type agent's promise keeping constraint $U_t^H(\Gamma_t^H) = u_H$ gives

$$0 = \max_{a \in [0,1]} \dot{u}_H + \phi(1-a) + \lambda_H a(R - u_H), \tag{PK$_H$}$$

where $\dot{u}_H \equiv du_H/dt$, the second term is the benefit from the leisure, and the last term is the expected additional payoff from completing the task.

Denote a value function $V_H(u_H)$ as a function that maximizes the principal's expected utility $P_t^H(\Gamma_t^H)$ subject to $U_t^H(\Gamma_t^H) = u_H$.[9] The HJB equation for the value function $V_H(u_H)$ is

$$0 = \max_{R \geq 0,\ a \geq 0} -c + (\Pi - R - V_H(u_H))\lambda_H a + V_H'(u_H)\ \dot{u}_H \tag{HJB$_H$}$$

subject to (PK$_H$). Since the outside option of the agent is zero, if the agent's promised utility is equal to zero, the project would not be operated and the principal would also end up getting zero, i.e., $V_H(0) = 0$. This equation serves as a boundary condition.

The above maximization problem is identical to the problem in the single-stage case of Green and Taylor (2016a), thus I can directly use their results. They show that to induce $a = 1$ from (PK$_H$), $\lambda_H(R - u_H) \geq \phi$ should hold and it eventually binds at the optimal contract. Then, they derive $V_H(u_H)$ as follows:

$$V_H(u_H) = \left(\Pi - \frac{c}{\lambda_H} - u_H\right) - \left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi} u_H}, \tag{4.1}$$

where the first term is the first best payoff minus the promised utility u_H and the second term is the agency cost. This payoff can be realized by the contract with a finite deadline, $T = u_H/\phi$, and a diminishing payoff, $R(\tau_m) = \phi(1/\lambda_H + T - \tau_m)$.[10]

[8] This is a typical approach in the dynamic contract literature (see, e.g., Spear and Srivastava, 1987).

[9] We can disregard the time subscript of the high type because no matter what the time is, the contract that maximizes the principal's utility subject to the same promised utility of the agent would be identical.

[10] See Proposition 1 in Green and Taylor (2016a).

4.2.2 When the agent is low type

The HJB equation for the agent's promise keeping constraint $U_0^L(\Gamma^L) = u_L$ gives

$$0 = \max_{\substack{a,b\geq 0,\\ a+b\leq 1}} \dot{u}_L + \phi(1-a-b) + (R-u_L)\lambda_L a + (u_H - u_L)\lambda_S b, \tag{PK$_L$}$$

where $\dot{u}_L \equiv du_L/dt$, the second term is the benefit from the leisure, the third term is the expected additional payoff from completing the task and the last term is the expected additional promised utility from improving the skill.

Denote a value function $V_L(u_L)$ as a function that maximizes the principal's expected utility $P_0^L(\Gamma^L)$ subject to $U_0^L(\Gamma^L) = u_L$. The HJB equation for $V_L(u_L)$ gives that

$$0 = \max_{\substack{R\geq 0,\ u_H\geq 0,\\ a,b\geq 0,\ 1\geq a+b}} -c + (\Pi - R - V_L(u_L))\lambda_L a + (V_H(u_H) - V_L(u_L))\lambda_S b + V_L'(u_L)\ \dot{u}_L \tag{HJB$_L$}$$

subject to (PK$_L$), and the boundary condition is $V_L(0) = 0$.

The above maximization problem solves the principal's problem under the constraint that the agent's promised utility is equal to u_L. Therefore, to derive the optimal contract, the principal solves

$$\max_{u_L\geq 0} V_L(u_L). \tag{MP$_L$}$$

The rest of the section is devoted to derive V_L and solve the above maximization problem.

4.3 Preview of the Main Result

4.3.1 Immediate Payment vs. Deadline Extension

Since (HJB$_L$) and (PK$_L$) are linear in both a and b, we can focus on the pure effort levels (see Lemma A.3 for the detailed argument). Now I introduce two contractual modes with the pure effort levels defined as follows:

1. **Inducing to wing it with a least incentive (Mode W)**

    - The agent is recommended to fully wing it, receives ϕ/λ_L (the least amount of incentive not to shirk) in addition to the current promised utility as a *payment* when the project succeeds, and receives nothing when the skill is improved, i.e., an instantaneous contractual term (a, b, R, u_H) is given as $(1, 0, u_L + \phi/\lambda_L, 0)$;
    - if the principal wants to induce the agent to fully wing it (i.e., $a = 1$), the inequalities $\lambda_L(R - u_L) \geq \phi$ and $\lambda_L(R - u_L) \geq \lambda_S(u_H - u_L)$ need to hold. Then,

we can see that $(a, b, R, u_H) = (1, 0, u_L + \phi/\lambda_L, 0)$ is incentive compatible and the first IC constraint binds.

2. **Inducing to study with a least incentive (Mode S)**

   - The agent is recommended to fully study, receives nothing when the project succeeds, and receives ϕ/λ_S (the least amount of incentive not to shirk) in addition to the current promised utility as a *promised utility* for the high type when the skill is improved, i.e., an instantaneous contractual term (a, b, R, u_H) is given as $(0, 1, 0, u_L + \phi/\lambda_S)$;
   - if the principal wants to induce the agent to fully study (i.e., $b = 1$), the inequalities $\lambda_S(u_H - u_L) \geq \phi$ and $\lambda_S(u_H - u_L) \geq \lambda_L(R - u_L)$ need to hold. Then, we can see that $(a, b, R, u_H) = (1, 0, u_L + \phi/\lambda_L, 0)$ is incentive compatible and the first IC constraint binds.

In the appendix, I show that at each instance of the time, one of the above contractual modes will be executed in the optimal contract. The contractual mode may switch over time in the optimal contract. To simplify the argument, in the main text of the paper, I will focus on comparing the above two contractual modes.

Mode W and Mode S mainly differ in the arrival rate and the form of compensation to the agent. Under Mode W, the agent succeeds with a lower arrival rate (λ_L) and he receives an immediate payment upon success. Under Mode S, the agent makes a skill improvement with a relatively higher arrival rate (λ_S) and he is compensated in the form of the promised utility for the high type agent upon the skill improvement. Note that $\dot{u}_L$ is equal to $-\phi$ under Mode W and Mode S. Then, if the agent's promised utility level for the low type is u_L, the principal employs one of the above contractual modes, and the agent has not made success or skill improvement for u_L/ϕ unit of time, the promised utility becomes zero and the contract is terminated, i.e., the deadline of the contract would be u_L/ϕ. Since $\dot{u}_H$ is also equal to $-\phi$ under the optimal contract for the high type agent and the updated promised utility is $u_L + \phi/\lambda_S$, the updated deadline becomes $u_L/\phi + 1/\lambda_S$, i.e., the deadline is extended by $1/\lambda_S$. Once the deadline is extended, the agent is expected to exert full efforts on completing the project with the advanced skill. In sum, the agent is compensated by an immediate payment with a lower probability under Mode W, whereas he is compensated by a deadline extension with a higher chance under Mode S.

At each instance of the time, the principal chooses the contractual mode by comparing the expected payoffs from the immediate payment and the deadline extension. In the appendix, it is shown that Mode W would be preferred to Mode S only if

1. the time is close to the deadline,

2. the payoff from the project (Π) is small enough, and

3. $\kappa \equiv \lambda_H/\lambda_S$ is low enough.

Under the assumption that both technologies are equally efficient, i.e., $1/\lambda_L = 1/\lambda_S + 1/\lambda_H$, note that λ_S and λ_H can be written as follows:

$$\lambda_S = (1 + 1/\kappa)\,\lambda_L \quad \text{and} \quad \lambda_H = (1+\kappa)\lambda_L. \tag{4.2}$$

Then, κ can be interpreted as the effectiveness of the skill improvement because low κ implies that it is easy to obtain the advanced skill but it is hard to complete the task with the advanced skill.

Intuitive explanations for the above three necessary conditions are as follows.

1. If the time is adequately far from the deadline, the principal would benefit from checking the intermediate progress and it would lessen the moral hazard problem, thus, Mode S might be preferred over Mode W. However, if it is close to the deadline, Mode S requires two breakthroughs for a relatively short period of time, whereas Mode W only requires a breakthrough. Therefore, it may be possible that Mode W is preferred to Mode S for the principal when the time is close to the deadline.

2. Since Mode S employs a deadline extension as a compensation method, the probability of making a success and the expected length of the contract under Mode S would be greater than that under Mode W. Note that the expected revenue is Π times the probability of making a success and the expected cost is c times the expected length of the contract. Then, since the expected cost under Mode S is larger than that under Mode W due to the deadline extension, if Π is small enough, Mode W might be preferred to Mode S.

3. Recall that if the skill is improved under Mode S, the deadline would be extended by $1/\lambda_S$. Then, the agent will have a chance to complete the task with the arrival rate λ_H. Note that a lower κ means that the deadline extension is shorter and the arrival rate is lower, i.e., it is less likely that the second breakthrough arrives. Therefore, when κ is low, the deadline extension might be a less appealing compensation method than the immediate payment.

4.3.2 Thresholds

In this subsection, I argue that the form of the optimal contract is determined by three factors: (i) the recommended action at the deadline, (ii) the length of the contract, and (iii) feasibility. From these factors, three thresholds of Π depending on κ would also be derived to characterize the optimal contract.

Firstly, the first and the second necessary condition of the previous subsection imply that if Π is very big and Mode S is preferred to Mode W at the deadline, Mode S would also be preferred to Mode W for any time during the contract. Then, for a given κ, let $\Pi_S(\kappa)$ be a threshold that characterizes the above property, i.e., if $\Pi > \Pi_S(\kappa)$, Mode S is preferred over Mode W at the deadline, whereas if $\Pi < \Pi_S(\kappa)$, Mode W is preferred over Mode S at the deadline.

Secondly, note that as Π increases, the principal would also want to increase the length of the contract to have more chances for a success to compensate the expected cost. Then, if Π is small, the length of the contract would be short and Mode W might be preferred over mode S during the contract. Define $\Pi_W(\kappa)$ to be the corresponding threshold for a given κ, i.e., if $\Pi < \Pi_W(\kappa)$, Mode W is preferred over Mode S while the contract is running, whereas if $\Pi > \Pi_W(\kappa)$, there is an instance such that Mode S is preferred over Mode W.

Lastly, if Π is very small, it might be optimal for the principal not to begin the contract in the first place. Denote that the project is feasible if contracting with the agent for a positive length of the time is profitable to the principal. Define $\Pi_F(\kappa)$ to be the threshold that determines feasibility for a given κ, i.e., if $\Pi > \Pi_F(\kappa)$, the project is feasible, whereas if $\Pi < \Pi_F(\kappa)$, the project is infeasible. A remark is that there might be some cases such that the project is infeasible even if it is optimal to have make a contract under the first best, i.e., $\Pi_F(\kappa) > c/\lambda_L$.

The next step is to compare the above three thresholds. In Lemma 4.3, it will be shown that there exists a threshold of κ (κ^*) such that the order of three thresholds would be determined according to whether κ is greater than or less than κ^*. Specifically, if κ is greater than κ^*, $\Pi_F(\kappa)$ would be greater than both $\Pi_S(\kappa)$ and $\Pi_W(\kappa)$, whereas if κ is smaller than κ^*, the inequality $\Pi_S(\kappa) > \Pi_W(\kappa) > \Pi_F(\kappa)$ holds. The result is mainly due to the third necessary condition of the previous subsection: if κ is high enough, Mode W may not be preferred to Mode S. These relationships among thresholds are illustrated in the left figure of Figure 1.

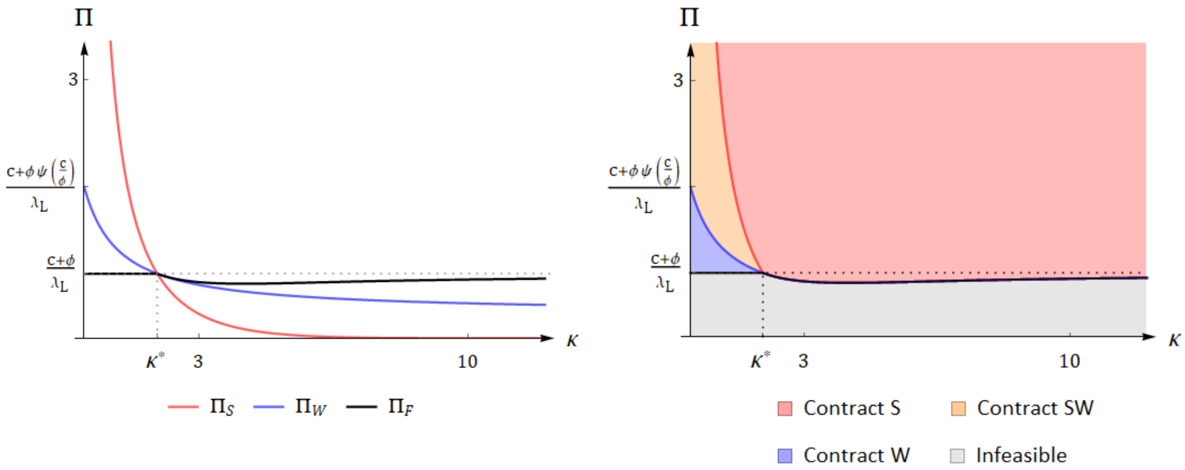

Figure 1: Thresholds and optimal contracts when $(c, \phi, \lambda_L) = (1, 0.5, 1)$

4.3.3 Candidates for the Optimal Contract

I introduce three candidates for the optimal contract (with the deadline T^L): (i) executing Mode W always, (ii) executing Mode S always, (iii) executing Mode S before T^s and executing Mode W after T^s ($< T^L$), i.e., switching from Mode S to Mode W. I specify these three candidates in terms of the contract introduced in Section 4.1.

1. **Contract W**

   - for all $0 \leq t \leq T^L$,
     - $(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$, i.e., the recommended action is always to wing it and the agent's payment upon success diminishes over time;
     - if the skill is improved, the contract is terminated;

2. **Contract S**

   - for all $0 \leq t \leq T^L$,
     - $(a_t^L, b_t^L, R_t^L) = (0, 1, 0)$, i.e., the recommended action is always to study and the agent is not paid even if he succeeds by winging it;
     - the updated contract Γ_t^H upon skill improvement at time t is given as follows:
       * the deadline is extended by $1/\lambda_S$, i.e., $T^H = T^L + 1/\lambda_S$;
       * for all $t \leq s \leq T^H$, $(a_s^{H,t}, R_s^{H,t}) = (1, \phi(T^H - s + 1/\lambda_H))$;

3. **Contract SW**

- there is a switch of the recommended action at T^s ($< T^L$):
  - for all $0 \leq t < T^s$,
    - $(a_t^L, b_t^L, R_t^L) = (0, 1, 0)$, i.e., the recommended action is to study and the agent is not paid even if he succeeds by winging it;
    - the updated contract Γ_t^H upon skill improvement at time t is given as follows:
      - the deadline is extended by $1/\lambda_S$, i.e., $T^H = T^L + 1/\lambda_S$;
      - for all $t \leq s \leq T^H$, $(a_s^{H,t}, R_s^{H,t}) = (1, \phi(T^H - s + 1/\lambda_H))$;
  - for all $T^s \leq t \leq T^L$,
    - $(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$, i.e., the recommended action is to wing it and the agent is paid when he succeeds;
    - if the skill is improved, the contract is terminated.

The main result of the paper is that the optimal contract would take a form of one of above three contracts when both technologies are equally efficient. Moreover, the optimal contract would be determined as follows (Theorem 1):

1. when κ is above a threshold κ^*,

   (a) if Π is above a threshold $\Pi_F(\kappa)$, Π is also greater than $\Pi_S(\kappa)$ from $\Pi_F(\kappa) > \Pi_S(\kappa)$, thus Mode S is preferred to Mode W at the deadline and the optimal contract would take a form of Contract S;

   (b) if Π is below $\Pi_F(\kappa)$, the project is infeasible;

2. when κ is below a threshold κ^*,

   (a) if Π is very big ($\Pi > \Pi_S(\kappa)$), thus Mode S is preferred to Mode W at the deadline and the optimal contract would take a form of Contract S;

   (b) if Π is moderately big ($\Pi_S(\kappa) \geq \Pi > \Pi_W(\kappa)$), Mode W is preferred to Mode S at the deadline but the length of the contract is long enough to have a switch of the recommended action, thus the optimal contract would take a form of Contract SW;

   (c) if Π is moderately small ($\Pi_W(\kappa) \geq \Pi > \Pi_F(\kappa)$), Mode W is preferred to Mode S at the deadline and the length of the contract is so short that the only recommended action is to wing it, thus the optimal contract would take a form of Contract W;

   (d) if Π is below $\Pi_F(\kappa)$, the project is infeasible.

4.4 Derivation of the Optimal Contract

In this subsection, I characterize the optimal contract by solving the principal's maximization problem (MP_L). Five steps to derive the optimal contract are as follows:

1. determine the recommended action at the deadline and derive $\Pi_S(\kappa)$;
2. derive the value function;
3. check the feasibility of the project and derive $\Pi_F(\kappa)$;
4. check the length of the contract and derive $\Pi_W(\kappa)$;
5. specify the contract.

4.4.1 The Recommended Action at the Deadline

Since we have the boundary condition $V_L(0) = 0$, it is natural to begin with solving the value function near $u_L = 0$. In other words, I will determine which contractual mode would be selected at the deadline. To determine which contractual modes to recommend, we need to compare the right hand sides of (HJB_L) at $u_L = 0$ for two contractual modes:

$$\begin{aligned}
&\textbf{Mode W:} \quad && -c + \lambda_L \Pi - \phi - \phi V_L'(0), \\
&\textbf{Mode S:} \quad && -c + \lambda_S V_H\left(\frac{\phi}{\lambda_S}\right) - \phi V_L'(0) \\
& && = -c + \underbrace{\left(1 + \frac{1}{\kappa}\right)\left(1 - e^{-\kappa}\right)\lambda_L \Pi}_{\substack{\text{Expected payoff from} \\ \text{the deadline extension}}} - \underbrace{\frac{1}{\kappa}\left(1 - e^{-\kappa}\right)c}_{\substack{\text{Expected cost from} \\ \text{the deadline extension}}} - \phi - \phi V_L'(0).
\end{aligned}$$

In Mode W, the principal's expected instantaneous payoff is $\lambda_L \Pi$ and the principal incurs the operating cost c and pays ϕ to the agent in expectation. In Mode S, with the arrival rate λ_S, the agent's type becomes high and the promised utility is refueled to ϕ/λ_S (and the deadline is extended), thus the principal earns $\lambda_S V_H(\phi/\lambda_S)$ in expectation and incurs the operating cost c. $\lambda_S V_H(\phi/\lambda_S)$ can be decomposed to three parts: (i) the expected payoff from the deadline extension ($(1 + 1/\kappa)(1 - e^{-\kappa})\lambda_L \Pi$); (ii) the expected cost from the deadline extension ($-1/\kappa \cdot (1 - e^{-\kappa})c$); (iii) the incentive payment to the agent ($-\phi$).

Note that $\lambda_L < (1 + 1/\kappa)(1 - e^{-\kappa})\lambda_L$, equivalently $e^{\kappa} > \kappa + 1$, for all $\kappa > 0$. From this observation, we can see that Mode S would be preferred if Π is large enough and Mode W would be preferred if Π is small. The following lemma determines the threshold.

Lemma 4.1. *Define*

$$\Pi_S(\kappa) \equiv \frac{1-e^{-\kappa}}{1-(\kappa+1)e^{-\kappa}} \cdot \frac{c}{\lambda_L}.\text{[11]} \tag{4.3}$$

When $\Pi_S(\kappa) \geq \Pi > c/\lambda_L$, Mode W is preferred to Mode S at the deadline, i.e., $\lambda_L \Pi - \phi \geq \lambda_S V_H(\phi/\lambda_S)$. When $\Pi > \Pi_S(\kappa)$, Mode S is preferred to Mode W at the deadline, i.e., $\lambda_S V_H(\phi/\lambda_S) > \lambda_L \Pi - \phi$. In addition, $\Pi_S(\kappa)$ is decreasing in κ.

4.4.2 Value Function Derivation

In this subsection, I characterize the principal's value function for the low type agent. I guess the value function in the main text and verify in the appendix. The guess of the value function is based on the following intuition: (i) when it is close to the termination, it may be easier for the principal to induce the agent to wing it than to induce to study because winging it requires only one breakthrough for success; (ii) when the time is moderately far from the deadline, the principal may prefer inducing the agent to study because the principal can obtain an additional benefit by observing the agent's intermediate breakthrough. I introduce two value functions reflecting this intuition (derivation of the value functions is relegated to Appendix A.2.1).

1. Let $V_L^w : \mathbb{R}_+ \to \mathbb{R}$ be the value function that induces the agent to wing it ($a = 1$) with $R = u_L + \phi/\lambda_L$ (Mode W) for all $u_L \geq 0$. Then, (HJB_L) becomes

$$0 = -c + \lambda_L \left(\Pi - \frac{\phi}{\lambda_L} - u_L - V_L^w(u_L) \right) - \phi V_L^{w\prime}(u_L) \tag{4.4}$$

with the boundary condition $V_L^w(0) = 0$.

By solving the differential equation, we obtain

$$V_L^w(u_L) = \left(\Pi - \frac{c}{\lambda_L} \right) \left(1 - e^{-\frac{\lambda_L}{\phi} u_L} \right) - u_L. \tag{4.5}$$

2. Let $V_L^{ws}(\cdot|u_s) : [u_s, \infty) \to \mathbb{R}$ be the value function that induces the agent to wing it ($a = 1$) with $R = u_L + \phi/\lambda_L$ (Mode W) for $0 \leq u_L < u_s$ and to study ($b = 1$) with $u_H = u_L + \phi/\lambda_S$ (Mode S) for $u_L \geq u_s$. Then, (HJB_L) for $u_L \geq u_s$ becomes

$$0 = -c + \lambda_S \left(V_H \left(u_L + \frac{\phi}{\lambda_S} \right) - V_L^{ws}(u_L|u_s) \right) - \phi V_L^{ws\prime}(u_L|u_s) \tag{4.6}$$

[11] Note that $1 - e^{-\kappa} > 1 - (\kappa+1)e^{-\kappa} > 0$ for all $\kappa > 0$, thus, $\Pi_S(\kappa) > c/\lambda_L$.

with the boundary condition $V_L^{ws}(u_s|u_s) = V_L^w(u_s)$.

By solving the differential equation, we obtain

$$\begin{aligned} V_L^{ws}(u_L|u_s) = &\left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S}\right)\left(1 - e^{\frac{\lambda_S}{\phi}(u_s - u_L)}\right) + (V_L^w(u_s) + u_s)e^{\frac{\lambda_S}{\phi}(u_s - u_L)} \\ &- \left(\Pi - \frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi}u_s}}{\lambda_S - \lambda_H}\left(e^{\frac{\lambda_H}{\phi}(u_s - u_L)} - e^{\frac{\lambda_S}{\phi}(u_s - u_L)}\right) - u_L.\text{[12]} \end{aligned} \tag{4.7}$$

Note that V_L^w depends on λ_L and V_L^{ws} depends on λ_L and κ (Recall that λ_L and κ determine λ_S and λ_H by (4.2)). Throughout the paper, I will mostly omit that these value functions depend on λ_L or κ, but when I need to describe that V_L^{ws} depends on κ, I will denote it at $V_L^{ws}(\cdot|u_s, \kappa)$.

The following lemma is useful for the derivation of the value funcion.

Lemma 4.2. *Suppose that both technologies are equally efficient. Then, the followings hold:*

(a) $V_L^w(u_L) < \Pi - c/\lambda_L - u_L$, $V_L^{w\prime}(u_L) > -1$ *and* $V_L^{w\prime\prime}(u_L) < 0$*;*

(b) Suppose that u_s *satisfy* $V_L^{w\prime}(u_s) \le V_L^{ws\prime}(u_s|u_s)$*. Then, for all* $u_L \ge u_s$, $V_L^{ws}(u_L|u_s) < \Pi - c/\lambda_L - u_L$, $V_L^{ws\prime}(u_L|u_s) > -1$ *and* $V_L^{ws\prime\prime}(u_L|u_s) < 0$.

The first result means that contracting with Mode W cannot achieve the first best level $\Pi - c/\lambda_L - u_L$, but performs better than immediately paying out the promised utility ($V_L^{w\prime}(u_L) > -1$). Moreover, the principal's value would be strictly concave with respect to u_L. The second result means that if Mode W is executed from $[0, u_s)$ and the instantaneous benefit from Mode S at u_s is greater than that from Mode W, the same result as above holds for contracting with Mode S (for $u_L \ge u_s$). The assumption $V_L^{w\prime}(u_s) \le V_L^{ws\prime}(u_s|u_s)$ is essential for the result. If the assumption is violated, $V_L^{ws}(\cdot|u_s)$ may no longer be concave. For example, see Figure 2d. In the example, $V_L^{w\prime}(0) > 0 > V_L^{ws\prime}(0|0)$, so the assumption does not hold. We can easily see that $V_L^{ws}(\cdot|0)$ is not concave.

The principal's value function for the low type agent would consist of two parts: (i) when the promised utility u_L is lower than a threshold $u_s(\kappa)$, i.e., close to the deadline, Mode W is executed and the principal's value function is $V_L^w(u_L)$; (ii) when the promised utility u_L is

[12] When $\lambda_S = \lambda_H$, the penultimate term needs to be changed. Note that

$$\lim_{\lambda_S \to \lambda_H} \frac{e^{\frac{\lambda_H}{\phi}(u_s - u_L)} - e^{\frac{\lambda_S}{\phi}(u_s - u_L)}}{\lambda_S - \lambda_H} = \frac{(u_L - u_s)e^{\frac{\lambda_H}{\phi}(u_s - u_L)}}{\phi}.$$

Hence, the penultimate term becomes $-\left(\Pi - c/\lambda_H\right)(u_L - u_s)e^{-1-\lambda_H u_L/\phi}/\phi$.

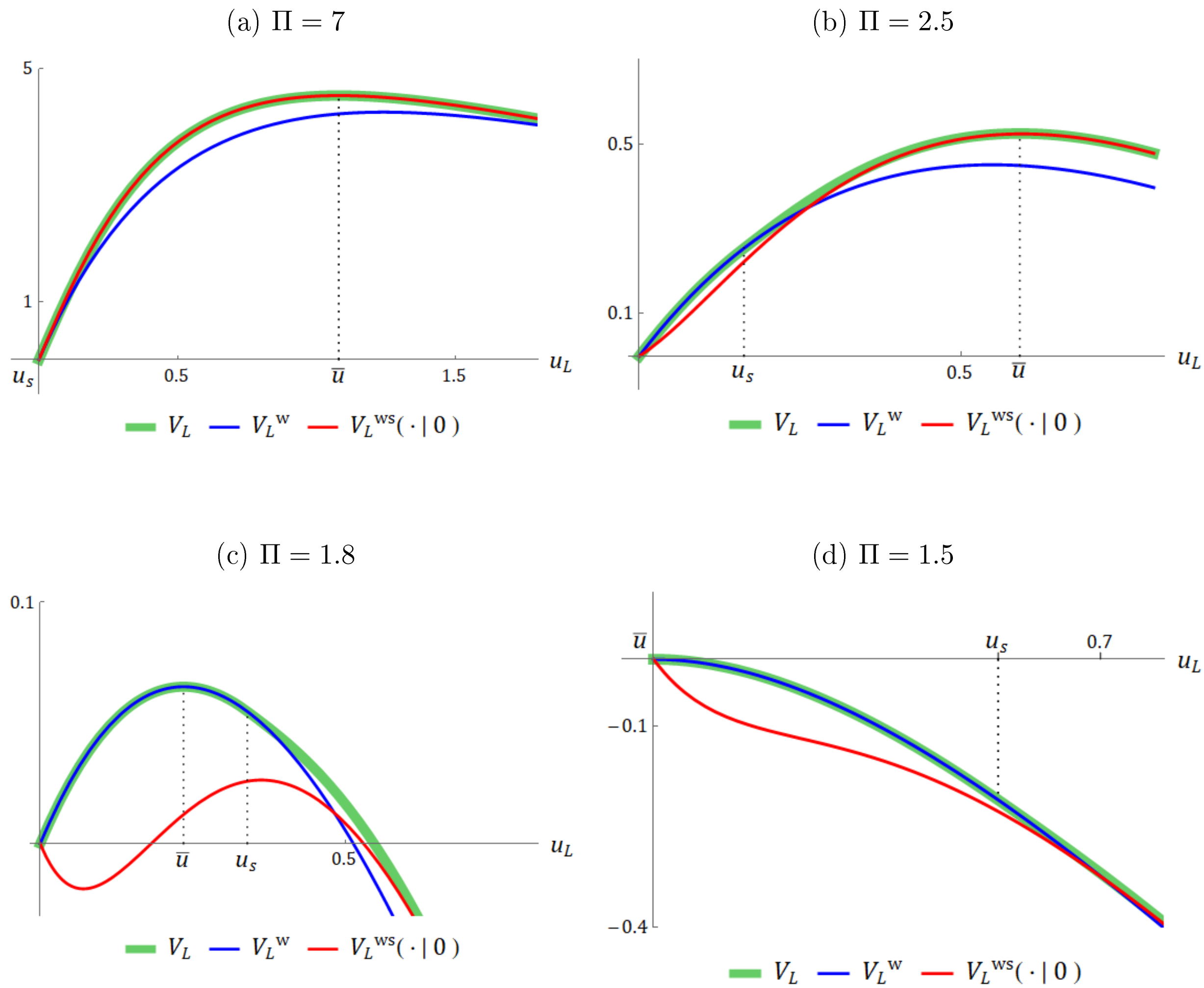

Figure 2: The value functions and benchmark value functions when parameter values are $\lambda_L = 1,\ \lambda_S = 4,\ \lambda_H = 4/3,\ c = 1,\ \phi = 0.5$

higher than $u_s(\kappa)$, i.e., far from the deadline, Mode S is executed and the principal's value function is $V_L^{ws}(u_L|u_s(\kappa))$. The threshold $u_s(\kappa)$ is identified by the smooth pasting condition $V_L^{w\prime}(u_s(\kappa)) = V_L^{ws\prime}(u_s|u_s)$, equivalently,

$$\lambda_L \left(\Pi - \frac{\phi}{\lambda_L} - u_s - V_L^w(u_s) \right) = \lambda_S \left(V_H \left(u_s + \frac{\phi}{\lambda_S} \right) - V_L^w(u_s) \right). \tag{4.8}$$

One caveat is that if Π is greater than $\Pi_S(\kappa)$, by the argument in the previous subsection, Mode S is preferred to Mode W at the deadline, so only second part of the conjecture contract would be executed in this case. The following proposition formally states the above arguments.

Proposition 4.1. *Assume that both technologies are equally efficient and λ_H/λ_S is equal to κ. Then, the principal's value function for the low-skilled agent are characterized as follows:*

(a) if $\Pi > \Pi_S(\kappa)$, the value function is $V_L(u_L) = V_L^{ws}(u_L|0,\kappa)$;

(b) if $\Pi_S(\kappa) \geq \Pi > c/\lambda_L$,

(i) define $u_s(\kappa) \equiv \frac{\phi}{\lambda_L}\left[\frac{1}{\kappa}\log\left(\frac{\lambda_H\Pi-c}{\lambda_L\Pi-c}\right)-1\right]$, then $u_s(\kappa) \geq 0$;

(ii) $V_L^{w\prime}(u_s(\kappa)) = V_L^{ws\prime}(u_s(\kappa)|u_s(\kappa),\kappa)$;

(iii) for all $u_L \leq u_s(\kappa)$, the value function is $V_L(u_L) = V_L^w(u_L)$;

(iv) for all $u_L > u_s(\kappa)$, the value function is $V_L(u_L) = V_L^{ws}(u_L|u_s(\kappa),\kappa)$.

Note that V_L also depends on κ and denote $V_L(\cdot|\kappa)$ if needed. To illustrate, consider a numerical example with $c = 1$, $\phi = .5$, $\lambda_L = 1$, $\lambda_S = 4$, $\lambda_H = 4/3$ and $\kappa = 1/3$. Graphs of the value function with two benchmark functions (V_L^w and $V_L^{ws}(\cdot|0)$) for several cases are shown in Figure 2. Note that $\Pi_S(1/3) = (1-e^{-1/3})/(1-(4/3)e^{-1/3}) \approx 6.35$. Since $7 > \Pi_S(1)$, Mode S would be executed for all u_L and the value function V_L would be exactly same as $V_L^{ws}(\cdot|0)$. This is illustrated in Figure 2a. For Figures 2b, 2c and 2d, since Π is less than $\Pi_S(\kappa)$, we can see that there is a switch of the contractual mode at $u_s(\kappa)$ in the value function.

An important property of the value function to note is concavity. This is because the two parts of the value function are strictly concave and they are smoothly pasted ($V_L^{w\prime}(u_s(\kappa)) = V_L^{ws\prime}(u_s(\kappa)|u_s(\kappa))$).

Corollary 4.2. *The principal's value function for the low type agent V_L is concave.*

4.4.3 Feasibility

The next step is to check the feasibility of the project. If the maximum of the value function V_L is greater than 0, the principal earns positive expected payoff from the contract, thus the project is feasible. If $V_L'(0) > 0$, there exists $u_L > 0$ such that $V(u_L) > 0$, thus the project is feasible. On the other hand, if $V_L'(0) \leq 0$, since V_L is concave (Corollary 4.2), the maximum of the value function is 0 at $u_L = 0$, so the project is infeasible. Note that from (HJB$_L$), $V_L'(0) > 0$ is equivalent to

$$\max\left[\lambda_L\Pi - \phi, \lambda_S V_H\left(\frac{\phi}{\lambda_S}\right)\right] > c, \tag{4.9}$$

i.e., the project is feasible if at least one of the instantaneous payoff at the deadline covers the operating cost c. Recall that $\lambda_H V_H(\phi/\lambda_S) = (1+1/\kappa)(1-e^{-\kappa})\lambda_L \Pi - (1-e^{-\kappa})c/\kappa$. Define a threshold $\Pi_F(\kappa)$ as

$$\Pi_F(\kappa) \equiv \min\left[\frac{c+\phi}{\lambda_L},\ \frac{1}{(\kappa+1)\lambda_L}\left(c+\frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right)\right]. \tag{4.10}$$

Then, it can be shown that $\Pi_F(\kappa) < \Pi$ is equivalent to (4.9).

In the previous numerical example, $(c+\phi)/\lambda_L = 1.5$ and $\left(c+\frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right)/((\kappa+1)\lambda_L) \approx 2.07$, thus, $\Pi_F(\kappa) = 1.5$. The graph of the value function when $\Pi = 1.5$ is shown in Figure 2d. We can see that the value function takes negative values for all positive promised utility levels. Therefore, the principal is better not to contract with the agent in the first place, i.e., the project is infeasible. For examples with Π higher than 1.5 (Figure 2a, 2b and 2c), there exist positive promised utility levels for the agent that yield positive expected payoffs for the principal, i.e., the project is feasible.

4.4.4 The Length of the Contract

The next step is to solve the maximization problem (MP_L), equivalently to check the length of the contract. Define $\bar{u}(\kappa)$ be the solution of (MP_L) when λ_H/λ_S is equal to κ. Since V_L is concave and differentiable, $\bar{u}(\kappa)$ is the solution of $V_L'(u_L|\kappa) = 0$ when the project is feasible. Then, at the beginning of the contract, the agent's promised utility is $\bar{u}(\kappa)$, and as time goes by, the promised utility diminishes conditional that the agent has made neither success nor skill improvement.

To check whether there is a switch of the recommended action during the contract, we need to compare $u_s(\kappa)$ and $\bar{u}(\kappa)$. If $\bar{u}(\kappa) \le u_s(\kappa)$, the principal always recommends the agent to wing it. Whereas, if $\bar{u}(\kappa) > u_s(\kappa)$, for u_L with $\bar{u}(\kappa) \ge u_L \ge u_s(\kappa)$, the principal's recommendation would be to study, while for u_L with $u_s(\kappa) > u_L$, the recommendation would be to wing it. Since V_L is concave and $V_L'(\bar{u}(\kappa)) = 0$, it is enough to check whether $V_L'(u_s(\kappa))$ is greater or smaller than 0.

By the definition of $u_s(\kappa)$ in Proposition 4.1(b)(i) and $V_L'(u_s(\kappa)) = V_L^{w\prime}(u_s(\kappa))$, $V_L'(u_s(\kappa)) = 0$ is equivalent to

$$\frac{(\lambda_L\Pi - c)^{1+\kappa}}{(1+\kappa)\lambda_L\Pi - c}\left(\frac{e}{\phi}\right)^{\kappa} = 1. \tag{4.11}$$

Moreover, $V_L'(u_s(\kappa)) > 0$ is equivalent to the inequality that the left hand side of the above equation is greater than 1. Note that the left hand side is increasing in Π, is equal to zero when Π is equal to c/λ_L, and diverges as Π goes to infinity. Therefore, there exists a unique solution that satisfy (4.11) and denote the solution as $\Pi_W(\kappa)$. Then, $\Pi > \Pi_W(\kappa)$ is equivalent to $u_s(\kappa) > \bar{u}(\kappa)$.

In the previous numerical example, by solving (4.11), we can obtain that $\Pi_W(1/3) \approx 1.91$. For Figure 2a and 2b, Π is greater than $\Pi_W(1/3)$ and we can see that $\bar{u}$ is greater than u_s.[13] In contrast, Π is smaller than $\Pi_W(1/3)$ and we can see that u_s is greater than $\bar{u}$.[14] In this example, $\Pi_S(1/3) \approx 6.35 > \Pi_W(1/3) \approx 1.91 > \Pi_F(1/3) = 1.5$. This order is preserved for κ smaller than some threshold κ^*. For κ greater than κ^*, Π_F would be greater than Π_S and Π_W. The following lemma characterizes these relationships.

Lemma 4.3. *Define κ^* as a positive solution of*

$$0 = \phi + (c + \phi)\kappa - \phi e^{\kappa}.$$

Then,

1. *if $\kappa > \kappa^*$,*
$$\Pi_F(\kappa) = \frac{1}{(\kappa+1)\lambda_L}\left(c + \frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right) > \max\left[\Pi_W(\kappa),\ \Pi_S(\kappa)\right],$$

2. *if $\kappa = \kappa^*$,*
$$\Pi_F(\kappa^*) = \frac{c+\phi}{\lambda_L} = \frac{1}{(\kappa^*+1)\lambda_L}\left(c + \frac{\kappa^*}{1-e^{-\kappa^*}}(c+\phi)\right) = \Pi_W(\kappa^*) = \Pi_S(\kappa^*);$$

3. *if $\kappa < \kappa^*$,*
$$\Pi_F(\kappa) = \frac{c+\phi}{\lambda_L} < \Pi_W(\kappa) < \Pi_S(\kappa);$$

4. *as $\kappa \to 0$,*
$$\lim_{\kappa\to 0}\Pi_W(\kappa) = \frac{c + \phi\cdot\psi\left(c/\phi\right)}{\lambda_L} \quad \text{and} \quad \lim_{\kappa\to 0}\Pi_S(\kappa) = \infty,$$
where $\psi : \mathbb{R}_+ \to [1,\infty)$ is the inverse function of $x\log(x)$ for $x \geq 1$.

The left panel of Figure 1 illustrates graphs of Π_S, Π_W and Π_F when $c = 1$, $\phi = 0.5$ and $\lambda_L = 1$. We can see that the three graphs coincide at $(\kappa^*, (c+\phi)/\lambda_L)$. When κ is smaller than κ^*, Π_S is greater than Π_W and Π_W is greater than Π_F. In this case, the form of the value function depends on the project's payoff (Π).

[13] When $\Pi = 7$, there is no switch, thus, we can consider $u_s = 0$. By solving the the maximum of V_L, we can derive that $\bar{u} \approx 1.08 > u_s$. When $\Pi = 2.5$, $u_s = \frac{\phi}{\lambda_L}\left[\frac{1}{\kappa}\log\left(\frac{\lambda_H\Pi - c}{\lambda_L\Pi - c}\right) - 1\right] \approx .16$ and we can derive that $\bar{u} \approx .59 > u_s$.

[14] When $\Pi = 1.8$, $u_s \approx .34$ and we can derive that $\bar{u} \approx .235 < u_s$. When $\Pi = 1.5$, the value function is maximized at the origin, i.e., $\bar{u} = 0$ and $u_s \approx .54 > \bar{u}$.

1. If Π is very big ($\Pi > \Pi_S$), the principal's value function corresponds to $V_L^{ws}(u_L|0)$, i.e., the recommended action is always to study and the agent is compensated by the deadline extension. In other words, if the advanced skill is effective enough, the winging it technology would never be recommended.

2. If Π is moderately big ($\Pi_S \geq \Pi > \Pi_W$), the principal's value function corresponds to $V_L^w(u_L)$ for $u_L \leq u_s(\kappa)$ and $V_L^{ws}(u_L|u_s(\kappa))$ for $u_L > u_s(\kappa)$, moreover, $\bar{u}(\kappa) > u_s(\kappa)$. At the beginning of the contract, the agent's promised utility level is $\bar{u}(\kappa)$ and the agent is recommended to study and compensated by the deadline extension. Conditional on no skill improvement, the promised utility decreases as time goes on. If the promised utility reaches $u_s(\kappa)$, the recommended action switches to winging it from then on and the agent is compensated by the immediate payment upon success.

3. If Π is moderately small ($\Pi_W \geq \Pi > \Pi_F$), $u_s(\kappa) \geq \bar{u}(\kappa)$, thus, the principal's value function corresponds to $V_L^w(u_L)$ for all $u_L \leq \bar{u}(\kappa)$. Therefore, the agent is always recommended to wing it and compensated by the immediate payment upon success.

4. If Π is very small ($\Pi_F \geq \Pi$), the project is infeasible.

When κ is bigger than κ^*, Π_F is greater than both Π_S and Π_W. It means that whenever the project is feasible, the principal's value function corresponds to $V_L^{ws}(u_L|0)$. It is same as the case of low κ ($\kappa < \kappa^*$) and the very big project payoff ($\Pi > \Pi_S(\kappa)$).

4.4.5 Implementation

The last step to characterize the optimal contract is to specify a contract that implements $(\bar{u}(\kappa), V_L(\bar{u}(\kappa)))$, i.e., a pair of the agent's promised utility level and the principal's expected payoff, of which the expected payoff is maximized. According to the discussion in the last subsection, there would be three candidates for the optimal contract: (i) executing Mode S always (Contract S), (ii) switching from Mode S to Mode W (Contract SW), and (iii) executing Mode W always (Contract W). Since $\dot{u}_L = -\phi$ for all candidates of the optimal contract, the deadline would be $T^L = \bar{u}(\kappa)/\phi$. The optimal contract would be determined by comparing $\bar{u}(\kappa)$ and $u_s(\kappa)$: (i) if $u_s(\kappa) = 0$, Mode S would be always executed, (ii) if $u_s(\kappa) < \bar{u}(\kappa)$, the contractual mode would be switched at the time of which the promised utility is equal to $u_s(\kappa)$, i.e., $T^s = (\bar{u}(\kappa) - u_s(\kappa))/\phi$, and (iii) if $u_s(\kappa) > \bar{u}(\kappa)$, Mode W would be always executed. The next theorem summarizes the above discussion.

Theorem 1. *Suppose that both technologies are equally efficient. The optimal contract is characterized as follows:*

1. *when* $\kappa \geq \kappa^*$ & $\Pi > \Pi_F(\kappa)$ *or* $\kappa < \kappa^*$ & $\Pi > \Pi_S(\kappa)$*, Contract S with* $T^L = \bar{u}(\kappa)/\phi$ *is the optimal contract;*

2. *when* $\kappa < \kappa^*$ *and* $\Pi_S(\kappa) \geq \Pi > \Pi_W(\kappa)$*, Contract SW with* $T^L = \bar{u}(\kappa)/\phi$ *and* $T^s = (\bar{u}(\kappa) - u_s(\kappa))/\phi$ *is the optimal contract;*

3. *when* $\kappa < \kappa^*$ *and* $\Pi_W(\kappa) \geq \Pi > \Pi_F(\kappa)$*, Contract W with* $T^L = \bar{u}(\kappa)/\phi$ *is the optimal contract;*

4. *when* $\Pi_F(\kappa) \geq \Pi$*, the project is infeasible.*

5 Examples for Unequally Efficient Technologies

In this section, I relax the assumption that technologies are equally efficient. Then, it is no longer true that one of three contracts introduced in the previous section should be the optimal contract. Since there are too many subcases for the unequally efficient technologies case, rather than characterizing optimal contracts for every cases, I provide two numerical examples that are somewhat different from the optimal contracts under the equally efficient technologies case.

5.1 The Role of Efficiency

Before presenting the numerical examples, I will illustrate that the efficiency determines the recommended action at the time far from the deadline, i.e., large enough u_L. By (4.5) and (4.7), we can derive that

$$\lim_{u_L \to \infty} V_L^w(u_L) + u_L = \Pi - \frac{c}{\lambda_L},$$
$$\lim_{u_L \to \infty} V_L^{ws}(u_L|u_s) + u_L = \Pi - \frac{c}{\lambda_S} - \frac{c}{\lambda_H},$$

for all $u_s \geq 0$.

When both technologies are equally efficient, both $V_L^w(u_L)$ and $V_L^{ws}(u_L|u_s)$ converge to an asymptotic line $\Pi - c/\lambda_L - u_L$. Nevertheless, Proposition 4.1 suggests that the recommended action is to study when it is far from the deadline.

When technologies are unequally efficient, $V_L^w(u_L)$ and $V_L^{ws}(u_L|u_s)$ no longer converge to the same asymptotic line and the result of Proposition 4.1 may not hold . Graphs in Figure 3a and 3b illustrate that as u_L increases, $V_L^w(u_L)$ converges to $\Pi - c/\lambda_L - u_L$ and $V_L^{ws}(u_L|0)$

(a) Winging it is more efficient (b) Studying is more efficient

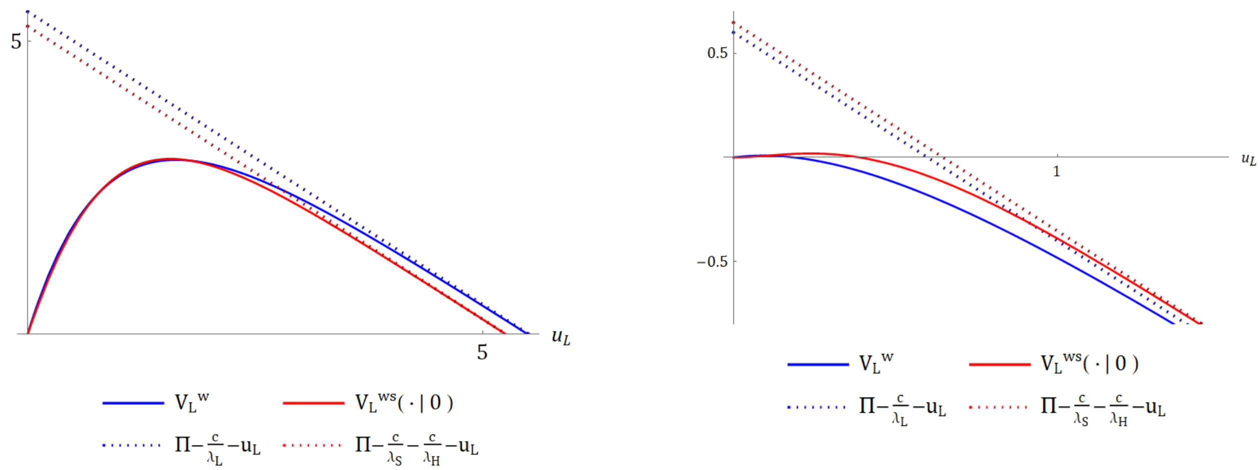

Figure 3: The benchmark value functions and asymptotic lines

converges to $\Pi - c/\lambda_S - c/\lambda_H - u_L$. When the winging it technology is more efficient, $V_L^w(u_L)$ would be greater than $V_L^{ws}(u_L|u_s)$ for large enough u_L. It means for large enough u_L, switching the recommended action from studying (for $u_s \leq u \leq u_L$) to winging it (for $0 \leq u < u_s$) may not be optimal because it gives less expected payoff than winging it always (for $0 \leq u \leq u_L$). Likewise, when the studying technology is more efficient, $V_L^{ws}(u_L|u_s)$ would be greater than $V_L^w(u_L)$ for large enough u_L. In this case, winging it always may not be preferred to switching from studying to winging it. From these observations, we can guess that the recommended technology would be more efficient one when the time is far from the deadline.

5.2 The Winging It Technology is more Efficient

In this subsection, I present a numerical example of which optimal contract involves two switches of contractual modes as follows:

1. Mode W is executed when the time is far from the deadline ($0 \leq t \leq T^{s,1}$);

2. Mode S is executed when the time is moderately far from the deadline ($T^{s,1} < t < T^{s,2}$);

3. Mode W is executed when the time is close to the deadline ($T^{s,2} \leq t \leq T^L$).

The contract that has the above property can be described as follows:

- for all $0 \leq t \leq T^{s,1}$ and $T^{s,2} \leq t \leq T^L$,

- $(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$, i.e., the recommended action is to wing it and the agent is paid when he succeeds;
- if the skill is improved, the contract is terminated;

- for all $T^{s,1} < t < T^{s,2}$,
  - $(a_t^L, b_t^L, R_t^L) = (0, 1, 0)$, i.e., the recommended action is to study and the agent is not paid even if he succeeds by winging it;
  - the updated contract Γ_t^H upon skill improvement at time t is given as follows:
    * the deadline is extended by $1/\lambda_S$, i.e., $T^H = T^L + 1/\lambda_S$;
    * for all $t \leq s \leq T^H$, $(a_s^{H,t}, R_s^{H,t}) = (1, \phi(T^H - s + 1/\lambda_H))$.

Define some points of the agent's promised utility as follows:

$$\bar{u} = \phi \cdot T^L, \; u_{s,1} \equiv \phi \cdot (T^L - T^{s,1}), \quad \text{and } u_{s,2} \equiv \phi \cdot (T^L - T^{s,2}).$$

The value functions for $0 \leq u_L \leq u_{s,1}$ would be identical to those in the equally efficient technologies case ((4.5) for $0 \leq u_L \leq u_{s,2}$ and (4.7) for $u_{s,2} \leq u_L \leq u_{s,1}$). For $u_{s,1} < u_L$, the value function $V_L^{wsw}(\cdot|u_{s,2}, u_{s,1}) : [u_{s,1}, \infty) \to \mathbb{R}$, is derived by solving the differential equation identical to (4.4) with the boundary condition $V_L^{wsw}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{ws}(u_{s,1}|u_{s,2})$:

$$\begin{aligned} V_L^{wsw}(u_L|u_{s,2}, u_{s,1}) \equiv & \left(\Pi - \frac{c}{\lambda_L}\right)\left(1 - e^{\frac{\lambda_L}{\phi}(u_{s,1} - u_L)}\right) - u_L \\ & + \left(V_L^{ws}(u_{s,1}|u_{s,2}) + u_{s,1}\right) e^{\frac{\lambda_L}{\phi}(u_{s,1} - u_L)}. \end{aligned}$$

In sum, the value function is guessed as follows:

$$V_L(u_L) = \begin{cases} V_L^w(u_L), & 0 \leq u_L \leq u_{s,2}, \\ V_L^{ws}(u_L|u_{s,1}), & u_{s,2} \leq u_L \leq u_{s,1}, \\ V_L^{wsw}(u_L|u_{s,1}, u_{s,2}), & u_{s,1} \leq u_L. \end{cases} \tag{5.1}$$

The next step is to identify $u_{s,1}$ and $u_{s,2}$ by the smooth pasting conditions $V_L^{w\prime}(u_{s,2}) = V_L^{ws\prime}(u_{s,2}|u_{s,2})$ and $V_L^{ws\prime}(u_{s,1}|u_{s,2}) = V_L^{wsw\prime}(u_{s,1}|u_{s,2}, u_{s,1})$. First, solve $V_L^{w\prime}(u_{s,2}) = V_L^{ws\prime}(u_{s,2}|u_{s,2})$, equivalently (4.8), to obtain $u_{s,2}$. If there are multiple positive solutions, the least positive solution should be chosen as $u_{s,2}$. Then, $u_{s,1}$ ($> u_{s,2}$) can be obtained by solving $V_L^{ws\prime}(u_{s,1}|u_{s,2}) = V_L^{wsw\prime}(u_{s,1}|u_{s,2}, u_{s,1})$, equivalently

$$\lambda_S\left(V_H\left(u_{s,1} + \frac{\phi}{\lambda_S}\right) - V_L^{ws}(u_{s,1}|u_{s,2})\right) = \lambda_L\left(\Pi - \frac{\phi}{\lambda_L} - u_{s,1} - V_L^{ws}(u_{s,1}|u_{s,2})\right). \tag{5.2}$$

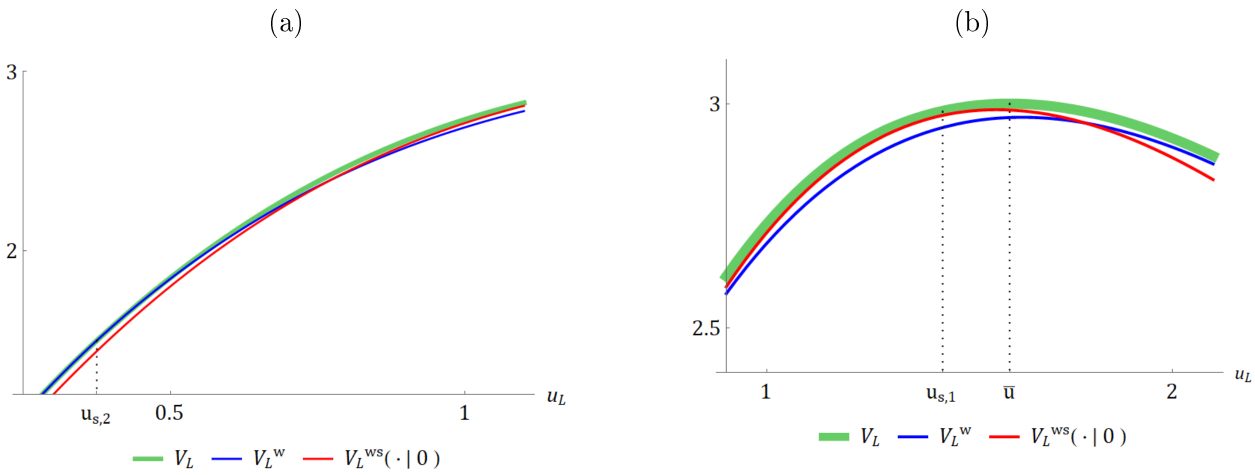

Figure 4: The value function and benchmark functions when parameters are $\lambda_L = 1,\ \lambda_S = \lambda_H = 1.6,\ \Pi = 6.5,\ c = 1,\ \phi = 0.9$

I remark that when both technologies are equally efficient the above equation would not have a solution, thus the value function of this case does not have the third part of (5.1).

Now I present a numerical example of which value function takes a form of (5.1). Let parameter values be $(\lambda_L, \lambda_S, \lambda_H, \Pi, c, \phi) = (1, 1.6, 1.6, 6.5, 1, 0.9)$. Note that $1/\lambda_L < 1/\lambda_S + 1/\lambda_H$, i.e., the winging it technology is more efficient than the studying technology. First, by solving (4.8), we can derive that $u_{s,2} \approx 0.375$.[15] Second, by solving (5.2), we have $u_{s,1} \approx 1.342$. Lastly, we can derive that V_L is maximized at $\bar{u} \approx 1.599$. Figure 4 illustrates the value function V_L and the benchmark value functions V_L^w and $V_L^{ws}(\cdot|0)$ for these parameter values.

From (5.2), we can derive that $T^{s,1} \approx .181$, $T^{s,2} \approx 1.360$ and $T^L \approx 1.776$. Then, at the beginning of the optimal contract, the agent's promised utility is $\bar{u}$ and he is recommended to wing it and compensated by the immediate payment. If the agent has not made success until time $T^{s,1}$, the recommended action is switched to study and if the agent improves skill, the deadline would be extended by $1/\lambda_S = .625$ and he is expected to complete the task with the advanced skill. If the agent has not improved skill until time $T^{s,2}$, the recommended action is switched again to wing it and he is compensated by the immediate payment.

5.3 The Studying Technology is more Efficient

In previous cases, the optimal contract consists of Mode W or Mode S, i.e., the incentive compatibility condition always binds. However, this may be no longer true when the studying

[15] The equation (4.8) has two solutions 0.375 and 1.436, and we need to choose the smaller one for $u_{s,2}$.

technology is more efficient. I present a numerical example of which value function is derived as follows:[16]

1. Mode S is executed when the agent's promised utility is large enough ($u_{s,1} \leq u_L$);
2. Studying is recommended and the updated promised utility upon the skill improvement u_H is equal to $u_{s,1} + \phi/\lambda_S$ when the agent's promised utility is moderately large ($u_{s,2} < u_L < u_{s,1}$);
3. Mode W is executed when the agent's promised utility is small enough ($0 \leq u_L \leq u_{s,2}$).

The value functions for $0 \leq u_L \leq u_{s,2}$ would be identical to the one in the equally efficient technologies case, i.e., $V_L(u_L) = V_L^w(u_L)$ as defined in (4.5). For $u_{s,2} < u_L < u_{s,1}$, the value function $V_L^{wn}(\cdot|u_{s,2}, u_{s,1}) : [u_{s,2}, u_{s,1}] \to \mathbb{R}$, is derived by plugging $b = 1$ and $u_{s,2} = u_{s,1} + \phi/\lambda_S$ into (HJB_L) and (PK_L), i.e.,

$$0 = -c + \left(V_H \left(u_{s,1} + \frac{\phi}{\lambda_S} \right) - V_L^{wn}(u_L|u_{s,2}, u_{s,1}) \right) \cdot \lambda_S - V_L^{wn\prime}(u_L|u_{s,2}, u_{s,1}) \cdot (\phi + (u_{s,1} - u_L) \cdot \lambda_S),$$

with the boundary condition $V_L^{wn}(u_{s,2}|u_{s,2}, u_{s,1}) = V_L^w(u_{s,2})$. Then the solution of the above differential equation $V_L^{wn}(\cdot|u_{s,2}, u_{s,1})$ can be derived as follows:

$$V_L^{wn}(u_L|u_{s,2}, u_{s,1}) \equiv V_L^w(u_{s,2}) + \left(V_H \left(u_{s,1} + \frac{\phi}{\lambda_S} \right) - V_L^w(u_{s,2}) - \frac{c}{\lambda_S} \right) \cdot \frac{\lambda_S(u_L - u_{s,2})}{\phi + (u_{s,1} - u_{s,2})\lambda_S}.$$

For $u_{s,1} \leq u_L$, the value function $V_L^{wns}(\cdot|u_{s,2}, u_{s,1}) : [u_{s,1}, \infty) \to \mathbb{R}$, is derived by solving the differential equation identical to (4.6) with the boundary condition $V_L^{wns}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{wn}(u_{s,1}|u_{s,2}, u_{s,1})$:

$$\begin{aligned} V_L^{wns}(u_L|u_{s,2}, u_{s,1}) \equiv & \left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} \right) \left(1 - e^{\frac{\lambda_S}{\phi}(u_{s,1} - u_L)} \right) + (V_L^{wn}(u_{s,1}|u_{s,2}, u_{s,1}) + u_{s,1}) e^{\frac{\lambda_S}{\phi}(u_{s,1} - u_L)} \\ & - \left(\Pi - \frac{c}{\lambda_H} \right) \frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi} u_{s,1}}}{\lambda_S - \lambda_H} \left(e^{\frac{\lambda_H}{\phi}(u_{s,1} - u_L)} - e^{\frac{\lambda_S}{\phi}(u_{s,1} - u_L)} \right) - u_L. \end{aligned}$$

In sum, the value function is guessed as follows:

$$V_L(u_L) = \begin{cases} V_L^w(u_L), & 0 \leq u_L \leq u_{s,2}, \\ V_L^{wn}(u_L|u_{s,2}, u_{s,1}), & u_{s,2} \leq u_L \leq u_{s,1}, \\ V_L^{wns}(u_L|u_{s,2}, u_{s,1}), & u_{s,1} \leq u_L. \end{cases}$$

[16] In this case, the contractual terms in the optimal contract are not easily interpreted, so I present the value function first and then derive the corresponding contract.

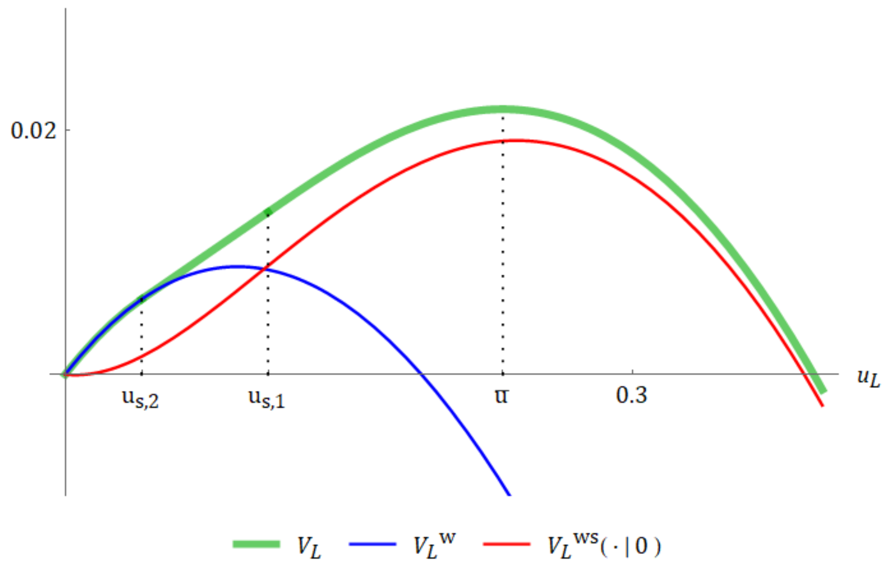

Figure 5: The value function and benchmark value functions when parameters are $\lambda_L = 1,\ \lambda_S = \lambda_H = 2.1,\ \Pi = 1.6,\ c = 1,\ \phi = 0.5$

The next step is to identify $u_{s,1}$ and $u_{s,2}$ by the smooth pasting conditions $V_L^{w\prime}(u_{s,2}) = V_L^{wn\prime}(u_{s,2}|u_{s,2}, u_{s,1})$ and $V_L^{wn\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{wns\prime}(u_{s,1}|u_{s,2}, u_{s,1})$. The detailed derivation is relegated to Appendix B.

Now I specify the contract that implements the above value function. In this case, $\dot{u}_L = -\phi$ does not hold any longer for $u_{s,2} \leq u_L \leq u_{s,1}$. For this region, we have

$$0 = \dot{u}_L + (u_{s,1} - u_L)\lambda_S + \phi \quad \Rightarrow \quad dt = -\frac{du_L}{\phi + \lambda_S(u_{s,1} - u_L)}.$$

Then, we can observe that the time t and the promised utility u_L conditional on no skill improvement and no success correspond as follows:

1. if $0 \leq t \leq T^{s,1}$ (and $\bar{u} \geq u_L \geq u_{s,1}$),

$$t = \frac{\bar{u} - u_L}{\phi} \quad \Longleftrightarrow \quad u_L = \bar{u} - \phi t,$$

2. if $T^{s,1} \leq t \leq T^{s,2}$ (and $u_{s,1} \geq u_L \geq u_{s,2}$),

$$t = T^{s,1} + \frac{1}{\lambda_S}\log\left[1 + \frac{\lambda_S(u_{s,1} - u_L)}{\phi}\right] \quad \Longleftrightarrow \quad u_L = u_{s,1} - \frac{\phi}{\lambda_S}\left[e^{\lambda_S(t - T^{s,1})} - 1\right],$$

3. if $T^{s,2} \leq t \leq T^L$ (and $u_{s,2} \geq u_L \geq 0$),

$$t = T^{s,2} + \frac{u_{s,2} - u_L}{\phi} \quad \Longleftrightarrow \quad u_L = u_{s,2} + \phi(T^{s,2} - t).$$

Define $D : [u_{s,2}, u_{s,1}] \to \mathbb{R}_+$ and $\bar{D}$ as follows:

$$D(u) \equiv \frac{u_{s,1} - u}{\phi} - \frac{1}{\lambda_S} \log\left[1 + \frac{\lambda_S(u_{s,1} - u)}{\phi}\right], \quad \bar{D} \equiv D(u_{s,2}).$$

Then, note that (i) $T^{s,1} = (\bar{u} - u_{s,1})/\phi$, (ii) $t = T^{s,1} + (u_{s,1} - u_L)/\phi - D(u_L) = (\bar{u} - u_L)/\phi - D(u_L)$ for $T^{s,1} \leq t \leq T^{s,2}$, and (iii) $t = T^{s,2} + (u_{s,2} - u_L)/\phi = (\bar{u} - u_L)/\phi - \bar{D}$ thus $T^L = \bar{u}/\phi - \bar{D}$.

Note that $\dot{u}_H$ is still equal to $-\phi$, thus, if the skill is improved at time t, the updated deadline would be $T^{H,t} = t + u_H/\phi$. For $0 \leq t \leq T^{s,1}$, $T^{H,t} = t + (u_L + \phi/\lambda_S)/\phi = \bar{u}/\phi + 1/\lambda_S = T^L + \bar{D} + 1/\lambda_S$. For $T^{s,1} \leq t \leq T^{s,2}$, $T^{H,t} = t + (u_{s,1} + \phi/\lambda_S)/\phi = T^L + t - T^{s,1} + \bar{D} + 1/\lambda_S$.

The contract that has the above property can be described as follows:

- for all $0 \leq t \leq T^{s,1}$,
  - $(a_t^L, b_t^L, R_t^L) = (0, 1, 0)$, i.e., the recommended action is to study and the agent is not paid even if he succeeds by winging it;
  - the updated contract Γ_t^H upon skill improvement at time t is given as follows:
    - the deadline is extended by $\bar{D} + 1/\lambda_S$, i.e., $T^H = T^L + \bar{D} + 1/\lambda_S$;
    - for all $t \leq s \leq T^H$, $(a_s^{H,t}, R_s^{H,t}) = (1, \phi(T^H - s + 1/\lambda_H))$.
- for all $T^{s,1} < t < T^{s,2}$,
  - $(a_t^L, b_t^L, R_t^L) = (0, 1, 0)$, i.e., the recommended action is to study and the agent is not paid even if he succeeds by winging it;
  - the updated contract Γ_t^H upon skill improvement at time t is given as follows:
    - the deadline is extended by $t - (\bar{u} - u_{s,1})/\phi + \bar{D} + 1/\lambda_S$, i.e., $T^H = T^L + t - (\bar{u} - u_{s,1})/\phi + \bar{D} + 1/\lambda_S$;
    - for all $t \leq s \leq T^H$, $(a_s^{H,t}, R_s^{H,t}) = (1, \phi(T^H - s + 1/\lambda_H))$.
- for all $T^{s,2} \leq t \leq T^L$,
  - $(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$, i.e., the recommended action is to wing it and the agent is paid when he succeeds;

– if the skill is improved, the contract is terminated.

Now I present a numerical example of which value function takes a form of (5.3). Let parameter values be $(\lambda_L, \lambda_S, \lambda_H, \Pi, c, \phi) = (1, 2.1, 2.1, 1.6, 1, 0.5)$. Note that $1/\lambda_L > 1/\lambda_S + 1/\lambda_H$, i.e., the studying technology is more efficient than the winging it technology. By solving the smooth pasting conditions $V_L^{w\prime}(u_{s,2}) = V_L^{wn\prime}(u_{s,2}|u_{s,2}, u_{s,1})$ and $V_L^{wn\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{wns\prime}(u_{s,1}|u_{s,2}, u_{s,1})$, we can derive that $u_{s,2} \approx .040$ and $u_{s,1} \approx .107$. Then, we can derive that V_L is maximized at $\bar{u} \approx .231$. Figure 4 illustrates the value function V_L and the benchmark value functions V_L^w and $V_L^{ws}(\cdot|0)$ for these parameter values.

From (5.2), we can derive that $T^{s,1} \approx .248$, $T^{s,2} \approx .366$ and $T^L \approx .447$. Then, at the beginning of the optimal contract, the agent's promised utility is $\bar{u}$ and he is recommended to study and if the agent improves skill, the deadline would be extended by $\bar{D}+1/\lambda_S \approx .492$ and he is expected to complete the task with the advanced skill. If the agent has not improved skill until time $T^{s,1}$, he is still recommended to study and if the agent improves skill, the deadline would be extended by $(t - T^{s,1}) + \bar{D} + 1/\lambda_S$. It means that whenever the agent improves the skill, the remaining time for completing the project with the advanced skill would be $T^L - T^{s,1} + \bar{D} + 1/\lambda_S \approx .691$. If the agent has not improved skill until time $T^{s,2}$, the recommended action is switched to wing it and he is compensated by the immediate payment.

6 Concluding Remarks

In this paper, I study economic tradeoffs between improving the skill and completing the project with the current skill and design an optimal incentive scheme in the dynamic principal agent setup. When the winging it technology and the studying technology are equally efficient and the skill improvement is observable to the principal and the agent, I show that the principal provides least incentive for the agent not to shirk and the recommended action schedule would be one of winging it always, studying always, or switching once from studying to winging it. The form of the optimal contract would be determined by the payoff of the project and the effectiveness of the advanced skill. For the cases that the technologies are unequally efficient, I provide numerical examples that do not have the above form of the optimal contract–there may be two switches (winging it $\rightarrow$ studying $\rightarrow$ winging it) or the incentive compatibility condition may not bind.

There are many possible extensions of this paper. For example, the skill improvement may not be observable to the principal. In this case, the principal also needs to consider an incentive for the agent to report the progress truthfully. Also, I assume that the advanced skill surely exists and the arrival rates of the skill improvement and the project under the advanced skill are publicly known. This assumption can be relaxed by adding uncertainty on the advanced skill, e.g., the advanced skill exists under certain states or the arrival rate of the project with the advanced skill depends on some state. Finally, I focus on the project such that only one ultimate breakthrough pays off. This assumption can be modified by considering projects which have flow payoffs or multiple paying off breakthroughs.

References

Dirk Bergemann and Juuso Valimaki. Bandit problems. Cowles Foundation Discussion Paper, 2006.

Bruno Biais, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78(1):73–118, 2010.

Alessandro Bonatti and Johannes Hörner. Career concerns with exponential learning. *Theoretical Economics*, 12(1):425–475, 2017. ISSN 1933-6837.

Agostino Capponi and Christoph Frei. Dynamic contracting: Accidents lead to nonlinear contracts. *SIAM Journal on Financial Mathematics*, 6(1):959–983, 2015. ISSN 1945-497X.

Gonzalo Cisternas. Career concerns and the nature of skills. *American Economic Journal: Microeconomics*, 10(2):152–89, 2018. ISSN 1945-7669.

Mathias Dewatripont, Ian Jewitt, and Jean Tirole. The economics of career concerns, part i: Comparing information structures. *The Review of Economic Studies*, 66(1):183–198, 1999a.

Mathias Dewatripont, Ian Jewitt, and Jean Tirole. The economics of career concerns, part ii: Application to missions and accountability of government agencies. *The Review of Economic Studies*, 66(1):199–217, 1999b.

Mathias Dewatripont, Ian Jewitt, and Jean Tirole. Multitask agency problems: Focus and task clustering. *European economic review*, 44(4-6):869–877, 2000.

Brett Green and Curtis R Taylor. Breakthroughs, deadlines, and self-reported progress: contracting for multistage projects. *American Economic Review*, 106(12):3660–99, 2016a.

Brett Green and Curtis R Taylor. On breakthroughs, deadlines, and the nature of progress: Contracting for multistage projects. 2016b. URL http://people.duke.edu/~crtaylor/breakthroughs.pdf.

Bengt Holmström. Managerial incentive problems: A dynamic perspective. *The review of Economic studies*, 66(1):169–182, 1999.

Bengt Holmstrom and Paul Milgrom. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7: 24, 1991.

C.R. Jaccard. Objectives and philosophy of public affairs education. Technical report, 1956. URL https://ageconsearch.umn.edu/record/17341/files/ar560012.pdf.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Bill Koenig. Automotive industry improves automation, adopts advanced tech, March 2019. URL https://advancedmanufacturing.org/automotive-industry-improves-automation/.

Christian Laux. Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics*, 32(3):514–527, 2001.

Gustavo Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011. ISSN 0022-1082.

Julien Prat and Boyan Jovanovic. Dynamic contracts when the agent's quality is unknown. *Theoretical Economics*, 9(3):865–914, 2014. ISSN 1933-6837.

Debraj Ray. The time structure of self-enforcing agreements. *Econometrica*, 70(2):547–582, 2002.

Stephen E Spear and Sanjay Srivastava. On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617, 1987.

Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973.

Peng Sun and Feng Tian. Optimal contract to induce continued effort. *Management Science*, 64(9):4193–4217, 2017. ISSN 0025-1909.

Martin Szydlowski. Incentives, project choice, and dynamic multitasking. *Theoretical Economics*, forthcoming.

Felipe Varas. Managerial short-termism, turnover policy, and the dynamics of incentives. *The Review of Financial Studies*, 31(9):3409–3451, 2017. ISSN 0893-9454.

A Proofs

A.1 Proofs for Section 3

Proof of Proposition 3.1. Let W_H (W_L) be the principal's maximum value when the agent's action and type are observable to the agent's type is high (low).

When the agent is high type, the action process (a, l) induces a probability distribution $\mathbb{P}_H^{a,l}$ over τ_m. Let $\mathbb{E}_H^{a,l}$ be the corresponding expectation operator. Then, W_H can be written as follows:

$$W_H = \max_{(a,l)} \mathbb{E}_H^{a,l} \left[\Pi - c \cdot \tau_m\right].$$

The HJB equation of W_H is derived as follows:

$$W_H = \max_{a\in[0,1]} \left(-c + \lambda_H a\Pi\right) dt + (1 - \lambda_H a dt) W_H$$

$$\Longrightarrow \qquad 0 = \max_{a\in[0,1]} -c + \lambda_H(\Pi - W_H)a$$

The right hand side is maximized at $a = 0$ or $a = 1$. If $a = 0$, RHS is equal to $-c < 0$. Hence, a should be equal to 1 and $W_H = \Pi - c/\lambda_H$. Also note that $\lambda_H(\Pi - W_H) = c > 0$, thus, $a = 1$ is induced in the maximization problem.

When the agent is low type, the action process (a, b, l) induces a probability distribution $\mathbb{P}_L^{a,b,l}$ over τ_m and τ_s. Let $\mathbb{E}_L^{a,b,l}$ be the corresponding expectation operator. Then, W_L can be written defined as follows:

$$W_L = \max_{(a,b,l)} \mathbb{E}_L^{a,b,l} \left[\Pi \cdot \mathbf{1}_{\tau_m \le \tau_s} + W_H \cdot \mathbf{1}_{\tau_m > \tau_s} - c \cdot (\tau_m \wedge \tau_s)\right].$$

The HJB equation of the social planner's value is derived as follows:

$$W_L = \max_{a,b\in[0,1],a+b\le 1} \left(-c + \lambda_L a\Pi + \lambda_S b W_H\right) dt + (1 - \lambda_L a dt - \lambda_S b dt) W_L$$

$$\Rightarrow \qquad 0 = \max_{a,b\in[0,1],a+b\le 1} -c + \lambda_L(\Pi - W_L)a + \lambda_S(W_H - W_L)b \tag{A.1}$$

Since the maximization problem is linear in a and b, the optimal solution pair (a, b) would be one of $(1, 0)$, $(0, 1)$, $(0, 0)$. $(0, 0)$ cannot be optimum because $-c < 0$.

When $a = 1$, solving (A.1) gives $W_L = \Pi - c/\lambda_L$ and $\lambda_L(\Pi - W_L) \ge \lambda_S(W_H - W_L)$ is required to induce $a = 1$. The inequality is equivalent to $1/\lambda_L \le 1/\lambda_H + 1/\lambda_S$.

When $b = 1$, solving (A.1) gives $W_L = W_H - c/\lambda_S = \Pi - c/\lambda_H - c/\lambda_S$ and $\lambda_L(\Pi - W_L) \le$

$\lambda_S(W_H - W_L)$ is required to induce $b = 1$. The inequality is equivalent to $1/\lambda_L \geq 1/\lambda_H + 1/\lambda_S$.

Therefore, when the winging it arm is more efficient ($1/\lambda_L < 1/\lambda_S + 1/\lambda_H$), the basic skill schedule maximizes the principal's payoff, and when the studying arm is more efficient ($1/\lambda_L > 1/\lambda_S + 1/\lambda_H$), the advanced skill schedule maximizes the principal's payoff. □

A.2 Proofs for Section 4

In this section, I provide proofs for Section 4. The section proceeds as follows:

1. derive candidates for the value function described in Section 4.4.2;
2. introduce functions that describe the agent's deviation from the recommended action and provide some properties of the functions;
3. provide proofs of main results;
4. provide proofs of lemmas (in both the main text and the appendix).

A.2.1 Derivation of Value Function Candidates

1. V_L^w takes the same form of V_H as in (4.1) except that the arrival rate is changed from λ_H to λ_L. Therefore, V_L^w should take a form as in (4.5).
2. By multiplying $e^{\lambda_S u_L/\phi}/\phi$ to the HJB equation, it can be rewritten as follows:

$$\frac{\lambda_S}{\phi} e^{\frac{\lambda_S}{\phi} u_L} V_L^{ws}(u_L|u_s) + e^{\frac{\lambda_S}{\phi} u_L} V_L^{ws}(u_L|u_s) = \frac{\lambda_S}{\phi} e^{\frac{\lambda_S}{\phi} u_L} \left(V_H \left(u_L + \frac{\phi}{\lambda_S} \right) - \frac{c}{\lambda_S} \right).$$

   The left hand side is equal to $\frac{d}{du_L}\left(e^{\lambda_S u_L/\phi} V_L^{ws}(u_L|u_s)\right)$. By plugging the closed form solution of V_H into the equation, the right hand side can be rewritten as follows:

$$\begin{aligned}
&\left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - u_L\right) \frac{\lambda_S}{\phi} e^{\frac{\lambda_S}{\phi} u_L} - e^{\frac{\lambda_S}{\phi} u_L} - \left(\Pi - \frac{c}{\lambda_H}\right) \frac{\lambda_S}{\phi} e^{\frac{\lambda_S - \lambda_H}{\phi} u_L - \frac{\lambda_H}{\lambda_S}} \\
=& \frac{d}{du_L} \left(\left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - u_L\right) e^{\frac{\lambda_S}{\phi} u_L} - \left(\Pi - \frac{c}{\lambda_H}\right) \frac{\lambda_S}{\lambda_S - \lambda_H} e^{\frac{\lambda_S - \lambda_H}{\phi} u_L - \frac{\lambda_H}{\lambda_S}} \right)
\end{aligned}$$

Then, by integrating the HJB equation from u_s to u_L gives

$$
\begin{aligned}
& e^{\frac{\lambda_S}{\phi}u_L} V_L^{ws}(u_L|u_s) - e^{\frac{\lambda_S}{\phi}u_s} V_L^{ws}(u_s|u_s) \\
=& \left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - u_L\right) e^{\frac{\lambda_S}{\phi}u_L} - \left(\Pi - \frac{c}{\lambda_H}\right) \frac{\lambda_S}{\lambda_S - \lambda_H} e^{\frac{\lambda_S - \lambda_H}{\phi}u_L - \frac{\lambda_H}{\lambda_S}} \\
& - \left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - u_s\right) e^{\frac{\lambda_S}{\phi}u_s} + \left(\Pi - \frac{c}{\lambda_H}\right) \frac{\lambda_S}{\lambda_S - \lambda_H} e^{\frac{\lambda_S - \lambda_H}{\phi}u_s - \frac{\lambda_H}{\lambda_S}}.
\end{aligned}
$$

By rearranging the above equation and using the boundary condition $V_L^{ws}(u_s|u_s) = V_L^w(u_s)$, V_L^{ws} is derived as in (4.7).

A.2.2 Some Useful Functions

In this subsection, I introduce functions that specify deviations from the given value functions.[17]

1. Functions for deviation given V_L^w

(a) Define

$$L_1^a(u, R) \equiv \lambda_L(\Pi - R - V_L^w(u)) - c - \lambda_L(R - u)V_L^{w\prime}(u).$$

Given $u = u_L$, maximizing this function with respect to $R \geq u + \phi/\lambda_L$ is equivalent to maximize the right hand side of (HJB$_L$) under the condition that $a = 1$ solves (PK$_L$). Note that $\frac{\partial}{\partial R} L_1^a = -\lambda_L(1 + V_L^{w\prime}(u)) < 0$ by Lemma 4.2. Therefore, for a fixed u, L_1^a is maximized at $R = u + \phi/\lambda_L$. By the definition of V_L^w, $L_1^a(u, u + \phi/\lambda_L) = 0$.

(b) Define

$$
\begin{aligned}
L_1^b(u, w) \equiv& \lambda_S(V_H(w) - V_L^w(u)) - c - \lambda_S(w - u)V_L^{w\prime}(u) && \text{(A.2)} \\
=& \lambda_S \left[\left(\Pi - \frac{c}{\lambda_H}\right) \left(1 - e^{-\frac{\lambda_H}{\phi}u - \frac{\lambda_H}{\phi}(w-u)}\right) \right. && \text{(A.3)} \\
& \left. - \left(\Pi - \frac{c}{\lambda_L}\right) \left(1 - e^{-\frac{\lambda_L}{\phi}u}\right) - (w - u) \left(\frac{\lambda_L \Pi - c}{\phi}\right) e^{-\frac{\lambda_L}{\phi}u} \right] - c.[18]
\end{aligned}
$$

Given $u = u_L$, maximizing this function with respect to $w \geq u + \phi/\lambda_S$ is equivalent to maximize the right hand side of (HJB$_L$) under the condition that $b = 1$ solves (PK$_L$).

[17] This approach is inspired by the tangible first breakthrough case of Green and Taylor (2016b). In their paper, they only need to consider the deviation from working to shirking. In this paper, we also need to consider the deviation from a technology to another technology, thus, we need to define two functions for each case.

[18] Note that L_1^b also depends on λ, Π, c and ϕ.

Then, it is enough to show that $L_1^b(u,w) \le 0$ for all $u \ge 0$ and $w \ge u + \phi/\lambda_S$.

Define $x \equiv e^{-\frac{\lambda_L}{\phi}u}$ and $y \equiv w - u$. Note that $u \ge 0$ and $w \ge u + \phi/\lambda_S$ imply that $1 \ge x > 0$ and $y \ge \phi/\lambda_S$. Then, L_1^b can be rewritten as follows:

$$\tilde{L}_1^b(x,y) \equiv -\lambda_S \left[\left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi}y} \cdot x^{\frac{\lambda_H}{\lambda_L}} - \left(1 - \frac{\lambda_L y}{\phi}\right)\left(\Pi - \frac{c}{\lambda_L}\right) x + \left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) c\right].$$

Also note that from $\phi V_L^{ws\prime}(0|0) = -c + \lambda_S V_H(\phi/\lambda_S)$ (by (4.6) and $V_L^{ws}(0|0) = 0$), we can see that

$$\tilde{L}_1^b\left(1, \frac{\phi}{\lambda_S}\right) = L_1^b\left(0, \frac{\phi}{\lambda_S}\right) = \phi V_L^{ws\prime}(0|0) - \phi V_L^{w\prime}(0). \tag{A.4}$$

2. Functions for deviation given V_L^{ws}

(a) Define

$$L_2^a(u,R|u_s) \equiv \lambda_L\left(\Pi - R - V_L^{ws}(u|u_s)\right) - c - \lambda_L(R-u)V_L^{ws\prime}(u|u_s).$$

Note that if $V_L^{w\prime}(u_s) \le V_L^{ws\prime}(u_s|u_s)$, $\frac{\partial}{\partial R}L_2^a = -\lambda_L(1 + V_L^{ws\prime}(u|u_s)) < 0$ by Lemma 4.2. Therefore, given u, L_2^a is maximized at $R = u + \phi/\lambda_L$.

$$\begin{aligned}
&L_2^a\left(u, u + \frac{\phi}{\lambda_L} \mid u_s\right) \\
&= \lambda_L\left(\Pi - \frac{\phi}{\lambda_L} - u - V_L^{ws}(u|u_s)\right) - c - \lambda_L \phi V_L^{ws\prime}(u|u_s) \\
&= \lambda_L c\left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) - \left(\Pi - \frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi}u_s}}{\lambda_S - \lambda_H}(\lambda_H - \lambda_L)e^{\frac{\lambda_H}{\phi}(u_s - u)} \\
&\quad - (\lambda_S - \lambda_L)\left[\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - (V_L^w(u_s) + u_s) - \left(\Pi - \frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi}u_s}}{\lambda_S - \lambda_H}\right] e^{\frac{\lambda_S}{\phi}(u_s - u)}
\end{aligned}$$

Define $x_1 \equiv e^{\frac{\lambda_S}{\phi}(u_s - u)}$. Then, L_2^a can be rewritten as follows:

$$\begin{aligned}
\tilde{L}_2^a(x_1|u_s) \equiv{} & \lambda_L c\left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) - \left(\Pi - \frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi}u_s}}{\lambda_S - \lambda_H}(\lambda_H - \lambda_L)x_1^{\frac{\lambda_H}{\lambda_S}} \\
& - (\lambda_S - \lambda_L)\left[\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} - (V_L^w(u_s) + u_s)\right. \\
& \left. \qquad - \left(\Pi - \frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi}u_s}}{\lambda_S - \lambda_H}\right] x_1.
\end{aligned} \tag{A.5}$$

Also note that from $\phi V_L^{w\prime}(u_s) = -c + \lambda_L (\Pi - \phi/\lambda_L - u_s - V_L^w(u_s))$ (by (4.4)) and $V_L^{ws}(u_s|u_s) = V_L^w(u_s)$, we can see that

$$\tilde{L}_2^a(1|u_s) = L_2^a\left(u_s, u_s + \frac{\phi}{\lambda_L} \mid u_s\right) = \phi V_L^{w\prime}(u_s) - \phi V_L^{ws\prime}(u_s|u_s). \tag{A.6}$$

(b) Define

$$L_2^b(u, w|u_s) \equiv \lambda_S \left(V_H(w) - V_L^{ws}(u|u_s)\right) - c - \lambda_S(w-u)V_L^{ws\prime}(u|u_s).$$

Note that $\frac{\partial}{\partial w}L_2^b = \lambda_S(V_H'(w) - V_L^{ws\prime}(u|u_s))$ and $\frac{\partial^2}{\partial w^2}L_2^b = \lambda_S V_H''(w) < 0$

The following lemmas give useful properties of the above functions.

Lemma A.1. *Suppose that λ, c, and Π are fixed and $\lambda_L\Pi > c$ is satisfied. Then, $\tilde{L}_1^b$ satisfies the following properties:*

(a) $\tilde{L}_1^b$ is strictly concave in x;

(b) Suppose that $1/\lambda_L \le 1/\lambda_S + 1/\lambda_H$ holds for (b) and (c). If $y \ge \phi/\lambda_L$ and $x \ge 0$, then $\tilde{L}_1^b(x, y) \le 0$;

(c) Define

$$x^*(y) \equiv \left[\frac{(\lambda_L\Pi - c)\left(1 - \frac{\lambda_L y}{\phi}\right) e^{\frac{\lambda_H}{\phi}y}}{(\lambda_H\Pi - c)}\right]^{\frac{\lambda_L}{\lambda_H - \lambda_L}}. \tag{A.7}$$

If $x^*(\phi/\lambda_S) \le 1$, for all $\phi/\lambda_L > y \ge \phi/\lambda_S$ and $1 \ge x > 0$,

$$\tilde{L}_1^b(x, y) \le \tilde{L}_1^b\left(x^*\left(\frac{\phi}{\lambda_S}\right), \frac{\phi}{\lambda_S}\right).$$

If $x^*(\phi/\lambda_S) > 1$, for all $\phi/\lambda_L > y \ge \phi/\lambda_S$ and $1 \ge x > 0$,

$$\tilde{L}_1^b(x, y) \le \tilde{L}_1^b\left(1, \frac{\phi}{\lambda_S}\right).$$

Lemma A.2. *Suppose that λ, c, and Π are fixed, $\lambda_L\Pi > c$. Then, $\tilde{L}_2^a$ and L_2^b satisfy the following properties:*

(a) $\tilde{L}_2^a$ is strictly convex in x_1;

(b) *Suppose that* $1/\lambda_L = 1/\lambda_S + 1/\lambda_H$ *and* $V_L^{w\prime}(u_s) \leq V_L^{ws\prime}(u_s|u_s)$ *hold for (b) and (c).* $\tilde{L}_2^a(x_1|u_s) < 0$ *for all* $x_1 \in (0,1)$*, thus,* $L_2^a(u, R|u_s) < 0$ *for all* $R \geq u + \phi/\lambda_L$*;*

(c) *For all* $u \geq u_s$ *and* $w \geq u + \phi/\lambda_S$*,* $L_2^b(u, w|u_s) \leq L_2^b(u, u + \phi/\lambda_S|u_s) = 0$*.*

A.2.3 Proofs of Main Results

Now we are ready to prove Proposition 4.1.

Proof of Proposition 4.1. (a) By Lemma 4.1, $V_L^{ws\prime\prime}(0|0) > V_L^{w\prime}(0)$. By (b) and (c) of Lemma A.2, for all $a, b \in [0,1]$ with $a + b \leq 1$, $R \geq u + \phi/\lambda_L$ and $w \geq u + \phi/\lambda_S$,

$$0 \geq a \cdot L_2^a(u, R|0) + b \cdot L_2^b(u, w|0),$$

whereas $L_2^b(u, u + \phi/\lambda_S|0) = 0$.

Note that $a \cdot L_2^a(u, R|0) + b \cdot L_2^b(u, w|0)$ is equal to the right hand side of (HJB$_L$) with $V_L(u) = V_L^{ws}(u|0)$. Therefore, $V_L^{ws}(u|0)$ solves (HJB$_L$).

(b) By Lemma 4.1 , $V_L^{ws\prime\prime}(0|0) \leq V_L^{w\prime}(0)$. and $\tilde{L}_1^b(1, \phi/\lambda_S) \leq 0$. Also note that $\tilde{L}_1^b(0, \phi/\lambda_S) = 0$ from $1/\lambda_H + 1/\lambda_S - 1/\lambda_L = 0$. Note that $\tilde{L}_1^b$ is strictly concave in x_1 ((a) of Lemma A.1) and it is maximized at $x^*(\phi/\lambda_S) > 0$ where x^* is defined as (A.7). Then, $\tilde{L}_1^b(x^*(\phi/\lambda_S), \phi/\lambda_S)$ has to be greater than 0 (if not, it contradicts $\tilde{L}_1^b(0, \phi/\lambda_S) = 0$) and $x^*(\phi/\lambda_S)$ has to be less than 1 (if not, it contradicts the strict concavity of $\tilde{L}_1^b(\cdot, \phi/\lambda_S)$ combined with $\tilde{L}_1^b(0, \phi/\lambda_S) = 0$ and $\tilde{L}_1^b(1, \phi/\lambda_S) \leq 0$). Then, $\tilde{L}_1^b(x, \phi/\lambda_S)$ is decreasing in x for $x \geq x^*(\phi/\lambda_S)$ and $\tilde{L}_1^b(x^*(\phi/\lambda_S), \phi/\lambda_S) > 0 \geq \tilde{L}_1^b(1, \phi/\lambda_S)$, thus, there exists a unique $1 \geq \bar{x} > x^*(\phi/\lambda_S)$ such that $\tilde{L}_1^b(\bar{x}, \phi/\lambda_S) = 0$.

Note that $\bar{x}$ satisfies

$$\left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\lambda_S}} \bar{x}^{\frac{\lambda_H}{\lambda_L}} = \left(\frac{1}{\lambda_L} - \frac{1}{\lambda_S}\right)(\lambda_L \Pi - c)\,\bar{x} + \left(\frac{1}{\lambda_L} - \frac{1}{\lambda_H} - \frac{1}{\lambda_S}\right) c,$$

and we can derive that $\bar{x} = \left(\frac{\lambda_L \Pi - c}{\lambda_H \Pi - c}\right)^{1/\kappa} e$. Then, for all $y \geq \phi/\lambda_S$,

$$\begin{aligned}\frac{\partial}{\partial y}\,\tilde{L}_1^b(\bar{x}, y) =& \lambda_S \left[\frac{\lambda_H}{\phi}\left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\lambda_S}} \bar{x}^{\frac{\lambda_H}{\lambda_L}} e^{\frac{\lambda_H}{\phi}\left(\frac{\phi}{\lambda_S} - y\right)} + \frac{1}{\phi}(\lambda_L \Pi - c)\,\bar{x}\right] \\ \leq & \frac{\lambda_S \lambda_H}{\phi}\left(\frac{1}{\lambda_L} - \frac{1}{\lambda_H} - \frac{1}{\lambda_S}\right)((\lambda_L \Pi - c)\bar{x} + c) = 0.\end{aligned}$$

Therefore, $\tilde{L}_1^b(\bar{x}, y) \le 0$ for all $y \ge \phi/\lambda_S$. Since $\tilde{L}_1^b(0, y) = 0$ and $\tilde{L}_1^b$ is concave in x, for all $x \ge \bar{x}$ and $y \ge \phi/\lambda_S$, $\tilde{L}_1^b(x, y) \le 0$.

Note that $\bar{x} = e^{-\lambda_L u_s(\kappa)/\phi}$. Then, the above observation implies that, for all $u_L \le u_s(\kappa)$ and $u_H \ge u + \phi/\lambda_S$, $L_1^b(u_L, u_H) \le 0$.

Now, we verify the value function for $u_L \ge u_s(\kappa)$. For convenience, denote u_s for $u_s(\kappa)$ hereafter. Note that from $L_1^b(u_s, u_s + \phi/\lambda_S) = 0$ and (A.2), we have

$$V_L^{w\prime}(u_s) = \frac{\lambda_S}{\phi}\left(V_H\left(u_s + \frac{\phi}{\lambda_S}\right) - V_L^w(u_s)\right) - \frac{c}{\phi}.$$

By (4.6) and $V_L^w(u_s) = V_L^{ws}(u_s|u_s)$, we also have

$$V_L^{ws\prime}(u_s|u_s) = \frac{\lambda_S}{\phi}\left(V_H\left(u_s + \frac{\phi}{\lambda_S}\right) - V_L^w(u_s)\right) - \frac{c}{\phi}.$$

Then, $V_L^{w\prime}(u_s) = V_L^{ws\prime}(u_s|u_s)$ and we can apply Lemma A.2 in the same manner as in (a).

□

A.2.4 Proofs of Lemmas

Lemma A.3. *For given u_L and $V_L(u_L)$, if a mixed effort level (i.e., $a, b > 0$) solves (HJB$_L$) subject to (PK$_L$), then the pure effort levels (i.e., $a = 1$ or $b = 1$) also solve (HJB$_L$) subject to (PK$_L$).*

Proof. Suppose that (a^*, b^*, R^*, u_H^*) with $a^*, b^* > 0$ solves (HJB$_L$) subject to (PK$_L$). Note that $(R^* - u_L)\lambda_L - \phi = (u_H^* - u_L)\lambda_H - \phi \ge 0$ from the maximization of (PK$_L$). Also note that $(\Pi - R^* - V_L(u_L))\lambda_L = (V_H(u_H^*) - V_L(u_L))\lambda_S \ge 0$ from the maximization of (HJB$_L$).

Then, $(a, b, R, u_H) = (1, 0, R^*, 0)$, $(a, b, R, u_H) = (0, 1, 0, u_H^*)$ and $(a, b, R, u_H) = (a^*, b^*, R^*, u_H^*)$ have the same values for the RHS of (HJB$_L$) and the RHS of (PK$_L$). Since (a^*, b^*, R^*, u_H^*) solves (HJB$_L$) subject to (PK$_L$), $(1, 0, R^*, 0)$ and $(0, 1, 0, u_H^*)$ also solve (HJB$_L$) subject to (PK$_L$). □

Proof of Lemma 4.1. Since $1 - (\kappa + 1)e^{-\kappa} > 0$ for all $\kappa > 0$, $\lambda_L \Pi - \phi \ge \lambda_S V_H(\phi/\lambda_S)$ is equivalent to

$$\Pi_S(\kappa) = \frac{1 - e^{-\kappa}}{1 - (\kappa + 1)e^{-\kappa}} \cdot \frac{c}{\lambda_L} \ge \Pi.$$

By differentiating Π_S, we have

$$\Pi'_S(\kappa) = -\frac{e^{\kappa}\left(e^{-\kappa}+\kappa-1\right)}{\left(1-(\kappa+1)e^{-\kappa}\right)^2}\cdot\frac{c}{\lambda_L}.$$

Since $e^{-\kappa} > -\kappa+1$, $\Pi'_S(\kappa) < 0$, thus, Π_S is decreasing in κ. $\square$

Proof of Lemma 4.2. (a) By (4.5) and $e^{-\lambda_L u_L/\phi} > 0$, $V_L^w(u_L) < \Pi - c/\lambda_L - u_L$. By differentiating (4.5), we have

$$V_L^{w\prime}(u_L) = \left(\Pi - \frac{c}{\lambda_L}\right)\frac{\lambda_L}{\phi}e^{-\frac{\lambda_L}{\phi}u_L} - 1 > -1.$$

By differentiating once again, we have

$$V_L^{w\prime\prime}(u_L) = -\left(\Pi - \frac{c}{\lambda_L}\right)\frac{\lambda_L^2}{\phi^2}e^{-\frac{\lambda_L}{\phi}u_L} < 0.$$

(b) Note that for all $u_s \geq u_L$

$$\frac{e^{\frac{\lambda_H}{\phi}(u_s-u_L)} - e^{\frac{\lambda_S}{\phi}(u_s-u_L)}}{\lambda_S-\lambda_H} \geq 0. \tag{A.8}$$

Then, by (4.7), $\Pi - c/\lambda_H - c/\lambda_S = \Pi - c/\lambda_L$, $V_L^w(u_s) + u_s < \Pi - c/\lambda_L - u_L$ and the above inequality, we have

$$V_L^{ws}(u_L|u_s) < \Pi - \frac{c}{\lambda_L} - u_L.$$

Note that

$$\begin{aligned} V_L^{w\prime}(u_s) &= -\frac{c}{\phi} - 1 + \frac{\lambda_L}{\phi}\left(\Pi - u_s - V_L^w(u_s)\right), \\ V_L^{ws\prime}(u_s|u_s) &= -\frac{c}{\phi} - 1 + \frac{\lambda_S}{\phi}\left(\left(\Pi - \frac{c}{\lambda_H}\right)\left(1 - e^{-\frac{\lambda_H}{\phi}u_s - \frac{\lambda_H}{\lambda_S}}\right) - V_L^w(u_s) - u_s\right). \end{aligned}$$

Then, $V_L^{w\prime}(u_s) \leq V_L^{ws\prime}(u_s|u_s)$ is equivalent to

$$-\Pi + \frac{\lambda_S\ c}{(\lambda_S-\lambda_L)\lambda_H} + \frac{\lambda_S}{\lambda_S-\lambda_L}\left(\Pi - \frac{c}{\lambda_H}\right)e^{-\frac{\lambda_H}{\phi}u_s - \frac{\lambda_H}{\lambda_S}} \leq -\left(u_s + V_L^w(u_s)\right). \tag{A.9}$$

By differentiating (4.7) twice, we have

$$
\begin{aligned}
V_L^{ws\prime\prime}(u_L|u_s) = & -\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_s-u_L)}\left[\left(\Pi-\frac{c}{\lambda_H}-\frac{c}{\lambda_S}\right)-(V_L^w(u_s)+u_s)\right] \\
& -\left(\Pi-\frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S}-\frac{\lambda_H}{\phi}u_s}}{\lambda_S-\lambda_H}\left(\left(\frac{\lambda_H}{\phi}\right)^2 e^{\frac{\lambda_H}{\phi}(u_s-u_L)}-\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_s-u_L)}\right) \\
= & -\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_s-u_L)}\left[\left(\Pi-\frac{c}{\lambda_H}-\frac{c}{\lambda_S}\right)-(V_L^w(u_s)+u_s)\right. \\
& \left.-\left(1+\frac{\lambda_H}{\lambda_S}\right)\left(\Pi-\frac{c}{\lambda_H}\right)e^{-\frac{\lambda_H}{\lambda_S}-\frac{\lambda_H}{\phi}u_s}\right] \\
& -\left(\Pi-\frac{c}{\lambda_H}\right)\lambda_S e^{-\frac{\lambda_H}{\lambda_S}-\frac{\lambda_H}{\phi}u_s}\left(\frac{\lambda_H}{\phi}\right)^2\frac{e^{\frac{\lambda_H}{\phi}(u_s-u_L)}-e^{\frac{\lambda_S}{\phi}(u_s-u_L)}}{\lambda_S-\lambda_H}
\end{aligned}
$$

By using (A.9),

$$
\begin{aligned}
& \Pi-\frac{c}{\lambda_H}-\frac{c}{\lambda_S}-(V_L^w(u_s)+u_s)-\left(1+\frac{\lambda_H}{\lambda_S}\right)\left(\Pi-\frac{c}{\lambda_H}\right)e^{-\frac{\lambda_H}{\lambda_S}-\frac{\lambda_H}{\phi}u_s} \\
\geq & \frac{\lambda_L}{\lambda_S-\lambda_L}\left(\frac{1}{\lambda_S}+\frac{1}{\lambda_H}-\frac{1}{\lambda_L}\right)\left(c+(\lambda_H\Pi-c)e^{-\frac{\lambda_H}{\phi}u_s-\frac{\lambda_H}{\lambda_S}}\right)=0.
\end{aligned}
$$

Then, plugging the above equation and the inequality (A.8) into the equation of $V_L^{ws\prime\prime}(u_L|u_s)$, we can derive that $V_L^{ws\prime\prime}(u_L|u_s) \leq 0$.

Note that

$$
\begin{aligned}
V_L^{ws\prime}(u_L|u_s) = & \frac{\lambda_S}{\phi}e^{\frac{\lambda_S}{\phi}(u_s-u_L)}\left[\left(\Pi-\frac{c}{\lambda_H}-\frac{c}{\lambda_S}\right)-(V_L^w(u_s)+u_s)\right] \\
& +\left(\Pi-\frac{c}{\lambda_H}\right)\frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S}-\frac{\lambda_H}{\phi}u_s}}{\lambda_S-\lambda_H}\left(\frac{\lambda_H}{\phi}e^{\frac{\lambda_H}{\phi}(u_s-u_L)}-\frac{\lambda_S}{\phi}e^{\frac{\lambda_S}{\phi}(u_s-u_L)}\right)-1, \\
\lim_{u_L\to\infty} & V_L^{ws\prime}(u_L|u_s) = -1.
\end{aligned}
$$

Then, by the concavitiy of $V_L^{ws}(u_L|u_s)$, $V_L^{ws\prime}(u_L|u_s) > -1$.

$\square$

Proof of Lemma 4.3. Note that for all $\kappa > 0$ and $\lambda_L > 0$,

$$\begin{aligned}
& & \frac{c+\phi}{\lambda_L} > \frac{1}{(\kappa+1)\lambda_L}\left(c + \frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right) \\
\Leftrightarrow & & (c+\phi)(\kappa+1)(e^{\kappa}-1) > c(e^{\kappa}-1) + (c+\phi)(\kappa+1-e^{\kappa}) \\
\Leftrightarrow & & 0 > g(\kappa) \equiv \phi + (c+\phi)\kappa - \phi e^{\kappa}.
\end{aligned}$$

Also note that $g(\kappa)$ is concave in κ, $\lim_{\kappa\to 0} g(\kappa) = 0$, $\lim_{\kappa\to\infty} g(\kappa) = -\infty$ and $\lim_{\kappa\to 0} g'(\kappa) = c > 0$. Then, there exists a unique positive solution of $g(\kappa) = 0$, which is κ^*. Then, $g(\kappa) < 0$ is equivalent to $\kappa > \kappa^*$. Therefore,

$$\Pi_F(\kappa) = \begin{cases} \dfrac{c+\phi}{\lambda_L}, & \text{if } \kappa < \kappa^*, \\ \dfrac{c+\phi}{\lambda_L} = \dfrac{1}{(\kappa^*+1)\lambda_L}\left(c + \dfrac{\kappa^*}{1-e^{-\kappa^*}}(c+\phi)\right), & \text{if } \kappa = \kappa^*, \\ \dfrac{1}{(\kappa+1)\lambda_L}\left(c + \dfrac{\kappa}{1-e^{-\kappa}}(c+\phi)\right), & \text{if } \kappa > \kappa^*. \end{cases}$$

For $i \in \{F, W, S\}$, note that $\Pi_i(\kappa)$ can be considered as a unique solution (greater than c/λ) of the equation

$$L(\Pi) = R_i(\Pi|\kappa),$$

where

$$\begin{aligned}
L(\Pi) &= (\kappa+1)\lambda_L\Pi - c, \\
R_F(\Pi|\kappa) &= \begin{cases} \dfrac{\kappa}{1-e^{-\kappa}}(c+\phi) & \text{if } \kappa \geq \kappa^* \\ \kappa(c+\phi)+\phi & \text{if } \kappa \leq \kappa^* \end{cases}, \text{[19]} \\
R_W(\Pi|\kappa) &= \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_L\Pi - c}{\phi}\right)^{\kappa+1}, \\
R_S(\Pi|\kappa) &= \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_L\Pi - c}{\phi}\right).
\end{aligned}$$

Note that $L(c/\lambda_L) < R_i(c/\lambda_L|\kappa)$, $\lim_{\Pi\to\infty} L(\Pi) > \lim_{\Pi\to\infty} R_i(\Pi|\kappa)$ and L and $R_i(\cdot|\kappa)$ cross only once for all $i \in \{F, W, S\}$ and $\kappa > 0$.

[19] Note that $\kappa^*(c+\phi)/(1-e^{-\kappa^*}) = \kappa^*(c+\phi)+\phi$ by the definition of κ^*.

If $R_i(\Pi_i(\kappa)|\kappa) > R_j(\Pi_i(\kappa)|\kappa)$,

$$L(\Pi_i(\kappa)) = R_i(\Pi_i(\kappa)|\kappa) > R_j(\Pi_i(\kappa)|\kappa),$$

and it implies that $\Pi_j(\kappa)$ is smaller than $\Pi_i(\kappa)$. Similarly, $R_i(\Pi_i(\kappa)|\kappa) = R_j(\Pi_i(\kappa)|\kappa)$ implies that $\Pi_j(\kappa)$ is equal to $\Pi_i(\kappa)$ and $R_i(\Pi_i(\kappa)|\kappa) < R_j(\Pi_i(\kappa)|\kappa)$ implies that $\Pi_j(\kappa)$ is greater than $\Pi_i(\kappa)$.

1. When $\kappa > \kappa^*$, to prove that $\Pi_F(\kappa) > \max[\Pi_W(\kappa), \Pi_S(\kappa)]$, it is enough to show that $R_F(\Pi_F(\kappa)|\kappa) < R_W(\Pi_F(\kappa)|\kappa)$ and $R_F(\Pi_F(\kappa)|\kappa) < R_S(\Pi_F(\kappa)|\kappa)$.

   Define $x(\kappa)$ as follows:
$$x(\kappa) = \frac{\kappa}{e^\kappa - 1}\left(\frac{c+\phi}{\phi}\right).$$
   Then, $x(\kappa) < 1$ is equivalent to $g(\kappa) < 0$, i.e., $\kappa > \kappa^*$. Also note that
$$\frac{\lambda_L \Pi_F(\kappa) - c}{\phi} = \frac{\kappa}{\kappa+1}\left(\frac{c + e^\kappa\phi}{e^\kappa - 1}\right) = \frac{x(\kappa)+\kappa}{\kappa+1}.$$

   By using the definition of $x(\kappa)$ and the above equation, we can see that
$$\begin{aligned} R_F(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot x(\kappa), \\ R_W(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right)^{\kappa+1}, \\ R_S(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right). \end{aligned} \tag{A.10}$$

   Consider a function $h(x) = \left(\frac{x+\kappa}{1+\kappa}\right)^{\kappa+1}$. Note that $h'(x) = \left(\frac{x+\kappa}{1+\kappa}\right)^{\kappa}$ and $h''(x) = \frac{\kappa}{1+\kappa}\left(\frac{x+\kappa}{1+\kappa}\right)^{\kappa-1} > 0$. Then, $h(x) > h(1) + h'(1)(x-1) = x$ for $x < 1$. Hence, $R_W(\Pi_F(\kappa)|\kappa) > R_F(\Pi_F(\kappa)|\kappa)$. Also, we can easily see that $\dfrac{x+\kappa}{\kappa+1} > x$ is equivalent to $x < 1$, i.e., $R_S(\Pi_F(\kappa)|\kappa) > R_F(\Pi_F(\kappa)|\kappa)$.

2. When $\kappa = \kappa^*$, to prove that $\Pi_F(\kappa) = \Pi_W(\kappa) = \Pi_S(\kappa)$, it is enough to show that $R_F(\Pi_F(\kappa)|\kappa) = R_W(\Pi_F(\kappa)|\kappa) = R_S(\Pi_F(\kappa)|\kappa)$.

   Note that $x(\kappa^*) = 1$. Hence, by (A.10), $R_F(\Pi_F(\kappa)|\kappa) = R_W(\Pi_F(\kappa)|\kappa) = R_S(\Pi_F(\kappa)|\kappa)$.

3. When $\kappa < \kappa^*$, to prove that $\Pi_S(\kappa) > \Pi_W(\kappa) > \Pi_F(\kappa)$, it is enough to show that $R_F(\Pi_F(\kappa)|\kappa) > R_W(\Pi_F(\kappa)|\kappa)$ and $R_W(\Pi_S(\kappa)|\kappa) > R_S(\Pi_S(\kappa)|\kappa)$.

In this case, $\Pi_F(\kappa) = (c+\phi)/\lambda_L$. Then, by the definition of R_F and R_W, $R_F(\Pi_F(\kappa)|\kappa) = \kappa(c+\phi)+\phi$ and $R_W(\Pi_F(\kappa)|\kappa) = \phi \cdot e^{\kappa}$. Since $\kappa < \kappa^*$ is equivalent to $\kappa(c+\phi)+\phi > \phi e^{\kappa}$, $R_F(\Pi_F(\kappa)|\kappa) > R_W(\Pi_F(\kappa)|\kappa)$.

Also note that

$$\frac{\lambda_L \Pi_S(\kappa) - c}{\phi} = \frac{\frac{1-e^{-\kappa}}{1-(\kappa+1)e^{-\kappa}} c - c}{\phi} = \frac{\kappa \cdot c}{(e^{\kappa} - (\kappa+1))\phi} > 1.$$

Then, since $R_W(\Pi_S(\kappa)|\kappa) = R_S(\Pi_S(\kappa)|\kappa) \cdot \left(\frac{\lambda_L \Pi - c}{\phi}\right)^{\kappa}$, $R_W(\Pi_S(\kappa)|\kappa) > R_S(\Pi_S(\kappa)|\kappa)$.

4. When $\kappa \to 0$, by L'Hospital's Rule,

$$\lim_{\kappa \to 0} \Pi_S(\kappa) = \lim_{\kappa \to 0} \frac{1 - e^{-\kappa}}{1 - (\kappa+1)e^{-\kappa}} \cdot \frac{c}{\lambda_L} = \lim_{\kappa \to 0} \frac{e^{-\kappa}}{\kappa\ e^{-\kappa}} \cdot \frac{c}{\lambda_L} = \infty.$$

Define $y(\kappa) \equiv \left(\lambda_L \Pi_W(\kappa) - c\right)/\phi > 0$. Then, from (4.11), $y(\kappa)$ satisfies the following equations for all $\kappa > 0$:

$$\begin{aligned} & y(\kappa)^{1+\kappa} \cdot e^{\kappa} = (1+\kappa) y(\kappa) + \frac{c}{\phi}\kappa \\ \Rightarrow \quad & (1+\kappa)\log\left[y(\kappa)\right] + \kappa = \log\left[(1+\kappa)y(\kappa) + \frac{c}{\phi}\kappa\right]. \end{aligned}$$

By differentiating the above equation by κ, we have

$$\log\left[y(\kappa)\right] + 1 + \frac{1+\kappa}{y(\kappa)} y'(\kappa) = \frac{y(\kappa) + \frac{c}{\phi}}{(1+\kappa)y(\kappa) + \frac{c}{\phi}\kappa} + \frac{1+\kappa}{(1+\kappa)y(\kappa) + \frac{c}{\phi}\kappa} y'(\kappa).$$

By sending $\kappa \to 0$, we have

$$y(0) \cdot \log\left[y(0)\right] = \frac{c}{\phi},$$

i.e., $y(0) = \psi(c/\phi)$. Then, we have

$$\lim_{\kappa \to 0} \Pi_W(\kappa) = \frac{c + \phi \cdot y(0)}{\lambda_L} = \frac{c + \phi \cdot \psi\left(\frac{c}{\phi}\right)}{\lambda_L}.$$

□

Proof of Lemma A.1. (a) By $\lambda_H > \lambda_L$, $\Pi > c/\lambda_L > c/\lambda_H$, and $x > 0$,

$$\frac{\partial^2 \tilde{L}_1^b}{\partial x^2} = -\lambda_S \left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi} y} \left(\frac{\lambda_H}{\lambda_L}\right) \left(\frac{\lambda_H}{\lambda_L} - 1\right) x^{\frac{\lambda_H}{\lambda_L} - 2} < 0,$$

thus, $\tilde{L}_1^b$ is strictly concave in x.

(b) By differentiating $\tilde{L}_1^b$ once by x,

$$\frac{\partial \tilde{L}_1^b}{\partial x} = -\lambda_S \left[\frac{\lambda_H}{\lambda_L} \left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi} y} x^{\frac{\lambda_H - \lambda_L}{\lambda_L}} - \left(1 - \frac{\lambda_L y}{\phi}\right) \left(\Pi - \frac{c}{\lambda_L}\right)\right].$$

If $y \geq \phi/\lambda_L$ and $x \geq 0$, $\tilde{L}_1^b$ is decreasing in x, thus, $\tilde{L}_1^b$ is maximized at $x = 0$. Then, since $1/\lambda_L \leq 1/\lambda_S + 1/\lambda_H$, for all $1 \geq x > 0$ and $y \geq \phi/\lambda_L$, the following inequalities hold:

$$\tilde{L}_1^b(x, y) \leq \tilde{L}_1^b(0, y) = -\lambda_S \left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) c \leq 0.$$

(c) When $\phi/\lambda_L > y \geq \phi/\lambda_S$ and y is fixed, since $\tilde{L}_1^b(x, y)$ is concave in x, $\tilde{L}_1^b(\cdot, y)$ is maximized at

$$x^*(y) \equiv \left[\frac{(\lambda_L \Pi - c) \left(1 - \frac{\lambda_L y}{\phi}\right) e^{\frac{\lambda_H}{\phi} y}}{(\lambda_H \Pi - c)}\right]^{\frac{\lambda_L}{\lambda_H - \lambda_L}}.$$

Define $g(y) \equiv (1 - \lambda_L y/\phi)\, e^{(\lambda_H/\phi) y}$. Then, differentiating $g(y)$ gives

$$\begin{aligned} g'(y) =& -\frac{\lambda_L}{\phi} e^{\frac{\lambda_H}{\phi} y} + \frac{\lambda_H}{\phi} \left(1 - \frac{\lambda_L y}{\phi}\right) e^{\frac{\lambda_H}{\phi} y} \\ =& \frac{\lambda_L \lambda_H}{\phi} e^{\frac{\lambda_H}{\phi} y} \left(-\frac{1}{\lambda_H} + \frac{1}{\lambda_L} - \frac{y}{\phi}\right). \end{aligned}$$

Note that since $y \geq \phi/\lambda_S$ and $1/\lambda_L \leq 1/\lambda_S + 1/\lambda_H$, $g(y)$ is decreasing in y, hence, $x^*(y)$ is also decreasing in y.

Now, restrict attention to $1 \geq x > 0$. If $x^*(y) < 1$, the maximum value of $\tilde{L}_1^b(\cdot, y)$ is

$$\begin{aligned} \tilde{L}_1^b(x^*(y), y) =& \lambda_S \left[\left(\frac{\lambda_H}{\lambda_L} - 1\right) \left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi} y} x^*(y)^{\frac{\lambda_H}{\lambda_L}} - \left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) c\right] \\ =& \lambda_S \left[\left(\frac{1}{\lambda_L} - \frac{1}{\lambda_H}\right) \left[(\lambda_H \Pi - c)^{-\frac{\lambda_L}{\lambda_H}} (\lambda_L \Pi - c) \left(1 - \frac{\lambda_L y}{\phi}\right) e^{\frac{\lambda_L y}{\phi}}\right]^{\frac{\lambda_H}{\lambda_H - \lambda_L}} \right. \\ & \left. - \left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L}\right) c\right]. \end{aligned}$$

Note that $(1 - \lambda_L y/\phi)e^{\lambda_L y/\phi}$ is decreasing in y,[20] thus, $\tilde{L}_1^b(x^*(y), y)$ is also decreasing in y.

If $x^*(y) \geq 1$, the maximum value of $\tilde{L}_1^b(\cdot, y)$ is

$$\tilde{L}_1^b(1, y) = -\lambda_S \left[\left(\Pi - \frac{c}{\lambda_H} \right) e^{-\frac{\lambda_H}{\phi} y} - \left(1 - \frac{\lambda_L y}{\phi} \right) \left(\Pi - \frac{c}{\lambda_L} \right) + \left(\frac{1}{\lambda_H} + \frac{1}{\lambda_S} - \frac{1}{\lambda_L} \right) c \right].$$

Note that $x^*(y) \geq 1$ implies that $0 > -\lambda_L y/\phi \geq (\lambda_H \Pi - c)e^{-\frac{\lambda_H}{\phi} y} - (\lambda_L \Pi - c)$. Also note that

$$\frac{\partial \, \tilde{L}_1^b(1, y)}{\partial y} = \frac{\lambda_S}{\phi} \left[(\lambda_H \Pi - c) \, e^{-\frac{\lambda_H}{\phi} y} - (\lambda_L \Pi - c) \right] < 0.$$

Therefore, $\tilde{L}_1^b(1, y)$ is decreasing in y.

When $x^*(\phi/\lambda_S) \leq 1$, $x^*(y) \leq 1$ holds for all $\phi/\lambda_L > y \geq \phi/\lambda_S$ since $x^*(y)$ is decreasing in y. Then,

$$\tilde{L}_1^b(x, y) \leq \tilde{L}_1^b(x^*(y), y) \leq \tilde{L}_1^b \left(x^* \left(\phi/\lambda_S \right), \phi/\lambda_S \right)$$

from the optimality of $x^*(y)$ and decreasingness of $\tilde{L}_1^b(x^*(y), y)$ in y.

When $x^*(\phi/\lambda_S) > 1$, note that $x^*(\phi/\lambda_L) = 0$, thus, there exists $y^* \in (\phi/\lambda_S, \phi/\lambda_L)$ such that $x^*(y^*) = 1$. Then, $x^*(y) < 1$ for $y > y^*$ and $x^*(y) > 1$ for $y < y^*$. When $y < y^*$, by using the decreasingness of $\tilde{L}_1^b(1, y)$ for $x^*(y) > 1$,

$$\tilde{L}_1^b(x, y) \leq \tilde{L}_1^b(1, y) \leq \tilde{L}_1^b(1, \phi/\lambda_S).$$

When $y > y^*$,

$$\tilde{L}_1^b(x, y) \leq \tilde{L}_1^b(x^*(y), y) \leq \tilde{L}_1^b(x^*(y^*), y^*) = \tilde{L}_1^b(1, y^*) \leq \tilde{L}_1^b(1, \phi/\lambda_S).$$

By combining the above results, we can show that

$$\max_{\substack{1 \geq x > 0, \\ \frac{\phi}{\lambda_L} > y \geq \frac{\phi}{\lambda_S}}} \tilde{L}_1^b(x, y) = \begin{cases} \tilde{L}_1^b \left(x^* \left(\frac{\phi}{\lambda_S} \right), \ \frac{\phi}{\lambda_S} \right) & \text{if } \ x^* \left(\frac{\phi}{\lambda_S} \right) \leq 1, \\ \tilde{L}_1^b \left(1, \ \frac{\phi}{\lambda_S} \right) & \text{if } \ x^* \left(\frac{\phi}{\lambda_S} \right) > 1. \end{cases}$$

□

[20] Differentiating the term by y gives $-(\lambda_L^2 y/\phi^2)e^{\lambda_L y/\phi} < 0$.

Proof of Lemma A.2. (a) By differentiating (A.5) twice, we have

$$\tilde{L}_2^{a\prime\prime}(x_1) = \left(\Pi - \frac{c}{\lambda_H}\right) \frac{e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi} u_s}(\lambda_H - \lambda_L)}{\lambda_S} x_1^{\frac{\lambda_H}{\lambda_S} - 2} > 0,$$

from $\Pi > c/\lambda_L > c/\lambda_H$ and $\lambda_H - \lambda_L > 0$. Therefore, $\tilde{L}_2^a$ is strictly convex in x_1.

(b) Note that $\tilde{L}_2^a(0) = 0$ from $1/\lambda_L = 1/\lambda_S + 1/\lambda_H$ and $\tilde{L}_2^a(1) = \phi(V_L^{w\prime}(u_s) - V_L^{ws\prime}(u_s|u_s)) \leq 0$ from (A.6) and $V_L^{w\prime}(u_s) \leq V_L^{ws\prime}(u_s|u_s)$. Then, by the convexity of $\tilde{L}_2^a$, for all $x_1 \in (0,1)$,

$$\tilde{L}_2^a(x_1) < x_1 \, \tilde{L}_2^a(1) + (1 - x_1) \, \tilde{L}_2^a(0) \leq 0.$$

(c) By differentiating (4.6) once, we can derive that

$$\phi V_L^{ws\prime\prime}(u|u_s) = \lambda_S \left(V_H' \left(u + \frac{\phi}{\lambda_S} \right) - V_L^{ws\prime}(u|u_s) \right).$$

By (b) of Lemma 4.2, $V_L^{ws\prime\prime}(u|u_s) < 0$, thus, the following inequality holds:

$$0 \geq \frac{\partial}{\partial w} L_2^b(u, u + \phi/\lambda_S | u_s) = \lambda_S \left(V_H' \left(u + \frac{\phi}{\lambda_S} \right) - V_L^{ws\prime}(u|u_s) \right).$$

Since $\frac{\partial^2}{\partial w^2} L_2^b < 0$, for a given u, $L_2^b(u, w|u_s)$ subject to $w \geq u + \phi/\lambda_S$ is maximized at $w = u + \phi/\lambda_S$. Also note that $L_2^b(u, u + \phi/\lambda_S|u_s) = 0$ holds by (4.6).

□

B Details in Section 5.3

Will be added