Completing a Project on Time: When to Study It and When to Go for It

Yonggyun Kim*

Duke University

yonggyun.kim@duke.edu

September 5, 2019

Abstract

I study a dynamic agency problem where the agent has two possible technologies to complete a task: the *go-for-it* and the *study-it* technologies. The go-for-it technology requires only one breakthrough but it arrives at a low rate, whereas the study-it technology has high arrival rates but requires two breakthroughs. The intermediate breakthrough of the study-it technology is assumed to be observable and contractually verifiable. To derive the optimal incentive scheme, the principal compares an immediate payoff (to induce going for it) and a deadline extension (to induce studying it) to maximize her expected payoff. When the go-for-it and study-it technologies are equally efficient, the recommended action for the agent in the optimal contract is one of three types: (i) always going for it; (ii) always studying it; (iii) switching once from studying it to going for it. Which of these is optimal depends on the payoff of the project and the effectiveness of the study-it technology.

JEL Classification: J41, L14, D86

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

*I am indebted to Fei Li and Curtis Taylor for their guidance and support in this project. I am also grateful to Attila Ambrus, Arjada Bardhi, Yeon-Koo Che, Jinwoo Kim, Pino Lopomo, Huseyin Yildirim, and seminar participants at Duke and Seoul National University for comments and suggestions. All remaining errors are my own.

A woodsman was once asked, "What would you do if you had just five minutes to chop do 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 it would cost 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 axe. A research assistant (RA), who needs to complete an empirical project, may learn advanced skills such as programming in 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 other hand, 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 the current mediocre skill. Automobile companies can divert resources from investment to production. The woodman can try to chop down the tree with the dull axe rather than sharpening the axe. 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 it) and completing tasks with the current basic skill (going for 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 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 at a low rate. The agent can also invest in an advanced skill, and then the agent can acquire the skill at a higher arrival rate. Once the advanced skill is obtained, the breakthrough arrives at 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 *go-for-it* technology (completing the project with the basic skill), the *study-it* technology (investing in the advanced skill), and *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 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 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 the recommended actions change over time.

I start by analyzing the first best case, i.e., the case where the agent's action is observable to the principal. In this case, there are two candidates for the optimal recommended action schedule: the go-for-it only schedule and the study-it only schedule. The go-for-it only schedule recommends the agent to choose going for it until the ultimate breakthrough is made, whereas the study-it only schedule recommends studying it until 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 goal of this paper is to solve the case where 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 point in time, the recommended action is determined by comparing the principal's expected payoffs from each technology. Unlike 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 different ways. If the principal wants to induce the agent to go for 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, 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. This is because the skill improvement does not give an immediate benefit to the principal.

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

1. **Going for it always**: The agent is recommended to go for it until the deadline and immediately receive a payment upon the ultimate breakthrough. If the agent does not make the breakthrough by the deadline, the contract is terminated.

2. **Studying it always**: The agent is recommended to study it 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 it to going for it**: The agent is recommended to study it until the intermediate deadline and if the agent improves the skill before the intermediate deadline, the final 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 go for it until the final deadline and immediately receive a payment upon the ultimate breakthrough. If the agent does not make the breakthrough until the final 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 study-it 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 where 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 the recent developments in 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 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) consider an economic situation where 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 in 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 in these technologies is 'how' the ultimate breakthrough is made–by the basic skill or by the advanced skill.

[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 where 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.

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 project requires an ultimate breakthrough and I denote it as 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 of 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., they 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 point in time t, the agent allocates 1 unit of effort to the completion of the main task (a_t), the skill improvement (b_t), and 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 skill improvement. The skill advancement for the low type agent arrives at 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 go-for-it technology (a_t) where the success requires one breakthrough at a low arrival rate (λ_L); (ii) the study-it technology (b_t) where the success requires two breakthrough at high arrival rates ($\lambda_S,\ \lambda_H$); (iii) leisure (l_t) where the success never arrives. For the high type agent, an action can be understood as a choice between the working technology (a_t) and leisure (l_t).

[6]Since the skill improvement is assumed to be arrived at 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 a case where a past effort affects the chance of the current skill improvement, the arrival rate would need to be proportional to the cumulative effort level.

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.

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 when the skill is improved and τ_m be the (random) date when the main task is completed. Then, the two schedules are defined as follows:

Go-for-it Only Schedule

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

Study-it Only Schedule

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

The go-for-it only schedule means that the agent tries to complete the task only with the basic skill, i.e., only chooses the go-for-it technology. On the other hand, the study-it 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 study-it 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 going for it only schedule is given by $\lambda_L e^{-\lambda_L \tau_m}$. On the other hand, for the study-it 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 study-it only schedule is given by $\lambda_S e^{-\lambda_S \tau_s}$. Then, the expected profits for both schedules are given as follows:

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

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

By comparing the expected profits, we can easily see that the go-for-it only schedule is indifferent to the study-it only schedule if and only if

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

Denote that going for it is more efficient if the left hand side of (3.1) is less than the right hand side, while studying it is more efficient if the left hand side is greater than the right hand side. Also denote that both technologies are equally efficient if the equality holds. Note that $1/\lambda_L$ is the expected duration of the go-for-it only schedule and $1/\lambda_S + 1/\lambda_H$ is the expected duration of the study-it only schedule. Therefore, the efficiency relation gives clear intuition for comparing two schedules: the shorter the expected duration, the greater the expected payoff. Furthermore, the following proposition shows that one of the 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 going for it is more efficient ($1/\lambda_L < 1/\lambda_S + 1/\lambda_H$), the go-for-it only schedule is the first best schedule, i.e., it gives the highest expected profit to the principal. If studying it is more efficient ($1/\lambda_L > 1/\lambda_S + 1/\lambda_H$), the study-it only schedule is 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 payoff. 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 $\Gamma^L \equiv \left\{a^L, b^L, R^L, \Gamma^H, T^L\right\}$, where each variable is defined as follows at the calendar time t:

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

2. $b_t^L \in [0,1]$: the recommended effort to the skill improvement conditional on no success and no skill improvement;

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

4. $\Gamma_t^H \equiv \{a^{H,t}, R^{H,t}, T^{H,t}\}$: an updated contract for skill improvement at time t conditional on no success;

   (a) $a_s^{H,t} \in [0,1]$: 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;

   (b) $R_s^{H,t} \geq 0$: the 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;

   (c) $T^{H,t} \geq t$: the date at which the project is terminated conditional on skill improvement at time t and no success;

5. $T^L \geq 0$: 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 the skill improvement and the success dates (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 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 first 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 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.,

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

   where $\mathcal{A}_t^H \equiv \{\{a_s\}_{t\le s\le T^{H,t}}\ :\ a_s\in[0,1]\}$;

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

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

   where $\mathcal{A}^L \equiv (\tilde{a},\tilde{b}) \in \{\{a_t,b_t\}_{0\le t\le T^L}\ :\ (a_t,b_t,a_t+b_t)\in[0,1]^3\}$.

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 a 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 the promised utility for the high (low) type agent.

4.2.1 When the agent is high type

The high type agent's promised utility can be written as follows:

$$u_{H,t} = \max_{a_t \in [0,1]} R_t \lambda_H a_t dt + \phi(1-a_t)dt + (1-\lambda_H a_t dt) u_{H,t+dt}.^{9}$$

By using a Taylor expansion $u_{H,t+dt} = u_{H,t} + \dot{u}_{H,t}dt + o(dt)$ (where $\dot{u}_{H,t} = du_{H,t}/dt$) and sending $dt \to 0$, we can derive the Hamilton-Jacobi-Bellman (HJB) equation for the high type agent's promise keeping (PK) constraint as follows:

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

where the second term is the benefit from leisure, and the last term is the expected additional payoff from completing the task.

Denote a value function $V_H(u_H)$ as the function that maximizes the principal's expected utility $P_t^H(\Gamma_t^H)$ subject to $U_t^H(\Gamma_t^H) = u_H$ and the incentive compatibility condition.[10] Then, the value function can be written as follows:

$$V_H(u_{H,t}) = \max_{R_t \geq 0,\ a_t \in [0,1]} (\Pi - R_t)\lambda_H a_t dt + (1 - \lambda_H a_t dt) V_H(u_{H,t+dt}).$$

By using $V_H(u_{H,t+dt}) = V_H(u_{H,t}) + V_H'(u_{H,t})\dot{u}_{H,t}dt + o(dt)$, the HJB equation for the value function $V_H(u_H)$ can be derived as follows:

$$0 = \max_{R \geq 0,\ a \in [0,1]} -c + (\Pi - R - V_H(u_H))\lambda_H a + V_H'(u_H)\ \dot{u}_H. \tag{HJB$_H$}$$

Note that since the outside option of the agent is zero, if the agent's promised utility is equal to zero, he would not work on the project and the principal would also end up getting zero. Then, the principal's problem is to solve (HJB$_H$) subject to (PK$_H$) with the boundary condition $V_H(0) = 0$.

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

[9] For simplicity, I take off the superscripts H from R^H and a^H.

[10] 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.

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 in 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, $R(\tau_m) = \phi(1/\lambda_H + T - \tau_m)$.[11]

4.2.2 When the agent is low type

The low type agent's promised utility can be written as follows:

$$u_{L,t} = \max_{\substack{a_t, b_t \geq 0 \\ a_t + b_t \leq 1}} R_t \lambda_L a_t dt + u_{H,t} \lambda_S b_t dt + \phi(1 - a_t - b_t)dt + (1 - \lambda_L a_t dt - \lambda_S b_t dt) u_{L,t+dt}.$$

Then, the HJB equation for the agent's promise keeping constraint can be derived as follows:

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

where $\dot{u}_L \equiv du_L/dt$, the second term is the benefit from 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 the function that maximizes the principal's expected utility $P_0^L(\Gamma^L)$ subject to $U_0^L(\Gamma^L) = u_L$ and the incentive compatibility condition. The value function can be written as follows:

$$V_L(u_{L,t}) = \max_{\substack{R_t \geq 0, \ u_{H,t} \geq 0, \\ a_t, b_t \geq 0, \ 1 \geq a_t + b_t}} (\Pi - R_t) \lambda_L a_t dt + V_H(u_{H,t}) \lambda_S b_t dt + (1 - \lambda_L a_t dt - \lambda_S b_t dt) V_L(u_{L,t+dt}).$$

The HJB equation for the value function $V_L(u_L)$ can be derived as follows:

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

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

Then, the principal's problem is to solve (HJB$_L$) subject to (PK$_L$) with the boundary condition $V_L(0) = 0$. The following lemma shows that the solution of the problem maximizes the principal's expected payoff subject to a promise keeping constraint $U_0^L(\Gamma^L) = u_L$.

Lemma 4.1 (Verification Lemma). *Suppose that a concave $\mathcal{C}^1$ function $\hat{V}$ solves* (HJB$_L$) *subject to* (PK$_L$) *with the boundary condition $\hat{V}(0) = 0$. Then, for any feasible contract Γ^L with $U_0^L(\Gamma^L) = u_L$,*

$$\hat{V}(u_L) \geq P_0^L(\Gamma^L).$$

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 deriving V_L and solving the above maximization problem.

4.3 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 B.6 for the detailed argument). Now I introduce two contractual modes with the pure effort levels defined as follows:

1. **Inducing to go for it with the minimum incentive (Mode G)**

   The agent is recommended to fully go for it and 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. He receives nothing when the skill is improved. Thus, 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 go for 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 the minimum incentive (Mode S)**

   The agent is recommended to fully study 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. He receives nothing when the project succeeds. Thus, 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 that the first IC constraint binds.

In the appendix, I show that one of the above contractual modes will be executed at each point in time in the duration of the optimal contract. The contractual mode may switch over time in the optimal contract. To simplify the argument, I will focus on comparing the above two contractual modes in the main text of the paper. Given that the agent's promised utility is u_L, define $w_G(u_L)$ and $w_S(u_L)$ as the principal's instantaneous payoffs by executing Mode G and Mode S, i.e.,

$$
\begin{aligned}
w_G(u_L) &\equiv \lambda_L \left(\Pi - \phi/\lambda_L - u_L - V_L(u_L)\right) - c, \\
w_S(u_L) &\equiv \lambda_S \left(V_H(u_L + \phi/\lambda_S) - V_L(u_L)\right) - c.
\end{aligned}
$$

At each point in time, the principal chooses the contractual mode by comparing $w_G(u_L)$ and $w_S(u_L)$.

Mode G and Mode S mainly differ in the arrival rate and the form of compensation to the agent. Under Mode G, the agent succeeds at a lower arrival rate (λ_L) and receives an immediate payment upon success. Under Mode S, the agent makes a skill improvement at 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 G and Mode S. Then, consider the case where 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 in this case 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 effort on completing the project with the advanced skill. In sum, the agent is compensated by an immediate payment with a lower probability under Mode G, whereas he is compensated by a deadline extension with a higher chance of success under Mode S.

1. Mode S is preferred to Mode G when the deadline is far away;

2. Mode G 'may' be preferred to Mode S when the deadline is close by.

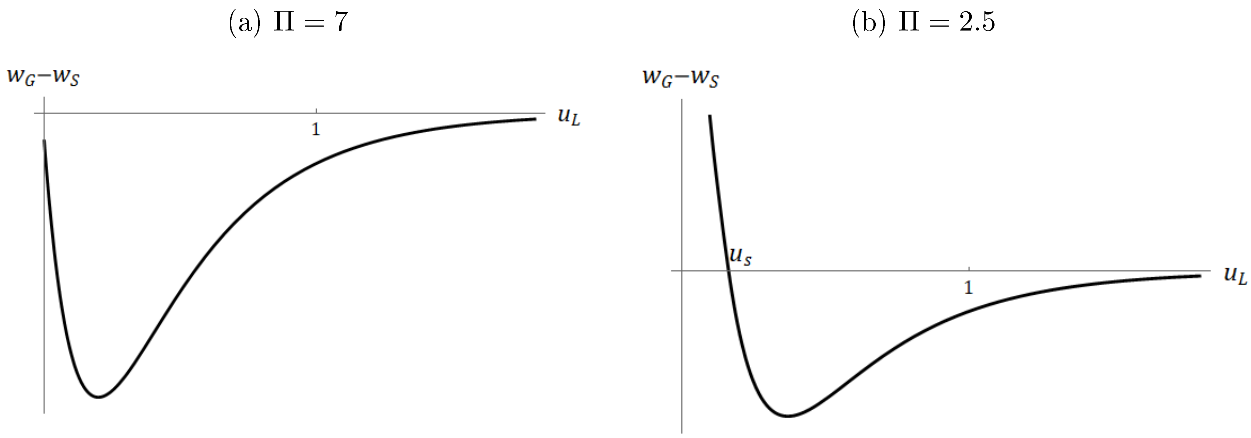

Figure 1: The differences between instantaneous payoffs of Mode G and Mode S when parameter values are $\lambda_L = 1,\ \lambda_S = 4,\ \lambda_H = 4/3,\ c = 1,\ \phi = 0.5$

If the deadline is adequately far away, 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 G. However, if the deadline is close by, Mode S requires two breakthroughs in a relatively short period of time, whereas Mode G only requires one breakthrough. Therefore, it may be possible that the principal prefers Mode G to Mode S when the deadline is close by.

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, I derive three thresholds of Π depending on κ to characterize the optimal contract.

Firstly, the first and the second conditions of the previous subsection imply that if Π is very big and Mode S is preferred to Mode G at the deadline, Mode S would also be preferred to Mode G at any time during the contract. For a given κ, let $\Pi_S(\kappa)$ be a threshold that characterizes the above property. If $\Pi > \Pi_S(\kappa)$, Mode S is preferred over Mode G at the deadline, whereas if $\Pi < \Pi_S(\kappa)$, Mode G is preferred over Mode S at the deadline. See Appendix B.1.2 for details.

Secondly, note that as Π increases, the principal would also want to increase the length of the contract to have a better chance of a successful project to compensate the expected cost. Then, if Π is small, the length of the contract would be short and Mode G might be preferred over mode S during the contract. Define $\Pi_G(\kappa)$ as the corresponding threshold for

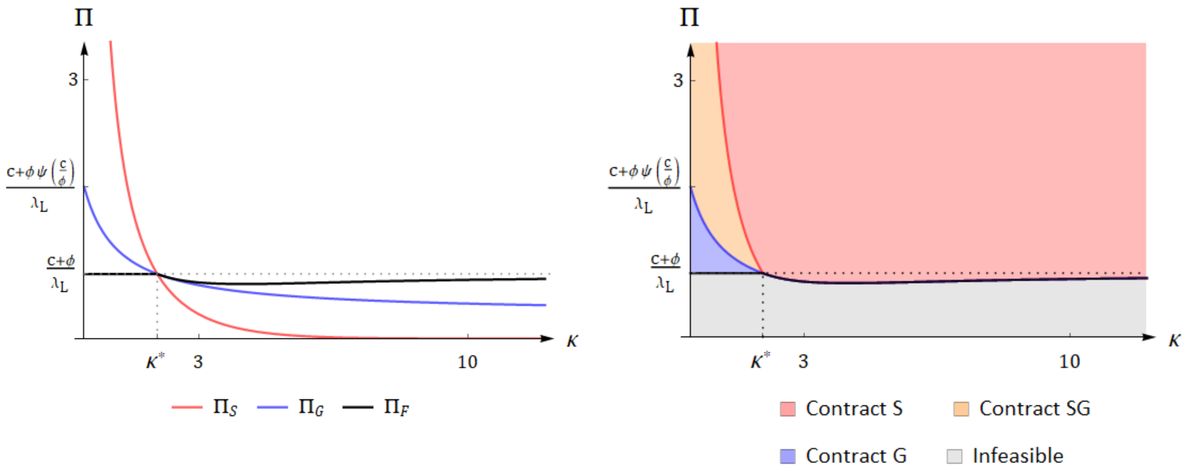

Figure 2: Thresholds and optimal contracts when $(c, \phi, \lambda_L) = (1, 0.5, 1)$

a given κ. If $\Pi < \Pi_G(\kappa)$, Mode G is preferred over Mode S while the contract is running, whereas if $\Pi > \Pi_G(\kappa)$, there is an instance at which Mode S is preferred over Mode G. See Appendix B.2.2 for details.

Lastly, if Π is very small, it might be optimal for the principal not to initiate 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. There might be some cases where the project is infeasible even if it is optimal to have made a contract under the first best, i.e., $\Pi_F(\kappa) > c/\lambda_L$. See Appendix B.2.1 for details.

The next step is to compare the above three thresholds. In Lemma B.5, I show that there exists a threshold of κ (κ^*) such that the order of three thresholds of Π 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 $\Pi_G(\kappa)$, whereas if κ is smaller than κ^*, the inequality $\Pi_S(\kappa) > \Pi_G(\kappa) > \Pi_F(\kappa)$ holds. The result is mainly due to the third condition of the previous subsection: if κ is high enough, Mode G may not be preferred to Mode S. These relationships among thresholds are illustrated in the left panel of Figure 2.

4.3.3 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 it in the appendix. The guess of the value function is based on the intuition from the previous subsections: (i) executing Mode G when it is close to the deadline; (ii) executing Mode S when the time is moderately far from the deadline. I introduce two value functions reflecting this intuition (the derivation of the value functions is relegated to Appendix B.1.1).

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

$$0 = -c + \lambda_L \left(\Pi - \frac{\phi}{\lambda_L} - u_L - V_L^g(u_L) \right) - \phi V_L^{g\prime}(u_L) \tag{4.2}$$

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

   By solving the differential equation, we obtain

$$V_L^g(u_L) = \left(\Pi - \frac{c}{\lambda_L} \right) \left(1 - e^{-\frac{\lambda_L}{\phi} u_L} \right) - u_L. \tag{4.3}$$

2. Let $V_L^{gs}(\cdot|u_s) : [u_s, \infty) \to \mathbb{R}$ be the value function that induces the agent to go for it ($a = 1$) with $R = u_L + \phi/\lambda_L$ (Mode G) 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

$$0 = -c + \lambda_S \left(V_H \left(u_L + \frac{\phi}{\lambda_S} \right) - V_L^{gs}(u_L|u_s) \right) - \phi V_L^{gs\prime}(u_L|u_s) \tag{4.4}$$

   with the boundary condition $V_L^{gs}(u_s|u_s) = V_L^g(u_s)$.

   By solving the differential equation, we obtain

$$\begin{aligned} V_L^{gs}(u_L|u_s) = & \left(\Pi - \frac{c}{\lambda_H} - \frac{c}{\lambda_S} \right) \left(1 - e^{\frac{\lambda_S}{\phi}(u_s - u_L)} \right) + (V_L^g(u_s) + u_s) e^{\frac{\lambda_S}{\phi}(u_s - u_L)} \\ & - \left(\Pi - \frac{c}{\lambda_H} \right) \frac{\lambda_S e^{-\frac{\lambda_H}{\lambda_S} - \frac{\lambda_H}{\phi} u_s}}{\lambda_S - \lambda_H} \left(e^{\frac{\lambda_H}{\phi}(u_s - u_L)} - e^{\frac{\lambda_S}{\phi}(u_s - u_L)} \right) - u_L.\text{[12]} \end{aligned} \tag{4.5}$$

The principal's value function for the low type agent consists 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 G is executed and the principal's value function is $V_L^g(u_L)$; (ii) when the promised utility u_L is higher than $u_s(\kappa)$, i.e., far from the deadline, Mode S is executed and the principal's value

[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 $-\left(\Pi - c/\lambda_H\right)(u_L - u_s) e^{-1-\lambda_H u_L/\phi}/\phi$.

function is $V_L^{gs}(u_L|u_s(\kappa))$. The threshold $u_s(\kappa)$ is identified by the smooth pasting condition $V_L^{g\prime}(u_s(\kappa)) = V_L^{gs\prime}(u_s|u_s)$, or equivalently,

$$\lambda_L \left(\Pi - \frac{\phi}{\lambda_L} - u_s - V_L^g(u_s) \right) = \lambda_S \left(V_H \left(u_s + \frac{\phi}{\lambda_S} \right) - V_L^g(u_s) \right). \tag{4.6}$$

One caveat is that if Π is greater than $\Pi_S(\kappa)$, Mode S is preferred to Mode G at the deadline by the argument in the previous subsection. Thus, only the 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 is characterized as follows:*

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

(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) $V_L^{g\prime}(u_s(\kappa)) = V_L^{gs\prime}(u_s(\kappa)|u_s(\kappa))$;

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

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

An important property of the value function to note is concavity. This is because the two parts of the value function are strictly concave (see Lemma B.1) and they are smoothly pasted ($V_L^{g\prime}(u_s(\kappa)) = V_L^{gs\prime}(u_s(\kappa)|u_s(\kappa))$).

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

4.3.4 Implementation of the Optimal Contract

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

1. **Contract G**

   For all $0 \leq t \leq T^L$, the recommended action is to go for it and the agent's payment upon success diminishes linearly over time, i.e., $(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$. If the skill is improved, the contract is terminated without payment.

2. **Contract S**

   For all $0 \le t \le T^L$, the recommended action is to study and the agent is not paid even if he succeeds by going for it, i.e., $(a^L_t, b^L_t, R^L_t) = (0, 1, 0)$. The updated contract Γ^H_t upon skill improvement at time t is given as follows: (a) the deadline is extended by $1/\lambda_S$, i.e., $T^H = T^L + 1/\lambda_S$; (b) for all $t \le s \le T^H$, $(a^{H,t}_s, R^{H,t}_s) = (1, \phi(T^H - s + 1/\lambda_H))$.

3. **Contract SG**

   There is a switch of the recommended action at T^s $(< T^L)$.

   (a) For all $0 \le t < T^s$, the contract is the same as Contract S ($(a^L_t, b^L_t, R^L_t) = (0, 1, 0)$, $T^H = T^L + 1/\lambda_S$, and $(a^{H,t}_s, R^{H,t}_s) = (1, \phi(T^H - s + 1/\lambda_H))$ for all $t \le s \le T^H$).

   (b) For all $T^s \le t \le T^L$, the contract is the same as Contract G ($(a^L_t, b^L_t, R^L_t) = (1, 0, \phi(T^L - t + 1/\lambda_L))$ and termination without payment after the skill improvement).

The main result of this paper is that the optimal contract would take a form of one of above three contracts when both technologies are equally efficient. Since $\dot{u}_L = -\phi$ holds for all candidates for 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 always be 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 G would always be executed. The following theorem summarizes the above discussion.

Theorem 1. *Suppose that both technologies are equally efficient. Define κ^* as the positive solution of*

$$0 = \phi + (c + \phi)\kappa - \phi e^{\kappa}. \tag{4.7}$$

The optimal contract is characterized as follows:

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

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

This theorem shows that the form of the optimal contract is determined by the payoff of the project (Π) and the effectiveness of the advanced skill (κ). First, consider a case where κ is above the threshold κ^*. If Π is greater than $\Pi_F(\kappa)$, Mode S is always executed, i.e., the agent is recommended to study with the minimum incentive. If Π is less than $\Pi_F(\kappa)$, the project is infeasible. Second, consider a case where κ is below κ^*. In this case, there are four possible subcases. If Π is very big ($\Pi > \Pi_S(\kappa)$), the agent is recommended to study with the minimum incentive. If Π is moderately big ($\Pi_S(\kappa) \geq \Pi > \Pi_G(\kappa)$), Mode S is executed until the intermediate deadline T^s. Conditional on no skill improvement by T^s, Mode G is executed until the deadline T^L. Thus, there is one switch of recommended action in the optimal contract. If Π is moderately small ($\Pi_G(\kappa) \geq \Pi > \Pi_F(\kappa)$), Mode G is always executed, i.e., the agent is recommended to go for it with the minimum incentive. If Π is very small ($\Pi < \Pi_F(\kappa)$), the proejct is infeasible. The right panel of Figure 2 illustrates the above discussion.

4.4 A Numerical Example

As an example, consider a numerical case with $c = 1$, $\phi = .5$, $\lambda_L = 1$, $\lambda_S = 4$, $\lambda_H = 4/3$ and $\kappa = 1/3$. By solving (4.7), we can derive that $\kappa^* \approx 1.90$, thus, $\kappa < \kappa^*$. In this case, we can also derive the thresholds of Π as follows:[13]

$$\Pi_S(1/3) \approx 6.35, \quad \Pi_G(\kappa) \approx 1.91, \quad \Pi_F(\kappa) = 1.5.$$

I consider four subcases: $\Pi = 7$, $\Pi = 2.5$, $\Pi = 1.8$ and $\Pi = 1.5$. For each case, graphs of the value function with two benchmark functions (V_L^g and $V_L^{gs}(\cdot|0)$) are shown in Figure 3.

When $\Pi = 7$, Π is greater than $\Pi_S(1/3)$. Then, by Proposition 4.1, the value function V_L is exactly same as $V_L^{gs}(\cdot|0)$. This is shown in Figure 3a. In this case, Mode S is executed for all u_L, i.e., Contract S is the optimal contract as stated in Theorem 1.

When $\Pi = 2.5$, Π is less than $\Pi_S(1/3)$. Then, by Proposition 4.1, the value function V_L is $V_L^{gs}(\cdot|u_s)$ where $u_s \approx 0.16$. The value function V_L is maximized at $\bar{u} \approx .59$. This is shown in Figure 3b Note that $V_L(\bar{u}) > 0$. This is because Π is greater than $\Pi_F(1/3)$. Also note that $\bar{u} > u_s$. This is because Π is greater than $\Pi_G(1/3)$. In this case, the optimal contract is Contract SG with a deadline $T^L = \bar{u}/\phi \approx 1.18$ and a switch point $T^s = (\bar{u} - u_s)/\phi \approx .86$.

When $\Pi = 1.8$, Π is less than $\Pi_S(1/3)$. Then, by Proposition 4.1, the value function V_L is $V_L^{gs}(\cdot|u_s)$ where $u_s \approx 0.34$. The value function V_L is maximized at $\bar{u} \approx .235$. This is shown in Figure 3c. Note that $V_L(\bar{u}) > 0$. This is because Π is greater than $\Pi_F(1/3)$. Also note

[13] For the formula of Π_S, see (B.1) in Lemma B.2. For the formula of Π_G, see (B.10) in Appendix B.2.2. For the formula of Π_F, see (B.9) in Appendix B.2.1.

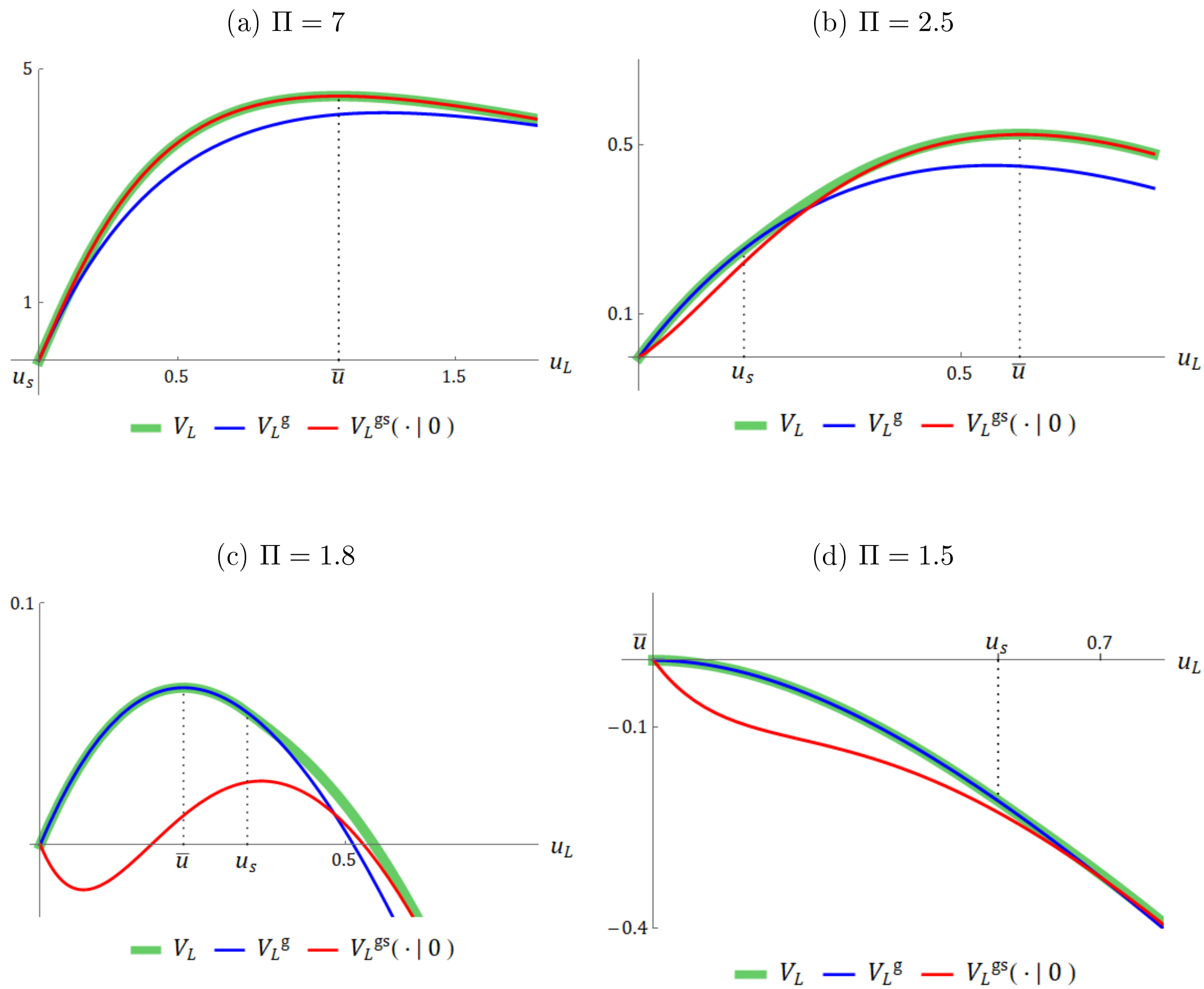

Figure 3: The value functions and benchmark value functions when parameter values are $\lambda_L = 1,\ \lambda_S = 4,\ \lambda_H = 4/3,\ c = 1,\ \phi = 0.5$

that $\bar{u} < u_s$. This is because Π is less than $\Pi_G(1/3)$. In this case, Mode G is executed for all $u_L < \bar{u}$, i.e., Contract G with the deadline $T^L \approx .47$ is the optimal contract as stated in Theorem 1.

When $\Pi = 1.5$, Π is equal to $\Pi_F(1/3)$. In this case, the value function is maximized at $\bar{u} = 0$. This is shown in Figure 3d. We can see that the value function takes negative values for all positive promised utility levels. Therefore, the principal is better off not to contract with the agent in the first place, i.e., 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. There are too many subcases for the unequally efficient technologies case. Thus, 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 show that the efficiency determines the recommended action at a time far from the deadline, i.e., large enough u_L. By (4.3) and (4.5), we can derive that

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

for all $u_s \geq 0$.

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

When technologies are unequally efficient, $V_L^g(u_L)$ and $V_L^{gs}(u_L|u_s)$ no longer converge to the same asymptotic line and the result of Proposition 4.1 may not hold . Graphs in Figure 4a and 4b show that as u_L increases, $V_L^g(u_L)$ converges to $\Pi - c/\lambda_L - u_L$ and $V_L^{gs}(u_L|0)$ converges to $\Pi - c/\lambda_S - c/\lambda_H - u_L$. When going for it is more efficient, $V_L^g(u_L)$ would be greater than $V_L^{gs}(u_L|u_s)$ for large enough u_L. It means that for large enough u_L, switching

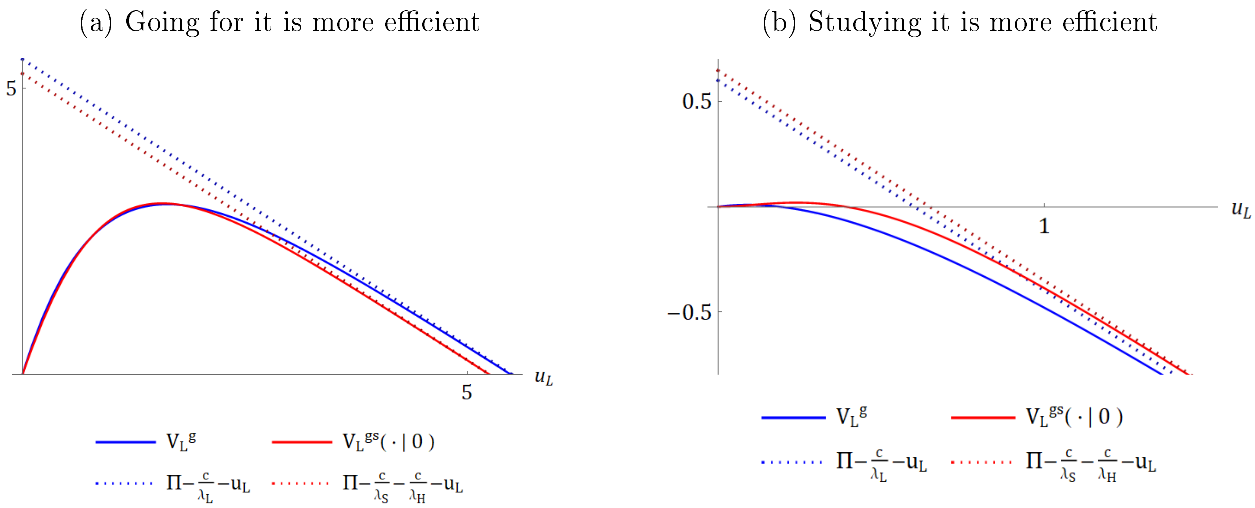

Figure 4: The benchmark value functions and asymptotic lines

the recommended action from studying (for $u_s \leq u \leq u_L$) to going for it (for $0 \leq u < u_s$) may not be optimal because it gives less expected payoff than going for it always (for $0 \leq u \leq u_L$). Likewise, when studying it is more efficient, $V_L^{gs}(u_L|u_s)$ would be greater than $V_L^g(u_L)$ for large enough u_L. In this case, going for it always may not be preferred to switching from studying to going for it. From these observations, we can guess that the recommended technology would be the more efficient one when the deadline is far away.

5.2 Going for it is more Efficient

In this subsection, I present a numerical example where the optimal contract involves the following two switches of contractual modes:

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

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

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

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

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

The value functions for $0 \leq u_L \leq u_{s,1}$ would be identical to those in the equally efficient technologies case ((4.3) for $0 \leq u_L \leq u_{s,2}$ and (4.5) for $u_{s,2} \leq u_L \leq u_{s,1}$). For $u_{s,1} < u_L$, the value function $V_L^{gsg}(\cdot|u_{s,2},u_{s,1}) : [u_{s,1},\infty) \to \mathbb{R}$, is derived by solving the differential equation identical to (4.2) with the boundary condition $V_L^{gsg}(u_{s,1}|u_{s,2},u_{s,1}) = V_L^{gs}(u_{s,1}|u_{s,2})$:

$$V_L^{gsg}(u_L|u_{s,2},u_{s,1}) \equiv \left(\Pi - \frac{c}{\lambda_L}\right)\left(1 - e^{\frac{\lambda_L}{\phi}(u_{s,1}-u_L)}\right) - u_L \\ + \left(V_L^{gs}(u_{s,1}|u_{s,2}) + u_{s,1}\right) e^{\frac{\lambda_L}{\phi}(u_{s,1}-u_L)}.$$

In sum, I guess the value function as follows:

$$V_L(u_L) = \begin{cases} V_L^g(u_L), & 0 \le u_L \le u_{s,2}, \\ V_L^{gs}(u_L|u_{s,1}), & u_{s,2} \le u_L \le u_{s,1}, \\ V_L^{gsg}(u_L|u_{s,1},u_{s,2}), & u_{s,1} \le u_L. \end{cases} \tag{5.1}$$

The next step is to identify $u_{s,1}$ and $u_{s,2}$ by the smooth pasting conditions $V_L^{g\prime}(u_{s,2}) = V_L^{gs\prime}(u_{s,2}|u_{s,2})$ and $V_L^{gs\prime}(u_{s,1}|u_{s,2}) = V_L^{gsg\prime}(u_{s,1}|u_{s,2},u_{s,1})$. First, solve $V_L^{g\prime}(u_{s,2}) = V_L^{gs\prime}(u_{s,2}|u_{s,2})$, or equivalently (4.6), to obtain $u_{s,2}$. If there are multiple positive solutions, the minimum positive solution should be chosen as $u_{s,2}$. Then, $u_{s,1}$ ($> u_{s,2}$) can be obtained by solving $V_L^{gs\prime}(u_{s,1}|u_{s,2}) = V_L^{gsg\prime}(u_{s,1}|u_{s,2},u_{s,1})$, or equivalently

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

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

Now I present a numerical example where the value function takes the form of (5.1). Let the 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., going for it is more efficient than the study-it technology. First, by solving (4.6), we can derive that $u_{s,2} \approx 0.375$.[14] Second, by solving (5.2), we have $u_{s,1} \approx 1.342$. Lastly, we can show that V_L is maximized at $\bar{u} \approx 1.599$. Figure 5 illustrates the value function V_L and the benchmark value functions V_L^g and $V_L^{gs}(\cdot|0)$ for these parameter values.

From (5.2), we can show 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 go for it and receive the immediate payment as compensation. 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 in skill by time $T^{s,2}$, the

[14] The equation (4.6) has two solutions 0.375 and 1.436, and we need to choose the smaller one for $u_{s,2}$.

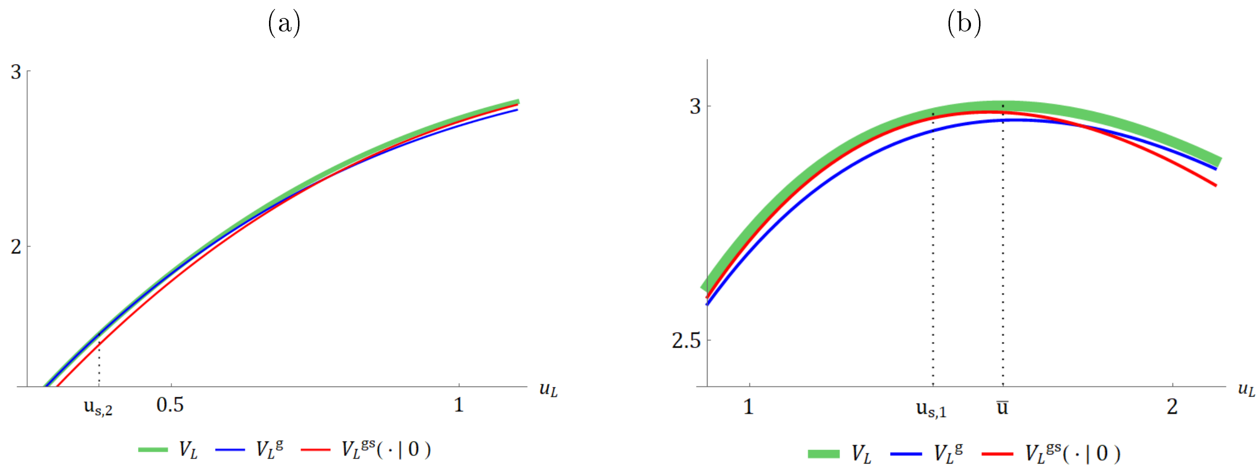

Figure 5: The value function and benchmark functions when parameters are $\lambda_L = 1,\ \lambda_S = \lambda_H = 1.6,\ \Pi = 6.5,\ c = 1,\ \phi = 0.9$

recommended action is switched again to go for it and he is compensated by the immediate payment.

5.3 Studying it is more Efficient

Green and Taylor (2016b) explore the possibility of the incentive compatibility condition not binding. When there is no go-for-it technology, they show that the incentive compatibility condition may not bind if Π is relatively small (Proposition 4.3). Moreover, this case arises when the deadline is close by. However, in the previous cases, the optimal contract consists of Mode G or Mode S, i.e., the incentive compatibility condition always binds. This is because the option to choose the go-for-it technology near the deadline makes the IC bind. Nevertheless, this statement may no longer be true when studying it is more efficient. In this case, the principal would like to recommend going for it for a very short period of time since the go-for-it technology is less appealing. Then, this period of time may be shorter than the period of time that the principal needs to make the IC bind. Based on this intuition, I present a numerical example where the value function is derived as follows:[15]

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 ($u_{s,2} < u_L < u_{s,1}$);

[15]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.

3. Mode G 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., $V_L(u_L) = V_L^g(u_L)$ as defined in (4.3). For $u_{s,2} < u_L < u_{s,1}$, the value function $V_L^{gn}(\cdot|u_{s,2}, u_{s,1}) : [u_{s,2}, u_{s,1}] \to \mathbb{R}$, is derived by plugging $b = 1$ and $u_{s,2} = u_{s,1} + \phi/\lambda_S$ into (HJB_L) and (PK_L), i.e.,

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

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

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

For $u_{s,1} \leq u_L$, the value function $V_L^{gns}(\cdot|u_{s,2}, u_{s,1}) : [u_{s,1}, \infty) \to \mathbb{R}$, is derived by solving the differential equation identical to (4.4) with the boundary condition $V_L^{gns}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{gn}(u_{s,1}|u_{s,2}, u_{s,1})$:

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

In sum, I guess the value function as follows:

$$V_L(u_L) = \begin{cases} V_L^g(u_L), & 0 \leq u_L \leq u_{s,2}, \\ V_L^{gn}(u_L|u_{s,2}, u_{s,1}), & u_{s,2} \leq u_L \leq u_{s,1}, \\ V_L^{gns}(u_L|u_{s,2}, u_{s,1}), & u_{s,1} \leq u_L. \end{cases}$$

The next step is to identify $u_{s,1}$ and $u_{s,2}$ by the smooth pasting condition $V_L^{g\prime}(u_{s,2}) = V_L^{gn\prime}(u_{s,2}|u_{s,2}, u_{s,1})$ and the super contact condition $V_L^{gn\prime\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{gns\prime\prime}(u_{s,1}|u_{s,2}, u_{s,1})$. The detailed derivation is relegated to Appendix C.

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

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

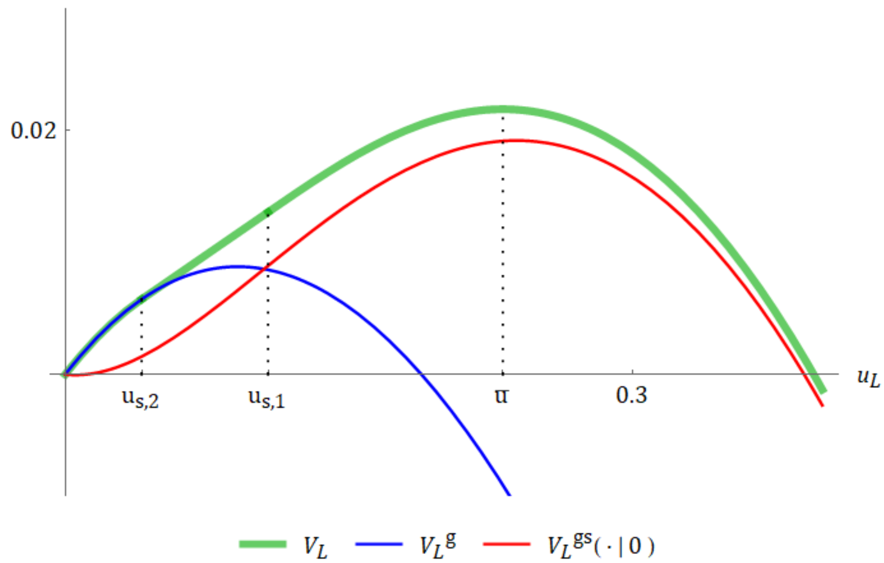

Figure 6: The value function and benchmark value functions when parameters are $\lambda_L = 1,\ \lambda_S = \lambda_H = 2.1,\ \Pi = 1.6,\ c = 1,\ \phi = 0.5$

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 $0 \le t \le T^{s,1}$ (and $\bar{u} \ge u_L \ge u_{s,1}$),

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

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

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

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

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

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:

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

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

3. For all $T^{s,2} \leq t \leq T^L$, the contract is same as in Contract G ($(a_t^L, b_t^L, R_t^L) = (1, 0, \phi(T^L - t + 1/\lambda_L))$, termination after the skill improvement).

Now I present a numerical example where the value function takes the form of (5.3). Let the 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., studying it is more efficient than the go-for-it technology. By solving the smooth pasting conditions $V_L^{g\prime}(u_{s,2}) = V_L^{gn\prime}(u_{s,2}|u_{s,2}, u_{s,1})$ and $V_L^{gn\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{gns\prime}(u_{s,1}|u_{s,2}, u_{s,1})$, we can show that $u_{s,2} \approx .040$ and $u_{s,1} \approx .107$. Then, we can show that V_L is maximized at $\bar{u} \approx .231$. Figure 5 shows the value function V_L and the benchmark value functions V_L^g and $V_L^{gs}(\cdot|0)$ for these parameter values.

From (5.2), we can show 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. If the agent improves in 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 in skill by time $T^{s,1}$, he is still recommended to study, but if the agent improves in 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 in skill by time $T^{s,2}$, the recommended action is switched to go for 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. I design an optimal incentive scheme in the dynamic principal agent setup. When the go-for-it technology and the study-it technology are equally efficient and the skill improvement is observable to the principal and the agent, I show that the principal provides the minimum incentive for the agent not to shirk and the recommended action schedule would be one of always going for it, always studying, or switching once from studying to going for 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 where 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 (going for it $\rightarrow$ studying $\rightarrow$ going for 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 his progress truthfully. Also, I assume that the advanced skill surely exists and that 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 where only one ultimate breakthrough pays off. This assumption can be modified by considering projects that have flow payoffs or multiple breakthroughs that pay off.

References

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.

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 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.

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.

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 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 principal and 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:

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

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 defined as follows:

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

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

$$\begin{aligned} W_L &= \max_{a,b \in [0,1], a+b \leq 1} \left(-c + \lambda_L a \Pi + \lambda_S b W_H \right) dt + (1 - \lambda_L a dt - \lambda_S b dt) W_L \\ \Rightarrow \qquad 0 &= \max_{a,b \in [0,1], a+b \leq 1} -c + \lambda_L (\Pi - W_L) a + \lambda_S (W_H - W_L) b \end{aligned} \tag{A.1}$$

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 optimal because $-c < 0$.

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

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

Therefore, when the go-for-it technology is more efficient ($1/\lambda_L < 1/\lambda_S + 1/\lambda_H$), the going for it only schedule maximizes the principal's payoff, and when the study-it technology is more efficient ($1/\lambda_L > 1/\lambda_S + 1/\lambda_H$), the studying only schedule maximizes the principal's payoff. $\square$

B Proofs for Section 4

B.1 Derivation of the Value Function

B.1.1 Derivation of Value Function Candidates

1. V_L^g takes the same form as V_H in (4.1) except that the arrival rate goes from λ_H to λ_L. Therefore, V_L^w should take the same form as (4.3).

2. By multiplying $e^{\lambda_S u_L/\phi}/\phi$ to the HJB equation (4.4), the equation can be rewritten as follows:

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

   The left hand side is equal to $\frac{d}{du_L}\left(e^{\lambda_S u_L/\phi} V_L^{gs}(u_L|u_s)\right)$. 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, integrating the HJB equation from u_s to u_L gives

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

   By rearranging the above equation and using the boundary condition ($V_L^{gs}(u_s|u_s) = V_L^g(u_s)$), we can show that V_L^{gs} takes the same form as (4.5).

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

Lemma B.1. *Suppose that both technologies are equally efficient. Then, the following statements hold:*

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

(b) Suppose that u_s *satisfies* $V_L^{g\prime}(u_s) \leq V_L^{gs\prime}(u_s|u_s)$. *Then, for all* $u_L \geq u_s$, $V_L^{gs}(u_L|u_s) < \Pi - c/\lambda_L - u_L$, $V_L^{gs\prime}(u_L|u_s) > -1$ *and* $V_L^{gs\prime\prime}(u_L|u_s) < 0$.

The first result means that contracting with Mode G cannot achieve the first best level $\Pi - c/\lambda_L - u_L$, but performs better than immediately paying out the promised utility ($V_L^{g\prime}(u_L) > -1$). Moreover, the principal's value would be strictly concave with respect to u_L. The second result means that if Mode G is executed from $[0, u_s)$ and the instantaneous benefit from Mode S at u_s is greater than that from Mode G, the same result as above holds for contracting with Mode S (for $u_L \geq u_s$). The assumption $V_L^{g\prime}(u_s) \leq V_L^{gs\prime}(u_s|u_s)$ is essential for the result. If the assumption is violated, $V_L^{gs}(\cdot|u_s)$ may no longer be concave. As an example, see Figure 3d. In the example, $V_L^{g\prime}(0) > 0 > V_L^{gs\prime}(0|0)$, so the assumption does not hold. We can easily see that $V_L^{gs}(\cdot|0)$ is not concave.

B.1.2 The Recommended Action at the Deadline (Π_S)

Since we have the boundary condition $V_L(0) = 0$, it is natural to begin by solving the value function near $u_L = 0$. In other words, I will determine which contractual mode is 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:

$$\textbf{Mode G:} \qquad -c + \lambda_L \Pi - \phi - \phi V_L'(0),$$

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

In Mode G, 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 into 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 $\lambda_L < (1+1/\kappa)(1-e^{-\kappa})\lambda_L$, or 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 G would be preferred if Π is small. The following lemma determines the threshold.

Lemma B.2. *Define*

$$\Pi_S(\kappa) \equiv \frac{1-e^{-\kappa}}{1-(\kappa+1)e^{-\kappa}} \cdot \frac{c}{\lambda_L}.\text{[16]} \tag{B.1}$$

When $\Pi_S(\kappa) \geq \Pi > c/\lambda_L$, Mode G 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 G 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.

B.1.3 Some Useful Functions

In this subsection, I introduce functions that specify deviations from the given value functions. Then I present some properties of these functions.[17]

1. Functions for deviation given V_L^g

(a) Define

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

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

(b) Define

$$L_1^b(u,w) \equiv \lambda_S(V_H(w) - V_L^g(u)) - c - \lambda_S(w-u)V_L^{g\prime}(u) \tag{B.2}$$

$$= \lambda_S \left[\left(\Pi - \frac{c}{\lambda_H}\right)\left(1 - e^{-\frac{\lambda_H}{\phi}u - \frac{\lambda_H}{\phi}(w-u)}\right) \right. \tag{B.3}$$

$$\left. - \left(\Pi - \frac{c}{\lambda_L}\right)\left(1 - e^{-\frac{\lambda_L}{\phi}u}\right) - (w-u)\left(\frac{\lambda_L\Pi - c}{\phi}\right)e^{-\frac{\lambda_L}{\phi}u} \right] - c.\text{[18]}$$

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

[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.

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

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

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

From $\phi V_L^{gs\prime}(0|0) = -c + \lambda_S V_H(\phi/\lambda_S)$ (by (4.4) and $V_L^{gs}(0|0) = 0$), we can derive that

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

2. Functions for deviation given V_L^{gs}

(a) Define

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

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

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

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

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

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

From $\phi V_L^{g\prime}(u_s) = -c+\lambda_L(\Pi-\phi/\lambda_L-u_s-V_L^g(u_s))$ (by (4.2)) and $V_L^{gs}(u_s|u_s) = V_L^g(u_s)$, we can derive that

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

(b) Define

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

Note that $\frac{\partial}{\partial w}L_2^b = \lambda_S(V_H'(w)-V_L^{gs\prime}(u|u_s))$ 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 B.3. *Suppose that λ, c, and Π are fixed, $\lambda_L\Pi > c$ is satisfied, and both technologies are equally efficient ($1/\lambda_L = 1/\lambda_S + 1/\lambda_H$). Then, $\tilde{L}_1^b$ satisfies the following properties:*

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

(b) If $y \geq \phi/\lambda_L$ and $x \geq 0$, then $\tilde{L}_1^b(x,y) \leq 0$;

(c) Define

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

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

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

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

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

Lemma B.4. *Suppose that λ, c, and Π are fixed, $\lambda_L \Pi > c$ and both technologies are equally efficient ($1/\lambda_L = 1/\lambda_S + 1/\lambda_H$). 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 $V_L^{g\prime}(u_s) \leq V_L^{gs\prime}(u_s|u_s)$ 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 $w \geq u + \phi/\lambda_S$, $L_2^b(u, w|u_s) \leq L_2^b(u, u + \phi/\lambda_S|u_s) = 0$.

B.1.4 Proofs of Proposition 4.1

Proof of Proposition 4.1. (a) By construction, $V_L^{gs}(u_L|0)$ is the principal's payoff under Contract S with $T^L = u_L/\phi$.

By Lemma B.2, $V_L^{gs\prime\prime}(0|0) > V_L^{g\prime}(0)$. By (b) and (c) of Lemma B.4, 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 $V_L(u) = V_L^{gs}(u|0)$. Therefore, $V_L^{gs}(u|0)$ solves (HJB$_L$).

(b) By construction, $V_L^{gs}(u_L|u_s)$ is the principal's payoff under Contract SG with $T^L = u_L/\phi$ and $T^s = (u_L - u_s)/\phi$.

By Lemma B.2 , $V_L^{gs\prime\prime}(0|0) \leq V_L^{g\prime}(0)$ and $\tilde{L}_1^b(1, \phi/\lambda_S) \leq 0$. Also note that $\tilde{L}_1^b(0, \phi/\lambda_S) = 0$. Note that $\tilde{L}_1^b$ is strictly concave in x_1 ((a) of Lemma B.3) and it is maximized at $x^*(\phi/\lambda_S) > 0$ where x^* is defined as (B.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

$$\left(\Pi - \frac{c}{\lambda_H}\right) e^{-\kappa} \bar{x}^{\kappa+1} - \frac{\lambda_L}{\lambda_H} \left(\Pi - \frac{c}{\lambda_L}\right) \bar{x} = 0,$$

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

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

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

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

Now, we verify the value function for $u_L \geq u_s(\kappa)$. 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 (B.2), we have

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

By (4.4) and $V_L^g(u_s) = V_L^{gs}(u_s|u_s)$, we also have

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

Then, $V_L^{g\prime}(u_s) = V_L^{gs\prime}(u_s|u_s)$ and we can apply Lemma B.4 in the same manner as in (a).

$\square$

B.2 Derivation of the Thresholds

B.2.1 Feasibility (Π_F)

In this subsection, I 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 $V_L'(0) \leq 0$, the maximum of the value function is 0 at $u_L = 0$ since V_L is concave (Corollary 4.2). Thus, the project is infeasible. Note that from (HJB$_L$), $V_L'(0) > 0$ is equivalent to

$$\max\left[\lambda_L \Pi - \phi, \lambda_S V_H\left(\frac{\phi}{\lambda_S}\right)\right] > c, \tag{B.8}$$

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

$$\Pi_F(\kappa) \equiv \min\left[\frac{c+\phi}{\lambda_L},\ \frac{1}{(\kappa+1)\lambda_L}\left(c+\frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right)\right]. \tag{B.9}$$

Then, it can be shown that $\Pi_F(\kappa) < \Pi$ is equivalent to (B.8).

B.2.2 The Length of the Contract (Π_G)

The next step is to solve the maximization problem (MP_L), which is the same problem as checking the length of the contract. Define $\bar{u}(\kappa)$ as 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 $\bar{u}(\kappa)$. As time goes by, the promised utility diminishes, conditional on the agent not having made success or skill improvement.

To check whether there is a switch in the recommended action during the contract, we need to compare $u_s(\kappa)$ and $\bar{u}(\kappa)$. If $\bar{u}(\kappa) \leq u_s(\kappa)$, the principal always recommends the agent to go for it. But, if $\bar{u}(\kappa) > u_s(\kappa)$, for u_L with $\bar{u}(\kappa) \geq u_L \geq u_s(\kappa)$, the principal's recommendation would be to study, while for u_L with $u_s(\kappa) > u_L$, the recommendation would be to go for 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 $V_L'(u_s(\kappa)) = V_L^{g\prime}(u_s(\kappa))$, $V_L'(u_s(\kappa)) = 0$ is equivalent to

$$\frac{(\lambda_L\Pi - c)^{1+\kappa}}{(1+\kappa)\lambda_L\Pi - c}\left(\frac{e}{\phi}\right)^{\kappa} = 1. \tag{B.10}$$

Moreover, $V_L'(u_s(\kappa)) > 0$ is equivalent to the inequality where 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 satisfies (B.10) and I denote the solution as $\Pi_G(\kappa)$. Then, $\Pi > \Pi_G(\kappa)$ is equivalent to $\bar{u}(\kappa) > u_s(\kappa)$.

Lemma B.5. *Let κ^* be the positive solution of (4.7). Then,*

1. *if $\kappa > \kappa^*$,*

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

2. if $\kappa = \kappa^*$,

$$\Pi_F(\kappa^*) = \frac{c+\phi}{\lambda_L} = \frac{1}{(\kappa^*+1)\lambda_L}\left(c + \frac{\kappa^*}{1-e^{-\kappa^*}}(c+\phi)\right) = \Pi_G(\kappa^*) = \Pi_S(\kappa^*);$$

3. if $\kappa < \kappa^*$,

$$\Pi_F(\kappa) = \frac{c+\phi}{\lambda_L} < \Pi_G(\kappa) < \Pi_S(\kappa);$$

4. as $\kappa \to 0$,

$$\lim_{\kappa\to 0} \Pi_G(\kappa) = \frac{c + \phi \cdot \psi\,(c/\phi)}{\lambda_L} \quad \textit{and} \quad \lim_{\kappa\to 0} \Pi_S(\kappa) = \infty,$$

where $\psi : \mathbb{R}_+ \to [1,\infty)$ is the inverse function of $x\log(x)$ for $x \geq 1$.

The left panel of Figure 2 illustrates graphs of Π_S, Π_G and Π_F when $c = 1$, $\phi = 0.5$ and $\lambda_L = 1$. We can see that the three graphs coincide at $(\kappa^*, (c+\phi)/\lambda_L)$.

B.3 Implementation

Proof of Theorem 1. 1. By Lemma B.5, when $\kappa \geq \kappa^*$, $\Pi_F(\kappa) \geq \Pi_S(\kappa)$. Thus, when $\kappa \geq \kappa^*$ & $\Pi > \Pi_F(\kappa)$ or $\kappa < \kappa^*$ & $\Pi > \Pi_S(\kappa)$, Π is greater than $\Pi_S(\kappa)$. Then, by Proposition 4.1(a), the value function is $V_L(u_L) = V_L^{gs}(u_L|0)$. The value function is maximized at $\bar{u}(\kappa)$. Thus, the optimal contract is to execute Mode S for all $0 \leq u_L \leq \bar{u}(\kappa)$. Therefore, Contract S with $T^L = \bar{u}(\kappa)/\phi$ implements $(\bar{u}(\kappa), V_L(\bar{u}(\kappa)))$.

2. Suppose that $\kappa < \kappa^*$ and $\Pi_S(\kappa) \geq \Pi > \Pi_G(\kappa)$. By Proposition 4.1(a), the value function is $V_L(u_L) = V_L^{gs}(u_L|u_s(\kappa))$ for $u_s(\kappa) < u_L$ and $V_L(u_L) = V_L^g(u_L)$ for $u_L \leq u_s(\kappa)$. By the argument of Section B.2.2, $\bar{u}(\kappa)$ is greater than $u_s(\kappa)$. Thus, the optimal contract is to execute Mode S for $u_s(\kappa) < u_L \leq \bar{u}(\kappa)$ and Mode G for $0 \leq u_L \leq u_s(\kappa)$. Therefore, Contract SG with $T^L = \bar{u}(\kappa)/\phi$ and $T^s = (\bar{u}(\kappa) - u_S(\kappa))/\phi$ implements $(\bar{u}(\kappa), V_L(\bar{u}(\kappa)))$.

3. Suppose that $\kappa < \kappa^*$ and $\Pi_G(\kappa) \geq \Pi > \Pi_F(\kappa)$. By the argument of Section B.2.2, $\bar{u}(\kappa)$ is less than or equal to $u_s(\kappa)$. Note that $\Pi_S(\kappa) > \Pi_G(\kappa) \geq \Pi$. By Proposition 4.1(a), the value function is $V_L(u_L) = V_L^g(u_L)$ for $u_L \leq \bar{u}(\kappa) \leq u_s(\kappa)$. Thus, the optimal contract is to execute Mode G for all $u_L \leq \bar{u}(\kappa)$. Therefore, Contract G with $T^L = \bar{u}(\kappa)/\phi$ implements $(\bar{u}(\kappa), V_L(\bar{u}(\kappa)))$.

4. When $\Pi \leq \Pi_F(\kappa)$, V_L is maximized at $u_L = 0$. Therefore, it is better off not to initiate the contract with the agent, i.e., the project is infeasible.

□

B.4 Proofs of Lemmas

Lemma B.6. *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 $(R^* - u_L)\lambda_L - \phi = (u_H^* - u_L)\lambda_H - \phi \geq 0$ from the maximization of (PK$_L$). Also note that $(\Pi - R^* - V_L(u_L))\lambda_L = (V_H(u_H^*) - V_L(u_L))\lambda_S \geq 0$ 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. Consider an arbitrary (deterministic) contract Γ^L where the agent's expected payoff is given by . Let $A_t \equiv \int_0^t a_s ds$ and $B_t = \int_0^t b_s ds$. The payoff to the principal under Γ^L

$$
\begin{aligned}
P_0^L(\Gamma^L) &= \int_0^T \lambda_L a_t e^{-\lambda_L A_t - \lambda_S B_t} (\Pi - R_t - c\,t)\, dt \\
&\quad + \int_0^T \lambda_S b_t e^{-\lambda_L A_t - \lambda_S B_t} (V_H(u_{H,t}) - c\,t) + e^{-\lambda_L A_T - \lambda_S B_T} (-c\,T) \\
&= \int_0^T e^{-\lambda_L A_t - \lambda_S B_t} ((\Pi - R_t)\lambda_L a_t + V_H(u_{H,t})\lambda_S b_t - c)\, dt
\end{aligned}
$$

where $u_{H,t} = U_t^H(\Gamma_t^H)$.[19]

Since $\hat{V}$ solves the HJB equation, we have

$$0 \geq -c + (\Pi - R_t - \hat{V}(u_{L,t}))\lambda_L a_t + (V_H(u_{H,t}) - \hat{V}(u_{L,t}))\lambda_S b_t + \hat{V}'(u_{L,t})\dot{u}_{L,t}.$$

By rearranging and multiplying by $e^{-\lambda_L A_t - \lambda_S B_t}$,

$$
\begin{aligned}
&(\lambda_L a_t + \lambda_S b_t) e^{-\lambda_L A_t - \lambda_S B_t} \hat{V}(u_{L,t}) - e^{-\lambda_L A_t - \lambda_S B_t} \hat{V}'(u_{L,t})\dot{u}_{L,t} \\
\geq & e^{-\lambda_L A_t - \lambda_S B_t} ((\Pi - R_t)\lambda_L a_t + V_H(u_{H,t})\lambda_S b_t - c)
\end{aligned}
\tag{B.11}
$$

Note that

$$\frac{d}{dt}\left(-e^{-\lambda_L A_t - \lambda_S B_t}\hat{V}(u_{L,t})\right) = (\lambda_L a_t + \lambda_S b_t) e^{-\lambda_L A_t - \lambda_S B_t} \hat{V}(u_{L,t}) - e^{-\lambda_L A_t - \lambda_S B_t} \hat{V}'(u_{L,t})\dot{u}_{L,t}.$$

[19] For simplicity, I take off L from T^L.

Then, by integrating (B.11) over $[0, T]$ and noting that $u_{L,T} = 0$, we have

$$\begin{aligned}
&\hat{V}(u_{L,0}) = \hat{V}(u_{L,0}) - e^{-\lambda_L A_T - \lambda_S B_T}\hat{V}(u_{L,T}) \\
\geq &\int_0^T e^{-\lambda_L A_t - \lambda_S B_t}\left((\Pi - R_t)\lambda_L a_t + V_H(u_{H,t})\lambda_S b_t - c\right) dt = P_0^L(\Gamma^L).
\end{aligned}$$

Therefore, $\hat{V}(u_{L,0})$ is greater than or equal to any deterministic contract where the agent's expected payoff is equal to $u_{L,0}$. Since $\hat{V}$ is assumed to be concave, it is greater than or equal to any randomized contract. □

Proof of Lemma B.1. (a) By (4.3) and $e^{-\lambda_L u_L/\phi} > 0$, $V_L^g(u_L) < \Pi - c/\lambda_L - u_L$. By differentiating (4.3), we have

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

By differentiating once again, we have

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

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

$$\frac{e^{\frac{\lambda_H}{\phi}(u_s - u_L)} - e^{\frac{\lambda_S}{\phi}(u_s - u_L)}}{\lambda_S - \lambda_H} \geq 0. \tag{B.12}$$

Then, by (4.5), $\Pi - c/\lambda_H - c/\lambda_S = \Pi - c/\lambda_L$, $V_L^g(u_s) + u_s < \Pi - c/\lambda_L - u_L$ and the above inequality, we have

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

Note that

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

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

$$-\Pi + \frac{\lambda_S\ c}{(\lambda_S-\lambda_L)\lambda_H} + \frac{\lambda_S}{\lambda_S-\lambda_L}\left(\Pi - \frac{c}{\lambda_H}\right)e^{-\frac{\lambda_H}{\phi}u_s - \frac{\lambda_H}{\lambda_S}} \leq -\left(u_s + V_L^g(u_s)\right). \quad \text{(B.13)}$$

By differentiating (4.5) twice, we have

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

By using (B.13),

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

Then, plugging the above equation and the inequality (B.12) into the equation of $V_L^{gs\prime\prime}(u_L|u_s)$, we can derive that $V_L^{gs\prime\prime}(u_L|u_s) \leq 0$.

Note that

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

$$\lim_{u_L\to\infty} V_L^{gs\prime}(u_L|u_s) = -1.$$

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

$\square$

Proof of Lemma B.2. Since $1-(\kappa+1)e^{-\kappa} > 0$ for all $\kappa > 0$, $\lambda_L \Pi - \phi \geq \lambda_S V_H(\phi/\lambda_S)$ is equivalent to

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

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 $e^{-\kappa} > -\kappa+1$, we have $\Pi_S'(\kappa) < 0$. Thus, Π_S is decreasing in κ. □

Proof of Lemma B.3. (a) By $\lambda_H > \lambda_L$, $\Pi > c/\lambda_L > c/\lambda_H$, and $x > 0$,

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

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

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

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

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, for all $1 \geq x > 0$ and $y \geq \phi/\lambda_L$, the following inequalities hold:

$$\tilde{L}_1^b(x,y) \leq \tilde{L}_1^b(0,y) = 0.$$

(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

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

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 $1/\lambda_L = 1/\lambda_S + 1/\lambda_H$, $g(y)$ 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

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

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

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

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,

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

since $x^*(y)$ is optimal and $\tilde{L}_1^b(x^*(y), y)$ is decreasing 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 $y > y^*$,

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

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

By combining the above results, we can show that

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

□

Proof of Lemma B.4. (a) By differentiating (B.5) twice, we have

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

from $\Pi > c/\lambda_L > c/\lambda_H$. 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 $\tilde{L}_2^a(1) = \phi(V_L^{g\prime}(u_s) - V_L^{gs\prime}(u_s|u_s)) \leq 0$ from (B.6) and $V_L^{g\prime}(u_s) \leq V_L^{gs\prime}(u_s|u_s)$. 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 (4.4) once, we can derive that

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

By (b) of Lemma B.1, $V_L^{gs\prime\prime}(u|u_s) < 0$. Thus, the following inequality holds:

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

Since $\frac{\partial^2}{\partial w^2}L_2^b < 0$, $L_2^b(u,w|u_s)$ subject to $w \geq u + \phi/\lambda_S$ is maximized at $w = u + \phi/\lambda_S$ for a given u. Also note that $L_2^b(u, u+\phi/\lambda_S|u_s) = 0$ holds by (4.4).

□

Proof of Lemma B.5. 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 to $g(\kappa) = 0$, 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) to the equation

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

where

$$
\begin{aligned}
L(\Pi) &= (\kappa+1)\lambda_L\Pi - c, \\
R_F(\Pi|\kappa) &= \begin{cases} \dfrac{\kappa}{1-e^{-\kappa}}(c+\phi) & \text{if } \ \kappa \geq \kappa^* \\ \kappa(c+\phi)+\phi & \text{if } \ \kappa \leq \kappa^* \end{cases}, [21] \\
R_W(\Pi|\kappa) &= \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_L\Pi - c}{\phi}\right)^{\kappa+1}, \\
R_S(\Pi|\kappa) &= \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_L\Pi - c}{\phi}\right).
\end{aligned}
$$

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 $\kappa > 0$.

[21] 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 $\Pi_F(\kappa) > \max[\Pi_G(\kappa), \Pi_S(\kappa)]$, 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
$$\begin{aligned} R_F(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot x(\kappa), \\ R_W(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right)^{\kappa+1}, \\ R_S(\Pi_F(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right). \end{aligned} \tag{B.14}$$

   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 $\Pi_F(\kappa) = \Pi_G(\kappa) = \Pi_S(\kappa)$, 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 (B.14), $R_F(\Pi_F(\kappa)|\kappa) = R_W(\Pi_F(\kappa)|\kappa) = R_S(\Pi_F(\kappa)|\kappa)$.

3. When $\kappa < \kappa^*$, to prove that $\Pi_S(\kappa) > \Pi_G(\kappa) > \Pi_F(\kappa)$, 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'Hôpital'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 $y(\kappa) \equiv (\lambda_L \Pi_G(\kappa) - c)/\phi > 0$. Then, from (B.10), $y(\kappa)$ satisfies the following equations for all $\kappa > 0$:

$$y(\kappa)^{1+\kappa} \cdot e^\kappa = (1+\kappa) y(\kappa) + \frac{c}{\phi}\kappa$$

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

By differentiating the above equation by κ, we have

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

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

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

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

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

□

C Details in Section 5.3

In this section, I present how to pin down $u_{s,1}$ and $u_{s,2}$. Note that $V_L^{gn}(u_L|u_{s,2}, u_{s,1})$ is linear and denote the slope as A. For simplicity, define

$$\begin{aligned}
\Delta &\equiv u_{s,1} - u_{s,2}, \\
B &= \left(\Pi - \frac{c}{\lambda_L}\right) e^{-\frac{\lambda_L}{\phi} u_{s,2}}, \\
C &= \left(\Pi - \frac{c}{\lambda_H}\right) e^{-\frac{\lambda_H}{\phi}\left(u_{s,1} + \frac{\phi}{\lambda_S}\right)}.
\end{aligned}$$

Then, from the definition of $V_L^{gn}(u_L|u_{s,2}, u_{s,1})$, we have

$$\begin{aligned}
A &= \frac{\lambda_S}{\phi + \lambda_S \Delta}\left(V_H\left(u_{s,1} + \frac{\phi}{\lambda_S}\right) - V_L^g(u_{s,2}) - \frac{c}{\lambda_S}\right) \\
&= \frac{\lambda_S}{\phi + \lambda_S \Delta}\left[\left(\frac{1}{\lambda_L} - \frac{1}{\lambda_S} - \frac{1}{\lambda_H}\right) c + (B - C)\right] - 1
\end{aligned} \tag{C.1}$$

The smooth pasting condition $V_L^{g\prime}(u_{s,2}) = V_L^{gn\prime}(u_{s,2}|u_{s,2}, u_{s,1})$ gives

$$A = \frac{\lambda_L}{\phi} B - 1. \tag{C.2}$$

The super contact condition $V_L^{gn\prime\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{gns\prime\prime}(u_{s,1}|u_{s,2}, u_{s,1})$ is equivalent to $V_L^{gns\prime\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = 0$.[22] By differentiating the HJB equation, we can derive that

$$V_L^{gns\prime\prime}(u_L|u_{s,2}, u_{s,1}) = \lambda_S\left(V_H'\left(u_L + \frac{\phi}{\lambda_S}\right) - V_L^{gns\prime}(u_L|u_{s,2}, u_{s,1})\right).$$

Therefore, the super contact condition is equivalent to

$$A = V_H'\left(u_L + \frac{\phi}{\lambda_S}\right) = \frac{\lambda_H}{\phi} C - 1. \tag{C.3}$$

From (C.2) and (C.3), we have $\lambda_L B = \lambda_H C$, equivalently,

$$\log\left(\frac{\lambda_H \Pi - c}{\lambda_L \Pi - c}\right) - \frac{\lambda_H}{\lambda_S} = \left(\frac{\lambda_H - \lambda_L}{\phi}\right) u_{s,2} + \frac{\lambda_H}{\phi}\Delta. \tag{C.4}$$

[22]Note that the smooth pasting condition $V_L^{gn\prime}(u_{s,1}|u_{s,2}, u_{s,1}) = V_L^{gns\prime}(u_{s,1}|u_{s,2}, u_{s,1})$ is equivalent to (C.1), thus we need to consider the super contact condition to identify the thresholds.

By plugging $\lambda_L B = \lambda_H C$ and (C.2) into (C.1), we can also derive that

$$\left(\frac{\lambda_L \Pi - c}{\phi}\right) e^{-\frac{\lambda_L}{\phi} u_{s,2}} \Delta = \left(\frac{1}{\lambda_L} - \frac{1}{\lambda_S} - \frac{1}{\lambda_H}\right) \left(\left(\lambda_L \Pi - c\right) e^{-\frac{\lambda_L}{\phi} u_{s,2}} + c\right). \tag{C.5}$$

By solving (C.4) and (C.5), we can pin down $u_{s,2}$ and Δ, as well as $u_{s,1} = u_{s,2} + \Delta$.