Essays on Information and Dynamic Incentives

by

Yonggyun Kim

Department of Economics
Duke University

Date: ______________________

Approved:

____
Curtis Taylor, Co-Supervisor

____
Fei Li, Co-Supervisor

____
Arjada Bardhi

____
Todd Sarver

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2023

Abstract

Essays on Information and Dynamic Incentives

by

Yonggyun Kim

Department of Economics
Duke University

Date: ______________________
Approved:

____
Curtis Taylor, Co-Supervisor

____
Fei Li, Co-Supervisor

____
Arjada Bardhi

____
Todd Sarver

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2023

Abstract

This dissertation studies information and dynamic incentives. I analyze the value of information in decision problems with multidimensional action spaces in Chapter 1, developing a condition called "monotone quasi-garbling" that involves adding reversely monotone noise to an information structure. This condition is necessary and sufficient for obtaining a higher expected payoff and is applied to problems in nonlinear monopoly pricing and optimal insurance, refining the previous garbling condition by Blackwell (1951, 1953).

In Chapter 2, I examine a dynamic principal-agent problem involving two project completion routes: directly attacking it or splitting it into two subprojects. When the project is split, the principal can better monitor the agent by verifying the completion of the first subproject, but the inflexible nature of this approach may generate inefficiencies. The optimal contract is determined by the interplay of monitoring, efficiency, and the endogenous deadline.

In Chapter 3, I collaborate with Francisco Poggi to explore firms' incentives to conceal intermediate research discoveries in innovation races. We consider an innovation game where two firms allocate resources between two R&D paths towards a final innovation, and fully characterize equilibrium behavior in cases of public and private research progress information. Additionally, we find that firms may conceal discoveries in private information settings, which can result in inefficiently slower innovation rates.

Contents

List of Figures

Acknowledgements

I am deeply grateful to my co-advisors, Curtis Taylor and Fei Li, for their invaluable support, guidance, and mentorship throughout my PhD journey. Their combined expertise, insightful feedback, and consistent encouragement have been integral to the successful completion of my dissertation. I am especially thankful for the fact that they were not just my advisors, but also my friends during this challenging journey.

I would also like to extend my heartfelt thanks to the members of my dissertations committee, Arjada Bardhi and Todd Sarver, for their thoughtful critiques and suggestions that helped me to refine and improve this dissertation. I also want to express my gratitude to the other faculty members in the microeconomic theory group in Duke and UNC, including Attila Ambrus, David McAdams, Philipp Sadowski, Huseyin Yildirim, Jim Anton, Pino Lopomo, Gary Biglaiser, Peter Norman, and Can Tian. I would also like to acknowledge Francisco Poggi, the coauthor of Chapter 3 of this dissertation, which served as my job market paper. It was fortunate to have the opportunity to collaborate with him, and I have learned a lot from this experience.

Finally, I would like to express my deepest gratitude to my wife, Shinah, and my son, Jayden. Their unwavering love, support, and understanding have been my source of strength throughout this challenging journey. I am grateful for their patience, encouragement, and sacrifices, which have allowed me to pursue my academic goals.

1

Introduction

This dissertation consists of three essays on information and dynamic incentives. In Chapter 2, I study the value of information in monotone decision problems where the action spaces are potentially multidimensional. As a criterion for comparing information structures, I develop a condition called *monotone quasi-garbling* meaning that an information structure is obtained by adding *reversely monotone noise* (more likely to return a higher signal in a lower state and a lower signal in a higher state) to another. It is shown that monotone quasi-garbling is a necessary and sufficient condition for decision makers to get a higher ex-ante expected payoff. Under the monotone likelihood ratio property, this new criterion is equivalent to the accuracy condition by Lehmann (1988) and refines the garbling condition by Blackwell (1951, 1953). To illustrate, I apply the result to problems in nonlinear monopoly pricing and optimal insurance.

In Chapter 3, I study a dynamic principal-agent problem where there are two routes of completing a project: directly attacking it or splitting it into two subprojects. When the project is split, the principal can better monitor the agent by verifying the completion of the first subproject. However, the inflexible nature of this approach may generate inefficiencies. To mitigate moral hazard, the principal needs to commit to a deadline, which also affects her choice of project management strategy. The optimal contract is determined by the interplay of these three factors: monitoring, efficiency, and an endogenous deadline.

Chapter 4, which is a joint work with Francisco Poggi, investigates firms' incentives to conceal intermediate research discoveries in innovation races. To study this, we introduce an innovation game where two racing firms dynamically allocate their resources between two distinct research and development (R&D) paths towards a final innovation: (i) developing it with the currently available but slower technology; (ii) conducting research to discover a faster new technology for developing it. We fully characterize the equilibrium behavior of the firms in the cases where their research progress is public and private information. Then, we extend the private information setting by allowing firms to conceal or license their intermediate discoveries. We show that when the reward of winning the race is high enough, firms would conceal their interim discoveries, which inefficiently retards the pace of innovation.

2

Comparing Information in General Monotone Decision Problems

2.1 Introduction

Consider a pair of information structures that provide signals about uncertain states. When can we say that one is superior to the other? This fundamental question has numerous economic applications including investment, monopoly pricing, and auctions. A common feature in such settings is that the decision maker would like to take a higher action when a higher signal is realized, that is, the decision problem is often *monotone*.

The classical way of comparing information structures is to use the *garbling condition* developed by Blackwell (1951, 1953). By this criterion an information structure (G) is worse than another (F) if G can be obtained from F by adding some noise—in other words, G is a garbling of F. Intuitively, the added noise reduces the value of information. Indeed, Blackwell's condition implies that for *any* preferences that satisfy the von Neumann-Morgenstern axioms, the expected payoff under F is higher than under G. This is a powerful result because once a pair of information structures are ranked by the garbling order, the rank is preserved in every decision problem. Unfortunately, it is hard to satisfy the garbling order, so Blackwell's criterion has limited applicability.

If we restrict attention to monotone decision problems, it is sometimes possible to compare information structures that are unrankable by Blackwell's condition. One well-known criterion for doing so is the *accuracy condition* by Lehmann (1988). Lehmann considers a specific class of monotone decision problems and shows that his criterion refines Blackwell's garbling condition. Although the accuracy condition has been applied widely in economic settings, its precise meaning remains underexplored. Specifically, the accuracy criterion does not deliver as clear an economic interpretation as Blackwell's garbling condition does. This leads to the first main question of this paper: can we understand Lehmann's condition by using the garbling notion?

To answer this question, I introduce a novel concept called *monotone quasi-garbling* by modifying Blackwell's garbling order. Specifically, I relax the assumption of the garbling condition that the added noise is independent of the state. I define the monotone quasi-garbling order by (i) allowing noise to be state-dependent; (ii) but restricting noise to be *reversely monotone* in the sense that it is more likely to return a higher signal in a lower state and a lower signal in a higher state. In other words, G is a monotone quasi-garbling of F if G can be obtained from F by adding reversely monotone noise. I show that if the monotone likelihood ratio property (MLRP) holds, then Lehmann's order and the monotone quasi-garbling order are equivalent (Theorem 2.4.1). This result provides a fresh view for interpreting Lehmann's condition by using the noise notion: under MLRP, F is more Lehmann accurate than G if and only if G is generated by adding reversely monotone noise to F.

The second main question of this paper is whether we can extend Lehmann's analysis to general classes of monotone decision problems. In the seminal paper on monotone comparative statics, Milgrom and Shannon (1994) posit a partial order on a multidimensional action set to study the direction of change of the optimal actions in response to exogenous changes in parameter values. Nevertheless, in most extant work studying information in monotone decision problems, the common assumption is that the action set is unidimensional and the order of actions is simply inherited from the real line (Quah and Strulovici, 2009; Chi, 2015; Athey and Levin, 2017; Li and Zhou, 2020). Because of this restriction, we cannot directly apply these earlier results to problems involving multidimensional actions such as an investment decision with Arrow-Debreu securities or a nonlinear monopoly pricing mechanism involving a menu of tariffs and quantities. I show that the introduction of the monotone quasi-garbling order helps extend information comparisons to general monotone decision problems exhibiting multidimensional actions in applications such as these.

As a first step towards this characterization, I formally define 'general monotone decision problems' where the action space can be multidimensional. Since there is no generic order in a multidimensional set, the action set needs to have a partial order to identify which actions are higher or lower. To guarantee that this order is sensible, I impose a condition called the dominated decreasing decision rule (DDDR). This condition means that for any state-contingent decision rule decreasing in states with respect to the given partial order, there exists an action dominating the decision rule independent of states. It is shown that this condition is even weaker than the interval dominance order condition by Quah and Strulovici (2007, 2009) (Section 2.5.1 and Appendix A.2) and can be satisfied in some economic applications (Section 2.6 and Appendix A.3). In addition, given an information structure, to say that a decision problem is monotone, an optimal action under a higher signal realization is higher in the partial order. If this holds for any pair of signal realizations, I say that the monotone comparative statics (MCS) condition is satisfied. When a partial order on the action set satisfies the DDDR condition and an information structure satisfies the

MCS condition, I then say the problem is in the class of general monotone decision problems.[1]

One of the key results of this paper is that if G is a monotone quasi-garbling of F, for any general monotone decision problem, the decision maker obtains a higher ex-ante expected payoff under F than under G (Theorem 2.5.1). That is, the monotone quasi-garbling order is a sufficient condition for informativeness on general monotone decision problems. Intuitively, under monotone decision problems, the information structure G is degraded by adding reversely monotone noise to F, thus, the expected payoff under G will be less than under F. In addition, when a class of decision problems includes a basic class of monotone decision problems, I also show that monotone quasi-garbling can serve as a necessary condition for informativeness on the class of decision problems (Theorem 2.5.2).

I apply this result to analyze the value of information in monopoly pricing with second-degree price discrimination as in Maskin and Riley (1984). At the beginning of the game, the seller chooses between two sources of information about the buyer's type. Unlike a standard monopoly pricing problem where the seller makes a unidimensional choice of price or quantity, in a nonlinear pricing problem, the seller needs to provide a menu of tariffs and quantities–which is multidimensional. Although this problem is usually modeled as a two-player game, it can be cast as a multidimensional decision problem for a seller constrained by incentive compatibility and individual rationality by regarding the buyer's type as a state. Moreover, with fairly mild assumptions, I show that this application satisfies the DDDR condition and the MCS condition, thus the main result of this paper is also applicable in this setting. In Appendix A.3, I also provide another application on optimal insurance with Arrow-Debreu securities.

[1] My analysis differs from Milgrom and Shannon (1994) in that they identify conditions of utility functions that make decision problems monotone whereas I assume that monotone comparative statics are already present and focus on the comparison of information structures.

The rest of the paper is organized as follows. I review the literature in Section 2.2. Section 2.3 sets up the preliminary notions. In Section 2.4, I introduce the monotone quasi-garbling order and compare it to other orderings such as Blackwell's garbling condition and Lehmann's accuracy condition. In Section 2.5, I define general monotone decision problems and show that the monotone quasi-garbling order is a necessary and sufficient condition for informativeness. I apply this result to a nonlinear pricing problem in Section 2.6. Proofs and examples omitted from the text are in the Appendix.

2.2 Related Literature

The comparison of information structures has been applied in numerous economic situations: Bayesian games (Gossner, 2000; Mekonnen and Leal Vizcaíno, 2021), investment decisions (Cabrales et al., 2013), auctions (Persico, 2000; Ganuza and Penalva, 2010), matching markets (Roesler, 2015), principal agent models with moral hazard (Kim, 1995), competitive markets with adverse selection (Levin, 2001), strategic sampling (Di Tillio et al., 2021), and monopoly pricing (Athey and Levin, 2017; Ottaviani and Prat, 2001).

Since Blackwell (1951, 1953) introduced a criterion to compare information structures in general decision problems, subsequent studies have refined this criterion by restricting it to monotone decision problems. In a seminal paper, Lehmann (1988) mentions location experiments to point out the limit of Blackwell's condition. Then, he restricts decision problems to have the MLRP for information structures and monotone utility introduced by Karlin and Rubin (1956) and establishes a theorem showing that his accuracy condition is a necessary and sufficient condition for informativeness. Persico (2000) utilizes this accuracy condition for decision problems with single crossing utility.[2] Quah and Strulovici (2009) introduce the interval dominance order property which is weaker than the single crossing property and shows that Lehmann's condition can also be utilized as a criterion for decision problems with this property. These studies are well summarized by a unified framework provided by Chi (2015). He establishes an equivalence of accuracy, informativeness, and posterior dispersion under decision problems with supermodular, single crossing, and interval dominance order preferences. Last but not least, Li and Zhou (2020) show that Lehmann's ordering is robust under monotone decision problems with ambiguity-averse decision makers. However, as mentioned in the introduction, all of these studies assume that the action set is unidimensional whereas this paper allows the decision maker to choose a multidimensional action.[3]

To my knowledge, Jewitt (2007) is the only paper that links Blackwell's condition and Lehmann's condition. He shows that under the MLRP condition, Lehmann's condition is equivalent to Blackwell on Dichotomies, which means that for any pair of states, an information structure restricted to those states is more Blackwell sufficient than the other. Although this characterization helps us to understand better the relationship between the conditions by Blackwell and Lehmann, it still does not provide a 'garbling' interpretation of Lehmann's condition.

A property of the monotone quasi-garbling criterion is that the ordering is independent of prior beliefs. It is also true for the conditions of Blackwell and Lehmann. On the other hand, several recent studies exploit prior beliefs to refine Lehmann's condition. Athey and Levin (2017) restrict utility functions to satisfy certain conditions such as supermodularity and fix a prior belief, then introduce a criterion called monotone information order. Ganuza and Penalva (2010) apply integral and super-

[2] Jewitt (2007) compares Karlin Rubin monotonicity and the single crossing property.

[3] Quah and Strulovici (2007) extend their comparative statics results in Quah and Strulovici (2009) to the multidimensional action case, but they do not provide an information comparison result. In Section 2.5.1, I discuss how their extended results can be applied in this paper's framework.

modular orders, which are defined on probability measures on expectations of states, to auction problems. Note that the expectation of states largely depends on the prior belief. Last, Cabrales et al. (2013) restricts attention to an investment decision problem and shows that an entropy ordering, which is prior dependent, gives a complete informative ordering. Since these orderings utilize prior beliefs as additional sources for information comparisons, they perform better than prior independent orderings if the prior is known. Nevertheless, it is still important to establish prior independent orderings because it is essential for applications with unknown or multiple prior beliefs. For example, in the presence of ambiguity, Li and Zhou (2016, 2020) emphasize the role of prior-free criteria such as the conditions of Blackwell and Lehmann.

2.3 Preliminaries

Let $\Omega \equiv [\underline{\omega}, \overline{\omega}] \subset \mathbb{R}$ be the set of states of nature. Denote $\omega \in \Omega$ as a generic state. The decision maker, hereafter the DM, has a prior belief $\Lambda \in \Delta(\Omega)$, where $\Delta(Z)$ is the set of cumulative distribution functions on any compact set $Z \subset \mathbb{R}$.[4] An information structure is composed of (i) a closed interval of signals $X \equiv [\underline{x}, \overline{x}] \subset \mathbb{R}$; and (ii) a collection of (cumulative) distribution functions $\{F(\cdot|\omega)\}_{\omega \in \Omega}$ with $F(\cdot|\omega) \in \Delta(X)$ for all $\omega \in \Omega$. Simply denote this information structure as F. For another information structure G, let the set of signals be $Y \equiv [\underline{y}, \overline{y}] \subset \mathbb{R}$, and the collection of signal distributions be $\{G(\cdot|\omega)\}_{\omega \in \Omega}$.

Without loss of generality, we assume that Λ is continuous in $\omega \in \Omega$, and F and G are continuous in $x \in X$ and $y \in Y$ for all $\omega \in \Omega$, i.e., there is no mass point. To achieve this, we can use the construction introduced in Theorem 5.1 of Lehmann (1988): if $H \in \Delta([\underline{z}, \overline{z}])$ is discontinuous at z_0, i.e., $p \equiv H(z_0) - H(z_0^-) > 0$, we can

[4] By Theorem 1.2.1 of Durrett (2019), for any $H \in \Delta(Z)$, the following statements hold: (i) $H(\cdot)$ is (weakly) increasing; (ii) $H(\cdot)$ is right continuous, i.e., $\lim_{y \downarrow z} H(y) = H(z)$; and (iii) $H(\overline{z}) = 1$ where $\overline{z} \equiv \max Z$.

define another distribution $\tilde{H} \in \Delta([\underline{z}, \overline{z}+1])$ such that

$$\tilde{H}(z) = \begin{cases} H(z), & \text{if } z < z_0, \\ H(z_0^-) + p \cdot (z - z_0), & \text{if } z \in [z_0, z_0+1], \\ H(z-1), & \text{if } z > z_0 + 1. \end{cases}$$

Then, $\tilde{H}$ is continuous on $[z_0, z_0+1]$, and from the DM's perspective, H and $\tilde{H}$ are statistically equivalent. From the continuity, there exist the probability distribution functions corresponding to Λ, $F(\cdot|\omega)$, and $G(\cdot|\omega)$: i.e., there exist λ, $f(\cdot|\omega)$ and $g(\cdot|\omega)$ such that $\Lambda(\omega) = \int_{\underline{\omega}}^{\omega} \lambda(z)dz$, $F(x|\omega) = \int_{\underline{x}}^{x} f(z|\omega)dz$, and $G(y|\omega) = \int_{\underline{y}}^{y} g(z|\omega)dz$.[5]

When a prior belief is Λ and an information structure is F, the marginal signal distribution, denoted by $F_\Lambda \in \Delta(X)$, is defined as $F_\Lambda(x) = \int_{\underline{\omega}}^{\overline{\omega}} F(x|\omega)\lambda(\omega)d\omega$ for all $x \in X$. Suppose that a signal x is realized, i.e., $f_\Lambda(x) > 0$ where f_Λ is the probability distribution function corresponding to F_Λ. Then, the posterior belief, $\Lambda_F^x \in \Delta(\Omega)$, is defined as $\Lambda_F^x(\omega) = f(x|\omega) \cdot \Lambda(\omega)/f_\Lambda(x)$, or equivalently $d\Lambda_F^x(\omega)dF_\Lambda(x) = dF(x|\omega)d\Lambda(\omega)$.

Let the set of actions or feasible decisions be $A \in \mathcal{A}$ where $\mathcal{A}$ is a collection of multidimensional real-valued compact sets, i.e., A is closed, bounded, and $A \subset \mathbb{R}^n$ for some $n \in \mathbb{N}$. Note that I allow an action to be multidimensional. Assume that the DM's payoff function $u : A \times \Omega \to \mathbb{R}$ is continuous in $a \in A$. Let $\overline{\mathcal{U}}_A$ denote the set of all continuous real-valued payoff functions on $A \times \Omega$.

The decision making process is given as follows.

1. The DM chooses between information structures F and G.

2. From the chosen info structure (say F), the DM receives a signal $x \in X$.

[5] By using the similar construction, the case with discrete state and signals can be transformed to the continuous state and signal case. For example, when a discrete set of states $\Omega = \{\omega_1, \cdots, \omega_N\}$ with $0 < \omega_1 < \cdots < \omega_N$ and the probability mass function $\lambda : \Omega \to [0,1]$ with $\sum_{n=1}^{N} \lambda(\omega_n) = 1$ are given, we an construct the corresponding continuous state space and the probability distribution function as follows: if $\omega \in (\omega_{n-1}, \omega_n]$ for some $n \geqslant 2$, $\lambda(\omega) = \lambda(\omega_n)/(\omega_n - \omega_{n-1})$, or if $\omega \in [0, \omega_1]$, $\lambda(\omega) = \lambda(\omega_1)/\omega_1$.

3. The DM updates a belief on the state to $\Lambda_F^x(\omega)$ and chooses $a \in A$.

4. The state ω is revealed and the payoff $u(a, \omega)$ is realized.

If the payoff function is u, when a signal is realized in the third stage, the DM chooses $a \in A$ to maximize the interim expected payoff function $U : A \times X$ defined as follows:

$$U(a; x) \equiv \int_\Omega u\,(a, \omega)\, d\Lambda_F^x(\omega). \tag{2.3.1}$$

Let $A^*_{F,\Lambda}(x) \subseteq A$ denote the set of solutions for the interim maximization problem, i.e., $A^*_{F,\Lambda}(x) \equiv \arg\max_{a \in A} U(a, x)$. Note that $A^*_{F,\Lambda}(x)$ is nonempty since u is continuous and A is compact. The expected utility of the DM given the information structure F at the first stage is that for any $a^*(x) \in A^*_{F,\Lambda}(x)$,

$$\begin{aligned} V(F; u, \Lambda) &= \int_X U(a^*(x); x) dF_\Lambda(x) \\ &= \int_X \int_\Omega u\,(a^*(x), \omega)\, d\Lambda_F^x(\omega) dF_\Lambda(x) \\ &= \int_\Omega \int_X u\,(a^*(x), \omega)\, dF(x|\omega) d\Lambda(\omega). \end{aligned} \tag{2.3.2}$$

Then, in the first stage, the DM chooses an information structure by comparing $V(F; u, \Lambda)$ and $V(G; u, \Lambda)$. Based on this comparison, the informativeness for decision problems in $\mathcal{U} \subseteq \bigcup_{A \in \mathcal{A}} \overline{\mathcal{U}}_A \equiv \overline{\mathcal{U}}$ is defined as follows.

Definition 2.3.1 (Informativeness on $\mathcal{U}$)**.** An information structure F is *more informative* than another information structure G on $\mathcal{U} \subseteq \overline{\mathcal{U}}$ if and only if for all priors $\Lambda \in \Delta(\Omega)$ and all $u \in \mathcal{U}$, $V(F; u, \Lambda) \geqslant V(G; u, \Lambda)$ holds.

Remark 1. Though I focus on the case where the state space Ω is a continuum, the main analysis can also be applied to the discrete state case by substituting integration for summation. For the applications in Section 2.5.1 and 2.6, I assume that the state space is discrete to facilitate the analysis.

2.4 Information Ranking Criteria

In this section, I introduce a novel condition called the monotone quasi-garbling order. Then, I explore how it is related to classical information ranking criteria such as Blackwell's garbling condition and Lehmann's accuracy condition. Since the monotone quasi-garbling order is motivated by Blackwell's garbling notion, I begin the section by reviewing the garbling condition.

Definition 2.4.1 (Garbling Condition by Blackwell (1951, 1953))**.** An information structure G is said to be a *garbling* of another information structure F, denoted $F \succeq_B G$, if there exists a function $\Gamma : X \to \Delta(Y)$ such that $G(y|\omega) = \int_X \Gamma(y|x) dF(x|\omega)$ for all $\omega \in \Omega$.

Blackwell shows that an information structure F is more informative than G for all decision problems if and only if G is a garbling of F. In other words, $V(F; u, \Lambda) \geqslant V(G; u, \Lambda)$ for any set of actions A, payoff function $u \in \overline{\mathcal{U}}_A$ and prior belief $\Lambda \in \Delta(\Omega)$ if and only if $F \succeq_B G$. Blackwell's garbling condition is widely used by economists not only because it has a powerful equivalence condition but also because it has a nice interpretation. We can imagine that a signal y from information G is constructed by adding *noise* Γ to signals in X from information F and the noise deteriorates the quality of information so that F gives better information than G.

Even though the garbling condition is useful, this criterion is quite restrictive because it requires preserving the order for 'all' decision problems. If we restrict attention to 'monotone' decision problems, there is room to improve Blackwell's ordering. I now introduce the monotone quasi-garbling order by adding a monotone structure to the garbling condition.

Definition 2.4.2 (Monotone Quasi-Garbling)**.** An information structure G is said to be a *monotone quasi-garbling* of another information structure F, denoted $F \succeq_{MQG}$

G, if there exists a function $\Gamma : X \times \Omega \to \Delta(Y)$ such that

1. $G(y|\omega) = \int_X \Gamma(y|x,\omega)dF(x|\omega)$ for all $y \in Y$, $x \in X$ and $\omega \in \Omega$,

2. $\Gamma(\cdot|x,\omega) \geqslant_{FOSD} \Gamma(\cdot|x,\omega')$ for all $\omega' > \omega$ and $x \in X$, i.e., $\Gamma(y|x,\omega) \leqslant \Gamma(y|x,\omega')$ for all $y \in Y$.

We can easily observe that the garbling order implies the monotone quasi-garbling order. If G is a garbling of F via a function $\hat{\Gamma} : X \to \Delta(Y)$, we can see that G is also a monotone quasi-garbling of F by setting $\Gamma(y|x,\omega) = \hat{\Gamma}(y|x)$ for all $\omega \in \Omega$, because a probability distribution is first order stochastic dominant over itself. However, the converse is not true. In Appendix A.1.2, I provide an example of a pair of information structure (F, G) such that $F \geq_{MQG} G$ but $F \not\geq_B G$. This observation is formally stated in the following proposition.

Proposition 2.4.1. *For any pair of information structures (F, G), $F \geq_B G$ implies $F \geq_{MQG} G$, but $F \geq_{MQG} G$ does not imply $F \geq_B G$.*

The key difference between these two criteria is that the noise may depend on the state ω under the monotone quasi-garbling order. Note that if there were no restriction on Γ, any information structure G could be generated by adding state-contingent noise to F and it would be a meaningless criterion.[6] The second condition restricts state-contingent noise to be *reversely FOSD-ordered*. This means that the noise is more likely to return a higher y for a lower state and lower y for a higher state, thus, it can be interpreted as a *reversely monotone* noise. We can guess that the reversely monotone noise would deteriorate the quality of information under monotone decision problems.

[6] For example, when there is no restriction on Γ, we can use $\Gamma(y|x,\omega) = G(y|\omega)$ for all $x \in X$, $y \in Y$ and $\omega \in \Omega$ to satisfy the first condition.

Next, I review the accuracy condition by Lehmann (1988), which is widely used in monotone decision problems. Then, I explore how it is related to the monotone quasi-garbling order.

Definition 2.4.3 (Accuracy Condition by Lehmann (1988))**.** An information structure F is said to be *more Lehmann accurate* than another information structure G, denoted $F \succeq_L G$, if for all $y \in Y$, a function $\Phi : \Omega \times Y \to X$ defined by $\Phi(\omega; y) \equiv \sup\{x \mid F(x|\omega) \leqslant G(y|\omega)\}$ is increasing in ω.[7]

An assumption that is usually paired with Lehmann's condition is the monotone likelihood ratio property.

Definition 2.4.4 (Monotone Likelihood Ratio Property)**.** An information structure F is said to satisfy the monotone likelihood ratio property (MLRP) if and only if $f(x_0|\omega_0)f(x_1|\omega_1) \geqslant f(x_1|\omega_0)f(x_0|\omega_1)$ holds for all $x_1 > x_0$ and $\omega_1 > \omega_0$.

First, I show that Lehmann's accurate condition implies the monotone quasi-garbling order, i.e., $F \succeq_L G$ implies $F \succeq_{MQG} G$. Observe that $\Phi(\omega; y)$ is increasing in y since $G(y|\omega)$ is increasing in y for any $\omega \in \Omega$. Construct a function $\Gamma : X \times \Omega \to \Delta(Y)$ as follows:

$$\Gamma(y|x, \omega) \equiv \begin{cases} 1, & \text{if } x \leqslant \Phi(\omega; y), \\ 0, & \text{if } x > \Phi(\omega; y). \end{cases}$$

Then, we can easily check that $\Gamma(y|x, \omega)$ is well defined.[8] In addition, since $\Phi(\omega; y) \leqslant \Phi(\omega'; y)$ for all $\omega' > \omega$ and $y \in Y$, we have $\Gamma(y|x, \omega) \leqslant \Gamma(y|x, \omega')$ for all $x \in X$, i.e., Γ is reversely monotone.

[7] In Lehmann (1988), he implicitly assumes that $F(x|\omega)$ is continuous and strictly increasing in x, thus $F^{-1}(\cdot|\omega)$ is properly defined and set $\Phi(\omega; y) \equiv F^{-1}(G(y|\omega)|\omega)$. Under the continuity and strict increasingness assumptions on $F(x|\omega)$, we have $F(\Phi(\omega; y)|\omega) = G(y|\omega)$, thus, Definition 2.4.3 coincides with Lehmann's original condition.

[8] First, $\Gamma(y|x, \omega)$ is nondecreasing in y because $\Phi(\omega; y)$ is nondecreasing in y. Next, $\Gamma(y|x, \omega)$ is right-continuous in y because (i) if $\Phi(\omega; y) \geqslant x$, since $\Phi(\omega; y)$ is nondecreasing in y and $\Phi(\omega; y') \geqslant x$ for all $y' > y$, thus $\lim_{\tilde{y} \downarrow y} \Gamma(\tilde{y}|x, \omega) = 1 = \Gamma(y|x, \omega)$; (ii) if $\Phi(\omega; y) < x$, by the definition of Φ, we have $G(y|\omega) < x$, then since $G(y|\omega)$ is continuous in y, there exist $\epsilon > 0$ and $\delta > 0$ such that $G(y + \epsilon|\omega) < x - \delta$ and $\Phi(\omega; y + \epsilon) < x$, thus, we have $\lim_{\tilde{y} \downarrow y} \Gamma(\tilde{y}|x, \omega) = 0 = \Gamma(y|x, \omega)$.

Last, observe that for all $y \in Y$ and $\omega \in \Omega$,

$$G(y|\omega) = F(\Phi(\omega; y)|\omega) = \int_{\underline{x}}^{\Phi(\omega;y)} dF(x|\omega) = \int_X \Gamma(y|x,\omega) dF(x|\omega).$$

Therefore, G is a monotone quasi-garbling of F. The following proposition formally states this result.

Proposition 2.4.2. *Suppose that $F \succeq_L G$. Then, $F \succeq_{MQG} G$.*

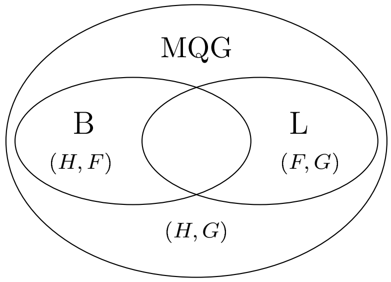
Next, I explore whether $F \succeq_{MQG} G$ implies $F \succeq_L G$. The following proposition shows that the monotone quasi-garbling order implies Lehmann's accuracy condition when F satisfies MLRP. The proof is relegated to Appendix A.1.1.

Proposition 2.4.3. *Suppose that an information structure F satisfies the MLRP. Then, $F \succeq_{MQG} G$ implies $F \succeq_L G$.*

Based on the above results, we can establish the equivalence between the monotone quasi-garbling order and Lehmann's order under the MLRP condition.

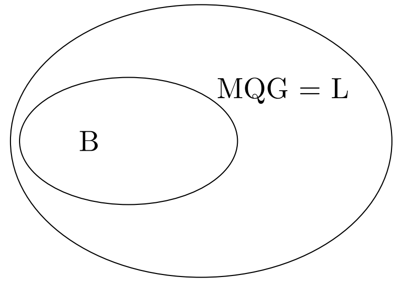
Theorem 2.4.1. *Suppose that an information structure F satisfies the MLRP. Then, $F \succeq_{MQG} G$ and $F \succeq_L G$ are equivalent.*

Last, I show that the MLRP plays a crucial role in the equivalence between the monotone quasi-garbling order and Lehmann's accuracy condition. In Appendix A.1.2, in addition to the information structures F and G which are used in Proposition 2.4.1, I introduce an information structure H which does not satisfy the MLRP. In this example, I show that (i) $H \succeq_B F$ but $H \not\succeq_L F$; and (ii) $H \succeq_{MQG} G$ but $H \not\succeq_B G$ and $H \not\succeq_L G$. The first result implies that when the MLRP is not assumed, Lehmann's accuracy condition is no longer a sufficient condition for Blackwell's garbling condition. The second result means that there exists a pair of information structures that can be comparable by the monotone quasi-garbling order but neither



(a)

(a) Without the MLRP



(b)

(b) With the MLRP

FIGURE 2.1: The relationship among criteria

by Blackwell's garbling condition nor by Lehmann's accuracy condition. Therefore, the monotone quasi-garbling order generally permits more comparisons than the conditions by Blackwell and Lehmann. The following proposition formally states these arguments.

Proposition 2.4.4. *Without the assumption of the MLRP on F, $F \succeq_B G$ does not imply $F \succeq_D G$, and $F \succeq_{MQG} G$ does not imply $F \succeq_L G$.*

The results in this section are summarized in Figure 2.1. Figure 2.1a illustrates the relationship among Blackwell's garbling condition, Lehmann's accuracy condition, and the monotone quasi-garbling condition when the MLRP is not assumed on the supposedly higher information structure. On the other hand, if the MLRP is assumed on the supposedly better information structure, the Lehmann order and the monotone quasi-garbling order are equivalent and refine the Blackwell garbling condition as shown in Figure 2.1b.

2.5 General Monotone Decision Problems

In this section, I show that the monotone quasi-garbling order can serve as a criterion for comparing information structures under monotone decision problems. Specifically, I allow the set of feasible decisions, A, to be multidimensional. Unlike the unidimensional action space case, there is no generic order in this setup. Hence, to establish a monotone decision problem under this general action space, an order on the set of actions needs to be properly defined. Since we will only use this order to compare the optimal actions, it does not need to be a complete order over A. Rather, a partial order, which can compare the potential optimal actions, would be enough. To say that a decision problem is monotone, (i) this partial order needs to be sensible; and (ii) an optimal action under a higher signal realization needs to be higher in this order.

Definition 2.5.1 (Dominated Decreasing Decision Rule Condition)**.** A payoff function $u : A \times \Omega \to \mathbb{R}$ satisfies the *dominated decreasing decision rule (DDDR) condition* with respect to a partial order $\geqslant_A$ of the set of feasible decisions A if for any (weakly) decreasing (in terms of $\geqslant_A$) state-contingent decision rule $h : \Omega \to A$,[9] there exists $\hat{a} \in A$ such that

$$u(\hat{a}, \omega) \geqslant u(h(\omega), \omega), \quad \forall \omega \in \Omega. \tag{2.5.1}$$

This condition is about whether the partial order $\geqslant_A$ is sensible or not. Loosely speaking, under a given payoff function, a highly ranked action needs to be better for a higher state and worse for a lower state and vice versa for a lowly ranked action. According to this interpretation, a state-contingent decision rule, which is decreasing in states, should not be an optimal way of establishing a decision rule. This is because a highly ranked action (which is worse for a lower state) corresponds to a lower state and a lowly ranked action (which is worse for a higher state) corresponds to a higher state. The DDDR condition implies that any decreasing state-contingent decision rule is dominated by a constant decision rule. In this sense, the order $\geqslant_A$ can be considered sensible.

[9] For $\omega' > \omega$ and $h : \Omega \to A$, h is decreasing in terms of $\geqslant_A$ if $h(\omega) \geqslant_A h(\omega')$.

Definition 2.5.2 (Monotone Comparative Statics Condition)**.** An information structure F satisfies the *monotone comparative statics (MCS) condition* with respect to a payoff function $u : A \times \Omega \to \mathbb{R}$ and a partial order $\geqslant_A$, if for any prior belief $\Lambda \in \Delta(\Omega)$, there exists a function $a^* : X \to A$ such that $a^*(x) \in A^*_{F,\Lambda}(x)$ and $a^*(x) \geqslant_A a^*(x')$ for all $x, x' \in X$ with $x \geqslant x'$.

This condition simply implies that there exists an optimal decision rule, which is nondecreasing (in terms of the given partial order) with respect to the signal. I call a class of decision problems $\mathcal{U}$ is *generally monotone* with respect to an information structure F, if for any $u \in \mathcal{U}$, there exists a partial order $\geqslant_A$ such that u satisfies the DDDR condition with respect to $\geqslant_A$ and the information structure F satisfies the MCS condition with respect to u and $\geqslant_A$. I also call a decision problem u is generally monotone with respect to F if $\mathcal{U} = \{u\}$ is generally monotone with respect to F.

The following theorem stipulates that the monotone quasi-garbling order can serve as a sufficient condition for informativeness on general monotone decision problems.

Theorem 2.5.1. *Suppose that $\mathcal{U}$ is generally monotone with respect to an information structure G and G is a monotone quasi-garbling of F. Then, F is more informative than G on $\mathcal{U}$.*

The formal proof is in Section 2.5.2, but I briefly delineate the idea of this result. Consider any increasing decision rule $\psi : Y \to A$, i.e., $\psi(y'') \geqslant_A \psi(y')$ for all $y'' \geqslant_A y'$. When G is a monotone quasi-garbling of F, i.e., G is generated from F by adding a reversely monotone noise, I show that there exists $\tilde{\psi} : (0, 1) \times X \times \Omega \to A$ such that $\tilde{\psi}$ is decreasing in ω and for any $\omega \in \Omega$,

$$\int_Y u(\psi(y), \omega) dG(y|\omega) = \int_0^1 \int_X u(\tilde{\psi}(t; x, \omega), \omega) dF(x|\omega) dt. \tag{2.5.2}$$

Then, by applying the DDDR condition, we can construct a mixed-action decision rules $\phi : X \times (0,1) \to A$ on the information structure F which dominates ψ on G (Proposition 2.5.2). Now assume that u is generally monotone with respect to G, i.e., the optimal decision rule under G, $a^*_G : Y \to A$, is increasing in y. By applying the above result, there exists a mixed-action decision rules dominating a^*_G, but this decision rule is dominated by a^*_F which eventually gives the expected payoff under F. Therefore, the expected payoff under F will be greater than that under G.

Next, to establish the necessary condition for informativeness, I introduce a bare-bones class of decision problems. Consider a binary action set $\underline{A} = \{0,1\}$ and the generic order on $\underline{A}$: $1 > 0$. I also define a class of utility functions $\underline{\mathcal{U}}$ consisting of following functions: for some $0 < \kappa < 1$ and $\hat{\omega} \in \Omega$,

$$u(a,\omega) = \begin{cases} -a \cdot \kappa, & \text{if } \omega < \hat{\omega}. \\ a \cdot (1-\kappa), & \text{if } \omega \geqslant \hat{\omega}.^{10} \end{cases} \tag{2.5.3}$$

The following theorem shows that the monotone quasi-garbling order is a necessary condition for informativeness when the class of decision problems $\mathcal{U}$ includes the bare-bones class of decision problems $\underline{\mathcal{U}}$. The proof is provided in Section 2.5.3.

Theorem 2.5.2. *Suppose that $\underline{\mathcal{U}} \subseteq \mathcal{U}$, $\mathcal{U}$ is generally monotone with respect to information structures F, and F is more informative than G on $\mathcal{U}$. Then, (i) F satisfies the MLRP; and (ii) F is more Lehmann accurate than G, thus, G is a monotone quasi-garbling of F.*

In this theorem, I show that the informativeness implies the monotone quasi-garbling order by showing that an information structure is more Lehmann accurate than the other (recall that $F \geq_L G$ implies $F \geq_{MQG} G$ by Proposition 2.4.2). As an

[10] This choice of utility function is inspired by Chi (2015). He shows that the class of monotone decision problems with certain properties, such as super modular, single crossing, or interval dominance order preferences, includes $\underline{\mathcal{U}}$.

intermediate step for the result, it is also shown that F satisfies the MLRP when all decision problems in the bare-bones class $\underline{\mathcal{U}}$ are generally monotone with respect to F.

Note that I impose different assumptions on the above results: in Theorem 2.5.1, I assume that the decision problems with a supposedly 'worse' information structure are generally monotone, whereas, in Theorem 2.5.2, I assume that the decision problems with a supposedly 'better' information structure are generally monotone and the class of decision problems includes $\underline{\mathcal{U}}$. Therefore, if we assume that the decision problems with both information structures are monotone (and include $\underline{\mathcal{U}}$), we can see that the monotone quasi-garbling order can serve as a necessary and sufficient condition for informativeness. In addition, under these assumptions, both information structures will satisfy the MLRP and the monotone quasi-garbling condition and Lehmann accuracy condition are equivalent.

Corollary 2.5.1. *Suppose that $\underline{\mathcal{U}} \subseteq \mathcal{U}$ and $\mathcal{U}$ is generally monotone with respect to information structures F and G. Then, F is more informative than G on $\mathcal{U}$ if and only if G is a monotone quasi-garbling of F, or equivalently, F is more Lehmann accurate than G.*

Remark 2. Corollary 2.5.1 extends the results of Chi (2015) that Lehmann's accuracy condition can serve as a necessary and sufficient condition for the informativeness in some monotone decision problems. First, he assumes that both information structures have the MLRP, but I derive the MLRP as a result. Next, he restricts attention to the 'unidimensional' class of decision problems with well-known preferences such as supermodular, single crossing, and interval dominance order preferences. On the other hand, Corollary 2.5.1 can be applied to any (potentially 'multidimensional') monotone decision problems including $\underline{\mathcal{U}}$.

2.5.1 General Interval Dominance Order

Before presenting the proofs of the main theorems, I demonstrate how the concept of the general monotone decision problem is connected to the general version of the interval dominance order condition introduced in Quah and Strulovici (2007).

I assume that the state space is discrete as described in Footnote 5: $\Omega = \{\omega_1, \cdots, \omega_N\}$ with $0 < \omega_1 < \cdots < \omega_N$. Suppose that a compact set of action $A \subset \mathbb{R}^n$ has a partial order $\geqslant_A$ with (i) transitivity ($a'' \geqslant_A a'$ & $a' \geqslant_A a \;\Rightarrow a'' \geqslant_A a$); (ii) reflexivity ($a \geqslant_A a$); and (iii) anti-symmetry ($a' \geqslant_A a$ & $a \geqslant_A a' \;\Rightarrow a = a'$). I also assume that the partial order is continuous: for all sequences $\{a_m\}_{m=1}^{\infty}$ and $\{b_m\}_{m=1}^{\infty}$ with $\lim_{m\to\infty} a_m = a$, $\lim_{m\to\infty} b_m = b$, and $a_m \geqslant_A b_m$ for all $m \geqslant 1$, then, $a \geqslant_A b$. Let $[a', a'']$ denote the set $\{a \in A \mid a' \leqslant_A a \leqslant_A a''\}$. Then, if a sequence $\{a_m\}_{m=1}^{\infty}$ in $[a', a'']$ converges to a, by the continuity of the partial order, $a' \leqslant_A a \leqslant_A a''$, i.e., $a \in [a', a'']$. Therefore, $[a', a'']$ is compact since it is a closed subset of A.

Now I briefly review the interval dominance order condition, then show that it is a sufficient condition for the DDDR condition in Proposition 2.5.1.

Definition 2.5.3 (Interval Dominance Order (IDO))**.** Let v' and v'' be two real-valued functions defined on A. We say that v'' dominates v' by the *interval dominance order* (or, for short, v'' I-dominates v') if, for any $a'' >_A a'$, whenever $v'(a'') \geqslant v'(a)$ for all $a \in [a', a'']$,

$$v'(a'') \;\geqslant (>)\; v'(a') \quad \Longrightarrow \quad v''(a'') \;\geqslant (>)\; v''(a'). \tag{2.5.4}$$

In addition, a payoff function $u : A \times \Omega \to \mathbb{R}$ is said to be *IDO-ordered* with respect to $\geqslant_A$ if $u(\cdot; \omega'')$ I-dominates $u(\cdot; \omega')$ for all $\omega'' > \omega'$.

Proposition 2.5.1. *Suppose that the set of actions A has a transitive, reflexive, anti-symmetric, and continuous partial order $\geqslant_A$, a payoff function $u : A \times \Omega \to \mathbb{R}$*

is continuous and IDO-ordered. Then, u satisfies the DDDR condition with respect to $\geqslant_A$.

Proof of Proposition 2.5.1. First, I show that the payoff function u satisfies the DDDR condition with respect to the partial order. Consider a decreasing state-contingent decision rule $h : \Omega \to A$. I inductively define a sequence $\{a_i\}_{i=1}^N$ as follows: $a_N = h(\omega_N)$, and for all $1 \leqslant n \leqslant N-1$,

$$a_n = \underset{a \in [a_{n+1}, h(\omega_n)]}{\arg\max}\; u(a, \omega_n). \tag{2.5.5}$$

Note that a_n exists since u is continuous and $[a_{n+1}, h(\omega_n)]$ is compact.

Observe that $u(a_n, \omega_n) \geqslant u(a, \omega_n)$ for all $a \in [a_{n+1}, a_n]$. Then, by the IDO condition, we have

$$u(a_n, \omega_{n'}) \geqslant u(a_{n+1}, \omega_{n'}) \tag{2.5.6}$$

for all $n' > n$.

Now set $\hat{a} = a_1$. Then, for all $1 \leqslant n \leqslant N$, by (2.5.6), we have $u(\hat{a}, \omega_n) \geqslant u(a_n, \omega_n)$. In addition, by (2.5.5), $u(a_n, \omega_n) \geqslant u(h(\omega_n), \omega_n)$. Therefore, $u(\hat{a}, \omega_n) \geqslant u(h(\omega_n), \omega_n)$ for all $1 \leqslant n \leqslant N$, i.e., u satisfies the DDDR condition with respect to the partial order. □

A natural subsequent question is whether the IDO condition can serve as a necessary condition. In Appendix A.2, I provide a simple (even unidimensional) example of a decision problem such that it is a general monotone decision problem but it does not satisfy the IDO condition. Thus, the IDO condition cannot be a necessary condition for the DDDR condition.

Next, to establish the MCS condition, additional conditions are needed. Following Topkis (1998), let $a' \vee a''$ denote (if exists) the least upper bound of $\{a', a''\}$, that is, $(a' \vee a'') \geqslant_A a'$, $(a' \vee a'') \geqslant_A a''$, and $a \geqslant_A (a' \vee a'')$ for all a with $a \geqslant_A a'$ and $a \geqslant_A a''$.

Likewise, let $a' \wedge a''$ denote the greatest lower bound of $\{a', a''\}$, i.e., $a' \geqslant_A (a' \wedge a'')$, $a'' \geqslant_A (a' \wedge a'')$, and $(a' \wedge a'') \geqslant_A a$ for all a with $a' \geqslant_A a$ and $a'' \geqslant_A a$. A partially ordered set A is called to be a *lattice* if $a' \vee a''$ and $a' \wedge a''$ exist for all $a', a'' \in A$. A payoff function $u : A \times \Omega \to \mathbb{R}$ is *supermodular* in A if for all $a', a'' \in A$ and $\omega \in \Omega$,

$$u(a' \vee a'', \omega) - u(a', \omega) \geqslant u(a'', \omega) - u(a' \wedge a'', \omega).$$

When an information structure F satisfies the MLRP, for any $x'' > x'$, $\Lambda_F^{x''}$ is the monotone likelihood ratio (MLR) shift of $\Lambda_F^{x'}$, that is, $\lambda_F^{x''}(\omega)/\lambda_F^{x'}(\omega)$ is nondecreasing in ω where λ_F^{x} is the probability mass function of Λ_F^{x}. Also note that $A^*_{F,\Lambda}(x) = \arg\max_{a \in A} U(a; x)$ is nonempty since U is continuous in a and A is compact. Under the IDO conditions on u and the MLRP of F, by Proposition 2 of Quah and Strulovici (2007), we have $U(\cdot; x'')$ I-dominates $U(\cdot; x')$. Observe that U is supermodular in A when u is supermodular in A. Then, by Theorem 1 of Quah and Strulovici (2007), we have $A^*_{F,\Lambda}(x'')$ dominates $A^*_{F,\Lambda}(x')$ in the strong set order, i.e., for any $a'' \in A^*_{F,\Lambda}(x'')$ and $a' \in A^*_{F,\Lambda}(x')$, $a'' \vee a' \in A^*_{F,\Lambda}(x'')$ and $a'' \wedge a' \in A^*_{F,\Lambda}(x')$.[11] Therefore, there exists $a^* : X \to A$ such that $a^*(x'') \geqslant_A a^*(x')$ for all $x'' > x'$, i.e., the MCS condition holds, and u is generally monotone with respect to F. The following corollary formally states these findings.

Corollary 2.5.2. *Assume that the set of actions A is compact and a lattice with a transitive, reflexive, anti-symmetric, and continuous partial order $\geqslant_A$. In addition, a payoff function $u : A \times \Omega \to \mathbb{R}$ is continuous, IDO-ordered, and supermodular in A, and an information structure F satisfies the MLRP. Then, u is generally monotone with respect to F.*

---

[11] Quah and Strulovici (2007) actually impose I-quasisupermodular condition which is weaker than the supermodular condition.

2.5.2 Proof of Theorem 2.5.1

I begin by stating a useful lemma.

Lemma 2.5.1. *For any reversely monotone noise* $\Gamma : X \times \Omega \to \Delta(Y)$ *and random variable* τ *uniformly distributed on* $(0, 1)$*, there exists a mapping* $m : (0, 1) \times X \times \Omega \to Y$ *such that the cumulative distribution induced by* $m(\tau|x, \omega)$ *is* $\Gamma(\cdot|x, \omega)$ *and* $m(t|x, \omega) \geqslant m(t|x, \omega')$ *for all* $\omega' > \omega$ *and* $t \in (0, 1)$.[12]

Proof of Lemma 2.5.1. The proof is a slight modification of Theorem 1.2.2 of Durrett (2019). For any $t \in (0, 1)$, $x \in X$ and $\omega \in \Omega$, define $m(t|x, \omega) \equiv \sup \mathcal{M}(t|x, \omega)$ where

$$\mathcal{M}(t|x, \omega) \equiv \left\{ y \in Y = [\underline{y}, \overline{y}] \ : \ \Gamma(y|x, \omega) < t \right\} \cup \{\underline{y}\}.$$

Note that $\mathcal{M}(t|x, \omega)$ is nonempty and bounded above, thus, $m(t|x, \omega)$ is properly defined.

Now suppose that a random variable τ is uniformly distributed on $(0, 1)$. Observe that $\Pr(m(\tau|x, \omega) \leqslant \hat{y}) = \Gamma(\hat{y}|x, \omega)$ for any $\hat{y} \in Y$ if and only if $L(\hat{y}|x, \omega) = R(\hat{y}|x, \omega)$ where

$$L(\hat{y}|x, \omega) \equiv \{t \in (0, 1) \ : \ m(t|x, \omega) \leqslant \hat{y}\},$$
$$R(\hat{y}|x, \omega) \equiv \{t \in (0, 1) \ : \ t \leqslant \Gamma(\hat{y}|x, \omega)\}.$$

First, when $\hat{y} = \overline{y}$, by the definition of $\mathcal{M}(t|x, \omega)$ and $\Gamma(\overline{y}|x, \omega) = 1$, we have $L(\overline{y}|x, \omega) = R(\overline{y}|x, \omega) = (0, 1)$.

Next, consider the case where $\hat{y} \in [\underline{y}, \overline{y})$. Suppose that $t \in L(\hat{y}|x, \omega)$, i.e., $m(t|x, \omega) \leqslant \hat{y}$. If $\Gamma(\hat{y}|x, \omega) < t$, by the right continuity of Γ, there exists $y \in (\hat{y}, \overline{y})$ such that $\Gamma(y|x, \omega) < t$. Then, $y \in \mathcal{M}(t|x, \omega)$ and it contradicts $\hat{y} = \sup \mathcal{M}(t|x, \omega)$. Therefore, we have $\Gamma(\hat{y}|x, \omega) \geqslant t$, which implies $t \in R(\hat{y}|x, \omega)$, and $L(\hat{y}|x, \omega) \subseteq R(\hat{y}|x, \omega)$. Conversely, if $t \in R(\hat{y}|x, \omega)$, we have $\Gamma(y|x, \omega) \geqslant \Gamma(\hat{y}|x, \omega) \geqslant t$ for all

[12] I thank an anonymous referee for providing this lemma.

$y \geqslant \hat{y}$. Thus, the upper bound of $\mathcal{M}(t|x,\omega)$ is at most $\hat{y}$, i.e., $\hat{y} \geqslant m(t|x,\omega)$. Hence, we have $t \in L(\hat{y}|x,\omega)$, and $R(\hat{y}|x,\omega) \subseteq L(\hat{y}|x,\omega)$. Therefore, $R(\hat{y}|x,\omega) = L(\hat{y}|x,\omega)$ for any $\hat{y} \in Y$ and the distribution induced by $m(\tau|x,\omega)$ is $\Gamma(\cdot|x,\omega)$.

By the reverse monotonicity, $\Gamma(y|x,\omega) \leqslant \Gamma(y|x,\omega')$ for all $y \in Y$, $x \in X$ and $\omega,\omega' \in \Omega$ with $\omega' > \omega$. Therefore, we have $\mathcal{M}(t|x,\omega') \subseteq \mathcal{M}(t|x,\omega)$, and it implies $m(t|x,\omega) \geqslant m(t|x,\omega')$ for all $t \in (0,1)$ and $\omega' > \omega$. □

Next, by using this lemma, for any increasing decision rule $\psi : Y \to A$ on G, I construct a mixed-action decision rules $\phi : X \times (0,1) \to A$ on F which gives a higher expected payoff than ψ for any state.

Proposition 2.5.2. *Suppose that G is a monotone quasi-garbling of F and u satisfies the DDDR condition with respect to $\geqslant_A$. For any increasing decision rule $\psi : Y \to A$, there exists a mixed-action decision rules $\phi : X \times (0,1) \to A$ such that for all $\omega \in \Omega$,*

$$\int_X \left[\int_0^1 u(\phi(x;t),\omega)dt\right] dF(x|\omega) \geqslant \int_Y u(\psi(y),\omega)dG(y|\omega). \tag{2.5.7}$$

Proof of Proposition 2.5.2. By the definition of $F \succeq_{MQG} G$, we have

$$\int_Y u(\psi(y),\omega)dG(y|\omega) = \int_X \int_Y u(\psi(y),\omega)d\Gamma(y|x,\omega)dF(x|\omega).$$

From Lemma 2.5.1, the right hand side is equal to

$$\int_X \left[\int_0^1 u(\psi(m(t|x,\omega)),\omega)dt\right] dF(x|\omega).$$

Since ψ is increasing and m is decreasing in ω, $\psi(m(t|x,\omega))$ is decreasing in ω. (Therefore, $\tilde{\psi}(t;x,\omega) \equiv \psi(m(t|x,\omega))$ satisfies (2.5.2) for all $\omega \in \Omega$ and is decreasing in ω.)

Given t and x, by the DDDR condition, there exists an action $a \in A$ such that $u(a, \omega) \geqslant u(\psi(m(t|x, \omega)), \omega)$ for all ω, and denote $\phi(x; t)$. Then, by using the inequalities for each t and x, we can derive (2.5.7). □

Now we are ready to prove Theorem 2.5.1.

Proof of Theorem 2.5.1. Fix a prior belief $\Lambda \in \Delta(\Omega)$ and a payoff function $u \in \mathcal{U}$. Let $a_G^* : Y \to A$ be the optimal decision function defined in the MCS condition: $a_G^*(y) \in A_{G,\Lambda}^*(y)$ and $a_G^*(y) \geqslant_A a_G^*(y')$ for all $y \geqslant y'$. By Proposition 2.5.2, there exists a mixed-action decision rules $\phi : X \times (0, 1)$ satisfying (2.5.7). Then, we have

$$\begin{aligned}
V(G; u, \Lambda) &= \int_\Omega \int_Y u(a_G^*(y), \omega) dG(y|\omega)\, d\Lambda(\omega) \\
&\leqslant \int_\Omega \int_X \left[\int_0^1 u(\phi(x; t), \omega) dt \right] dF(x|\omega)\ d\Lambda(\omega) \\
&= \int_0^1 \left[\int_X \int_\Omega u(\phi(x; t), \omega) d\Lambda_F^x(\omega) dF_\Lambda(x) \right] dt \\
&\leqslant \int_X \int_\Omega u(a_F^*(x), \omega) d\Lambda_F^x(\omega) dF_\Lambda(x) = V(F; u, \Lambda).
\end{aligned}$$

The third equality is simply from rearranging the integration order. The fourth inequality holds from the optimality of $a_F^*(x)$, and the last equality is straightforward. Therefore, F is more informative than G on $\mathcal{U}$. □

Remark 3. Proposition 2.5.2 is closely related to Proposition 9 of Quah and Strulovici (2009). Their result is stronger in the sense that they show that there exists an increasing pure-action decision rule $\phi : X \to A$ (rather than a mixed-action decision rules which are not necessarily increasing) such that $\int_X u(\phi(x), \omega) dF(x|\omega) \geqslant \int_Y u(\psi(y), \omega) dG(y|\omega)$ for all $\omega \in \Omega$. This is because their assumptions are stronger than the ones on Proposition 2.5.2: they impose the interval dominance order condition on the payoff function which is stronger than the DDDR condition; and they also assume that F is more Lehmann accurate than G, which is also stronger than the monotone quasi-garbling order in the absence of the MLRP.

2.5.3 Proof of Theorem 2.5.2

I focus on information structures with which decision problems in $\mathcal{U}$ are generally monotone. To qualify that a decision problem in $\underline{\mathcal{U}}$ is generally monotone, we first need to check whether the DDDR condition holds. The following lemma confirms this and the proof is relegated to Appendix A.4.1. The proof of Theorem 2.5.2 follows the lemma.

Lemma 2.5.2. *The generic order on $\underline{A}$ satisfies the DDDR condition with respect to $\underline{\mathcal{U}}$.*

Proof of Theorem 2.5.2. First, I show that F satisfies the MLRP. I fix a pair of signals $x_0 < x_1$ and a pair of states $\omega_0 < \omega_1$ and show that $f(x_1|\omega_1) \cdot f(x_0|\omega_0) \geqslant f(x_1|\omega_0) \cdot f(x_0|\omega_1)$. Consider a prior belief $\Lambda \in \Delta(\Omega)$ such that there are probability masses $1/2$ on ω_0 and ω_1. Observe that the posterior belief has probability masses $\tilde{\lambda}_F(x) \equiv f(x|\omega_1)/(f(x|\omega_0) + f(x|\omega_1))$ on ω_1 and $1 - \tilde{\lambda}_F(x)$ on ω_0.[13] For the utility function, set $\hat{\omega} = \omega_1$. Then, upon receiving a signal x, the DM solves the following problem:

$$\begin{aligned}\max_{a\in\{0,1\}} \int_\Omega u(a,\omega) d\Lambda_F^x(\omega) &= \max_{a\in\{0,1\}} a\left[-\kappa \cdot (1-\tilde{\lambda}_F(x)) + (1-\kappa)\cdot\tilde{\lambda}_F(x)\right] \\ &= \max_{a\in\{0,1\}} a\left(\tilde{\lambda}_F(x) - \kappa\right).\end{aligned}$$

Note that $a = 1$ solves the above maximization problem if and only if $\kappa \leqslant \tilde{\lambda}_F(x)$. Consider $\kappa = \tilde{\lambda}_F(x_0) - \epsilon$ for some $\epsilon > 0$. Then, $a = 1$ is the only optimal choice

[13] This is the case when there is no probability mass at x neither at the state ω_0 nor at the state ω_1. If there is a probability mass at either state, $\tilde{\lambda}_F$ should be defined by $\tilde{\lambda}_F(x) = \Pr(x|\omega_1)/(\Pr(x|\omega_0) + \Pr(x|\omega_0))$ where $\Pr(x|\omega) \equiv F(x|\omega) - F(x^-|\omega)$.

under a signal x_0. To satisfy the MCS condition, $a = 1$ also needs to be a solution under a signal x_1. It implies that $\tilde{\lambda}_F(x_1) \geqslant \kappa = \tilde{\lambda}_F(x_0) - \epsilon$. By sending ϵ to zero, we can derive that $\tilde{\lambda}_F(x_1) \geqslant \tilde{\lambda}_F(x_0)$. By rearranging the terms, we can derive that $f(x_1|\omega_1) \cdot f(x_0|\omega_0) \geqslant f(x_0|\omega_1) \cdot f(x_1|\omega_0)$. Therefore, F satisfies the MLRP.

Now I show that $F \succeq_L G$, i.e., $\Phi(\omega_1; \hat{y}) \geqslant \Phi(\omega_0; \hat{y})$ for all $\omega_1 > \omega_0$ and $\hat{y} \in Y$. Set $\hat{\omega} = \omega_1$, and $\kappa = \tilde{\lambda}_F(\Phi(\omega_0; \hat{y}))$. By the MLRP of F, we have

$$\begin{cases} \tilde{\lambda}_F(x) \geqslant \kappa, & \text{if } x \geqslant \Phi(\omega_0; \hat{y}), \\ \tilde{\lambda}_F(x) < \kappa, & \text{otherwise.} \end{cases}$$

Recall that $a = 1$ will be chosen if and only if $\kappa \leqslant \tilde{\lambda}_F(x)$, i.e., $x \geqslant \Phi(\omega_0; \hat{y})$. Also note that $F_\Lambda(x) = (F(x|\omega_0) + F(x|\omega_1))/2$ and $\tilde{\lambda}_F(x) \cdot dF_\Lambda(x) = f(x|\omega_1)/2$. Then, we can derive that

$$\begin{aligned} V(F; u, \Lambda) =& \int_{\Phi(\omega_0; y)}^{\overline{x}} \left(\tilde{\lambda}_F(x) - \kappa\right) dF_\Lambda(x) \\ =& \frac{1-\kappa}{2} \left(1 - F(\Phi(\omega_0; \hat{y})|\omega_1)\right) - \frac{\kappa}{2} \left(1 - F(\Phi(\omega_0; \hat{y})|\omega_0)\right). \end{aligned}$$

Next, observe that

$$\begin{aligned} V(G; u, \Lambda) &= \int_{\underline{y}}^{\overline{y}} \left[\max_{a \in \{0,1\}} a \cdot \left(\tilde{\lambda}_G(y) - \kappa\right)\right] dG_\Lambda(y) \\ &\geqslant \int_{\hat{y}}^{\overline{y}} \left(\tilde{\lambda}_G(y) - \kappa\right) dG_\Lambda(y) \\ &= \frac{1-\kappa}{2} \left(1 - G(\hat{y}|\omega_1)\right) - \frac{\kappa}{2} \left(1 - G(\hat{y}|\omega_0)\right). \end{aligned}$$

Since F is more informative than G on $\mathcal{U}$ and $u \in \underline{\mathcal{U}} \subseteq \mathcal{U}$, we have $V(F; u, \Lambda) \geqslant V(G; u, \Lambda)$ which implies

$$(1-\kappa)\left(G(\hat{y}|\omega_1) - F(\Phi(\omega_0; \hat{y})|\omega_1)\right) \geqslant \kappa \left(G(\hat{y}|\omega_0) - F(\Phi(\omega_0; \hat{y})|\omega_0)\right).$$

Note that the right hand side is equal to zero by the definition of $\Phi(\omega_0; \hat{y})$. Then, we have $G(\hat{y}|\omega_1) \geqslant F(\Phi(\omega_0; \hat{y})|\omega_1)$, which implies $\Phi(\omega_1; \hat{y}) \geqslant \Phi(\omega_0; \hat{y})$. Therefore, $F \succeq_L G$, and by Proposition 2.4.2, we have $F \succeq_{MQG} G$. □

2.6 Application: Nonlinear Monopoly Pricing

As an application of general monotone decision problems, I investigate information ranking in the context of monopoly pricing with second-degree price discrimination as in Maskin and Riley (1984). In this section, I review their model and basic results and rewrite them in the grammar of Section 2.5.[14]

A risk-neutral monopolistic seller that possesses full power of commitment wishes to sell output to a single buyer. The buyer has a privately known type $\omega \in \Omega = \{\omega_1, \cdots, \omega_N\}$ with $\omega_n < \omega_{n+1}$. In the first stage of the game, the seller chooses between two sources of information about the buyer's type. Next, the signal is received from the chosen information source and the seller updates the belief on the buyer's type. Then, the seller posts a menu of quantities and tariffs (non-linear prices). In the last stage, the buyer selects an element from the menu or decides not to participate. The outside option of both parties yields a payoff of zero.

Let the utility of a type ω buyer from consuming q units of the good and paying p units of money be $v(q, \omega) - p$. Let the cost function of the seller be some $c : \mathbb{R}_+ \to \mathbb{R}_+$, then the seller's payoff would be $p - c(q)$.[15] I assume some standard conditions on c and v.

---

[14] Comparison of information structures for monopoly pricing has also been studied by Athey and Levin (2017) and Ottaviani and Prat (2001). Athey and Levin applied their result in a setting where the seller receives a signal about the buyer's type and determines the simple monopoly price and there is no second-degree price discrimination. In the study by Ottaviani and Prat, both the seller and the buyer receive a signal about the buyer's type and the seller offers a nonlinear price. The setup in this section is a mixture of these two models: the buyer is fully aware of her type, the seller receives a signal about the buyer's type and then a menu of nonlinear prices is offered.

[15] Note that in this example, the seller's utility itself is not dependent on ω. However, different payoffs across states arise from price discrimination.

Assumption. The seller's cost function $c : \mathbb{R}_+ \to \mathbb{R}_+$ and the buyer's utility function $v : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ satisfy the following properties:

1. v is continuous, nondecreasing in ω, and supermodular in (q, ω), i.e., $v(q', \omega') - v(q, \omega') \geqslant v(q', \omega) - v(q, \omega)$ for all $q' \geqslant q \geqslant 0$ and $\omega' > \omega$. In addition, $v(0, \omega) = 0$ for all $\omega \in \Omega$.

2. c is continuous in q with $c(0) = 0$;

3. there exists $Q > 0$ such that $v(Q, \omega) - c(Q) < 0$ for all $\omega \in \Omega$.

The seller has a prior belief $\Lambda \in \Delta(\Omega)$ with a corresponding probability mass function λ. After receiving a signal about the buyer's type, the seller forms a posterior belief and sets pricing options $\{\vec{q}(\omega), \vec{p}(\omega)\}_{\omega \in \Omega}$ based on the updated belief. The set of actions can be viewed as a subset of $[0, Q]^N \times \mathbb{R}_+^N$.

A pricing option $(\vec{q}, \vec{p})$ can be implemented if and only if it satisfies the following familiar constraints: for all $\omega, \omega' \in \Omega$,

$$v(\vec{q}(\omega), \omega) - \vec{p}(\omega) \geqslant v(\vec{q}(\omega'), \omega) - \vec{p}(\omega'), \tag{IC}$$

$$v(\vec{q}(\omega), \omega) - \vec{p}(\omega) \geqslant 0. \tag{IR}$$

Therefore, the set of feasible decisions is

$$A_M = \left\{ (\vec{q}, \vec{p}) \in [0, Q]^N \times \mathbb{R}_+^N \,\middle|\, (\vec{q}, \vec{p}) \text{ satisfies ICs and IRs} \right\}. \tag{2.6.1}$$

I call that an information F has the full support if $f(x|\omega) > 0$ for all $x \in X$ and $\omega \in \Omega$. Likewise, I call that a prior belief Λ has the full support if $\lambda(\omega) > 0$ for all $\omega \in \Omega$. Then, upon receiving a signal x, the probability mass of the posterior belief is $\lambda_F^x(\omega) > 0$ for all $\omega \in \Omega$. This is essential for the further argument because if $\lambda_F^x(\omega) = 0$ for some $\omega \in \Omega$, the seller would ignore some ICs and IRs and suggest a menu of pricing which might not be in A_M. Under the full support assumption, the seller's problem is

$$\max_{(\vec{q},\vec{p})\in A_M} \sum_{n=1}^{N} \left[-c(\vec{q}(\omega_n)) + \vec{p}(\omega_n)\right] \cdot \lambda_F^x(\omega_n). \tag{2.6.2}$$

It is well known that an allocation $\vec{q}$ is implementable iff $\vec{q}$ is monotone, and all the local downward ICs and IR for the lowest type are binding at the solution of (2.6.2) (Maskin and Riley (1984), Guesnerie and Laffont (1984)). That is, the optimal pricing policy given the quantity allocation $\vec{q}$ is as follows: for all $1 \leqslant n \leqslant N$,

$$\vec{p}(\omega_n) = v(\vec{q}(\omega_n), \omega_n) - \sum_{i=1}^{n-1} \left[v(\vec{q}(\omega_i), \omega_{i+1}) - v(\vec{q}(\omega_i), \omega_i)\right].^{16} \tag{2.6.3}$$

In addition, we can restrict attention to the set of decisions $\mathcal{Q}$:

$$\mathcal{Q} = \left\{\vec{q} \in [0, Q]^N \;\middle|\; \text{for all } 1 \leqslant i \leqslant j \leqslant N,\ \vec{q}(\omega_j) \geqslant \vec{q}(\omega_i)\right\}. \tag{2.6.4}$$

By plugging (2.6.2) into the seller's payoff function, define

$$u(\vec{q}, \omega_n) \equiv v(\vec{q}(\omega_n), \omega_n) - \sum_{i=1}^{n-1} \left[v(\vec{q}(\omega_i), \omega_{i+1}) - v(\vec{q}(\omega_i), \omega_i)\right] - c(\vec{q}(\omega_n)). \tag{2.6.5}$$

Then, we can rewrite (2.6.2) as $\max_{\vec{q}\in\mathcal{Q}} U(\vec{q}; x)$ where

$$U(\vec{q}\,; x) = \sum_{n=1}^{N} u(\vec{q}, \omega_n) \cdot \lambda_F^x(\omega_n). \tag{2.6.6}$$

Under the optimal mechanism, no matter what the seller's belief is, quantities for higher types are almost as high as the first best quantity. However, quantities for lower types serve as a tool for incentivizing not only the lower types of the buyer but also higher types through the pricing policy (2.6.3). From the supermodularity of v, we observe that reducing a quantity for a low type buyer would induce a price

[16] Note that the IR for the lowest type binds, i.e., $\vec{p}(\omega_1) = v(\vec{q}(\omega_1), \omega_1)$.

increase for a high type buyer. Note that if the seller believes that the buyer is more likely to be a high type, then the seller would be less concerned about the possibility of distorting the quantities designed for low types in order to raise the prices for high types.

Using this intuition, consider a partial order $\geqslant_{\mathcal{Q}}$ on $\mathcal{Q}$ defined as follows: for any $\vec{q}_i,\ \vec{q}_j \in \mathcal{Q}$,

$$\vec{q}_j \geqslant_{\mathcal{Q}} \vec{q}_i \quad \Longleftrightarrow \quad \vec{q}_j(\omega) \leqslant \vec{q}_i(\omega) \ \ \forall\ \omega \in \Omega. \tag{2.6.7}$$

The next lemma justifies this partial order by showing that the payoff function satisfies the DDDR condition with respect to this partial order.

Lemma 2.6.1. *Assume that v is supermodular in (q, ω). The payoff function defined in* (2.6.5) *satisfies the DDDR condition with respect to the partial order $\geqslant_{\mathcal{Q}}$ defined in* (2.6.7)*.*

The next task is to check the MCS condition. I use a similar argument as in the latter part of Section 2.5.1: (i) show that $U(\vec{q}\,; x)$ is supermodular in $\mathcal{Q}$; (ii) also show that $U(\vec{q}\,; x'')$ I-dominates $U(\vec{q}\,; x')$ for all $x'' > x'$; then (iii) apply Theorem 1 of Quah and Strulovici (2007).[17][18]

I begin by showing that $\mathcal{Q}$ is a lattice. Observe that $\vec{q}_i \vee \vec{q}_j = (\min\{\vec{q}_i(\omega), \vec{q}_j(\omega)\})_{\omega\in\Omega}$ and $\vec{q}_i \wedge \vec{q}_j = (\max\{\vec{q}_i(\omega), \vec{q}_j(\omega)\})_{\omega\in\Omega}$ under this partial order.[19] When $\vec{q}_i$ and $\vec{q}_j$ are in $\mathcal{Q}$, i.e., they are nondecreasing in ω, $\min\{\vec{q}_i, \vec{q}_j\}$ and $\max\{\vec{q}_i, \vec{q}_j\}$ are nondecreasing in ω as well. Thus, $\vec{q}_i \vee \vec{q}_j$ and $\vec{q}_i \wedge \vec{q}_j$ are in $\mathcal{Q}$, i.e., $\mathcal{Q}$ is a lattice. For any function w :

[17] This argument is slightly different from Section 2.5.1 in the second step. In Section 2.5.1, I use the IDO condition for the payoff function u to show that u satisfies the DDDR condition and $U(\vec{q}\,; x'')$ I-dominates $U(\vec{q}\,; x')$. In this example, I directly show these properties in Lemma 2.6.1 and 2.6.2.

[18] The proof of the MCS result in the previous version of the paper was based on the first-order condition and the ironing technique with stronger conditions on the primitives. The interval-dominance-order-based argument presented in this version is due to an anonymous referee, which not only weakens the conditions in the previous version, but also substantially simplifies the proof.

[19] Note that it is opposite from the case with the usual product order of which inequality is reversed from (2.6.7).

$\mathbb{R} \to \mathbb{R}$, we have $w(a) + w(b) = w(\max\{a, b\}) + h(\min\{a, b\})$. By using this to $v(\cdot, \omega)$ and $c(\cdot)$, we can easily derive that $u(\vec{q}_i, \omega_n) + u(\vec{q}_j, \omega_n) = u(\vec{q}_i \vee \vec{q}_j, \omega_n) + u(\vec{q}_i \wedge \vec{q}_j, \omega_n)$, thus, u is supermodular in $\mathcal{Q}$. Since $U(\vec{q}\,; x)$ is a convex combination of $u(\vec{q}, \omega)$, $U(\vec{q}\,; x)$ is also supermodular in $\mathcal{Q}$.

Next, the following lemma shows that $U(\vec{q}\,; x'')$ I-dominates $U(\vec{q}\,; x')$ when the MLRP condition is imposed. The proof is provided in Appendix A.4.3.

Lemma 2.6.2. *Suppose that an information structure F satisfies the MLRP and has the full support, and the prior belief Λ has the full support. Then, for any $x'' > x'$, $U(\vec{q}\,; x'')$ I-dominates $U(\vec{q}\,; x')$.*

Now we are ready to show that the monotone quasi-garbling order can be served as a criterion for comparing information in the nonlinear monopoly pricing problem.

Proposition 2.6.1. *Suppose that assumptions 1-4 hold for (v, c), F and G have full support, G satisfies the MLRP, and G is a monotone quasi-garbling of F. Then, the seller prefers F to G for any prior belief Λ with full support.*

Proof. By Lemma 2.6.1, the payoff function u defined in (2.6.5) satisfies the DDDR condition. Next, under the assumption that G satisfies the MLRP, since $U(\vec{q}\,; y)$ is supermodular in $\mathcal{Q}$ and $U(\vec{q}\,; y'')$ I-dominates $U(\vec{q}\,; y')$ for all $y'' > y'$, by Theorem 1 of Quah and Strulovici (2007), the MCS condition holds: there exists $\vec{q}_* : Y \to \mathcal{Q}$ such that $\vec{q}_*(y'') \geqslant_A \vec{q}_*(y')$ for all $y'' > y'$. Therefore, the seller's problem is generally monotone with respect to G. Then, when G is a monotone quasi-garbling of F, by Theorem 2.5.1, we have $V(F; u, \Lambda) \geqslant V(G; u, \Lambda)$. □

3

Managing a Project by Splitting it into Pieces

3.1 Introduction

In project management, a work breakdown structure (WBS)—a step-by-step approach to complete projects—is widely used (Organ and Bottorff, 2022).[1] The applications of the WBS range from simple projects such as moving an office to complicated projects such as construction, engineering, or software development.[2] Although there are many advantages of employing the WBS (e.g., clarifying the goals, communicating better, etc.), a fundamental benefit is monitoring progress. As a project is broken down into smaller chunks, a manager can better audit a subordinate's progress, which may reduce the moral hazard issue. Nevertheless, decomposing a project into too small pieces may make the project rigid (Golany and Shtub, 2001). It may lead the manager to micromanage the project which in turn slows down project progress, i.e., generates inefficiencies. Thus, when a manager splits a project, she faces a tradeoff between *monitoring* and *efficiency*.

[1] The PMBOK guide provides a formal definition of a work breakdown structure: "A hierarchical decomposition of the total scope of work to be carried out by the project team to accomplish the project objectives and create the required deliverables." (Project Management Institute, 2017)

[2] See Project Management Institute (2006) for further examples.

In this article, I introduce a stylized model to study this tension between efficiency and monitoring in breaking a project down.[3] Specifically, I consider a dynamic principal-agent model with two routes to achieving success. One is to attack the project head-on (the direct approach). This method requires a breakthrough which arrives at a low rate. The alternative route is to divide the project into two subprojects and complete them one by one (the sequential approach). This method requires two breakthroughs that each arrive at higher rates. Each breakthrough in this approach can be understood as the completion of a subproject. At the beginning of the game, the principal offers a contract specifying: a schedule dictating which approach to use, how much reward to pay upon success, and a cancellation policy. If the agent accepts the contract, then at each point in time, he chooses whether to work on the specified approach or shirk for private benefit. The principal does not observe the agent's effort, generating the potential for moral hazard.

To highlight the tension between the direct and sequential approaches, I impose two key assumptions. First, the completion of the first subproject (in the sequential approach) is observable and contractually verifiable. This implies that the project breakdown can help the principal monitor progress. The second assumption is that the direct approach is more efficient than the sequential one. Hence, the sequential approach has an advantage in monitoring the agent but has a disadvantage in efficiency *vis a vis* the direct approach. To be clear, the probability of ultimate project completion is higher under the sequential approach, but the expected running time and concomitant operating cost dominate this effect.

In addition to these elements, there is a third important economic factor that affects the choice of methodology: a *deadline effect*. The principal needs to impose a deadline because, in the absence of a deadline, the agent could (and would) shirk

[3] I refer to splitting the project in two as "breaking it down" to conform with project management nomenclature. My model never involves a project breakdown in the sense of an accident or mishap.

forever without completing the project. Thus, the deadline is essential to mitigate moral hazard. This deadline generates a subtle difference between the incentive schemes implementing each approach. If the principal wants to induce the direct approach, the agent is compensated by an immediate payment upon project success. On the contrary, if the principal wants to induce the sequential approach, the principal slightly extends the deadline after observing the subproject completion and pays the agent only upon the successful completion of the entire project. Thus, the principal chooses between approaches by comparing the expected payoffs from the immediate payment scheme and the deadline extension scheme. When the principal uses the deadline extension scheme, the agent may work for a longer period of time. As noted above, this means that under the sequential approach the probability of ultimate success is higher, but the expected running cost is higher as well. This highlights another motive for the principal to employ the direct approach besides comparing the monitoring loss and efficiency gain. Specifically, when the deadline is close by, i.e., the probability of project completion is low, the principal may prefer saving the expected cost by choosing the direct approach rather than slightly raising the chance of project completion by employing the sequential approach.

To facilitate analysis, I begin by characterizing the optimal contract in the case where both approaches are equally efficient. Here, I can abstract from efficiency considerations and focus on the interaction between monitoring and the deadline. Monitoring gives a generic advantage to the sequential approach while the deadline effect—described in the previous paragraph—gives an advantage to the direct approach when the deadline is imminent. This suggests that the principal would prefer the sequential approach when the deadline is distant and the direct approach when it is near at hand. However, it is possible that the monitoring advantage is so strong that the principal chooses the sequential approach even near the deadline. On the contrary, it is also possible that the optimal deadline is short enough that the principal would want to choose the direct approach even at contract inception. Therefore, depending on the economic environment, different types of contracts may be optimal even in this simplified setting.

The main result of this analysis is that the form of the optimal contract depends crucially on the project return; i.e., the gross value to the principal from completing the project (given its operating cost). As described above, the sequential approach has a higher chance of project completion despite its higher expected cost. Because the project return must be scaled by the probability of successful completion, the sequential approach is preferred when the project return is large and the direct approach is preferred when the project return is small. Based on this observation, I show that the optimal contract is derived as follows:

(a) when the project return is low, the optimal deadline is short and the principal only chooses the direct approach;

(b) when the project return is high, monitoring is highly advantageous and the principal only chooses the sequential approach;

(c) when the project return is intermediate, there is a switching point such that the principal first chooses the sequential approach and then – absent a breakthrough – switches to the direct approach until the deadline is reached.

Next, I introduce the efficiency loss to the sequential approach, representing the idea that requiring milestones may slow down ultimate project development. When the efficiency loss is small enough, I show that a similar result as in the previous case holds: there are three regions of the project return that characterize the form of the optimal contract. In other words, the characterization of the optimal contract when the approaches are equally efficient is robust to a small efficiency loss. This is mainly because efficiency dominates monitoring only if the deadline is distant. I also consider the case where the efficiency loss is large. Here, even if the project return is moderately high, the principal prefers the direct approach over the sequential approach to avoid the efficiency cost. Nevertheless, if the project return is very high, the principal prefers to monitor it to some degree. In fact, there is a cutoff value for the project return such that the principal chooses the direct approach when the return is below the cutoff. Interestingly, if the return is above the threshold, then the principal begins by choosing the direct approach, switches to the sequential approach, then switches back to the direct approach until the deadline is reached.

My results are congruent with the observation that applied scientific research (e.g., development of a new drug, clinical trials) is typically staged. The magnitude of applied research projects is usually large, implying the superiority of the sequential approach. In contrast, the immediate value of basic research (e.g., chemistry, in-vitro experiments) is lower than applied research because "basic research is performed without thought of practical ends" (Bush, 1945).[4] My results suggest that the direct approach should be preferred for basic research because such projects tend to have lower returns than applied ones. For instance, the Research Project designation (R01) grant by the National Institute of Health (NIH) supports "a discrete, specified, circumscribed project" rather than a staged project.[5]

The remainder of this article is organized as follows. Related literature is discussed below. Section 3.2 introduces the basic setup of the model and analyzes the first-best case. Section 3.3 provides heuristic arguments on the derivations of the value function and the optimal contract. Then, Section 3.4 and 3.5 characterize the optimal contracts for the cases with and without the efficiency losses from the project

[4] Bush argues that although broad and basic studies seem to be less important than applied ones, they are essential to combat diseases because progress in the treatment "will be made as the result of fundamental discoveries in subjects unrelated to those diseases, and perhaps entirely unexpected by the investigator." However, since this article does not consider externalities, I abstract from this possibility and focus on the principal's return from the completed project.

[5] `https://grants.nih.gov/grants/funding/r01.htm`

breakdown. The formal analysis and the proofs are relegated to an Appendix.

Related Literature

There is a growing literature on contracting for multi-stage projects, e.g., Hu (2014); Green and Taylor (2016a); Wolf (2018); Moroni (2022). The most closely related study is Green and Taylor (2016a), who study a model in which multiple breakthroughs are needed to complete a project and in which an agent must be incentivized to exert unobservable effort. The sequential approach considered here comprises the baseline model with the tangible breakthrough in the working paper version of their paper (Green and Taylor, 2016c). However, the option to complete the project directly, which is not considered in their setup, allows the principal to face a choice problem between the two approaches. Moreover, this choice problem arises at every point in time. Therefore, the principal's problem becomes more complex from a dynamic perspective.

Another article that has a similar flavor is Carnehl and Schneider (2021). They consider a two-armed bandit problem where an arm requires one breakthrough (the doing arm) but another arm requires multiple breakthroughs (the thinking arm) to succeed. The arrival rates for the thinking arm are known to the agent whereas the arrival rate for the doing arm is not: the agent needs to infer whether the method is feasible or not by experimenting. The presence of this uncertainty is one key difference between their article and this one. Moreover, their main analysis focuses on a decision problem by a single agent whereas this article considers a principal-agent contracting setting. In addition, they have an exogenous deadline whereas the deadline is endogenously determined in this paper. Despite these differences, we share a common insight in that the chosen approaches may switch up to two times. However, the economic forces that drive choosing the thinking arm (or the sequential approach) are somewhat different. In their paper, as the agent pulls the doing arm and does not achieve any success, the belief that the initial method is feasible goes down. When the belief becomes sufficiently low, the thinking arm would be chosen because it may be more efficient than the doing arm. Thus, experimentation and efficiency are key driving forces for choosing the thinking arm. On the contrary, in this paper, the principal chooses the sequential approach to monitor the agent, not because beliefs about the direct approach have deteriorated.

This article is also somewhat related to the literature on monitoring in dynamic contracts, e.g., Orlov (2015); Piskorski and Westerfield (2016); Dilmé and Garrett (2019); Marinovic and Szydlowski (2019); Varas et al. (2020); Marinovic and Szydlowski (2020); Chen et al. (2020). In most of these papers, a monitoring process provides some information on the agent's current or past action. In this sense, the first breakthrough in the sequential approach can be considered as a monitoring device since it lets the principal know that the agent has worked. However, the completion of the first subproject gives more information than merely the agent's past actions. Before the subproject completion, the success requires one relatively hard breakthrough or two easier breakthroughs. After completing the subproject, it requires only one relatively easy breakthrough. Thus, the subproject completion is distinguished from standard monitoring processes since it also provides information about the subsequent procedure toward success.

The problem of choosing approaches is naturally related to multitasking in the sense that there are multiple options to pursue. In their seminal study, Holmstrom and Milgrom (1991) consider an economic situation where a production worker faces multiple tasks such as producing output and maintaining quality in a static environment. Dewatripont et al. (2000) and Laux (2001) also study multitasking problems in static environments. Several subsequent multitasking problems are also explored in dynamic settings (Manso, 2011; Capponi and Frei, 2015; Varas, 2017; Szydlowski, 2019). A common assumption in these studies is that each task has a different payoff structure.[6] 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 experiment (pulls the risky arm), the payoff is stochastic, and if the agent chooses to exploit (pulls the safe arm), the payoff is constant. In contrast, the two approaches in this article have the same ultimate payoff. The difference in the approaches is 'how' the main project is completed–via the direct approach or via the sequential approach.

This article is relevant to the literature studying complementary innovations, e.g., Green and Scotchmer (1995); Gilbert and Katz (2011); Bryan and Lemus (2017); Poggi (2021). Two subprojects in the sequential approach can be considered as 'perfect' complements in the sense that completing a subproject does not create any value but completing both of them does. However, to my knowledge, most of the studies in this literature focus on the problems with competing firms or a single decision maker, whereas this article studies an agency problem.

Last, from a technical viewpoint, the current article utilizes Poisson processes which are widely used to address dynamic moral hazard, e.g., Biais et al. (2010); Myerson (2015); Green and Taylor (2016a); Bonatti and Hörner (2017); Varas (2017); Sun and Tian (2017).

3.2 The Model

3.2.1 The Setting

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

[6] The only study 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.

$\Pi > 0$, dubbed the project return, 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 limited liability, i.e., the principal can only transfer nonnegative rewards to the agent.[7] The principal and the agent are both risk-neutral and patient, i.e., they do not discount the future.[8] Both players have outside options of zero.

There are two routes to achieving success. One is to attack the project directly, and I accordingly call this the *direct approach*. Another way is to break the main project into two subprojects and I call this the *sequential approach*. Completing the (first) subproject does not have any independent value for the principal or the agent.[9] However, the completion of the subproject is observable by both players and contractually verifiable by a court. Thus observing the completion of the subproject can be considered a type of monitoring.

At time 0, the principal offers the agent a contract consisting of the deadlines, at which the project is terminated; the reward schedules upon project completion; the approaches to be taken; and the agent's recommended effort. See Section B.1 for the formal definition of the contract. The principal can fully commit to these contractual terms. Note that the contract is contingent on the subproject completion. When the subproject has not been completed, at each point in time t, the contract specifies which approach to take: the direct approach ($a_t = 1$) or the sequential approach ($a_t = 0$). The agent allocates his 1 unit of effort to working ($\tilde{b}_t \in [0, 1]$), and shirking ($1 - \tilde{b}_t$).[10] The allocation of efforts is unobservable to the principal. Then, at time t, the project is completed at the arrival rate $\lambda_D a_t \tilde{b}_t$ (and the agent

[7] See Remark 5 for further discussion of limited liability.

[8] I explore the case of a positive discount factor in Online Appendix Section B.8.

[9] In Online Appendix Section B.7, I consider the case where the completion of the first subproject raises outside options for the agent and the principal.

[10] The agent's choice will be denoted with tilde, whereas the principal's choices are not.

receives the reward R_t), the subproject is completed at the rate $\lambda_S(1-a_t)\tilde{b}_t$, and the agent receives $\phi(1-\tilde{b}_t)$ as a private flow benefit from shirking. I assume that the marginal private benefit ϕ is positive but less than the principal's flow operating cost c. It is easier to complete the subproject than to complete the main project, i.e., λ_S is greater than λ_D. If neither the main project nor the subproject is completed by the deadline, the project is terminated. When the subproject succeeds, the deadline and the reward schedule are updated. In this case, the agent only needs to complete one more subproject (with the same arrival rate) to make the entire project succeed. Thus, the main project is completed at the rate $\lambda_S\tilde{b}_t$.[11]

3.2.2 The First-Best Case

As a benchmark, I consider the case where the agent's effort is observable to the principal, namely the first-best situation. Since the benefit from shirking is less than the flow cost, it is optimal for the principal to make the agent work in this case.

To see which approach is more efficient, we need to compare the expected payoffs of both approaches. Suppose that the principal takes the direct approach indefinitely, i.e., until the project is completed. Then, the probability distribution of the date of project completion (τ_m) is given by $\lambda_D e^{-\lambda_D \tau_m}$. From this, we can derive the expected payoff of (indefinitely employing) the direct approach as follows:

$$\int_0^\infty (\Pi - c\cdot\tau_m)\,\lambda_D e^{-\lambda_D\tau_m} d\tau_m = \Pi - \frac{c}{\lambda_D}.$$

Since it does not have a deadline and the agent never shirks, the project will be surely completed and the principal will receive Π (recall that both parties do not discount). The second term comes from the cumulative costs: since the expected duration of the project is $1/\lambda_D$, the expected cost is c/λ_D.

[11] I assume that two subprojects in the sequential approach are symmetric. In Online Appendix Section B.6, I relax this assumption and analyze the case where the two subprojects are asymmetric.

Next, consider the case where the principal takes the sequential approach indefinitely. In this case, in addition to τ_m, we need to consider the date of the first subproject completion–denoted by τ_s. Conditional on the completion of the subproject at the date τ_s, the probability distribution of τ_m is $\lambda_S e^{-\lambda_S(\tau_m-\tau_s)}$ for $\tau_m > \tau_s$ and 0 otherwise. Since the marginal probability distribution of τ_s is $\lambda_S e^{-\lambda_S \tau_s}$, the marginal probability distribution of τ_m can be derived as follows:

$$\int_0^{\tau_m} \lambda_S e^{-\lambda_S(\tau_m-\tau_s)} \cdot \lambda_S e^{-\lambda_S \tau_s} d\tau_s = \lambda_S^2 \tau_m e^{-\lambda_S \tau_m}.$$

Then, the expected payoff of (indefinitely employing) the sequential approach is given as follows:

$$\int_0^{\infty} (\Pi - c \cdot \tau_m)\, \lambda_S^2 \tau_m e^{-\lambda_S \tau_m} = \Pi - \frac{2c}{\lambda_S}.$$

Since the expected duration for each step is $1/\lambda_S$, the expected cost is $2c/\lambda_S$ in this case.

I assume that the sequential approach is profitable: $\Pi > 2c/\lambda_S$. I also assume that the sequential approach is not more efficient than the direct approach, i.e., monitoring may harm the efficiency of the project: $\lambda_D \geqslant \lambda_S/2$. Together with the previous assumption, this implies that the direct approach is also profitable: $\Pi > c/\lambda_D$.

I introduce a parameter $\eta \equiv \lambda_S/\lambda_D - 1$, which measures the efficiency of the sequential approach. If η is equal to one, or equivalently $\lambda_S = 2\lambda_D$, the sequential approach is equally efficient as the direct approach. As η decreases, the efficiency loss from monitoring increases. In Section 3.4 and 3.5, I derive the optimal contract respectively for the case of no efficiency loss from monitoring ($\eta = 1$) and the case of efficiency loss from monitoring ($\eta < 1$), and show that the form of the optimal contract depends crucially on η.

Remark 4. When the sequential approach is less efficient than the direct approach, even if the principal is allowed to switch methods, she will stick to the direct approach in the first-best case. On the other hand, when the sequential approach is equally efficient to the direct approach, she may switch back and forth.

Remark 5. If the agent is not protected by limited liability, the following contract will allow the principal to achieve the first-best outcome. The principal does not use any deadline and always chooses the direct approach. Then, at each point of time, if the project is not completed, she charges ϕ to the agent, i.e., the agent pays ϕ to the principal. If the project is completed, the principal pays $\phi/\lambda_D + \epsilon$ to the agent for some small $\epsilon > 0$. Observe that the agent will always work because the instantaneous payoff from working is $\lambda_D(\phi/\lambda_D + \epsilon) - \phi$ whereas that from shirking is zero. Then, the agent's expected payoff is ϵ, and the principal's expected payoff is $\Pi - c/\lambda_D - \epsilon$. By sending ϵ to zero, the principal is able to achieve the first-best outcome.

3.3 Minimum-Incentive Contracts and Deadlines

In this section, I provide heuristic arguments on the optimal contract derivation. See Appendix B.1 and B.2 for the formal analysis.

3.3.1 The Value Function

Following the standard approach in the dynamic contract literature, I consider the agent's promised utility as a state variable and write a contract recursively (see, e.g., Spear and Srivastava, 1987). For a contract Γ, let $P_0(\Gamma)$ and $U_0(\Gamma)$ be the expected payoffs of the principal and the agent at time 0 when the agent adheres to the recommended effort specified in the contract. The core of the analysis is the derivation of the principal's value function, denoted by $V(u)$, which represents her maximized expected utility $P_0(\Gamma)$ subject to the promise-keeping constraint $U_0(\Gamma) = u$ and the incentive compatibility condition—which will be demonstrated in Section 3.3.2. If a contract Γ satisfies $P_0(\Gamma) = V(u)$ and $U_0(\Gamma) = u$, Γ is said to *implement* a pair of expected payoffs $(u, V(u))$. Once the value function is characterized, the principal solves

$$\bar{u} \equiv \arg\max_{u \geqslant 0} V(u). \tag{MP}$$

Then, the optimal contract is the contract that implements $(\bar{u}, V(\bar{u}))$. In the rest of this section, I describe how to derive the value function $V(u)$.

I begin by solving the principal's problem given that the subproject is completed at time t. I also consider the agent's promised utility u_S^t as a state variable. Upon the subproject completion at time t, the contract is updated to $\hat{\Gamma}^t$. Define the value function $V_S(u_S^t)$ as the function that maximizes the principal's expected utility $\hat{P}^t(\hat{\Gamma}^t)$ subject to the promise-keeping constraint $\hat{U}^t(\hat{\Gamma}^t) = u_S^t$ and the incentive compatibility condition. Since this case only requires one more breakthrough, it is identical to the single-stage benchmark of Green and Taylor (2016a), thus I can directly use their results. They show that $V_S(u_S^t)$ is the expected profit from running the project for u_S^t/ϕ units of time net of u_S^t:

$$V_S(u_S^t) = \left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi} u_S^t}\right) - u_S^t. \tag{3.3.1}$$

In addition, they show that this can be implemented by a contract with a deadline $t + u_S^t/\phi$ and a linearly diminishing reward schedule $\{R_s^t\}_{t \leqslant s \leqslant t + u_S^t/\phi}$ where

$$R_s^t = u_S^t + \phi/\lambda_S - \phi(s - t) \tag{3.3.2}$$

for all $t \leqslant s \leqslant t + u_S^t/\phi$. The intuition is that when the agent's promised utility is u_S^t, the principal can incentivize the agent to work at most u_S^t/ϕ units of time. This is because if the principal requires him to work more than u_S^t/ϕ units of time, he can achieve higher payoffs than the promised utility by shirking.

3.3.2 The Agent's Problem

Now consider the agent's problem when the subproject has not been completed. Suppose that the promised utility is u_t at some time t. Under the direct approach, if the agent works for a small interval of time $[t, t+dt)$, the breakthrough occurs and the agent receives the reward R_t with a probability $\lambda_D dt$. However, in this event, he loses the continuation utility, thus, the expected payoff of working is $\lambda_D(R_t - u_t)dt$. On the other hand, if he shirks, his payoff is ϕdt. Therefore, to induce the agent to work, R_t should be greater than or equal to $u_t + \phi/\lambda_D$. Next, under the sequential approach, the agent is compensated in the form of the promised utility upon the subproject completion. Thus, the expected payoff of working for $[t, t+dt)$ is $\lambda_S(u_S^t - u_t)dt$. Then, to induce the agent to work, u_S^t should be greater than or equal to $u_t + \phi/\lambda_S$. To sum up, $R_t \geqslant u_t + \phi/\lambda_D$ and $u_S^t \geqslant u_t + \phi/\lambda_S$ serve as incentive compatibility conditions for the direct approach ($a_t = 1$) and the sequential approach ($a_t = 0$) respectively.

3.3.3 The Principal's Problem

Next, consider the principal's problem. At each point of time t, the principal needs to choose which approach to take (a_t), the reward upon project completion (R_t), the updated contract upon the subproject completion ($\tilde{\Gamma}^t$), and the agent's recommended effort (b_t). If the agent shirks, neither the whole project nor the subproject can be completed, but the flow cost incurs. Thus, the principal would recommend the agent to work unless the contract is terminated. Note that when the subproject is completed, the agent only cares about his updated promised utility, u_S^t, which can eventually determine the updated contract and the principal's expected payoff $V_S(u_S^t)$. Thus, the principal's problem reduces to the choice of a_t, R_t, and u_S^t subject to the incentive compatibility conditions.

We can naturally guess that the incentive compatibility conditions bind in the optimal contract, i.e., the principal provides the minimum incentive for the agent to work.[12] I now show that under the minimum-incentive contract, (i) R_t linearly diminishes over time; (ii) the deadline is extended by $1/\lambda_S$ when the subproject is completed. First, observe that in the absence of any completion, the agent's promised utility is consumed at the same rate with the benefit from shirking: $du/dt = \dot{u}_t = -\phi$. This is because the agent is indifferent between shirking and working under the minimum incentives and it gives the following promised utility dynamics:

$$u_t = u_{t+dt} + \phi dt = u_t + \dot{u}_t dt + \phi dt \quad \Longrightarrow \quad 0 = \dot{u}_t + \phi.$$

Note that $u_t = u_0 - \phi t$, which means that if any completion has not been made by time t, the agent is promised to the continuation utility $u_0 - \phi t$. Under the direct approach, to make the incentive compatibility condition bind, $R_t = u_t + \phi/\lambda_D = u_0 - \phi t + \phi/\lambda_D$, which linearly diminishes over time. Next, if the completion has not been made by u_0/ϕ, the promised utility is equal to the agent's outside option 0, thus, the contract is terminated, or equivalently, the deadline of the contract is u_0/ϕ. Under the sequential approach, to make incentive compatibility bind, $u_S^t = u_t + \phi/\lambda_S$. Because the updated deadline is $t + u_S^t/\phi = u_0/\phi + 1/\lambda_S$, the deadline is extended by $1/\lambda_S$. In addition, when the original deadline is T ($u_0 = \phi T$), by (3.3.2) and $u_S^t = u_0 - \phi t + \phi/\lambda_S$, the reward upon project completion at time t is determined as follows:

$$R_s^t = u_0 + 2\phi/\lambda_S - \phi s = \phi\left(T + 2/\lambda_S - s\right),$$

i.e., the reward is linearly diminishing over time.

The remaining step is to determine which approach to take. If the principal

[12] In the formal analysis, I show that it is indeed optimal for the principal to provide the minimum incentive to the agent. To do this, I use the 'guess and verify' method which is widely used in continuous-time analysis. That is, I first guess the principal's value function by letting her provide the minimum incentives to the agent, then verify that the principal does not have an incentive to deviate from it. The verification parts of the proof are long and tedious, so they are relegated to the latter part of the appendix (Appendix B.4). Specifically, I provide proof that the principal does not deviate from providing the minimum incentives in B.4.2.

chooses the direct approach for $[t, t+dt)$, the expected payoff is

$$\begin{aligned}&\lambda_D dt(\Pi - R_t) + (1-\lambda_D dt)V(u_{t+dt}) - cdt \\ &= \lambda_D dt(\Pi - R_t) + (1-\lambda_D dt)\left(V(u_t) + V'(u_t)\dot{u}_t dt\right) - cdt.\end{aligned}$$

When dt is small enough, we can take $dt^2 \to 0$, then, the above equation can be rewritten as follows:

$$V(u_t) + \{\lambda_D\left(\Pi - R_t - V(u_t)\right) - \phi V'(u_t) - c\}\,dt.$$

Likewise, if the principal chooses the sequential approach for $[t, t+dt)$, the expected payoff can be written as follows:

$$V(u_t) + \left\{\lambda_S(V(u_S^t) - V(u_t)) - \phi V'(u_t) - c\right\}dt.$$

By plugging in $R_t = u_t + \phi/\lambda_D$ and $u_S^t = u_t + \phi/\lambda_S$, the principal determines which approach to take as follows:

$$a_t = \begin{cases} 1, & \text{if } \alpha_D(u_t) > \alpha_S(u_t), \\ 0, & \text{if } \alpha_D(u_t) < \alpha_S(u_t),^{13} \end{cases} \tag{3.3.3}$$

where

$$\begin{aligned}\alpha_D(u) &\equiv \lambda_D\Pi - \lambda_D(V(u) + u) - c - \phi, \\ \alpha_S(u) &\equiv \lambda_S\left(V_S\left(u + \frac{\phi}{\lambda_S}\right) + u + \frac{\phi}{\lambda_S}\right) - \lambda_S(V(u) + u) - c - \phi.\end{aligned}$$

3.3.4 The Optimal Approach at the Deadline

I now explore how the principal chooses the approach at the deadline. Note that $V(0) = 0$ because the contract is terminated when the agent's promised utility is 0.

[13] If equality holds, the principal is indifferent between the two approaches.

Then, at the deadline, the principal chooses which approach to take by comparing the following two values:

$$\begin{aligned}
\alpha_D(0) &= \lambda_D \Pi, \\
\alpha_S(0) &= \lambda_S \left(V_S\left(\frac{\phi}{\lambda_S}\right) + \frac{\phi}{\lambda_S} \right) = \lambda_S \left(\Pi - \frac{c}{\lambda_S} \right) (1 - e^{-1}) \\
&= (1 - e^{-1})(\eta + 1)\lambda_D \Pi - (1 - e^{-1})c.
\end{aligned}$$

This can be interpreted as follows. If the direct approach is chosen, the project is completed at rate λ_D, and the contract is terminated otherwise. If the sequential approach is chosen, the subproject is completed at rate λ_S. If the subproject is completed, the deadline will be extended by $1/\lambda_S$ and the agent will try to complete the remaining subproject at rate λ_S. This deadline extension has a mixed effect: (i) it may increase the chance of project completion if $(1 - e^{-1})(\eta + 1)\lambda_D > \lambda_D$, or equivalently, $\eta > 1/(e - 1)$; (ii) it will incur running cost ($(1 - e^{-1})c$ in expectation) which the principal would not pay under the direct approach. Based on this comparison, the principal chooses which approach to take at the deadline. First, when $\eta \leqslant 1/(e - 1)$, due to high efficiency loss from the project breakdown, the deadline extension does not even deliver a higher chance of project completion, thus, the direct approach will be preferred at the deadline, i.e., $\alpha_D(0) > \alpha_S(0)$ always holds. Second, when $\eta > 1/(e - 1)$, the principal needs to take into account the benefit and the cost from the deadline extension. In this case, we have

$$\alpha_D(0) > \alpha_S(0) \quad \Longleftrightarrow \quad \Pi < \frac{e - 1}{\eta(e - 1) - 1} \cdot \frac{c}{\lambda_D} \equiv \Pi_S(\eta). \tag{3.3.4}$$

Therefore, at the deadline, the principal would split the project if and only if the efficiency loss from project breakdown is low ($\eta > 1/(e - 1)$) and the project return is high ($\Pi \geqslant \Pi_S(\eta)$).

3.4 Optimal Contracts under No Efficiency Loss

In this section, I characterize the optimal contract under the assumption that there is no efficiency loss from breaking down the project, i.e., $\eta = 1$.

If the principal uses the sequential approach, she can monitor the agent's intermediate progress. On the other hand, if the principal chooses the direct approach, the principal cannot observe progress until the main project is done. Therefore, the sequential approach has a comparative advantage in supervising the agent and can potentially lessen the moral hazard problem. Moreover, in this case, the sequential approach is even more attractive due to no efficiency loss. However, as argued in the previous section, when the deadline is close by, the direct approach can be more appealing despite the monitoring advantage.

Based on this intuition, I construct three types of minimum-incentive contracts which involve at most one switch of the employed approach.

Definition 3.4.1. (a) A contract is called a *direct-only contract with a deadline T* if (i) the direct approach is employed up to the deadline T, (ii) the reward upon project completion at time t is $R_t = \phi(T - t + 1/\lambda_D)$, and (iii) the contract is terminated if the project is not completed by the deadline T;

(b) A contract is called a *sequential-only contract with a deadline T* if (i) the sequential approach is employed up to the deadline T, (ii) the contract is terminated if the subproject is not completed by the deadline T, (iii) when the subproject is completed before T, the contract is extended to $T + 1/\lambda_S$, i.e., the contract is terminated if the project is not completed by $T + 1/\lambda_S$, and (iv) the reward upon project completion at time t is $R_t = \phi(T - t + 2/\lambda_S)$;

(c) A contract is called a *contract with a switch from the sequential approach to the direct approach at S and a deadline T* if (i) the sequential approach is employed up to the intermediate deadline S, (ii) when the subproject is completed before S, the contract is extended to $T+1/\lambda_S$ and the reward upon project completion at time t is $R_t = \phi(T-t+2/\lambda_S)$, (iii) if the subproject is not completed by S, the direct approach is employed up to the deadline T and the reward upon project completion at time t is $R_t = \phi(T-t+1/\lambda_D)$, and (iv) the contract is terminated if the project is not completed by the deadline T.

In the following proposition, I derive the benchmark value functions that can be implemented by the above contracts. The proof is relegated to Appendix B.2.3.

Proposition 3.4.1. *The following statements hold:*

(a) A direct-only contract with the deadline u/ϕ implements a pair of expected payoffs of the principal and the agent $(V^d(u), u)$ where

$$V^d(u) = \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{-\frac{\lambda_D}{\phi}u}\right) - u. \tag{3.4.1}$$

(b) When $0 \leqslant u_1 \leqslant u$, a contract with a switch from the sequential approach to the direct approach at $(u-u_1)/\phi$ and the deadline u/ϕ implements $(V^{ds}(u|u_1), u)$ where

$$\begin{aligned} V^{ds}(u|u_1) = {} & \left(\Pi - \frac{2c}{\lambda_S}\right)\left(1 - e^{\frac{\lambda_S}{\phi}(u_1-u)}\right) + (V^d(u_1)+u_1)e^{\frac{\lambda_S}{\phi}(u_1-u)} \\ & - \left(\Pi - \frac{c}{\lambda_S}\right)\frac{\lambda_S}{\phi}(u-u_1)e^{-\frac{\lambda_S}{\phi}u-1} - u. \end{aligned} \tag{3.4.2}$$

(c) A sequential-only contract with the deadline u/ϕ implements $(V^{ds}(u|0), u)$.

Next, I show that the principal's value function can be characterized by the above benchmark value functions. From (3.3.4), the sequential approach is preferred at the deadline if and only if $\Pi \geqslant \Pi_S(1)$. In this case, we can guess that the principal will use a sequential-only contract, i.e., the principal's value function will take the form $V^{ds}(u|0)$. On the other hand, if Π is less than $\Pi_S(1)$, we can also guess that there exists a cutoff promised utility level $u_1 > 0$ such that the direct approach is employed for $0 \leqslant u < u_1$ and the sequential approach is employed for $u > u_1$. That is, the value function would take forms of (3.4.1) when $0 \leqslant u < u_1$ and (3.4.2) when $u > u_1$. The following proposition confirms the conjecture. The proof is relegated to Section B.4.

Proposition 3.4.2. *Suppose that η is equal to 1. Then, the principal's value function is characterized as follows.*

(a) When $\Pi_S(1) > \Pi > \frac{c}{\lambda_D} = \frac{2c}{\lambda_S}$, there exists $u_1 > 0$ such that

$$V(u) = \begin{cases} V^d(u), & \textit{if } 0 \leqslant u \leqslant u_1, \\ V^{ds}(u|u_1), & \textit{if } u_1 < u. \end{cases} \tag{3.4.3}$$

(b) When $\Pi \geqslant \Pi_S(1)$, $V(u) = V^{ds}(u|0)$ for all $u \geqslant 0$.

This result is illustrated in Figure 3.1. I set parameters $\lambda_D = 2,\ \lambda_S = 4,\ c = 1,\ \phi = .5$, then plot the graphs of the value functions for $\Pi = .8,\ .85$, and 1.3. The horizontal axis represents the agent's promised utility u and the vertical axis represents the principal's value function $V(u)$. Recall that $\dot{u} = -\phi$ when the principal employs the minimum incentive contract. Thus, time flows along the horizontal axis from the right to the left. Observe that $\Pi_S(1) \approx 1.196$ and $\Pi < \Pi_S(1)$ in Figure 3.1a and 3.1b. In these figures, the level u_1 represents the promised utility level at which the approach switches. If u is in the blue region, i.e., below u_1, it is optimal for the principal to employ the direct approach ($a = 1$). If u is in the red region, i.e., above u_1, it is optimal for the principal to use the sequential approach ($a = 0$). Also note that $\Pi > \Pi_S(1)$ in Figure 3.1c. Thus, the sequential approach is always chosen over the direct approach.

(a) $\Pi = .8,\ \eta = 1$ (b) $\Pi = .85,\ \eta = 1$ (c) $\Pi = 1.3,\ \eta = 1$

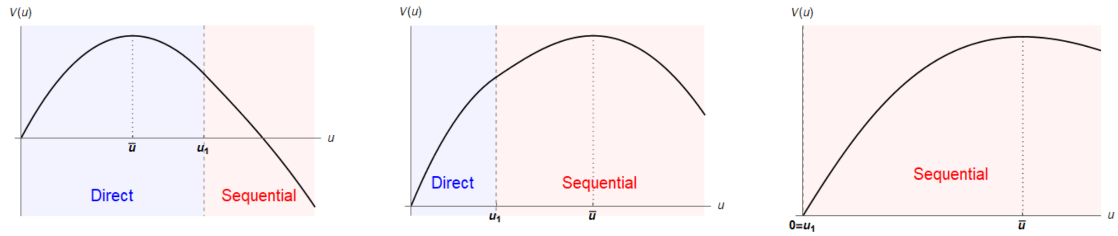


FIGURE 3.1: Value functions when there is no efficiency loss
$(\lambda_D = 2,\ \lambda_S = 4,\ c = 1,\ \phi = .5)$

Now that I have characterized the contract that maximizes the principal's expected payoff under the constraint that the agent's promised utility is equal to u, the next step is to pin down the optimal initial promised utility level. Recall that the optimal contract is the one that implements $(\bar{u}, V(\bar{u}))$ where $\bar{u}$ is the solution to (MP). this will determine where the principal will start the contract in Figure 3.1 and how long the deadline will be.

The feasibility of the project depends on whether $\bar{u}$ is greater than or equal to 0. When $\bar{u}$ is equal to zero, the principal's expected payoff is maximized at $u = 0$. In this case, it is optimal for the principal not to initiate the contract in the first place, i.e., the project is infeasible. This occurs when the project return is very low. In Appendix B.3.1, for any efficiency level η, I show that the project is feasible if and only if Π is greater than $\Pi_F \equiv (c + \phi)/\lambda_D$. Intuitively, at the deadline, a feasible project has to be profitable at least with the direct approach: $0 < \alpha_D(0) = \lambda_D \Pi - c - \phi$, or equivalently, $\Pi > \Pi_F$.

Next, the form of the optimal contract depends on whether $\bar{u}$ is greater than u_1 or not. For example, the value functions in Figure 3.1a and 3.1b both involve a switching point u_1, however, $\bar{u}$ is greater than u_1 in Figure 3.1a and less than u_1 in Figure 3.1b. Thus, the optimal contracts are a direct-only contract in Figure 3.1a and a contract with a switch from the sequential approach to the direct approach in Figure 3.1b. In Appendix B.3.2, I show that there exists a threshold $\Pi_D(1)$ such that $\bar{u} > u_1$ if and only if $\Pi > \Pi_D(1)$ (Lemma B.3.1).

Last but not least, as argued in Proposition 3.4.2, the form of the optimal contract depends on whether Π is greater than $\Pi_S(1)$ or not. This determines whether the direct approach is ever employed in the optimal contract. Based on the above arguments, the following theorem characterizes the optimal contract and shows that the project return Π determines which case will be applied. The formal proof is presented in Appendix B.3.3.

Theorem 3.4.1. *Suppose that there is no efficiency loss from monitoring ($\eta = 1$). There exists a threshold $\Pi_D(1)$ such that $\Pi_S(1) > \Pi_D(1) > \Pi_F \equiv (c+\phi)/\lambda_D$ and the optimal contract is determined as follows:*

(a) when $\Pi \leqslant \Pi_F$, the project is infeasible;

(b) when $\Pi_D(1) \geqslant \Pi > \Pi_F$, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with a deadline $\bar{u}/\phi$;

(c) when $\Pi_S(1) > \Pi > \Pi_D(1)$, there exists $u_1 \in (0, \bar{u})$ such that $(\bar{u}, V(\bar{u}))$ is implemented by a contract with a switch from the sequential approach to the direct approach at $(\bar{u} - u_1)/\phi$ and a deadline $\bar{u}/\phi$;

(d) when $\Pi \geqslant \Pi_S(1)$, $(\bar{u}, V(\bar{u}))$ is implemented by a sequential-only contract with a deadline $\bar{u}/\phi$.

This theorem provides a clear interpretation of the form of the optimal contract with respect to the project return. When the return is very low ($\Pi \leqslant \Pi_F$), the project is not feasible and it is optimal not to initiate the contract. As the return grows and Π is now in $(\Pi_F, \Pi_D(1))$, the project becomes feasible and the principal employs the direct approach near the deadline. In addition, the optimal length of the contract is too short to employ the sequential approach, thus, the direct-only contract is optimal. Next, when Π is in $(\Pi_D(1), \Pi_S(1))$, the optimal length of the contract is long enough to begin the contract with the sequential approach. In addition, the return is not too high so the direct approach will be preferred near the deadline. Thus, the optimal contract involves a switch from the sequential approach to the direct approach. Last, when Π is greater than $\Pi_S(1)$, the sequential approach is preferred even at the deadline, thus, the sequential-only contract is optimal.

Remark 6. A mixture of contracts also generates another contract. For example, a contract with a soft deadline—randomly terminating the agent after reaching the soft deadline—as in Green and Taylor (2016a) can be represented by a mixture of two contracts defined here. However, a mixed contract cannot improve the one characterized above. This follows because the value functions derived in Proposition 3.4.2 are shown to be concave (see Appendix B.4.3). Consider a set of contracts $\{\Gamma_i\}_{1\leqslant i\leqslant n}$ where the agent's expected utility of Γ_i is u_i and the weight is w_i with $\sum_{i=1}^n w_i = 1$ and $\sum_{i=1}^n u_i w_i = u$. The principal's expected payoff from this mixture is $\sum_{i=1}^n w_i P_0(\Gamma_i)$ and the agent's expected utility is u. By concavity, we have $V(u) \geqslant \sum_{i=1}^n w_i V(u_i)$. In addition, we have $V(u_i) \geqslant P_0(\Gamma_i)$ for all $1 \leqslant i \leqslant n$ because $V(u_i)$ is the principal's maximized expected profit given that the agent's expected payoff is u_i. Thus, $V(u)$ is greater than or equal to the expected payoff of the mixed contract. Hence, any mixed contract cannot improve the characterized contract.

3.5 Optimal Contracts under Efficiency Loss

I now introduce an efficiency loss from breaking down the project, that is, η is less than 1. In this case, we need to consider efficiency as another economic force determining the optimal contract in addition to monitoring and deadlines. I begin by showing that the value function from the previous section cannot be supported as a value function in the efficiency loss case when the promised utility is high enough. Assume the contrary, i.e., $V^{ds}(\cdot|u_1)$ is the principal's value function for all $u \geqslant u_1$. Note that $\lim_{u\to\infty} V^{ds}(u|u_1)+u = \Pi-\frac{2c}{\lambda_S}$ and $\lim_{u\to\infty} V_S\left(u+\frac{\phi}{\lambda_S}\right)+u+\frac{\phi}{\lambda_S} = \Pi-\frac{c}{\lambda_S}$. Also observe that

$$\lim_{u\to\infty} \alpha_D(u) = \lambda_D\Pi - \lambda_D\left(\Pi - \frac{2c}{\lambda_S}\right) - (c+\phi) = \frac{1-\eta}{1+\eta}c - \phi,$$

$$\lim_{u\to\infty} \alpha_S(u) = \lambda_S\left(\Pi - \frac{c}{\lambda_S}\right) - \lambda_S\left(\Pi - \frac{2c}{\lambda_S}\right) - (c+\phi) = -\phi.$$

From $\eta < 1$, $\alpha_D(u) > \alpha_S(u)$ for some high enough u, i.e., the (more efficient) direct approach is always preferred if u is high enough. Thus, $V^{ds}(u|u_1)$ cannot constitute the value function.

Intuitively, for longer time horizons, the sum of expected payoffs for both players from the contract converges to the first-best contract, that is, efficiency determines which approach should be chosen. Since we focus on the case where the sequential approach is less efficient than the direct approach, the principal would choose the direct approach when the deadline is far off. Based on this intuition, I construct a contract that provides a minimum incentive not to shirk and potentially has two switches from the direct approach to the sequential approach, then to the direct approach again.

Definition 3.5.1. A contract is called a*contract with two switches at S_1 and S_2 and a deadline T* if (i) the direct approach is employed up to the first intermediate deadline S_1 and the reward upon project completion is $R_t = \phi(T-t+1/\lambda_D)$, (ii) if the project is not completed by S_1, the sequential approach is employed up to the second intermediate deadline S_2, (iii) when the subproject is completed, the deadline is extended to $T+1/\lambda_S$ and the reward upon project completion at time t is $R_t = \phi(T-t+2/\lambda_S)$, (iv) if the subproject is not completed by S_2, the direct approach is employed up to the deadline T and the reward upon project completion at time t is $R_t = \phi(T - t + 1/\lambda_D)$, (v) the contract is terminated if the project is not completed by the deadline T.

The following proposition derives another benchmark value function that can be implemented by the above contract. The proof is relegated to Appendix B.2.3.

Proposition 3.5.1. *When $0 \leqslant u_1 < u_2 \leqslant u$, a contract with two switches at $(u - u_2)/\phi$ and $(u - u_1)/\phi$ and the deadline u/ϕ implements $(V^{dsd}(u|u_1, u_2), u)$ where*

$$V^{dsd}(u|u_1, u_2) = \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{\frac{\lambda_D}{\phi}(u_2 - u)}\right) + (V^{ds}(u_2|u_1) + u_2)e^{\frac{\lambda_D}{\phi}(u_2 - u)} - u. \tag{3.5.1}$$

As a next step, I show that the principal's value function can be characterized by the combination of (3.4.1), (3.4.2) and (3.5.1). Recall that the sequential approach is preferred at the deadline if and only if $\eta > 1/(e-1)$ and $\Pi \geqslant \Pi_S(\eta)$ (Section 3.3.4). In this case, for a low level of the agent's promised utility u, the principal's value function will take the form of $V^{ds}(u|0)$. However, as argued above, when u is large enough, the direct approach will be employed. Then, we can guess that there exists a cutoff promised utility level $u_2 > 0$ such that the principal's value function would take a form of $V^{ds}(u|0)$ for $0 \leqslant u < u_2$ and $V^{dsd}(u|0, u_2)$ for $u > u_2$.

In the remaining cases, the direct approach is preferred at the deadline. Unlike in the no efficiency loss case, due to the efficiency loss from the project breakdown, it is possible that the sequential approach is never chosen in the value function characterization. Specifically, there exists $\Pi_M(\eta) \geqslant 2c/\lambda_S$ such that the sequential approach is never employed if and only if $\Pi \leqslant \Pi_M(\eta)$ (Lemma B.4.2). Thus, when Π is less than or equal to $\Pi_M(\eta)$, the principal's value function is $V^d(u)$. On the other hand, when Π is greater than $\Pi_M(\eta)$, the sequential approach will be employed at some point, but the direct approach is preferred when u is very low (due to the

(a) $\Pi = .8,\ \eta = .9$ (b) $\Pi = .95,\ \eta = .9$ (c) $\Pi = 2,\ \eta = .9$

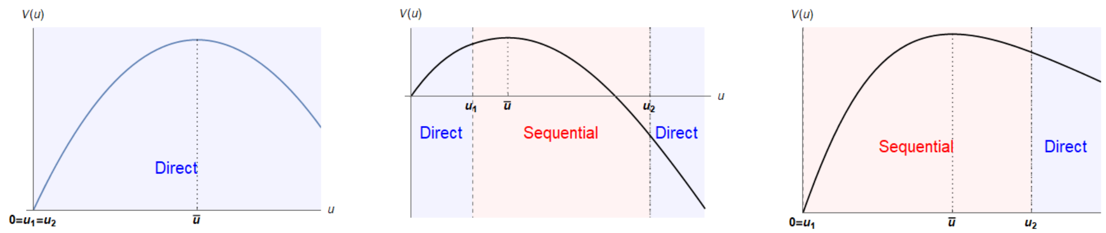


FIGURE 3.2: Value functions when the efficiency loss is small
$(\lambda_D = 2,\ \lambda_S = 3.8,\ c = 1,\ \phi = .5)$

deadline) or very high (due to efficiency). Therefore, we can guess that there exist $u_2 > u_1 > 0$ such that the value function takes a form of $V^d(u)$ when $u < u_1$, $V^{ds}(u|u_1)$ when $u_1 < u < u_2$, and $V^{dsd}(u|u_1, u_2)$ when $u_2 < u$. The following proposition confirms the conjecture. The proof is relegated to Section B.4.

Proposition 3.5.2. *Suppose that η is less than 1. Then, there exists $\Pi_M(\eta)$ such that the following statements hold.*

(a) When $\Pi_M(\eta) \geqslant \Pi > \frac{2c}{\lambda_S}$, $V(u) = V^d(u)$ for all $u \geqslant 0$.

(b) Suppose that one of the following hold: (i) $\eta \leqslant 1/(e-1)$ and $\Pi > \Pi_M(\eta)$; or (ii) $1/(e-1) < \eta < 1$ and $\Pi \in (\Pi_M(\eta), \Pi_S(\eta))$. There exist $u_2 > u_1 > 0$ such that

$$V(u) = \begin{cases} V^d(u), & \text{if } 0 \leqslant u \leqslant u_1, \\ V^{ds}(u|u_1), & \text{if } u_1 < u \leqslant u_2, \\ V^{dsd}(u|u_1, u_2), & \text{if } u_2 < u. \end{cases} \tag{3.5.2}$$

(c) When $1/(e-1) < \eta < 1$ and $\Pi \geqslant \Pi_S(\eta)$, there exists $u_2 > 0$ such that

$$V(u) = \begin{cases} V^{ds}(u|0), & \text{if } 0 \leqslant u \leqslant u_2, \\ V^{dsd}(u|0, u_2), & \text{if } u_2 < u. \end{cases} \tag{3.5.3}$$

This result is illustrated in the graphs in Figure 3.2 and 3.3. I set parameters $\lambda_D = 2$, $c = 1$, $\phi = .5$, and $\lambda_S = 3.8$ (or equivalently $\eta = .9$) for Figure 3.2 and $\lambda_S = 3.1$ (or equivalently $\eta = .55$) for Figure 3.3. Note that $.9 > 1/(e-1) \approx .582$. Then, there are three different cases for Figure 3.2: (a) when the project return is low, the sequential approach is never employed in the characterized value function (Figure 3.2a); (b) when the project return is intermediate, the sequential approach is employed for intermediate u ($u \in (u_1, u_2)$) and the direct approach is chosen for high and low u ($u > u_2$ or $u < u_1$) (Figure 3.2b); (c) when the project return is high, the sequential approach is utilized for low u ($u < u_2$) and the direct approach is employed for high u ($u > u_2$) (Figure 3.2c). Next, note that $.55 < 1/(e-1)$ and the third case in Proposition 3.5.2 will not be applied. Thus, there are two cases for Figure 3.3: (a) when the project return is low, the sequential approach is never employed in the characterized value function (Figure 3.3a); (b) when the project return is intermediate, the sequential approach is employed for intermediate u ($u \in (u_1, u_2)$) and the direct approach is chosen for high and low u ($u > u_2$ or $u < u_1$) (Figure 3.3b).

To obtain more precise results, we need to compare two switching points u_1 and u_2 (if they exist) with the profit maximizing promised utility level $\bar{u}$ as we did in the no efficiency loss case. In the efficiency loss case, the type of the optimal contract depends not only on Π but also on η. For the rest of this section, I characterize optimal contracts for two cases: (i) when η is above $\max\{\sqrt{c/(c+\phi)},\ 1/(e-1)\}$, i.e., the efficiency loss is small; (ii) when η is below $\min\{1/(e-1),\ c/(c+\phi)\}$, i.e., the efficiency loss is large.[14]

Theorem 3.5.1. *Suppose that η is greater than $\sqrt{c/(c+\phi)}$ and $1/(e-1)$, i.e., the*

[14] These cases do not cover the case where the efficiency loss is intermediate. In that case, the form of the optimal contract depends highly on parameter values η and Π and there are many subcases to consider.

(a) $\Pi = .8,\ \eta = .55$ (b) $\Pi = 4,\ \eta = .55$

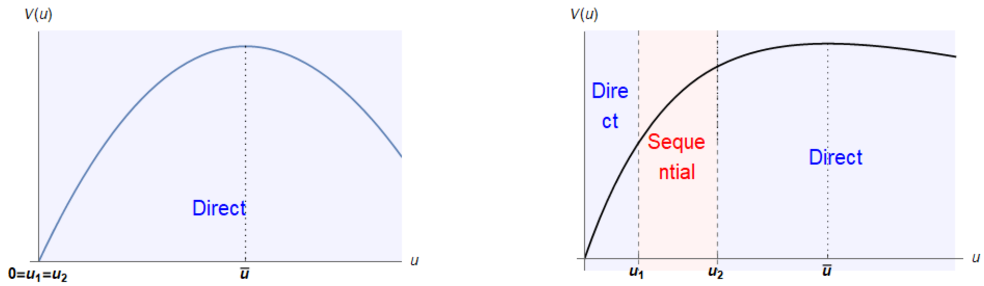


FIGURE 3.3: Value functions when the efficiency loss is large
($\lambda_D = 2,\ \lambda_S = 3.1,\ c = 1,\ \phi = .5$)

efficiency loss from monitoring is small. There exist thresholds $\Pi_D(\eta)$ *and* $\Pi_S(\eta)$ *with* $\Pi_S(\eta) > \Pi_D(\eta) > \Pi_F = (c+\phi)/\lambda_D$ *such that the optimal contract is determined as follows:*

(a) when $\Pi \leqslant \Pi_F$*, the project is infeasible;*

(b) when $\Pi_D(\eta) \geqslant \Pi > \Pi_F$*,* $(\bar{u}, V(\bar{u}))$ *is implemented by a direct-only contract with a deadline* $\bar{u}/\phi$*;*

(c) when $\Pi_S(\eta) \geqslant \Pi > \Pi_D(\eta)$*, there exists* $u_1 \in [0, \bar{u}]$ *such that* $(\bar{u}, V(\bar{u}))$ *is implemented by a contract with a switch from the sequential approach to the direct approach at* $(\bar{u} - u_1)/\phi$ *and a deadline* $\bar{u}/\phi$*;*

(d) when $\Pi > \Pi_S(\eta)$*,* $(\bar{u}, V(\bar{u}))$ *is implemented by a sequential-only contract with a deadline* $\bar{u}/\phi$*.*

This result is similar to Theorem 3.4.1. Intuitively, when the efficiency loss is small, it only induces the principal to choose the direct approach when the deadline is fairly far away and it may go further than the optimal length of the contract. Thus, the switch from the direct approach to the sequential approach may not occur in the optimal contract. The condition that η is greater than $\sqrt{c/(c+\phi)}$ ensures that the switching point u_2 would be always above the optimal initial promised utility level $\bar{u}$ (Figure 3.2b and 3.2c), or the switching point does not exist at all (Figure 3.2a).[15] Also, the condition that η is greater than $1/(e-1)$ makes the principal prefer the sequential approach even at the deadline when Π is large enough (Figure 3.2c).

These results confirm that the findings from the no efficiency loss case are robust to the introduction of small efficiency costs. However, – as might be expected – the results no longer hold when the efficiency loss is large.

Theorem 3.5.2. *Suppose that η is less than $c/(c+\phi)$ and $1/(e-1)$, i.e., the efficiency loss from monitoring is large. There exists a threshold $\Pi_M(\eta)$ with $\Pi_M(\eta) > \Pi_F = (c+\phi)/\lambda_D$ such that the optimal contract is determined as follows:*

(a) when $\Pi \leqslant \Pi_F$, the project is infeasible;

(b) when $\Pi_M(\eta) \geqslant \Pi > \Pi_F$, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with a deadline $\bar{u}/\phi$;

(c) when $\Pi > \Pi_M(\eta)$, there exist u_1 and u_2 such that $0 < u_1 < u_2 < \bar{u}$ and $(\bar{u}, V(\bar{u}))$ is implemented by a contract with two switches at $(\bar{u}-u_2)/\phi$ and $(\bar{u}-u_1)/\phi$ and a deadline $\bar{u}/\phi$.

This result is significantly different from the previous ones in that the optimal contract involves either zero or two switches. Intuitively, when there is a large efficiency loss from monitoring, the principal would choose the direct approach for the majority of the time and choose the sequential approach for only short periods. If Π is not that large, the principal would not employ the sequential approach at all (Figure 3.3a). However, if Π is large enough, the principal would choose the sequential approach in the middle of the optimal contract (Figure 3.3b). To get this result, we need to show that $\bar{u} > u_2 > u_1 > 0$. The condition that η is less than

[15] Refer to Lemma B.3.3 for details.

$c/(c+\phi)$ ensures that u_2 is always below $\bar{u}$.[16] Also, the condition that η is less than $1/(e-1)$ makes the principal choose the direct approach at the deadline no matter what Π is.[17]

Intuitively, the principal generally prefers the direct approach since there is a large efficiency loss from the sequential one. Nevertheless, when Π is large enough, the principal may take advantage of the monitoring benefit by choosing the sequential approach. If the principal decides to monitor at some point, it is optimal to monitor in the middle of the contract. This is because efficiency outweighs monitoring at the beginning of the contract and the deadline effect outweighs monitoring at the end of the contract. Hence, the optimal contract involves two switches when Π is large.

For a very high-return project, the theorem illustrates that a type of contract involving all three economic forces is optimal. At the beginning of the contract, the principal chooses the direct approach because it is more efficient (i.e., efficiency is initially the dominant concern). When the success is not delivered by a specified time, the principal begins to monitor the agent more closely by switching to the sequential approach (i.e., monitoring becomes the primary concern). She extends the deadline if the agent makes intermediate progress, but if he does not make progress before the deadline is near, the principal switches back to the direct approach in a "last-ditch" attempt at getting the job done (i.e., the deadline effect becomes the preeminent motivation).

[16] Refer to Lemma B.3.2 for details.

[17] Refer to Proposition 3.5.2 for details.

4

Strategic Concealment in Innovation Races

4.1 Introduction

In the course of research and development (R&D), firms often discover interim knowledge that brings them closer to successfully producing a final innovation. When multiple firms race towards such innovation, a firm's optimal R&D strategy is likely to be influenced by the information about whether its rivals have made intermediate breakthroughs. Thus, a firm may want to conceal intermediate discoveries in order to hinder its rivals from adjusting their R&D strategies. On the other hand, it may prefer to disclose an intermediate discovery because this can open the opportunity for monetization via licensing the technological breakthrough. In this paper, we introduce and study an innovation race model that captures the tradeoffs between licensing and concealing interim discoveries and characterize firms' equilibrium behavior.

We consider a situation where two firms race towards developing an innovative product, such as a COVID-19 vaccine or a full self-driving (FSD) vehicle. The first firm to develop the product receives a reward (e.g., a transitory flow of monopoly profit) and the other firm does not. At each point in time, the firms allocate their limited resources between two routes for developing the product and incur constant flow costs. One route is to conduct basic *research* to discover a new technology that does not directly deliver the product but makes developing it faster, e.g., messenger RNA (mRNA) or light detection and ranging (LIDAR) technology.[1][2] This route requires two breakthroughs: discovering the new technology and developing the product with it. The other route is to *develop* the product with a currently available but slow technology, namely the incumbent technology. For example, the viral vector method for developing a COVID vaccine and the camera-based vision technology for developing an FSD vehicle can be considered incumbent technologies.[3][4] This path requires a single breakthrough but the arrival rate is relatively low. We assume that the path with the new technology is more efficient: the total expected completion

[1] The mRNA technology was not utilized in practice before the COVID-19 outbreak. Thus, pharmaceutical firms had to first acquire basic knowledge in order to employ this new methodology. The advantage of possessing this intermediate technology is that firms can develop vaccines in a laboratory by using readily available materials. Hence vaccines can be developed faster with mRNA technology than with older methods. Moderna and Pfizer-BioNTech utilize mRNA technology to develop COVID-19 vaccines. For more information, see the web page of the Centers for Disease Control and Prevention (CDC): `https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/mrna.html`.

[2] LIDAR is a laser radar that can provide extensive and reliable information surrounding a vehicle including an object's distance, size, position, and velocity if it is moving. Most FSD vehicle developers including Waymo—formerly the Google self-driving car project—use LIDAR combined with cameras. The main drawback of LIDAR is its current high cost. Thus, to develop a commercializable FSD vehicle, firms first need to discover a way to make LIDAR less expensive. Once LIDAR becomes affordable, it will be relatively easy to develop a commercializable FSD vehicle. In this sense, successfully developing an FSD vehicle with the LIDAR technology can be understood as a route requiring two breakthroughs.

[3] The viral vector technology was used during recent disease outbreaks including the 2014-2016 Ebola outbreak in West Africa. Many pharmaceutical firms had access to this methodology when the COVID-19 outbreak began. Indeed, this technology was utilized to develop COVID-19 vaccines by Oxford-AstraZeneca and Janssen (Johnson&Johnson). For more information, see the web page of the CDC: `https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/viralvector.html`.

[4] Unlike other companies, Tesla's approach towards developing an FSD vehicle is to use only cameras without LIDAR (Templeton, 2019). Since camera technology is already very cheap, no cost-saving breakthrough is needed to implement it. However, the quality of information attained from cameras is inferior to that attained from LIDAR, thus it will take more time to develop an FSD vehicle utilizing only cameras.

time of doing research for the new technology and developing the product with it is shorter than that of developing with the incumbent strategy. Thus, the socially efficient policy is to have both firms allocate all their resources to research, and once one of them discovers the new technology, have it share the breakthrough with the other firm to prevent duplication of research costs.

We investigate three different settings in the context of this framework. First, we consider the case where it is public information whether a firm has discovered the new technology or not. In this setting, a firm can condition its strategy not only on its own technological breakthrough but also on its rival's progress. We show that there exists a unique equilibrium and its form is determined by the relative efficiency of the new technology. The efficiency measure is defined to be inversely proportional to the expected total completion time of the path with the new technology, i.e., doing research is more attractive when efficiency is high. It is shown that when efficiency is extreme (high or low), a firm's equilibrium strategy does not depend on its rival's progress. Specifically, when the new technology is highly efficient, both firms allocate all their resources to research (i.e., perform research only); and when the new technology is not much more efficient, both firms allocate all their resources to development (i.e., develop with the incumbent technology only) regardless of their rival's status. On the contrary, when efficiency is intermediate, the equilibrium strategy of each firm does depend on its rival's progress. In this case, both firms begin by conducting research, but once one firm makes the intermediate technological breakthrough, the other switches to developing with the incumbent technology, namely it pursues a *fall-back* strategy.

Next, we analyze the setting where technological discoveries are private information, i.e., a firm cannot observe its rivals' technological progress. As in the public information setting, when efficiency is high, each firm conducts research until it succeeds or its rival produces the final innovation. Similarly, when efficiency is low, both firms endeavor to develop with the incumbent technology. This invariance occurs because, in the extreme cases of very high and very low efficiency, firms do not use the information about their rival's progress even when it is observable. However, in the case of intermediate efficiency, the firms cannot use the fall-back strategy as in the public information setting since they are no longer able to make their resource allocations contingent on their rivals' state of technology. Instead, their resource allocations must depend on their 'beliefs' about their rivals' progress. We characterize the unique symmetric equilibrium that is Markov with respect to these beliefs. The equilibrium strategy has a cutoff structure: firms conduct research exclusively up to a certain date (belief), then they start allocating their resources between developing with the incumbent technology and researching the new one, namely they employ a *stationary fall-back* strategy. The most intriguing feature of this equilibrium is that beliefs remain constant once the allocation of resources to development begins. This stationarity derives from two conflicting effects in the belief evolution. First, as time passes, it becomes more likely that one's rival has found the new technology (the *duration effect*). On the other hand, the lack of one's rival producing the final innovation (which is publically observable) implies that it is less likely that the new technology has been discovered (the *still-in-the-race effect*).

Last, we extend the private information setting by allowing firms to protect their discoveries by using either a *patent* or a *trade secret*. First, when a firm treats the new technology as a trade secret, it conceals the discovery, i.e., its rival still cannot observe its progress. However, this does not prohibit the firm's rival from discovering the new technology independently. Second, when a firm files a patent, it discloses the discovery of the new technology. On the one hand, if its rival has not yet made the technological breakthrough, then the exclusive right to use the new technology is bestowed on the patenting firm. In addition, the patenting firm may *license* the new technology, i.e., it may permit its rival to use the new technology for a fee.

Once the licensee pays the fee, both firms race for the final innovation employing the new technology. On the other hand, if the rival firm has already discovered the new technology, i.e., it was protected as a trade secret. Then, the patenting firm cannot claim the exclusive right—rather, the new technology is now considered common property—and both firms can use it without making transfers.[5][6]

We first show that if a firm files a patent and the rival firm does not possess the new technology, the patenting firm always licenses. Thus, both firms develop the final innovation with the new technology, which is socially efficient. Once a firm files a patent, its rival can only try to develop the product with the old slow technology. Given this, the patenting firm can extract rent from its rival by allowing it to use the new technology for a fee. This is an application of the classical result of Coase (1960) in the sense that the socially efficient outcome can be achieved when the property right of the new technology is given to a firm and trade involves no transaction costs. Therefore, disclosing the new technology implies licensing it.

Finally, we explore whether a firm prefers to disclose or conceal the new technology. We show that this decision crucially depends on the size of the reward of winning the race: when the reward is high, firms may prefer to conceal their discoveries, whereas when the reward is low, they disclose and license them. Intuitively, this is because concealment involves a higher chance of winning the race, which is more attractive when the reward is high. Whereas, disclosure delivers an immediate payment from licensing, which is more appealing when the reward is low. More specifically, when a firm conceals a discovery, its rival does not know whether it possesses the new technology. Thus, per the results from the private information setting,

[5] When a firm files a patent, the firm with the trade secret can dispute the patent based on 35 U.S. Code §273 - Defense to infringement based on prior commercial use.

[6] For more information about trade secrets and patents, see the web page of the World Intellectual Property Organization: `https://www.wipo.int/about-ip/en/`. Also, see Lobel (2013) for examples.

the rival firm continues allocating some of its resources to researching the new technology. This is not desirable for the rival, especially when efficiency of the new technology is intermediate, because if it knew that the other firm already possessed the new technology, then its best response would be to allocate all its resources to development with the incumbent technology (i.e., to employ the fall-back strategy). In this sense, concealing the new technology hinders the rival firm from strategically responding to its discovery.

Concealment is detrimental not only to the rival firm but also to social surplus because it generates duplicate research efforts. This slows down the pace of innovation. On the contrary, the socially efficient outcome could be achieved by disclosing and licensing the new technology. These results on firms' incentives for concealment imply a simple policy intervention. Reducing the reward of winning the race (e.g., weakening the transitory monopoly power in the innovative product market by imposing a tax,) reduces incentives to conceal and promotes licensing, thus speeding up the pace of innovation.

Related Literature

This paper primarily contributes to the literature on patent vs. secrecy by introducing a novel incentive to conceal a firm's discovery: hindering its rival's strategic response. Previous studies mainly focused on the limited protection power of patents. For example, the seminal article by Horstmann et al. (1985) posits that "patent coverage may not exclude profitable imitation." Thus, in their framework, the main reason why a firm may choose secrecy over a patent is not to be imitated.[7] Another limitation of a patent is that it expires in a finite time. For instance, Denicolò and Franzoni (2004) consider a framework where a patent gives the patenting firm

[7] Many subsequent papers study the imitation threat and potential patent infringement, e.g., Gallini (1992); Takalo (1998); Anton and Yao (2004); Kultti et al. (2007); Kwon (2012); Zhang (2012).

monopoly power only for a certain period of time (and no profit after expiration), whereas secrecy can give indefinite monopoly power to a firm but it can be leaked or duplicated by a rival with some probability. On the contrary, in this paper, we abstract from the restrictions of patents and focus analysis on the potential advantages of concealment.

Another hallmark of this paper is its consideration of 'interim' discoveries. Therefore, it is naturally related to the literature on licensing of interim R&D knowledge, e.g., Bhattacharya et al. (1992); d'Aspremont et al. (2000); Bhattacharya and Guriev (2006); Spiegel (2008). In these papers it is assumed that firms already know which of them has superior knowledge, i.e., the firm that will license the technology is exogenously given. Unlike in those studies, we allow firms to choose when to license (and even allow them not to license), i.e., the licensing decision is endogenous.

We also contribute to the innovation literature by introducing a model with two characteristics. First, there are different avenues towards innovation: developing with the incumbent technology and doing research for the new technology. Second, one of the paths involves multiple stages: once a firm discovers the new technology, then the firm develops the innovative product with it.

With respect to the first characteristic, there is a recent branch of the literature that studies races where there are different routes to achieve a final objective. Das and Klein (2020) and Akcigit and Liu (2016) study a patent race where two firms compete for a breakthrough and there are two methods to get the breakthrough: a safe method and a risky method. In Das and Klein (2020) the safe method has a known constant arrival intensity while the risky method has an unknown constant arrival intensity. In Akcigit and Liu (2016), instead, the safe method has a known payoff associated with breakthrough arrival, while there is uncertainty about the payoff if the risky method is used. In this paper, firms face no uncertainty about whether the innovation is feasible. Instead, they are uncertain whether their rival possesses the new and faster technology.

The second characteristic, multi-stage innovation, is also widely studied in the literature, e.g., Scotchmer and Green (1990); Denicolò (2000); Green and Taylor (2016b); Song and Zhao (2021). Our paper shares the framework with these in that we use two sequential Poisson discovery processes and ask whether a firm would patent the first discovery or not. A feature setting apart from their works is that there is another path that only requires one but slower breakthrough toward innovation. This feature connects our model to Carnehl and Schneider (2022) and Kim (2022) in the sense that players can choose between a sequential approach—which requires two breakthroughs—and a direct approach, which requires only one breakthrough, but its riskier or slower.[8] Our model mainly differs from theirs in that multiple players compete by choosing between these approaches, whereas Carnehl and Schneider (2022) considers a problem by a single decision maker and Kim (2022) studies a contracting setup between a principal and an agent. In their studies, a key factor for a player to choose the direct approach is a deadline that is either exogenously given or endogenously determined to reduce moral hazard. In contrast to these, a deadline is not involved in our model. Rather, the race with the rival firm may induce a firm to develop with the incumbent technology, which can be considered as a direct approach.

Last, this paper is related to the recent literature on information disclosure in priority races, e.g., Hopenhayn and Squintani (2016); Bobtcheff et al. (2017). In those papers, once a firm makes a breakthrough, the innovation value grows as time passes until one of the firms files a patent. Thus, firms face a tradeoff between disclosing to claim the priority and delaying in order to grow the innovation value. On the

[8] In Carnehl and Schneider (2022), an agent is uncertain whether the direct approach is feasible or not, i.e., this approach is risky. On the other hand, in Kim (2022), there is no uncertainty on the feasibility of the direct approach, but its completion rate is slower than the ones for the sequential approach. In this sense, our framework is closer to Kim (2022).

contrary, in this paper, the value of innovation is fixed and the discovery of the new technology only allows the firm to develop the innovative product faster. Therefore, a firm may delay the disclosure purely to confound the rival's R&D decisions.

Roadmap

We introduce the model in the next section, then characterize equilibria in the private and the public information settings in Section 4.3 and 4.4. In Section 4.5, we extend the private information setting by allowing firms to disclose their discoveries. All proofs appear in the appendix.

4.2 Model

Two risk-neutral firms ($i \in \{A, B\}$) race to develop an *innovative product.* Time is continuous and infinite $t \in [0, \infty)$. Neither firm discounts the future and each owns a unit of resources per unit of time. The innovative good can be developed by either a slower incumbent technology or a faster new technology. At the beginning of the race, both firms have access to the incumbent technology, but not to the new technology. To utilize the new technology, firms need to make a discovery through conducting research.

At each time t, Firm i can use a fraction of its resources either to do 'research' to discover the new technology (σ_t^i) or to 'develop' the innovative product with the incumbent technology ($1 - \sigma_t^i$). With the incumbent technology, the firm develops the innovative product stochastically at rate $\lambda_L(1 - \sigma_t^i)$. From the research, firms discover the new technology stochastically at rate $\mu\sigma_t^i$. Hereafter, we simply denote that at time t, Firm i *develops* when $\sigma_t^i = 0$ and *does research* when $\sigma_t^i = 1$ unless stated otherwise. The new technology does not generate any direct benefit to a firm, but helps the development faster. That is, once the new technology is obtained by a firm, it will develop the product stochastically at rate $\lambda_H > \lambda_L$.

The firm that develops the innovative product first (either with the incumbent technology or the new one) receives a reward worth Π, i.e., it is a winner-takes-all competition.[9] The other firm does not capture any rents from innovating second. Firms pay a flow cost c at each point of time until making the innovation or quitting the race. Whether a firm makes the ultimate innovation is publicly observable, however, a firm cannot observe how its opponent allocates time between attempting to discover the new technology and attempting to develop the innovative product with the old technology.

We fix λ_L and c throughout the paper. To facilitate, we introduce two relevant parameters measuring the efficiency and the relative intensity of the new technology:

$$\eta \equiv \frac{\mathbb{E}[\text{ completion time with the incumbent technology }]}{\mathbb{E}[\text{ total completion time with the new technology }]} = \frac{\frac{1}{\lambda_L}}{\frac{1}{\mu} + \frac{1}{\lambda_H}},$$

$$\delta \equiv \frac{\mathbb{E}[\text{ research completion time with the new technology }]}{\mathbb{E}[\text{ total completion time with the new technology }]} = \frac{\frac{1}{\mu}}{\frac{1}{\mu} + \frac{1}{\lambda_H}}.$$

Most of the results will be presented in terms of these parameters and Π. Note that the more efficient technology will deliver a shorter expected completion time. Thus, η will increase as the new technology becomes more efficient. We assume that the new technology is more efficient than the incumbent technology: $\eta > 1$. Also note that when η is fixed, a higher δ implies that the new technology is more research-intensive (or less development-intensive). This is because when firms try to achieve innovation via the new technology, they are expected to spend more time in research as δ increases. Last, we assume that the incumbent technology is profitable: $\Pi \geqslant c/\lambda_L$ or equivalently $\pi \equiv \lambda_L \Pi / c \geqslant 1$.

[9] A winner-takes-all payoff structure has been commonly used in the innovation race literature, e.g., Loury (1979); Lee and Wilde (1980); Denicolò and Franzoni (2010). Following the literature, we can regard Π as the societal value of having the innovative product, and the first firm that introduces the innovative product becomes the monopolist and captures all the social value, e.g., by using the first-degree price discrimination. In this article, we abstract away from the market after the race and focus on the activities during the race.

We conclude the section by introducing the planner's problem, namely the first-best case. Consider a social planner who is able to control firms' resource allocations and observe progress. In addition, the planner can force a firm to share its technology to the other firm. The goal of the planner is to maximize the joint expected profit. Since there is no discounting in time and one of firms will receive the reward, the planner's problem is equivalent to the minimization problem of the expected completion time of the product development. Then, the planner's optimal solution is to take the more efficient path:

(i) both firms fully allocate their resources to research and the research is completed at rate 2μ;

(ii) if one of the firms discovers the new technology, and it immediately shares the new technology to the other firm, and then the firms develop the innovative product with the new technology—the development is completed at rate $2\lambda_H$.

Hence, the (ex-ante) expected completion time is given as follows:

$$L_{FB} = \frac{1}{2\mu} + \frac{1}{2\lambda_H} = \frac{1}{2\lambda_L \eta}. \tag{4.2.1}$$

4.3 Public Information Setting

We begin by exploring a setting where firms' research progress is public information, i.e., firms can observe whether their opponents have discovered the new technology. In this case, the list of firms that have obtained the new technology is common knowledge and we can regard it as a state: $\omega \in \Omega \equiv \{\{A, B\}, \{A\}, \{B\}, \varnothing\}$. We assume the firms employ Markov strategies, i.e., Firm i's strategy is defined by $\sigma^i : \Omega \to [0, 1]$. Since a firm possessing the new technology derives no value from research, we can restrict attention to the strategies such that $\sigma^i(\omega) = 0$ when $i \in \omega$.

A pair of strategies (σ^A, σ^B) constitutes a Markov perfect equilibrium if, for any state, each firm's strategy is the best response to the opponent's strategy.

Before characterizing the equilibrium of this game, we introduce three benchmark strategies.

Definition 4.3.1. (a) A *research strategy* is such that a firm fully allocates the resources to research regardless of the opponent's technology level ($\sigma^i(\varnothing) = \sigma^i(\{j\}) = 1$).

(b) A *fall-back strategy* is such that (i) a firm fully allocates the resources to doing research when neither firm has obtained the new technology ($\sigma^i(\varnothing) = 1$); (ii) a firm fully allocates the resources to developing with the incumbent technology if its opponent discovers the new technology ($\sigma^i(\{j\}) = 0$).

(c) An *incumbent strategy* is such that a firm fully allocates the resources to developing with the incumbent technology regardless of the opponent's technology level ($\sigma^i(\varnothing) = \sigma^i(\{j\}) = 0$).

The following proposition shows that it is a Markov perfect equilibrium for both firms to simultaneously use one of the above strategies, depending on parameter values. The proof is in Appendix C.1.

Proposition 4.3.1. *Suppose that firms can observe whether their opponents have made a technological breakthrough. Then, the Markov perfect equilibrium is uniquely characterized as follows.*

(a) If $\eta \geqslant \overline{\eta}(\delta) \equiv 1 + \delta$, both firms play the research strategies;

(b) If $\overline{\eta}(\delta) > \eta > \underline{\eta}(\delta) \equiv \frac{1}{2}\left(1 + \sqrt{1 + 4\delta(1-\delta)}\right)$, both firms play the fall-back strategies;

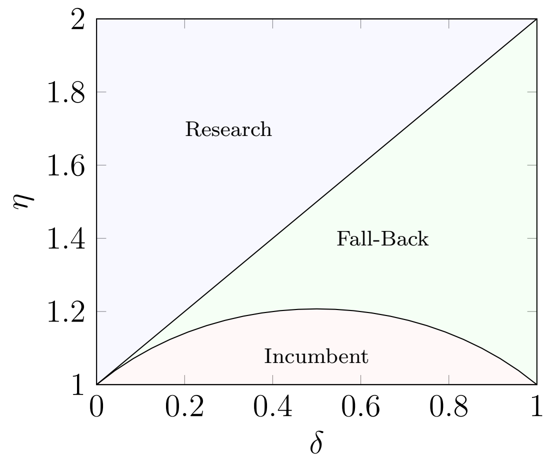


FIGURE 4.1: Markov Perfect Equilibrium under the Public Information Setting

(c) If $\underline{\eta}(\delta) \geqslant \eta$, both firms play the incumbent strategies.

The above proposition provides a clear picture of how the efficiency of the new technology (η) affects the firms' R&D decisions. When the new technology is sufficiently efficient and research is relatively easy ($\eta \geqslant \overline{\eta}(\delta)$), the firms do research regardless of whether their opponent has discovered the new technology. When the new technology is relatively inefficient ($\eta \leqslant \underline{\eta}(\delta)$), the firms do not engage in research at all. In the intermediate case ($\underline{\eta}(\delta) < \eta < \overline{\eta}(\delta)$), the firms' R&D decisions are affected by the opponent's progress: when neither firm has made the technological breakthrough, both firms do research; but once a firm obtains the new technology, the follower—the firm without the new technology—switches to develop with the incumbent technology.

The proposition also shows that the thresholds depend on δ, the relative intensity of the new technology. Figure 4.1 illustrates how these thresholds depend on δ. First, to determine the threshold for the equilibrium with the research strategy, we need to consider the case when one firm (the leader) has discovered the new technology and the other (the follower) has not. Say that Firm i is the follower and Firm j is the leader. Firm i needs to determine whether to follow j ($\sigma^i(\{j\}) = 1$) or to switch to the incumbent technology ($\sigma^i(\{j\}) = 0$). When it is difficult to attain the new technology, the follower would choose to follow only if the new technology is efficient enough. Thus, the threshold for the equilibrium with the research strategy is increasing as δ increases.

Second, why is the threshold for the equilibrium with the incumbent strategy hump-shaped? To answer this question, we need to consider the situation where neither firm yet possesses the new technology. The firms allocate resources by taking into account the difficulty and the advantage of becoming a leader. Consider the case with $\delta < 1/2$, i.e., attaining the new technology is relatively easy, but the leader's advantage is relatively weak (due to low λ_H). In this case, the major determinant is the difficulty of becoming a leader. Fix the efficiency (η) and marginally increase δ. Then it becomes more difficult to attain leadership, which makes the incumbent strategy more attractive. Next, consider the case with $\delta > 1/2$, i.e., attaining the new technology is relatively difficult, but it is more advantageous to become a leader (due to high λ_H). In this case, the major determinant is the leader's advantage. If δ decreases, the leader's advantage decreases, which again makes the incumbent strategy more attractive.

We conclude this section by deriving the expected completion time of developing the product in the public information setting. When $\underline{\eta}(\delta) \geqslant \eta$, the expected completion time is $\frac{1}{2\lambda_L}$ since both firms develop with the incumbent technology. Next, when $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$, the expected time until one of the firms discover new technology is $\frac{1}{2\mu}$. Then, a firm develops with the new technology and the other firm develops with the incumbent technology, thus, the expected completion time from then on is $\frac{1}{\lambda_H+\lambda_L}$. Therefore, the (total) expected completion time is $\frac{1}{2\mu} + \frac{1}{\lambda_H+\lambda_L}$. Last, when $\eta \geqslant \overline{\eta}(\delta)$, unlike in the previous case, the firm without the new technology keeps doing research. Then, the expected time until either the firm with the new technology develops the product or the firm without the new technology discovers it is $\frac{1}{\lambda_H+\mu}$. With the probability $\frac{\mu}{\lambda_H+\mu}$, the firm without the new technology discovers it earlier than the product development, then it takes an additional expected completion time $\frac{1}{2\lambda_H}$. Therefore, the total expected completion time is

$$\frac{1}{2\mu} + \frac{1}{\lambda_H+\mu} + \frac{\mu}{\lambda_H+\mu} \cdot \frac{1}{2\lambda_H} = \frac{1}{2}\left(\frac{1}{\mu} + \frac{1}{\lambda_H} + \frac{1}{\lambda_H+\mu}\right).$$

The expected completion time is summarized as follows:

$$L_{public} = \begin{cases} \frac{1}{2}\left(\frac{1}{\mu} + \frac{1}{\lambda_H} + \frac{1}{\lambda_H+\mu}\right), & \text{if } \eta \geqslant \overline{\eta}(\delta), \\ \frac{1}{2\mu} + \frac{1}{\lambda_H+\lambda_L}, & \text{if } \eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta)), \\ \frac{1}{2\lambda_L}, & \text{if } \eta \leqslant \underline{\eta}(\delta). \end{cases} \tag{4.3.1}$$

4.4 Private Information Setting

Now we assume that firms' research progress is private information, i.e., firms cannot observe whether their opponents have the new technology or not. In this case, the firms can only condition their strategies on their own progress and calendar time t. Again, a firm with the new technology will fully allocate its resources to develop with that technology. Thus, we focus on the dynamic resource allocation problem of a firm that has not discovered the new technology: a strategy for a player can be therefore described by a function $\sigma : \mathbb{R}_+ \to [0,1]$ that represents the allocation at a given time conditional on the new technology not being discovered. Let $\mathcal{S}$ be the set of such strategies.

4.4.1 Evolution of Beliefs and Recursive Formulation

Although firms cannot observe their opponents' technology levels, they do form beliefs about whether their opponents have acquired the new technology. Let p_t^i be

Firm i's belief that Firm j possesses the new technology given that neither firm has made the innovative product by t and that Firm j follows the resource allocation $\sigma^j : \mathbb{R}_+ \to [0,1]$.[10] Since it is common knowledge that neither firm possesses the new technology at the beginning of the race, we have $p_0^i = p_0^j = 0$, which will serve as an initial condition for the belief evolution. The following lemma characterizes how the belief p^i evolves conditional on the absence of innovation. The proof is provided in Appendix C.2.

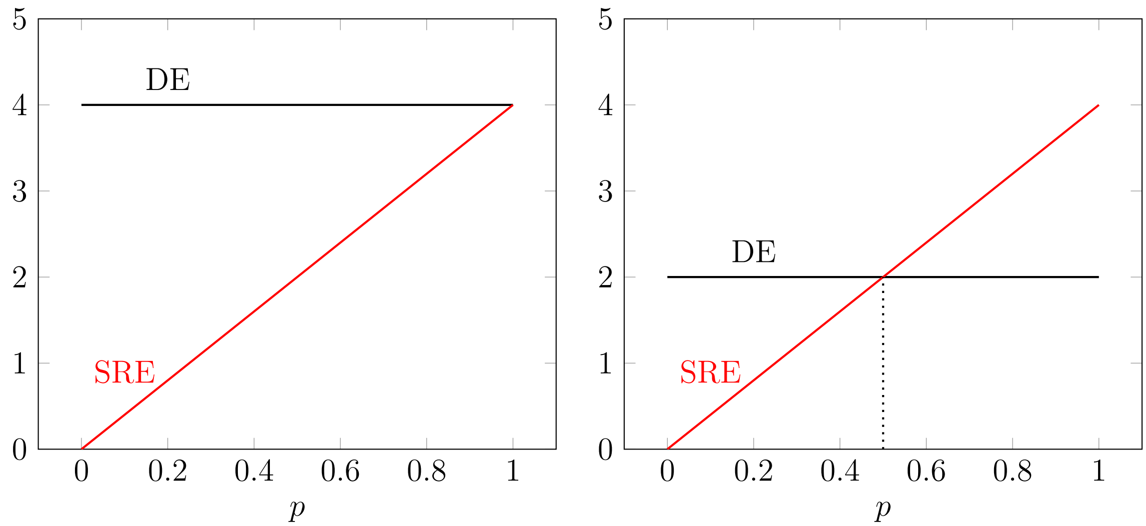
Lemma 4.4.1 (Evolution of Beliefs)**.** *Suppose that Firm i does not observe whether Firm j has made a technological breakthrough and believes that Firm j follows the resource allocation σ^j. Then, in the absence of innovation, Firm i's belief, p^i, evolves via the following differential equation:*

$$-\frac{d}{dt}\log(1-p_t^i) = \frac{\dot{p}_t^i}{1-p_t^i} = \mu \cdot \sigma_t^j - \left\{\lambda_H - (1-\sigma_t^j)\lambda_L\right\} \cdot p_t^i. \tag{4.4.1}$$

The left hand side of (4.4.1) is the opposite of the time derivative for the log belief that Firm j does not possess the new technology. The right hand side of (4.4.1) captures two distinct effects in updating the belief. First, given that Firm j has not yet attained the new technology by time t, the research succeeds at rate $\mu \cdot \sigma_t^j$ and it may raise the belief. The first term of (4.4.1) represents this effect, which we dub the duration effect (DE). On the other hand, the fact that firm j has not produced the innovative good indicates that it is less likely to have the new technology in hand. The second term of (4.4.1) reflects this effect, which we dub the still-in-the-race effect (SRE).[11] Notice that this term is proportional to $\lambda_H - (1-\sigma_t^j)\lambda_L$, which is the rate

[10] It is technically convenient to assume strategies depend on time rather than beliefs about the other firms breakthrough status because it is cumbersome to analyze all off-path beliefs. As we show below, in any symmetric equilibrium there is a unique mapping from calendar time to beliefs.

[11] Similar types of the belief updating can be found in the strategic experimentation literature, e.g., Keller et al. (2005); Bonatti and Hörner (2011). The main difference is that the agents form beliefs about whether the project is good or bad in those papers, whereas the firms form beliefs about whether the rival possesses the new technology or not in our model.



(a) $\sigma = 1$, $\mu = 4$, $\lambda_H = 4$ and $\delta = 1/2$.

(b) $\sigma = 1$, $\mu = 2$, $\lambda_H = 4$ and $\delta = 2/3$.

FIGURE 4.2: Duration Effect (Black) and Still-in-the-Race Effect (Red)

of successful innovation development given the new technology net of that without the new technology.

A natural benchmark is a firm's belief when the opponent fully allocates its resources to research. We characterize this belief in the following lemma. The proof is also in Appendix C.2.

Lemma 4.4.2. *Suppose that Firm i does not observe whether Firm j has made a technological breakthrough and believes that Firm j fully allocates the resources to research up to time T ($\sigma_t^j = 1$ for all $0 \leqslant t \leqslant T$). Then, $p_T^i = q_T$ where*

$$q_T \equiv \frac{\frac{1}{\lambda_H}\left(e^{-\mu T} - e^{-\lambda_H T}\right)}{\frac{1}{\mu}e^{-\mu T} - \frac{1}{\lambda_H}e^{-\lambda_H T}}. \tag{4.4.2}$$

In addition, q_T is increasing in T, $\lim_{T\to\infty} q_T = 1$ if $\mu > \lambda_H$ (or equivalently $\delta < 1/2$), and $\lim_{T\to\infty} q_T = \mu/\lambda_H = (1-\delta)/\delta$ if $\mu < \lambda_H$ (or equivalently $\delta > 1/2$).

This result is easily understood as a tradeoff between the duration effect and the still-in-the-race effect. In Figure 4.2, we illustrate these effects when Firm j fully allocates its resources to research ($\sigma_t^j = 1$ for all $t \geqslant 0$). Specifically, we provide the graphs of the terms of each effect divided by $(1-p)$: μ (DE), $\lambda_H p$ (SRE). In Figure 4.2a, we depict the case where $\mu = \lambda_H$, i.e., $\delta = 1/2$. Observe that the duration effect is always larger than the still-in-the-race effect here. If we fix λ_H and increase μ, we can easily see that the duration effect will still dominate the still-in-the-race effect. Hence, when $\delta = \lambda_H/(\lambda_H + \mu) < 1/2$, we can see that the belief keeps increasing up to 1 ($\lim_{T\to\infty} q_T = 1$). On the other hand, in Figure 4.2b, we illustrate the case where $\mu < \lambda_H$, i.e., $\delta < 1/2$. Observe that the duration effect is greater than the still-in-the-race effect only when $p < \mu/\lambda_H$. This induces the belief to converge to $(1-\delta)/\delta$ ($\lim_{T\to\infty} q_T = \mu/\lambda_H = (1-\delta)/\delta$).

We now explore the dynamics of the firms' expected payoffs. Suppose that Firm i possesses the new technology and Firm j's strategy is σ^j. Then, the expected payoff of Firm i at time T, denoted by $V_T^{1,i}$, can be written as follows.

$$\begin{aligned} V_T^{1,i} = &- cdt + \Pi \cdot \lambda_H dt + 0 \cdot \left(\lambda_H p_T^i + \lambda_L(1-p_T^i)(1-\sigma_T^j)\right) dt \\ &+ (1 - \lambda_H dt - \lambda_H p_T^i dt - \lambda_L(1-p_T^i)(1-\sigma_T^j)dt)(V_T^{1,i} + \dot{V}_T^{1,i} dt). \end{aligned}$$

Thus, we can derive the Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \dot{V}_T^{1,i} + \lambda_H(\Pi - V_T^{1,i}) - \left\{\lambda_H p_T^i + \lambda_L(1-p_T^i)(1-\sigma_T^j)\right\} V_T^{1,i} - c. \qquad (\mathrm{HJB}_1)$$

This HJB equation gives a clear interpretation on the evolution of $V_T^{1,i}$: at an instant T, (i) Firm i wins the race at rate λ_H and gets the rent Π but loses the continuation payoff $V_T^{1,i}$; (ii) Firm j wins the race at rate $\lambda_H p_T^i + \lambda_L(1-p_T^i)(1-\sigma_T^j)$ and Firm i loses the continuation payoff; (iii) the flow cost c is charged.

Next, let $V_T^{0,i}$ denote the expected payoff of Firm i at time T given that it has not succeeded in research. In this case, the firm needs to choose between doing research and developing with the incumbent technology:

$$
\begin{aligned}
V_T^{0,i} = \max_{\sigma_T^i \in [0,1]} & - cdt + \Pi \cdot \lambda_L(1-\sigma_T^i)dt + V_T^{1,i} \cdot \mu\sigma_T^i dt \\
& + 0 \cdot \left(\lambda_H p_T^i + \lambda_L(1-p_T^i)(1-\sigma_T^j)\right) dt \\
& + \left(1 - \lambda_L(1-\sigma_T^i)dt - \mu\sigma_T^i dt - \lambda_H p_T^i dt - \lambda_L(1-p_T^i)(1-\sigma_T^j)dt\right) \\
& \quad \times (V_T^{0,i} + \dot{V}_T^{0,i} dt).
\end{aligned}
$$

By using the linearity of σ_T^i, the corresponding HJB equation can be derived as follows:

$$
\begin{aligned}
0 = & \, \dot{V}_T^{0,i} - \left\{\lambda_H p_T^i + \lambda_L(1-p_T^i)(1-\sigma_T^j)\right\} V_T^{0,i} - c \\
& + \max_{\sigma_T^i \in [0,1]} \left[\sigma_T^i \cdot \mu(V_T^{1,i} - V_T^{0,i}) + (1-\sigma_T^i) \cdot \lambda_L(\Pi - V_T^{0,i})\right].
\end{aligned}
\tag{HJB$_0$}
$$

This HJB equation determines whether Firm i allocates the resources to research or development. If $\mu(V_T^{1,i} - V_T^{0,i}) > \lambda_L(\Pi - V_T^{0,i})$, Firm i allocates all the resources to doing research: $\sigma_T^i = 1$. If $\lambda_L(\Pi - V_T^{0,i}) > \mu(V_T^{1,i} - V_T^{0,i})$, Firm i allocates all the resources to developing with the incumbent technology: $\sigma_T^i = 0$. If $\mu(V_T^{1,i} - V_T^{0,i}) = \lambda_L(\Pi - V_T^{0,i})$, Firm i is indifferent between the research and the development: $\sigma_T^i \in [0,1]$.

4.4.2 Symmetric Markov Equilibrium

To characterize the equilibrium in the private information setting, we focus on strategy profiles that are symmetric and Markov with respect to the belief on the opponent's technology level. Formally, we call $\sigma : \mathbb{R}_+ \to [0,1]$ a *symmetric Markov strategy* if $\sigma^A = \sigma^B = \sigma$, and for $i \in \{A, B\}$, $p_t^i = p_{t'}^i$ implies $\sigma_t^i = \sigma_{t'}^i$ where p^i is the belief process derived from $p_0^i = 0$ and (4.4.1). A strategy profile σ constitutes a *symmetric Markov equilibrium* if σ is a symmetric Markov strategy and $\sigma^A = \sigma^B = \sigma$ solves (HJB$_0$) for all $T \geqslant 0$. The following proposition characterizes the symmetric Markov equilibria of the game.

Proposition 4.4.1. *Suppose that the firms cannot observe their opponent's technological level. The following statements hold.*

(a) ***(Cutoff Structure)*** *Any symmetric Markov equilibrium can be characterized by a cutoff time $T^* \in \mathbb{R} \cup \{\infty\}$ and a stationary strategy $\sigma^* \in [0, 1)$ where both firms fully conduct research up to time T^* ($\sigma_t = 1$ for all $t < T^*$), and choose σ^* from then on ($\sigma_t = \sigma^*$ for all $t > T^*$).*

(b) ***(Equilibrium Characterization)*** *The unique symmetric Markov equilibrium is characterized as follows.*

(i) If $\eta \geqslant \tilde{\eta}(\delta) \equiv \min\{\overline{\eta}(\delta), 2-\delta\}$, both firms play ***research strategy*** *($T^* = \infty$). In equilibrium, a firm's expected payoffs with and without the new technology are $V_0^t = \overline{V}_0(q_t)$ and $V_1^t = \overline{V}_1(q_t)$ where q is the belief defined in (4.4.2) and*

$$\overline{V}_1(q) \equiv \frac{1}{2}\left(\Pi - \frac{c}{\lambda_H}\right)(1 + \delta(1-q)), \tag{4.4.3}$$

$$\overline{V}_0(q) \equiv \frac{1}{2}\left(\Pi - \frac{c}{\mu} - \frac{c}{\lambda_H}\right)(1 - \delta q) - \frac{c}{2(\lambda_H + \mu)}. \tag{4.4.4}$$

(ii) If $\eta \leqslant \underline{\eta}(\delta)$, both firms play ***incumbent strategy*** *($T^* = 0$, $\sigma^* = 0$). In equilibrium, the expected payoff of each firm is $V_t^0 = \frac{\lambda_L \Pi - c}{2\lambda_L}$ for all $t \geqslant 0$.*[12]

(iii) If $\tilde{\eta}(\delta) > \eta > \underline{\eta}(\delta)$, both firms play ***stationary fall-back strategy*** *($T^* \in (0, \infty)$, $\sigma^* \in (0, 1)$). Moreover, for all $t \geqslant T^*$, $\sigma_t = \sigma^*$, $p_t = p^*$, $V_t^1 = V_1^*$ and $V_t^0 = V_0^*$ where*

$$\sigma^* = \frac{\eta}{1-\delta} - \frac{\eta+\delta}{\eta-\delta}, \qquad p^* = \frac{1}{2}\left\{\frac{\eta}{\delta} - \frac{1-\delta}{\eta-1}\right\},$$

$$V_1^* = \frac{\lambda_L}{\mu} \cdot \frac{\lambda_H \Pi - c}{\lambda_H - \lambda_L}, \qquad V_0^* = \frac{\lambda_L}{\mu - \lambda_L} \cdot \frac{\lambda_L \Pi - c}{\lambda_H - \lambda_L}.^{13} \tag{4.4.5}$$

[12] When a firm possesses the new technology—though it is off the equilibrium path—the expected payoff is $V_1^t = \frac{\lambda_H \Pi - c}{\lambda_H + \lambda_L}$.

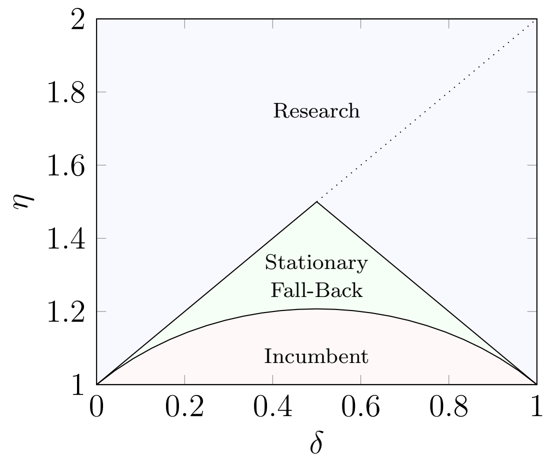


FIGURE 4.3: Symmetric Markov Equilibrium under the Private Information Setting

The formal proof of the proposition is relegated to Appendix C.3, but we provide a sketch of the proof here. We begin by showing that the belief derived from a symmetric Markov strategy is nondecreasing in time, i.e., $\dot{p}_t \geqslant 0$ (Lemma C.3.1). This is because if $\dot{p}_t < 0$ for some $t > 0$, by the Markov property, the belief cannot go above p_t which contradicts $\dot{p}_t < 0$. Note that $\dot{p}_t < 0$ if $p_t > 0$ and $\sigma_t = 0$. This allows us to focus on the following two cases: (i) $\sigma_t = 0$ for all $t \geqslant 0$; or (ii) $\sigma_t > 0$ for all $t \geqslant 0$ (Lemma C.3.2). The first case corresponds to incumbent strategy in the observable-breakthrough benchmark ($T^* = 0$ and $\sigma^* = 0$). Even though the strategy space is different from the benchmark, it is qualitatively equivalent since the firm always develops with the incumbent technology. Similarly, a special case of the second case—$\sigma_t = 1$ for all $t \geqslant 0$—corresponds to research strategy ($T^* = \infty$). In the remaining case, on the equilibrium path, $\sigma_S \in (0, 1)$ for some $S \geqslant 0$, i.e., firms are indifferent between researching to find the new technology and developing the innovation with the old technology at time S. In this case, we show that from then on ($t \geqslant S$), the firms continue to be indifferent between research and development (Lemma C.3.3). In addition, we show that to make firms indifferent for all $t \geqslant S$, the

[13] See Remark 7 for the expected payoffs for $t \in [0, T^*]$

firms' strategies and beliefs should be stationary: $\sigma_t = \sigma^*$ and $p_t = p^*$ for all $t \geqslant S$ (Lemma C.3.4). By identifying the earliest time T at which firms are indifferent, we can show that it corresponds to the stationary fall-back strategy: $\sigma_t = 1$ for all $t < T$ and $\sigma_t = \sigma^*$ for all $t > T$. Thus, we have three types of symmetric Markov equilibria: both firms play (i) the research strategy; (ii) the incumbent strategy; or (iii) the stationary fall-back strategy.

Next, we need to identify which regions of the parameter space give rise to each of the three types of equilibrium. First, consider parameter values under which the fall-back policy is not the second-best policy ($\eta \geqslant \overline{\eta}(\delta)$ or $\eta \leqslant \underline{\eta}(\delta)$). In this case, firms do not change their resource allocations even if they can observe their opponent's technological breakthroughs. Thus, the same strategy profiles (fully conducting research or fully developing with the incumbent technology) will constitute an equilibrium in the private information setting.

Now consider the remaining case ($\overline{\eta}(\delta) > \eta > \underline{\eta}(\delta)$). In this case, the new technology is efficient enough ($\eta > \underline{\eta}(\delta)$) for both firms to begin by doing only research. However, They need to determine whether to keep fully conducting research indefinitely ($T^* = \infty$) or to hedge their bets by switching to the stationary fall-back strategy at some point ($T^* < \infty$ and $\sigma^* \in (0, 1)$). The answer crucially depends on the relative intensity δ. When the new technology is more development-intensive ($\delta < 1/2$), if firms keep fully conducting research indefinitely, then the beliefs that their opponent has made a breakthrough converge to 1 by Lemma 4.4.2. Since they are in the parameter region where firms switch to development with the incumbent technology if they know that their opponent possesses the new technology (or equivalently $p = 1$), they will find it better to choose the stationary fall-back strategy when the belief is sufficiently close to 1. Next, consider the case where the new technology is more research-intensive ($\delta > 1/2$). In contrast to the previous case, there exists a region where both firms play the research strategy indefinitely in equilibrium

$(2-\delta<\eta<\overline{\eta}(\delta)=1+\delta)$. This result is also easily understood by considering the belief about technological status. By Lemma 4.4.2, this belief cannot exceed $(1-\delta)/\delta$ in this region of the parameter space. Since the belief is bounded strictly below 1, firms may find it preferable to keep fully conducting research in the private information setting, even though they would switch to development with the incumbent technology if they were able to observe a breakthrough by their opponent. As δ increases, the upper bound of the belief $(1-\delta)/\delta$ decreases, so the firms will keep fully conducting the research even with the relatively low η. Thus, the threshold between the equilibria with the stationary fall-back strategy and the research strategy decreases in δ.

Remark 7. In Lemma C.3.6, we characterize the generic forms of the expected payoffs when both firms use $\sigma_t=1$ for all $t\in[0,T]$. It allows us to derive the expected payoffs at each point in time under the equilibrium with the stationary fall-back policy. By using $V^1_{T*}=V^*_1$ and $V^0_{T*}=V^*_0$ in (4.4.5) as terminal conditions at time T^*, the constants C_0 and C_1 in (C.3.11) and (C.3.12) can be determined. Moreover, C_1 is negative (see the proof of Lemma C.3.8). Thus, the expected payoffs V^1_t and V^0_t for any $t\in[0,T^*]$ can be derived.

4.4.3 Comparison of Expected Completion Times

Now that we characterized the firms' equilibrium behavior in the private information settings, we compare it with the benchmarks. Recall that the (ex ante) expected completed times for the first-best case and the public information setting are characterized in equations (4.2.1) and (4.3.1). In Figure 4.4, we fix δ and display the curves of the expected completion times for the first-best case (black) and the second-best case (blue) with respect to the efficiency measure η. The gap between these two expected completion times is generated since the new technology is not shared.

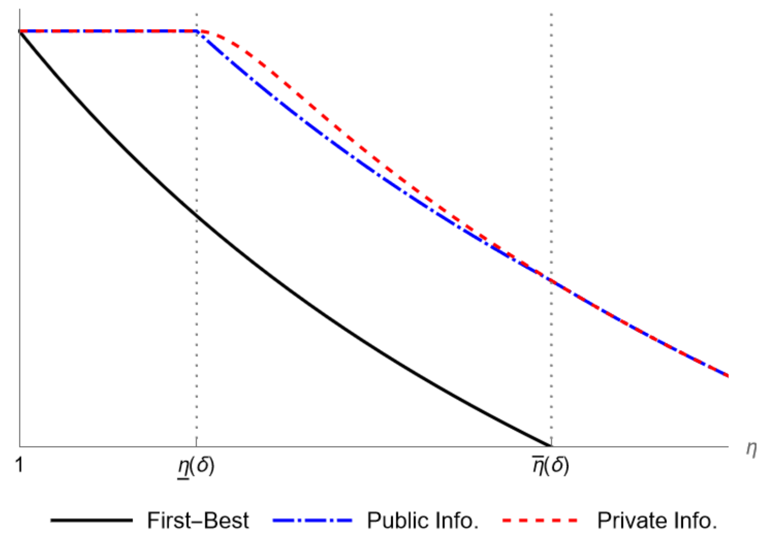


FIGURE 4.4: Expected completion times for the first-best case, private and public information settings

Observe that when $\eta \geqslant \overline{\eta}(\delta)$ or $\eta \leqslant \underline{\eta}(\delta)$, the equilibrium strategy in the private information setting is consistent with the public information setting, i.e., the expected completion times for these settings are same. When $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$, in the private information setting, the firms cannot use the fall-back strategy because they are not able to observe the rivals' technology levels. Rather, they use stationary fall-back strategies ($\eta \in (\underline{\eta}(\delta), \tilde{\eta}(\delta))$) or research strategies ($\eta \in [\tilde{\eta}(\delta), \overline{\eta}(\delta))$), which are suboptimal compared to the public information case. Therefore, the expected completion time would be longer, i.e., a lack of information transmission about the research status retards the pace of innovation. Figure 4.4 also illustrates these results: there is a gap between the expected completion times of the public and private information settings only if $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$.

4.5 Disclosure vs. Concealment

We now extend the private information setting by adding options to protect interim discoveries. Once a firm discovers the new technology, it can choose to conceal its discovery and treat it as a trade secret or disclose it to file a patent. In addition, if a firm files a patent, it can decide whether to license it or not. To facilitate the analysis, we assume that $\eta > \underline{\eta}(\delta)$ to rule out the case where firms do not engage in research at all.

4.5.1 The Game after Patenting

We begin by describing the game that takes place after a firm files a patent. First, consider the case where the rival already had the new technology and protected it as a trade secret. Then, the trade secret right allows the rival firm to dispute the patent. Thus, both firms have the right to use the new technology and the expected payoffs of them are $V_C = \frac{\lambda_H \Pi - c}{2\lambda_H}$.

Next, suppose that the rival firm does not possess the new technology. Then, the patenting firm has the exclusive right to use the new technology and the rival firm has to develop with the incumbent technology. Then, the expected payoffs of the patenting firm and the rival firm are $V_P = \frac{\lambda_H \Pi - c}{\lambda_H + \lambda_L}$ and $V_R = \frac{\lambda_L \Pi - c}{\lambda_H + \lambda_L}$.

Now we explore whether the patenting firm has an incentive to license the technology. Assume that the patenting firm can make a take-it-or-leave-it (TIOLI) offer $x \in \mathbb{R}_+$ to the rival firm for the right to use the new technology.[14] After licensing, both firms can use the new technology. Thus, the expected payoffs of the patenting firm and the rival firm after licensing are $V_C + x$ and $V_C - x$. Then, the optimal offer

[14] This assumption implies that the patenting firm has all the bargaining power. If the firm conceals the discovery even in this case, then it will conceal it in the less bargaining power cases.

x^* should satisfy $V_C - x^* = V_R$. With simple algebra, we can derive that

$$x^* = \frac{(\lambda_H - \lambda_L)(\lambda_H \Pi + c)}{2\lambda_H(\lambda_H + \lambda_L)} = \frac{\lambda_H - \lambda_L}{\lambda_H + \lambda_L}\left(V_C + \frac{c}{\lambda_H}\right) > 0. \tag{4.5.1}$$

Then, the expected payoff for the patenting firm after licensing is

$$V_L \equiv V_C + x^* = \left(1 + \frac{(\lambda_H - \lambda_L)c}{\lambda_H(\lambda_H \Pi - c)}\right) V_P > V_P. \tag{4.5.2}$$

This implies that the patenting firm can always be better off by licensing the new technology.

4.5.2 Immediate-Disclosure Equilibrium

We first explore whether the first-best outcome can be achieved by allowing the firms to disclose the new technology and license it. Recall that in the first-best case, both firms do research, and a firm's new technology is immediately spilled over to the rival. Thus, we consider a strategy profile such that a firm with the new technology employs the *immediate-disclosure strategy*—a firm discloses (and licenses) the new technology as soon as it discovers—and a firm without the new technology employs the research strategy ($\sigma_t = 1$ for all $t \geqslant 0$). Then, we ask whether both firms playing this strategy can be sustained as an equilibrium.

Suppose that a firm (say Firm A) just discovered the new technology and Firm B has not disclosed it yet. Given that Firm B sticks to the immediate disclosure and research strategy, Firm A's belief that Firm B has the new technology is zero. Then, by disclosing the new technology, Firm A expects to license it, i.e., the expected payoff for Firm A after disclosure is V_L. Now consider Firm A's deviation to delay the disclosure by time dt. With the probability $\lambda_H dt$, Firm A wins the race and receives Π. But with the probability μdt, Firm B will discover the new technology and files a patent, but it will be disputed by Firm A's trade secret right. Thus, both firms will race with the new technology from then on and the expected payoff is V_C. With the probability $(1 - \lambda_H dt - \mu dt)$, neither of the events happens and Firm A licenses, then the expected payoff is V_L. Last, the flow cost cdt will be paid. To sum up, Firm A's expected payoff from delaying the disclosure is

$$\begin{aligned}
&\Pi \cdot \lambda_H dt + V_C \cdot \mu dt + (1 - \lambda_H dt - \mu dt) \cdot V_L - cdt \\
=&V_L + \left[(\mu + 2\lambda_H)V_C - (\mu + \lambda_H)V_L\right] dt \\
=&V_L + \left[\lambda_H V_C - (\mu + \lambda_H)x^*\right] dt.
\end{aligned}$$

Then, from $\delta = \lambda_H/(\mu + \lambda_H)$, the immediate-disclosure and research strategy can be sustained as an equilibrium if and only if $\delta V_C \leqslant x^*$. By using (4.5.1) and some algebra, this inequality is equivalent to:

$$\frac{(1-\delta)(\overline{\eta}(\delta) - \eta)}{2(\eta - 1 + \delta)} \leqslant \frac{c}{\lambda_H \Pi - c}. \tag{4.5.3}$$

From the assumption that $\lambda_L \Pi \geqslant c$, observe that (4.5.3) always holds if $\eta \geqslant \overline{\eta}(\delta)$. Recall that firms do research regardless of the rival's progress. It implies that there does not exist any incentive for a firm to conceal its progress. Therefore, the firms would monetize the new technology by licensing it as soon as it discovers, and the first-best outcome would be achieved.

Next, suppose that $\underline{\eta}(\delta) < \eta < \overline{\eta}(\delta)$. Then, (4.5.3) is equivalent to:

$$\pi = \frac{\lambda_L \Pi}{c} \leqslant 1 + \frac{\eta^2 - (1-\delta)^2}{\eta(\overline{\eta}(\delta) - \eta)} \equiv \underline{\pi}(\eta, \delta). \tag{4.5.4}$$

Also note that $\underline{\pi}(\eta, \delta) > 1$ since $\eta > 1 - \delta > 0$. Therefore, when the reward of winning the race is sufficiently low ($1 \leqslant \pi \leqslant \underline{\pi}(\eta, \delta)$), the firms would license the new technology as soon as it discovers.

The following proposition formally summarizes the above results.

Proposition 4.5.1. *Suppose that one of the following conditions holds: (i) $\eta \geqslant \overline{\eta}(\delta)$; or (ii) $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$ and $1 \leqslant \pi \leqslant \underline{\pi}(\eta, \delta)$. Then, there exists an equilibrium such that the firms fully allocate their resources to research and license as soon as they find the new technology.*

4.5.3 No-Disclosure Equilibrium

We now explore whether the worst-case scenario for the planner can be realized, i.e., firms never disclose their discoveries and the expected completion time corresponds to the private information setting case.

First, we consider the case where $\tilde{\eta}(\delta) = \min\{2-\delta, \overline{\eta}(\delta)\} \leqslant \eta < \overline{\eta}(\delta) = 1+\delta$. By Proposition 4.4.1, in the equilibrium under the private information setting, firms do research until it succeeds ($T^* = \infty$ and $\sigma_t = 1$ for all $t \geqslant 0$). Suppose that both firms stick to this resource allocation strategy and never disclose their discoveries. When Firm A discovers the new technology at time t and never discloses it, the expected payoff of Firm A is $V_1^t = \{1 + \delta(1-q_t)\} \cdot V_C$ by Proposition 4.4.1 (b-i). If Firm A discloses the discovery at time t, Firm B has the new technology with the probability q_t. Thus, the expected payoff from the disclosure is $V_C \cdot q_t + V_L \cdot (1-q_t) = V_C + (1-q_t)x^*$. Therefore, the firm will not disclose if $x^* < \delta V_C$. We can also consider the case where Firm A discovers at time t but conceals until t' and decides to disclose or not at time t'. Even in this case, Firm A faces the same problem as before and will not disclose if $x^* < \delta V_C$. Recall that $x^* < \delta V_C$ is equivalent to $\pi > \underline{\pi}(\eta, \delta)$. Therefore, if $\pi > \underline{\pi}(\eta, \delta)$, there exists an equilibrium such that firms never disclose their discoveries and do research until it succeeds.

Next, we consider the case where $\underline{\eta}(\delta) < \eta < \tilde{\eta}(\delta)$. By Proposition 4.4.1 (b-iii), in the equilibrium under the private information setting, firms employ the stationary fall-back strategy (for some $T^* \in (0, \infty)$ and $\sigma^* \in (0,1)$, $\sigma_t = 1$ for all $0 \leqslant t < T^*$ and $\sigma_t = \sigma^*$ for all $t \geqslant T^*$). Suppose that Firm A discovers the new technology at $t \geqslant T^*$. If Firm A keeps the discovery secret, the expected payoff of Firm A is $V_1^* = \dfrac{2\delta}{\eta - 1 + \delta} V_C$. In addition, $V_1^* < \overline{V}_1(p^*) = \{1 + \delta(1 - p^*)\} V_C$ (see Lemma C.3.8). On the other hand, if Firm A discloses the discovery, the expected payoff from the disclosure is $V_C + (1 - p^*)x^*$. Then, Firm A does not disclose under the condition stronger than $x^* < \delta V_C$. In this case, there exists $\overline{\pi}(\eta, \delta) > \underline{\pi}(\eta, \delta)$ such that Firm A does not disclose when $\pi > \overline{\pi}(\eta, \delta)$. The following proposition formally states this result.

Proposition 4.5.2. *Suppose that* $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$. *Then, there exists* $\overline{\pi}(\eta, \delta) > \underline{\pi}(\eta, \delta)$ *such that for all* $\lambda_L \Pi / c > \overline{\pi}(\eta, \delta)$, *the following strategy can be sustained as an equilibrium: (i) firms never disclose their discoveries; (ii) and employ the equilibrium resource allocations in the private setting:* $\sigma_t = 1$ *for all* $t < T^*$ *and* $\sigma_t = \sigma^*$ *for all* $t > T^*$ *for some* $T^* \in (0, \infty]$ *and* $\sigma^* \in (0, 1)$.

5

Conclusion

In Chapter 1, I develop the concept of the monotone quasi-garbling order, which implies that one information structure is obtained from another by adding reversely monotone noise. I show that this ordering is equivalent to Lehmann's accuracy condition under the MLRP. I also show that this order is a necessary and sufficient condition for informativeness in general classes of monotone decision problems where the decision maker is allowed to choose a multidimensional action. I illustrate this result in decision problems of optimal insurance and nonlinear monopoly pricing. The general setup presented here can be applied to conduct research on information comparisons in many other economic contexts. For example, as in the nonlinear monopoly pricing application, many mechanism design problems can be recast as general monotone decision problems amenable to analysis via the methods presented in this paper.

In Chapter 2, I study the economic tradeoffs between a direct approach and a sequential approach for achieving a discrete goal in the context of a principal-agent setting. The optimal contract is determined by the interplay of monitoring, efficiency, and an endogenous deadline. I show that the form of the optimal contract depends on the project return. When the efficiency loss from splitting the project in two does not exist or is small, only the direct approach will be chosen if the project return is low, whereas only the sequential approach will be chosen if the project return is high. If the project return is intermediate, it is optimal to begin with the sequential approach and then switch to the direct approach. When the efficiency loss is large, the principal generally chooses the direct approach. However, if the project return is above a certain cutoff, she may choose the sequential approach for a short period of time in the middle of the contract (i.e., there may be two switches).

There are numerous avenues open for further research. For example, the principal may be able to design the approaches directly. In this article, I assume that the two approaches are exogenously given and the principal chooses between them. However, in practice, a project manager often designs how many milestones to partition the main project into and how difficult each subproject is. We could also introduce 'learning by doing' into the model. If we assume that the agent learns from early errors, the arrival rate of project completion would increase over time.[1] Finally, we might consider competition between firms. Many technology companies are often exposed to competition and this may significantly influence which approach project managers take. For instance, competitive pressure may manifest as increased time sensitivity tipping the choice of approach toward the more efficient direct methodology. I leave these intriguing questions–and others–for future work.

In Chapter 3, we study the long-lasting question of patent vs. secrecy from a different angle from the literature: a firm's concealing motive to hinder the rival's strategic response. We introduce an innovation race model with multiple paths and show that firms' disclosing decisions depend on the reward of winning the race. Based on this result, we can argue that reducing the reward for the winner may promote

[1] This possibility contrasts with the setting considered by Carnehl and Schneider (2021), where learning causes the expected arrival rate to fall.

licensing which would speed up the pace of innovation.

There are many avenues open for further research. For example, we assume that there are exogenously given two paths towards innovation, and one of the paths requires two breakthroughs. However, in practice, there are numerous ways to make an innovation, and it often requires more than two breakthroughs. We also assume that a firm's R&D resources are fixed over time, but we could also allow firms to endogenously choose how much effort to put in each point in time. Finally, we assume the contest structure is given by the winner-takes-all competition, but we might consider a contest designing problem. We leave these intriguing questions and others for future work.

Appendix A

Appendix for Chapter 2

A.1 Details for Relationship among Criteria

A.1.1 Proof of Proposition 2.4.3

Proof of Proposition 2.4.3. Suppose that $F \succeq_L G$ does not hold, i.e., there exists $\omega' > \omega$ and $y \in Y$ such that $\Phi(\omega'; y) < \Phi(\omega; y)$. Since $F \succeq_{MQG} G$, there exists Γ such that $G(y|\omega) = \int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega)$ and $\Gamma(\cdot|x,\omega) \geqslant_{FOSD} \Gamma(\cdot|x,\omega')$ for all $\omega' > \omega$. Also note that $G(y|\omega) = F(\Phi(\omega;y)|\omega) = \int_{\underline{x}}^{\Phi(\omega;y)} dF(x|\omega)$. Then, we can rewrite $G(y|\omega) = \int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega) = \int_{\underline{x}}^{\Phi(\omega;y)} dF(x|\omega)$ as follows:

$$\int_{\underline{x}}^{\Phi(\omega;y)} (1 - \Gamma(y|x,\omega))dF(x|\omega) = \int_{\Phi(\omega;y)}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega). \tag{A.1.1}$$

Suppose that (A.1.1) is equal to zero. Since $1 \geqslant \Gamma(y|x,\omega) \geqslant 0$ for all x, it implies that $\Gamma(y|x,\omega) = 1$ a.e. for $x \leqslant \Phi(\omega;y)$, and $\Gamma(y|x,\omega) = 0$ a.e. for $x > \Phi(\omega;y)$. Observe that, by $\Gamma(\cdot|x,\omega) \geqslant_{FOSD} \Gamma(\cdot|x,\omega')$, $\Gamma(y|x,\omega') = 1$ a.e. for $x \leqslant \Phi(\omega;y)$. Then, we have

$$G(y|\omega') = \int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega')dF(x|\omega') \geqslant F(\Phi(\omega;y)|\omega'),$$

which implies $\Phi(\omega'; y) = \sup\{x \mid F(x|\omega') \leqslant G(y|\omega')\} \geqslant \Phi(\omega; y)$ contradicting $\Phi(\omega'; y) < \Phi(\omega; y)$. Therefore (A.1.1) is nonzero.

Note that

$$\begin{aligned}
&\left(\int_{\Phi(\omega;y)}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega)\right) \cdot \left(\int_{\underline{x}}^{\Phi(\omega;y)} (1-\Gamma(y|x,\omega))dF(x|\omega')\right) \\
&= \int_{\Phi(\omega;y)}^{\overline{x}} \int_{\underline{x}}^{\Phi(\omega;y)} \Gamma(y|x,\omega)(1-\Gamma(y|x,\omega))dF(x|\omega)dF(x'|\omega') \\
&\leqslant \int_{\Phi(\omega;y)}^{\overline{x}} \int_{\underline{x}}^{\Phi(\omega;y)} \Gamma(y|x,\omega)(1-\Gamma(y|x,\omega))dF(x'|\omega)dF(x|\omega') \\
&= \left(\int_{\Phi(\omega;y)}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega')\right) \cdot \left(\int_{\underline{x}}^{\Phi(\omega;y)} (1-\Gamma(y|x,\omega))dF(x|\omega)\right)
\end{aligned}$$

where the inequality is derived by doubly integrating the MLRP conditions multiplied by $\Gamma(y|x,\omega)(1-\Gamma(y|x',\omega))$.

By canceling terms by using (A.1.1), we can derive that

$$F(\Phi(\omega;y)|\omega') = \int_{\underline{x}}^{\Phi(\omega;y)} dF(x|\omega') \leqslant \int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega').$$

Also from $\Gamma(y|x,\omega) \leqslant \Gamma(y|x,\omega')$ for all $x \in X$,

$$\int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega)dF(x|\omega') \leqslant \int_{\underline{x}}^{\overline{x}} \Gamma(y|x,\omega')dF(x|\omega') = G(y|\omega').$$

Then, the above two inequalities imply $F(\Phi(\omega;y)|\omega') \leqslant G(y|\omega')$, thus, $\Phi(\omega';y) \geqslant \Phi(\omega;y)$ contradicting $\Phi(\omega';y) < \Phi(\omega;y)$. Therefore, $F \succeq_L G$ holds. □

A.1.2 An Illustrative Example on Information Ranking

In this section, I provide an illustrative example that is discussed in Section 2.4. Let the state space be $\Omega \equiv [0, 4]$. I construct a triplet of information structures, F, G

and H where their signal spaces are $X = Y = [-1, 1]$ and $Z = [-2, 2]$, and their probability distribution functions are given as in Figure A.1. Observe that F and G satisfy the MLRP, but H does not.

	$h(z\|\omega)$				$f(x\|\omega)$		$g(y\|\omega)$	
$\omega\backslash$ signal	$[-2,-1)$	$[-1,0)$	$[0,1)$	$[1,2]$	$[-1,0)$	$[0,1]$	$[-1,0)$	$[0,1]$
$\omega \in [0,1)$	1/2	1/2	0	0	1	0	5/6	1/6
$\omega \in [1,2)$	1/2	0	1/2	0	1/2	1/2	2/3	1/3
$\omega \in [2,3)$	0	1/2	0	1/2	1/2	1/2	2/3	1/3
$\omega \in [3,4]$	0	0	1/2	1/2	0	1	1/3	2/3

FIGURE A.1: An example of information structures

In the subsequent subsections, I show the following statements which are summarized in Figure A.2.

1. G is neither a garbling of F nor H, but F is a garbling of G ($F \not\geq_B G$, $H \not\geq_B G$, but $F \geq_B G$);

2. H is more Lehmann accurate than neither F nor G, but F is more Lehmann accurate than G ($H \not\geq_L F$, $H \not\geq_L G$, but $F \geq_L G$);

3. G is a monotone quasi-garbling of H and F, and F is a monotone quasi-garbling of H ($H \geq_{MQG} F$, $H \geq_{MQG} G$, and $F \geq_{MQG} G$).

Blackwell's Garbling Condition

First, I show that G is a garbling of neither F nor H by showing that there exists a prior belief Λ and a utility function u such that $V(G; u, \Lambda) > V(F; u, \Lambda) = V(H; u, \Lambda)$. Consider a case where a prior Λ is uniformly distributed, $A = [0, 1]$ and utility function $u : A \times \Omega \to \mathbb{R}$ is linear in a, i.e., $u(a, \omega) = a \cdot u(1, \omega) + (1-a) \cdot u(0, \omega)$,

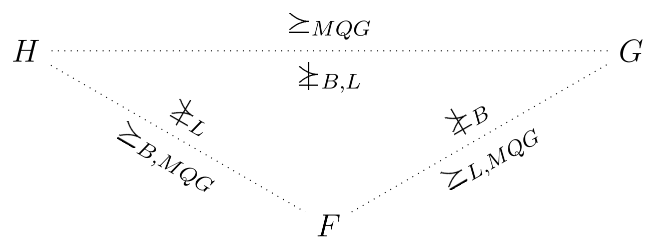


FIGURE A.2: Relationship among criteria in the example

and $u(1,\omega)$ and $u(0,\omega)$ are defined as follows:

$$u(1,\omega) = \begin{cases} 1, & \text{if } \omega \in [1,3), \\ 0, & \text{otherwise,} \end{cases} \qquad u(0,\omega) = \begin{cases} 0, & \text{if } \omega \in [1,3), \\ 1, & \text{otherwise.} \end{cases}$$

Observe that the DM's optimal decision is determined by the posterior belief that the state is in $[1,3)$. Under the information structure F and H, the posterior belief that the state is in $[1,3)$ is $1/2$ for any signal realization. Therefore, $V(F;u,\Lambda) = V(H;u,\Lambda) = 1/2$. Under the information structure G, if he receives a negative signal, which happens at probability $15/24$, the posterior belief that the state is in $[1,3)$ is $8/15$. In this case, it is optimal to choose $a = 1$ and the expected utility is $8/15$. If he receives a nonnegative signal, which happens at probability $9/24$, the posterior belief that the state is in $[1,3)$ is $4/9$. In this case, it is optimal to choose $a = 0$ and the expected utility is $5/9$. Then, we have

$$V(G;u,\Lambda) = \frac{15}{24}\cdot\frac{8}{15} + \frac{9}{24}\cdot\frac{5}{9} = \frac{13}{24} > \frac{1}{2} = V(F;u,\Lambda) = V(H;u,\Lambda).$$

Therefore, G is a garbling of neither F nor H, i.e., $F \not\succeq_B G$ and $H \not\succeq_B G$.

Next, I show that F is a garbling of H by constructing a noise $\Gamma : Z \to \Delta(X)$. Consider Γ derived from the following probability distribution function:

$$\gamma(x|z) = \begin{array}{r|cc} & x < 0 & x \geqslant 0 \\ \hline z < 0 & 1 & 0 \\ z \geqslant 0 & 0 & 1 \end{array}.$$

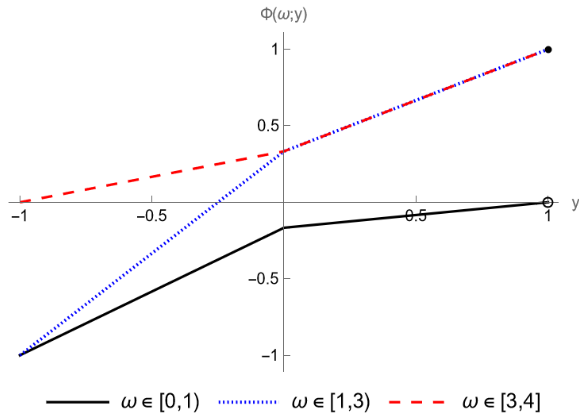


FIGURE A.3: Lehmann Effectiveness Condition for F and G

By simple algebra, we can see that $f(x|\omega) = \int_{-2}^{2} \gamma(x|z)h(z|\omega)dz$ for all $\omega \in \Omega$. Therefore, F is a garbling of H, i.e., $H \succeq_B F$.

Lehmann's Effectiveness Condition

I show that H is not more Lehmann accurate than F by finding a signal x and a pair of states (ω, ω') such that $\omega' > \omega$ and $\Phi(\omega; x) > \Phi(\omega'; x)$. Consider $x = 0$, $\omega = 0$, and $\omega' = 1$. Observe that

$$\Phi(0;0) = \sup\{z \mid H(z|0) \leqslant F(0|0) = 1\} = 2,$$

$$\Phi(1;0) = \sup\{z \mid H(z|1) \leqslant F(0|1) = 1/2\} = 0,$$

thus, $\Phi(0;0) > \Phi(1;0)$ and $H \not\succeq_L F$.

Likewise, I show that H is not more Lehmann accurate than G by considering $y = 0$, $\omega = 2$ and $\omega' = 3$:

$$\Phi(2;0) = \sup\{z \mid H(z|2) \leqslant G(0|2) = 2/3\} = 4/3,$$

$$\Phi(3;0) = \sup\{z \mid H(z|3) \leqslant G(0|3) = 1/3\} = 2/3.$$

Last, I show that F is more Lehmann accurate than G. With simple algebra, we can derive Φ as follows.

1. When $\omega \in [0, 1)$,

$$\Phi(\omega; y) = \begin{cases} \frac{5}{6}y - \frac{1}{6}, & \text{if } -1 \leqslant y < 0, \\ \frac{1}{6}y - \frac{1}{6}, & \text{if } 0 \leqslant y < 1, \\ 2, & \text{if } y = 1. \end{cases}$$

2. When $\omega \in [1, 3)$,

$$\Phi(\omega; y) = \begin{cases} \frac{4}{3}y + \frac{1}{3}, & \text{if } -1 \leqslant y < 0, \\ \frac{2}{3}y + \frac{1}{3}, & \text{if } 0 \leqslant y \leqslant 1. \end{cases}$$

3. When $\omega \in [3, 4]$,

$$\Phi(\omega; y) = \begin{cases} \frac{1}{3}y + \frac{1}{3}, & \text{if } -1 \leqslant y < 0, \\ \frac{2}{3}y + \frac{1}{3}, & \text{if } 0 \leqslant y \leqslant 1. \end{cases}$$

Figure A.3 illustrates the above results. As we can see in the figure, $\Phi(\omega; y)$ is nondecreasing in ω. Therefore, F is more Lehmann accurate than G.

Monotone Quasi-Garbling Order

Since Blackwell's garbling condition implies the monotone quasi-garbling order, we have $H \succeq_{MQG} F$. We also have $F \succeq_{MQG} G$ because Lehmann's accuracy condition also implies the monotone quasi-garbling order (Proposition 2.4.2). It remains to show that G is a monotone quasi-garbling of H. I construct a reversely monotone noise $\Gamma : Z \times \Omega \to \Delta(Y)$ derived from the following probability distribution functions:

1. if $\omega \in [0, 1)$,

$$\gamma(y|z, \omega) = \begin{array}{r|cc} & y < 0 & y \geqslant 0 \\ \hline z \in [-2, -1) & 1 & 0 \\ z \in [-1, 0) & 2/3 & 1/3 \\ z \in [0, 1) & 0 & 1 \\ z \in [1, 2] & 0 & 1 \end{array},$$

2. if $\omega \in [1, 2)$,

$$\gamma(y|z,\omega) = \begin{array}{c|cc} & y < 0 & y \geqslant 0 \\ \hline z \in [-2,-1) & 1 & 0 \\ z \in [-1,0) & 2/3 & 1/3 \\ z \in [0,1) & 1/3 & 2/3 \\ z \in [1,2] & 0 & 1 \end{array},$$

3. if $\omega \in [2, 4]$,

$$\gamma(y|z,\omega) = \begin{array}{c|cc} & y < 0 & y \geqslant 0 \\ \hline z \in [-2,-1) & 1 & 0 \\ z \in [-1,0) & 1 & 0 \\ z \in [0,1) & 1/3 & 2/3 \\ z \in [1,2] & 1/3 & 2/3 \end{array}.$$

With simple algebra, we can show that $g(y|\omega) = \int_{-2}^{2} \gamma(y|z,\omega)h(z|\omega)dz$ holds for all $y \in Y$ and $\omega \in \Omega$. In addition, we can also see that $\Gamma(y|x,\omega) \leqslant \Gamma(y|x,\omega')$ for all $y \in Y$, $x \in X$, and $\omega < \omega'$, i.e., Γ is a reversely monotone noise. Therefore, G is a monotone quasi-garbling of H, i.e., $H \geq_{MQG} G$.

A.2 An Example on IDO and DDDR Conditions

In this section, I provide an example of a decision problem that does not satisfy the IDO condition but is generally monotone with respect to any information structure with the MLRP. Let $\Omega = \{\omega_1, \omega_2\}$ and $A = [0, 6]$. Consider a usual order in A. Define a payoff function $u : A \times \Omega \to \mathbb{R}$ as follows:

$$u(a,\omega_1) \equiv \begin{cases} a+1, & \text{if } a \leqslant 2, \\ 5-a, & \text{if } 2 < a \leqslant 5, \\ a-5, & \text{if } 5 < a, \end{cases} \quad u(a,\omega_2) \equiv \begin{cases} 0, & \text{if } a \leqslant 1, \\ a-1, & \text{if } 1 < a \leqslant 4, \\ 5-a, & \text{if } 4 < a. \end{cases}$$

This function is illustrated in Figure A.4.

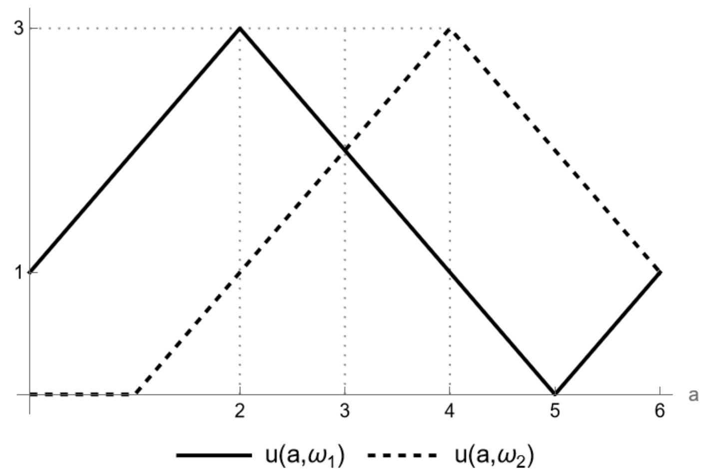


FIGURE A.4: Illustration of dominated decreasing decision rules, where the interval dominance property is violated.

First, observe that this function does not satisfy IDO condition: $u(6,\omega_1) > u(a,\omega_1)$ for all $5 \leqslant a < 6$, but $u(6,\omega_2) < u(5,\omega_2)$. Now I show that it satisfies the DDDR condition. Consider any decreasing function $h : \Omega \to A$. When $h(\omega_1) \leqslant 4$, set $\hat{a} = h(\omega_1)$. Then, $u(\hat{a},\omega_1) = u(h(\omega_1),\omega_1)$ by definition of $\hat{a}$, and $u(\hat{a},\omega_2) = u(h(\omega_1),\omega_2) \geqslant u(h(\omega_2),\omega_2)$ since $u(a|\omega_2)$ is increasing when $a < 4$. When $h(\omega_1) > 4$, set $\hat{a} = 4$. Then, $u(\hat{a},\omega_1) \geqslant u(h(\omega_1),\omega_1)$ since $u(4,\omega_1) \geqslant u(a,\omega_1)$ for all $a \geqslant 4$. In addition, $u(a,\omega_2)$ is maximized at $a = 4$, thus, $u(\hat{a},\omega_2) \geqslant u(h(\omega_2),\omega_2)$. Therefore, this decision problem satisfies the DDDR condition with respect to the generic order.

Now I show that it satisfies the MCS condition. Observe that for any $a < 2$ and ω, $u(a,\omega) < u(2,\omega)$. Similarly, for any $a \geqslant 4$, $u(a,\omega) \leqslant u(4,\omega)$. Therefore, it is without loss to focus on $a \in [2,4]$. Let the posterior belief given a signal x be $\psi_x = \Pr(\omega_2|x)$ and $1 - \psi_x = \Pr(\omega_1|x)$. Then, the expected payoff given the signal x is

$$\psi_x(a-1) + (1-\psi_x)(5-a) = (2\psi_x - 1)a + 5 - 6\psi_x.$$

Therefore, the optimal action given x is

$$a^*(x) = \begin{cases} 4, & \text{if } \psi_x \geqslant 1/2, \\ 2, & \text{if } \psi_x \leqslant 1/2. \end{cases}$$

Also observe that for any information structure with the MLRP, prior belief and pair of signals (x', x) with $x' > x$, the posterior belief on ω_2 is nondecreasing, i.e., $\psi_{x'} \geqslant \psi_x$. Therefore, $a^*(x)$ is nondecreasing in x, thus, the MCS condition holds.

A.3 Application: Optimal Insurance

I now present a familiar economic application that can be considered a general monotone decision problem. Specifically, I consider an optimal insurance problem similar to the investment problem studied by Cabrales et al. (2013). An agent (DM) privately acquires information about a future state $\omega \in \Omega = \{\omega_1, \cdots, \omega_N\}$, which will be his realized income, with $\omega_N > \cdots > \omega_1 > 0$. At the beginning of the investment decision, the agent chooses between two sources of information, say F and G. The agent receives a signal from the chosen source. After receiving a signal (and before the realization of the state), the agent can buy or sell Arrow securities. Assume that the Arrow security market is large enough so that the agent's demand does not affect prices. Let the price vector be $p = (p_1, \cdots, p_N)$ and assume that $\sum_{i=1}^N p_i = 1$ and $p_i > 0$.

In this example, an action by the agent is the number of securities bought (or sold) for each state, thus, the action space is multidimensional. For any $1 \leqslant i \leqslant N$, denote the amount of securities bought for a state ω_i by q_i. By allowing a short sale, q_i can be negative. Assume that the agent is not allowed to execute a short sale more than his realized income, i.e., $q_i + \omega_i \geqslant 0$ for all $1 \leqslant i \leqslant N$. Along with these constraints, the agent's budget constraint determines the set of feasible decisions. By assuming the agent's budget is zero, i.e., he must finance purchases through short sales, we get the set of feasible decisions as follows:

$$A = \{q = (q_1, \cdots, q_n) \in \mathbb{R}^n \,|\, p \cdot q \leqslant 0 \;\text{ and }\; \omega_i + q_i \geqslant 0 \;\; \forall i \in N\}\,. \tag{A.3.1}$$

Also note that the budget constraint will bind at the optimum and define the set of budget binding decisions as follows:

$$A' = \{q = (q_1, \cdots, q_n) \in \mathbb{R}^n \,|\, p \cdot q = 0 \;\text{ and }\; \omega_i + q_i \geqslant 0 \;\; \forall i \in N\}\,. \tag{A.3.2}$$

For the agent's utility function, assume constant relative risk aversion (CRRA) preferences with respect to the realized net income:

$$\mathcal{U} \equiv \left\{ u : A_2 \times \Omega \to \mathbb{R} \;:\; \begin{array}{l} u(q, \omega_i) = v(\omega_i + q_i) \;\; \forall 1 \leqslant i \leqslant N, \\ \exists\, \rho > 0 \;\text{ s.th. }\; v(z) = (z^{1-\rho} - 1)/(1 - \rho) \end{array} \right\}.^1 \tag{A.3.3}$$

Next I define a partial order $\geqslant_A$ on A. Recall that the purpose of this order is to compare the optimal actions, thus it only needs to be defined over $A' \subset A$. I construct the partial order as follows: for any $q, q' \in A'$,

$$q' \geqslant_A q \;\Leftrightarrow\; \frac{\omega_j + q'_j}{\omega_i + q'_i} \geqslant \frac{\omega_j + q_j}{\omega_i + q_i} \quad \text{for all } j > i. \tag{A.3.4}$$

This definition is inspired by the following equation holding at the optimum bundle: for any $i < j$,

$$\frac{\Pr(\omega_i) \cdot v'(\omega_i + q_i)}{\Pr(\omega_j) \cdot v'(\omega_j + q_j)} = \frac{p_i}{p_j}. \tag{A.3.5}$$

This equation says that the (expected) marginal rate of substitution of every pair of states is equal to the price ratio. Observe that when a bundle q' is higher than another bundle q in the order $\geqslant_A$, the left hand side of (A.3.5) increases. Suppose that the DM receives a higher signal that leads to a fall in the likelihood ratio $\Pr(\omega_i)/\Pr(\omega_j)$ in (A.3.5). Since the price ratio is assumed to be fixed, i.e., the right hand side of

[1] When $\rho = 1$, $v(z) = \log(z)$.

(A.3.5) is unchanged, the DM would choose a higher bundle in terms of $\geqslant_A$. In this sense, the order $\geqslant_A$ is reasonable in this example. The following proposition shows that this insurance problem is a general monotone decision problem with this order.

Proposition A.3.1. *The following statements hold:*

(a) The partial order $\geqslant_A$ defined on (A.3.4) *satisfies the DDDR condition for the set of payoff functions $\mathcal{U}$ defined on* (A.3.3)*;*

(b) Suppose that an information structure f satisfies the MLRP. Then, f satisfies the MCS condition with respect to $\mathcal{U}$ and $\geqslant_A$.

Proof of Proposition A.3.1. (a) Consider a decreasing decision rule $\{q^j\}_{j\in N} \in (A')^n$ where $q^1 \geqslant_A \cdots \geqslant_A q^n$. Define $\hat{q}_i \equiv q_i^i$ and $\hat{q} = (\hat{q}_1, \cdots, \hat{q}_n)$. To check whether $\geqslant_A$ satisfies the DDDR condition for $\mathcal{U}$, it is enough to show that $\hat{q} \in A$ since for any $u \in \mathcal{U}$, there exists a function v such that $u(\hat{q}, \omega_i) = v(\hat{q}_i + \omega_i) = v(a_i^i + \omega_i) = u(q^i, \omega_i)$ for all $1 \leqslant i \leqslant N$. Note that $\omega_i + \hat{q}_i = \omega_i + q_i^i \geqslant 0$ from $q^i \in A'$. Therefore, it suffices to show that $p \cdot \hat{q} \leqslant 0$.

For all $1 \leqslant i, j \leqslant N$, define $r_i^j \equiv p_i(\omega_i + q_i^j)/(\sum_{s=1}^n p_s\omega_s)$ and $r^j \equiv \{r_i^j\}_{1\leqslant i\leqslant N}$. Note that $\sum_{i\in N} r_i^j = 1$ from $q^j \in A'$, thus, r^j can be considered an element of $\Delta(\Omega)$. Moreover, the condition (A.3.4) makes $\{r^j\}_{j\in N}$ reversely MLRP ordered, i.e., $r^1 \geqslant_{MLRP} \cdots \geqslant_{MLRP} r^n$. Therefore, $r^1 \geqslant_{FOSD} \cdots \geqslant_{FOSD} r^n$. Then, the following inequalities hold:

$$\begin{aligned}
1 &= r_1^n + r_2^n + \cdots + r_{n-2}^n + r_{n-1}^n + r_n^n \\
&\geqslant (r_1^{n-1} + r_2^{n-1} + \cdots + r_{n-2}^{n-1} + r_{n-1}^{n-1}) + r_n^n \qquad (\text{by} \;\; r^{n-1} \geqslant_{FOSD} r^n) \\
&\geqslant (r_1^{n-2} + r_2^{n-2} + \cdots + r_{n-2}^{n-2}) + r_{n-1}^{n-1} + r_n^n \qquad (\text{by} \;\; r^{n-2} \geqslant_{FOSD} r^{n-1}) \\
&\;\;\vdots \\
&\geqslant r_1^1 + r_2^2 + \cdots r_{n-2}^{n-2} + r_{n-1}^{n-1} + r_n^n \qquad (\text{by} \;\; r^1 \geqslant_{FOSD} r^2) \\
&= \sum_{i \in N} \frac{p_i(\omega_i + \hat{q}_i)}{\sum_{s=1}^n p_s \omega_s}
\end{aligned}$$

From the last inequality, $0 \geqslant \sum_{i \in N} p_i \hat{q}_i = p \cdot \hat{q}$. Therefore, $\geqslant_A$ satisfies the DDDR condition for $\mathcal{U}$.

(b) Fix a prior belief $\Lambda \in \Delta(\Omega)$ and a utility function $u \in \mathcal{U}$ with $u(q, \omega_i) = v(\omega_i + q_i)$ and $v(z) = z^{1-\rho}/(1-\rho)$ for some $\rho > 0$. In the third stage with the signal x, DM's problem is

$$\max_{q \in A} \sum_{i \in N} \lambda_F^x(\omega_i) v(\omega_i + q_i). \tag{A.3.6}$$

When an information structure is $\{f(\cdot|\omega)\}_{\omega \in \Omega}$ and the agent receives x as a signal, let the solution of (A.3.6) be $\{q_i(x)\}_{i \in N}$. Note that the solution can be obtained as follows:

$$\omega_i + q_i(x) = \left(\frac{\lambda_F^x(\omega_i)}{p_i} \right)^{1/\rho} \cdot \frac{\sum_{s=1}^n \omega_s \cdot p_s}{\sum_{s=1}^n \lambda_F^x(\omega_s)^{1/\rho} \cdot {p_s}^{1-1/\rho}}, \tag{A.3.7}$$

for all $i \in N$.

Assume that an information structure F satisfies the MLRP. By Milgrom (1981), the posterior belief also shows the MLRP, that is, for all $x' > x$ and $j > i$,

$$\lambda_F^{x'}(\omega_j)\lambda_F^x(\omega_i) - \lambda_F^x(\omega_j)\lambda_F^{x'}(\omega_i) \geqslant 0 \quad \Leftrightarrow \quad \frac{\lambda_F^{x'}(\omega_j)}{\lambda_F^x(\omega_j)} \geqslant \frac{\lambda_F^{x'}(\omega_i)}{\lambda_F^x(\omega_i)}. \tag{A.3.8}$$

We see that this condition represents the intuition that with a lower signal, a lower state is more likely to happen and with a higher signal, a higher state is more likely to happen. Note that from (A.3.7),

$$\frac{\omega_j + q_j(x')}{\omega_j + q_j(x)} \geqslant \frac{\omega_i + q_i(x')}{\omega_i + q_i(x)} \quad \Leftrightarrow \quad \frac{\lambda_F^{x'}(\omega_j)^{1/\rho}}{\lambda_F^{x}(\omega_j)^{1/\rho}} \geqslant \frac{\lambda_F^{x'}(\omega_i)^{1/\rho}}{\lambda_F^{x}(\omega_i)^{1/\rho}}, \tag{A.3.9}$$

and it holds from (A.3.8). Therefore, we can see that

$$\{q_i(x')\}_{1\leqslant i\leqslant N} \geqslant_A \{q_i(x)\}_{1\leqslant i\leqslant N}$$

for all $x' > x$ and the MCS condition holds.

□

Proposition A.3.2 (Informativeness in the optimal insurance problem). *In the optimal insurance problem, if G satisfies the MLRP and G is a monotone quasi-garbling of F, then the decision maker obtains greater expected utility in the first stage from F than from G for all priors.*

Proof of Proposition A.3.2. When G satisfies the MLRP, by Proposition A.3.1, $\mathcal{U}$ is a general monotone decision problem with respect to G. Since G is a monotone quasi-garbling of F, by Theorem 2.5.1, the expected payoff from F is greater than that from G. □

A.4 Proofs of Lemmas

A.4.1 Proof of Lemma 2.5.2

Proof of Lemma 2.5.2 . Observe that any decreasing decision rule $h : \Omega \to A$ has a following structure:

$$h(\omega) = \begin{cases} 1, & \omega < \tilde{\omega}, \\ 0, & \omega \geqslant \tilde{\omega}. \end{cases}$$

When $\tilde{\omega} \leqslant \hat{\omega}$, observe that

$$u(h(\omega),\omega) = \begin{cases} u(1,\omega) = -\kappa \leqslant 0 = u(0,\omega), & \text{if } \omega < \tilde{\omega} \leqslant \hat{\omega}, \\ u(0,\omega) = 0 \leqslant 0 = u(0,\omega), & \text{if } \tilde{\omega} \leqslant \omega. \end{cases}$$

Thus, $h(\omega)$ is dominated by the constant action $a = 0$.

When $\tilde{\omega} > \hat{\omega}$, observe that

$$u(h(\omega),\omega) = \begin{cases} u(1,\omega) = -\kappa \leqslant -\kappa = u(1,\omega), & \text{if } \omega < \hat{\omega} < \tilde{\omega}, \\ u(1,\omega) = 1-\kappa \leqslant 1-\kappa = u(1,\omega), & \text{if } \hat{\omega} \leqslant \omega < \tilde{\omega}, \\ u(0,\omega) = 0 \leqslant 1-\kappa = u(1,\omega), & \text{if } \hat{\omega} < \tilde{\omega} \leqslant \omega. \end{cases}$$

Thus, $h(\omega)$ is dominated by the constant action $a = 1$. □

A.4.2 Proof of Lemma 2.6.1

Proof of Lemma 2.6.1. Consider any $\{q(\omega; s)\}_{(\omega,s)\in\Omega\times\Omega} \in Q^N$ such that for all $s' > s$

$$\{q(\cdot; s)\} \geqslant_Q \{q(\cdot; s')\},$$

i.e., $q(\omega; s') \geqslant q(\omega; s)$ for all $s' > s$ and $\omega \in \Omega$. Then, I consider $\hat{q}(\cdot)$ defined by $\hat{q}(\omega) \equiv q(\omega; \omega)$. Observe that for any $\omega' > \omega$, we have

$$\hat{q}(\omega') = q(\omega'; \omega') \geqslant q(\omega'; \omega) \geqslant q(\omega; \omega) = \hat{q}(\omega).$$

The first inequality holds from the definition of $\geqslant_Q$ and the second inequality holds since $q(\cdot; \omega)$ is nondecreasing. Therefore, $\hat{q}(\cdot)$ is nondecreasing and $\hat{q}(\cdot) \in Q$.

Next, I show that $u(\hat{q}(\cdot),\omega) \geqslant u(q(\cdot;\omega),\omega)$ for all $\omega \in \Omega$. From the definition of u and $\hat{q}(\omega) = q(\omega;\omega)$, we can derive that

$$\begin{aligned} & u(\hat{q}(\cdot),\omega_n) - u(q(\cdot;\omega_n),\omega_n) \\ &= \sum_{j=1}^{n-1} \left[v(q(\omega_j;\omega_n),\omega_{j+1}) - v(q(\omega_j;\omega_j),\omega_{j+1}) - v(q(\omega_j;\omega_n),\omega_j) + v(q(\omega_j;\omega_j),\omega_j) \right]. \end{aligned}$$

Then, from $q(\omega_j;\omega_n) \geqslant q(\omega_j;\omega_j)$ for $j < n$ and the supermodularity of v, we have $u(\hat{q}(\cdot),\omega) \geqslant u(q(\cdot;\omega),\omega)$. Therefore, $\geqslant_Q$ satisfies the DDDR condition for u. □

A.4.3 Proof of Lemma 2.6.2

Proof of Lemma 2.6.2 . Assume the contrary, that is, $U(\vec{q}\,;x'')$ does not I-dominate $U(\vec{q}\,;x')$. Then, there exist $\vec{q}_a, \vec{q}_b \in \mathcal{Q}$ such that $\vec{q}_a \geqslant_{\mathcal{Q}} \vec{q}_b$ and $U(\vec{q}_a\ ;x') \geqslant U(\vec{q}\,;x')$ for all $\vec{q} \in [\vec{q}_b, \vec{q}_a]$, but (i) $U(\vec{q}_a\ ;x'') < U(\vec{q}_b\ ;x'')$; or (ii) $U(\vec{q}_a\ ;x') > U(\vec{q}_b\ ;x')$ and $U(\vec{q}_a\ ;x'') = U(\vec{q}_b\ ;x'')$.

Consider the first case: $U(\vec{q}_a\ ;x'') < U(\vec{q}_b\ x'')$. Let $\vec{q}_c \in \arg\max_{\vec{q}\in[\vec{q}_b,\vec{q}_a]} U(\vec{q}\,;x'')$. Then, we have $U(\vec{q}_c\,;x'') > U(\vec{q}_a\,;x'')$. Now I show that for each $n \leqslant N$, the following inequality holds:

$$S_n \equiv \sum_{i=1}^{n} \left[u(\vec{q}_c, \omega_i) - u(\vec{q}_a, \omega_i)\right] \cdot \lambda_F^{x''}(\omega_i) \geqslant 0. \tag{A.4.1}$$

Define $\vec{q}_n \in [\vec{q}_c, \vec{q}_a]$ as follows:

$$\vec{q}_n(\omega_i) = \begin{cases} \vec{q}_a(\omega_i), & \text{if } i \leqslant n, \\ \vec{q}_c(\omega_i), & \text{if } i > n. \end{cases}$$

Observe that for any $i \leqslant n$,

$$\begin{aligned} u(\vec{q}_n(\omega_i), \omega_i) &= v(\vec{q}_a(\omega_i), \omega_i) - \sum_{k=1}^{i-1} \left(v(\vec{q}_a(\omega_k), \omega_{k+1}) - v(\vec{q}_a(\omega_k), \omega_k)\right) - c(\vec{q}_a(\omega_i)) \\ &= u(\vec{q}_a(\omega_i), \omega_i). \end{aligned}$$

Next, for any $i > n$, we have

$$\begin{aligned} &u(\vec{q}_n(\omega_i), \omega_i) - u(\vec{q}_c(\omega_i), \omega_i) \\ &= \sum_{k=1}^{n} \left\{\left(v(\vec{q}_c(\omega_k), \omega_{k+1}) - v(\vec{q}_a(\omega_k), \omega_{k+1})\right) - \left(v(\vec{q}_c(\omega_k), \omega_k) - v(\vec{q}_a(\omega_k), \omega_k)\right)\right\}. \end{aligned}$$

Since $\vec{q}_c(\omega_k) \geqslant \vec{q}_a(\omega_k)$ for all $1 \leqslant k \leqslant n$ (from $\vec{q}_a \geqslant_{\mathcal{Q}} \vec{q}_c$) and v is supermodular, we have $u(\vec{q}_n(\omega_i), \omega_i) \geqslant u(\vec{q}_c(\omega_i), \omega_i)$. From the optimality of $\hat{q}_c$, we have $U(\vec{q}_c\ ;x'') \geqslant$

$U(\vec{q}_n\ ; x'')$, which is equivalent to

$$0 \leqslant S_n + \sum_{l=n+1}^{N} \left[u(\vec{q}_c(\omega_k), \omega_k) - u(\vec{q}_n(\omega_k), \omega_k)\right] \lambda_F^{x''}(\omega_l).$$

By using $u(\vec{q}_n(\omega_l), \omega_l) \geqslant u(\vec{q}_c(\omega_l), \omega_l)$ for all $l \geqslant n+1$, we have $S_n \geqslant 0$.

Observe that

$$U(\vec{q}_c\ ; x') - U(\vec{q}_a\ ; x') = \frac{\lambda_F^{x'}(\omega_N)}{\lambda_F^{x''}(\omega_N)} S_N + \sum_{n=1}^{N-1} \left(\frac{\lambda_F^{x'}(\omega_n)}{\lambda_F^{x''}(\omega_n)} - \frac{\lambda_F^{x'}(\omega_{n+1})}{\lambda_F^{x''}(\omega_{n+1})} \right) \cdot S_n.$$

By the MLRP of F, we have $\lambda_F^{x'}(\omega_n)/\lambda_F^{x''}(\omega_n) \geqslant \lambda_F^{x'}(\omega_{n+1})/\lambda_F^{x''}(\omega_{n+1})$ for all $1 \leqslant n \leqslant N-1$. Then, from (A.4.1), we have

$$U(\vec{q}_c\ ; x') - U(\vec{q}_a\ ; x') \geqslant \frac{\lambda_F^{x'}(\omega_N)}{\lambda_F^{x''}(\omega_N)} S_N = \frac{\lambda_F^{x'}(\omega_N)}{\lambda_F^{x''}(\omega_N)} \left(U(\vec{q}_c\ ; x'') - U(\vec{q}_a\ ; x'')\right) > 0,$$

which contradicts the assumption that $\vec{q}_c \in [\vec{q}_b, \vec{q}_a]$ and $U(\vec{q}_a\ ; x') \geqslant U(\vec{q}\ ; x')$ for all $\vec{q} \in [\vec{q}_b, \vec{q}_a]$.

Consider the second case: $U(\vec{q}_a\ ; x') > U(\vec{q}_b\ ; x')$ and $U(\vec{q}_a\ ; x'') = U(\vec{q}_b\ ; x'')$. If there exists $\vec{q}_c \in [\vec{q}_b, \vec{q}_a]$ such that $U(\vec{q}_c\ ; x'') > U(\vec{q}_a\ ; x'')$, we can use the same argument as in the first case. If not, for all $\vec{q} \in [\vec{q}_b, \vec{q}_a]$, $U(\vec{q}_a\ ; x'') = U(\vec{q}_b\ ; x'') \geqslant U(\vec{q}\ ; x'')$, i.e., $\vec{q}_b \in \arg\max_{\vec{q} \in [\vec{q}_b, \vec{q}_a]} U(\vec{q}\ ; x'')$. Then, we can use the similar argument as in the first case by substituting $\vec{q}_c$ to $\vec{q}_b$. The only difference is that the last inequality in (A.4.3) should be replaced to equality, thus we have $U(\vec{q}_b\ ; x') \geqslant U(\vec{q}_a\ ; x')$. But it contradicts the assumption that $U(\vec{q}_a\ ; x') > U(\vec{q}_b\ ; x')$. Therefore, $U(\vec{q}\ ; x'')$ I-dominates $U(\vec{q}\ ; x')$. $\square$

Appendix B

Appendix for Chapter 3

B.1 Contracts

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 payoffs. Note that if the agent has not completed either the main project or the subproject, the calendar time is the only relevant variable summarizing the public history.

A (deterministic) contract is denoted by $\Gamma \equiv \left\{T, \{a_t, b_t, R_t, \hat{\Gamma}^t\}_{0\leqslant t\leqslant T}\right\}$, where each variable is defined as follows at the calendar time t:[1]

1. $T \in \mathbb{R}_+ \cup \{\infty\}$: the deadline date at which the project is terminated absent the completion of the main project or the subproject. $T = \infty$ means that no deadline is included in the contract;

2. $a_t \in \{0, 1\}$: the principal's choice of an approach at t;

3. $b_t \in [0, 1]$: the agent's recommended effort at t;

[1] See Remark 6 for discussion on deterministic and mixed contracts.

4. $R_t \geqslant 0$: the monetary payment from the principal to the agent for the success of the main project at t;[2]

5. $\hat{\Gamma}^t \equiv \{T^t, \{b_s^t, R_s^t\}_{t \leqslant s \leqslant T^t}\}$: an updated contract when the subproject is completed at t;

   (a) $T^t \in \{\check{T} \; : \; \check{T} \geqslant t\} \cup \{\infty\}$: the deadline date at which the project is terminated;

   (b) $b_s^t \in [0, 1]$: the agent's recommended effort at time $s \geqslant t$;

   (c) $R_s^t \geqslant 0$: the monetary payment from the principal to the agent for the completion of the main project at time s.

Consider the case where the subproject is completed at time t. Then, the updated contract $\hat{\Gamma}^t$ will be executed. In this case, the agent's admissible action space is $\hat{\mathcal{B}}^t \equiv \{\{\check{b}_s\}_{t \leqslant s \leqslant T^t} \; : \; \check{b}_s \in [0, 1]\}$. The agent's action profile $\tilde{b}^t \equiv \{\tilde{b}_s^t\}_{t \leqslant s \leqslant T^t} \in \hat{\mathcal{B}}^t$ induces a probability distribution $\mathbb{P}^{\tilde{b}^t}$ over a main project completion date τ_m. Let $\mathbb{E}^{\tilde{b}^t}$ denote the corresponding expectation operator. When the agent adheres to the recommended action of $\hat{\Gamma}^t$, the principal's expected utility at time t is given by

$$\hat{P}^t(\hat{\Gamma}^t) = \mathbb{E}^{b^t}\left[\left(\Pi - R_{\tau_m}^t\right) \cdot \mathbf{1}_{\{t \leqslant \tau_m \leqslant T^t\}} - \int_t^{T^t \wedge \tau_m} c \, ds\right],^3$$

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

$$\hat{U}^t(\hat{\Gamma}^t) = \mathbb{E}^{b^t}\left[R_{\tau_m}^t \cdot \mathbf{1}_{\{t \leqslant \tau_m \leqslant T^t\}} + \int_t^{T^t \wedge \tau_m} \phi(1 - b_s^t) ds\right],$$

[2] Since both the principal and the agent are risk neutral and do not discount the future, without loss of generality, all monetary payments to the agent can be backloaded (see, e.g., Ray, 2002). The nonnegativity of R_t is due to limited liability.

[3] For each x and y, let $x \wedge y$ denote the minimum of x and y, and let $x \vee y$ denote the maximum of x and y.

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

Now consider the problem at time 0. The agent's admissible action space (prior to any completion) is $\mathcal{B} \equiv \{\{\check{b}_t\}_{0 \leqslant t \leqslant T} \ : \ \check{b}_t \in [0,1]\}$. In this case, any completion depends not only on the agent's effort ($\tilde{b}_t$) but also the principal's choice of approach (a_t). Then, a pair of actions by the principal and the agent, $(a, \tilde{b})$, induces a probability distribution $\mathbb{P}^{a,\tilde{b}}$ over a pair of completion dates for the main project and the subproject (τ_m, τ_s). Let $\mathbb{E}^{a,\tilde{b}}$ denote the corresponding expectation operator. If the agent adheres to the recommended actions of Γ, the principal's (ex ante) expected utility is given by

$$P_0(\Gamma) = \mathbb{E}^{a,b}\left[(\Pi - R_{\tau_m}) \cdot \mathbf{1}_{\{\tau_m < \tau_s \wedge T\}} + \hat{P}_{\tau_s}(\hat{\Gamma}_{\tau_s}) \cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T\}} - \int_0^{T \wedge \tau_m \wedge \tau_s} c \ dt\right], \tag{B.1.1}$$

where the first term is the net profit from the main project completion, the second term is the expected payoff from the subproject completion at time τ_s, and the last term is the cumulative operating cost. The agent's expected utility is given by

$$U_0(\Gamma) = \mathbb{E}^{a,b}\left[R_{\tau_m} \cdot \mathbf{1}_{\{\tau_m \leqslant T\}} + \hat{U}^{\tau_s}(\hat{\Gamma}^{\tau_s}) \cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T\}} + \int_0^{T \wedge \tau_m \wedge \tau_s} \phi(1 - b_t) \ dt\right], \tag{B.1.2}$$

where the first term is the payoff from the main project completion, the second term is the expected payoff from the subproject completion at time τ_s, and the last term is the benefit from shirking. By using the agent's expected payoffs, I define incentive compatibility (IC) of contracts as follows.

Definition B.1.1. A contract $\Gamma = \left\{T, \{a_t, b_t, R_t, \hat{\Gamma}^t\}_{0 \leqslant t \leqslant T}\right\}$ is *incentive compatible* if

1. for all $t \leqslant T$, the recommended effort profile $\{b_s^t\}_{t \leqslant s \leqslant T^t}$ in the updated contract $\hat{\Gamma}^t$ maximizes the agent's expected utility at time t, i.e.,

$$\hat{U}^t(\hat{\Gamma}^t) = \max_{\tilde{b} \in \hat{\mathcal{B}}^t} \mathbb{E}^{\tilde{b}} \left[R_{\tau_m}^t \cdot \mathbf{1}_{\{\tau_m \leqslant T^t\}} + \int_t^{T^t \wedge \tau_m} \phi(1 - \tilde{b}_s) ds \right].$$

2. the recommended action profile $\{b_t\}_{0 \leqslant t \leqslant T}$ maximizes the agent's expected utility at time 0, i.e.,

$$U_0(\Gamma) = \max_{\tilde{b} \in \mathcal{B}} \mathbb{E}^{a,\tilde{b}} \left[R_{\tau_m} \cdot \mathbf{1}_{\{\tau_m \leqslant T\}} + \hat{U}^{\tau_s}(\hat{\Gamma}^{\tau_s}) \cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T\}} + \int_0^{T \wedge \tau_m \wedge \tau_s} \phi(1 - \tilde{b}_t)\ dt \right].$$

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

B.2 Recursive Formulation

B.2.1 The Agent's Problem

I now consider the agent's problem when he has not yet completed the first subproject. Given a contract Γ, let $U_t(\Gamma)$ denote the agent's maximized continuation utility at time t, that is,

$$U_t(\Gamma) \equiv \max_{\tilde{b} \in \mathcal{B}_t} \mathbb{E}^{a,\tilde{b}} \left[\begin{array}{c} R_{\tau_m} \cdot \mathbf{1}_{\{\tau_m \leqslant \tau_s \wedge T\}} + \hat{U}^{\tau_s}(\hat{\Gamma}^{\tau_s}) \cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T\}} \\ + \int_t^{T \wedge \tau_m \wedge \tau_s} \phi(1 - \tilde{b}_s)\ ds \end{array} \middle| \ t \leqslant \tau_m \wedge \tau_s \right], \tag{B.2.1}$$

where $\mathcal{B}_t \equiv \{\{\tilde{b}_s\}_{t \leqslant s \leqslant T} \ : \ \tilde{b}_s \in [0,1]\}$.

Observe that the agent's continuation utility can be heuristically rewritten as follows:

$$u_t = \max_{\tilde{b}_t \in [0,1]} \phi(1-\tilde{b}_t)dt + R_t \cdot \lambda_D a_t \tilde{b}_t dt + u_S^t \cdot \lambda_S (1-a_t)\tilde{b}_t dt$$

$$+ (1 - \lambda_D a_t \tilde{b}_t dt - \lambda_S(1-a_t)\tilde{b}_t dt) \cdot u_{t+dt}$$

where $u_t = U_t(\Gamma)$ and $u_S^t = \hat{U}^t(\hat{\Gamma}^t)$. Using a Taylor expansion $u_{t+dt} = u_t + \dot{u}_t dt + o(dt)$ where $\dot{u}_t \equiv du_t/dt$, canceling u_t on both sides, and taking the limit as $dt \to 0$, we obtain a Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \max_{\tilde{b}_t \in [0,1]} \dot{u}_t + \phi(1-\tilde{b}_t) + (R_t - u_t)\lambda_D a_t \tilde{b}_t + (u_S^t - u_t)\lambda_S(1-a_t)\tilde{b}_t. \qquad (\mathrm{HJB}_{PK})$$

Also note that $U_T(\Gamma) = 0$ since the contract is terminated at time T. The following lemma shows that the HJB equation (HJB_{PK}) with a boundary condition $u_T = 0$ characterizes the evolution of the continuation utility $U_t(\Gamma)$. The proof is relegated to Appendix B.2.1.

Lemma B.2.1. *Given a contract* Γ*, suppose that a continuous and differentiable process* $\{u_t\}_{0 \leqslant t \leqslant T}$ *satisfies* $u_T = 0$ *and* (HJB_{PK})*. Then,* $u_t = U_t(\Gamma)$*.*

The HJB equation (HJB_{PK}) provides a clear interpretation of the agent's behavior. The first term is the drift term of the agent's continuation utility from no success and the second term is the benefit from shirking. When the main project is completed at rate $\lambda_D a_t \tilde{b}_t$, the agent receives the immediate payment R_t but he loses the continuation utility since the contract is terminated. When the subproject is completed at rate $\lambda_S(1-a_t)\tilde{b}_t$, the new phase of the contract with the promised utility u_S^t begins and he loses the continuation utility since the current phase of the contract is over.

Proof of Lemma B.2.1

In this subsection, I formally derive the agent's continuation utility and prove Lemma B.2.1.

I begin by specifying the probability distribution functions for possible events given a profile of approaches $a = \{a_s\}_{t \leqslant s \leqslant T}$ and an admissible action profile $b \in \mathcal{B}_t$ conditional on no completion has made by time t. The probability that neither the main project nor the subproject is completed by time T is $f(a, b; t, T)$ where

$$f(a, b; x, y) \equiv e^{-\lambda_D \int_x^y a_l b_l dl} \cdot e^{-\lambda_S \int_x^y (1-a_l) b_l dl}.$$

Next, the probability density that the main project is completed at time s and the subproject is not completed by that time ($s = \tau_m < \tau_s$) is $\lambda_D a_s b_s \cdot f(a, b; t, s)$. Similarly, the probability density that the subproject is completed at time s and the main project is not completed by that time ($s = \tau_s < \tau_m$) is $\lambda_S(1-a_s)b_s \cdot f(a, b; t, s)$. Last, the probability density that either the main project or the subproject is completed at time s and the other has not arrived by then, i.e., $\tau_s \wedge \tau_m = s$, is $(\lambda_D a_s + \lambda_S(1 - a_s))b_s \cdot f(a, b; t, s)$.

Based on the above results, we can derive that

$$\mathbb{E}^{a,b}\left[R_{\tau_m} \cdot \mathbf{1}_{\{\tau_m \leqslant \tau_s \wedge T\}} \mid t \leqslant \tau_m \wedge \tau_s\right] = \int_t^T R_s \cdot \lambda_D a_s b_s \cdot f(a, b; t, s) ds,$$

$$\mathbb{E}^{a,b}\left[\hat{U}_{\tau_s}(\hat{\Gamma}^{\tau_s}) \cdot \mathbf{1}_{\{\tau_s < \tau_m \wedge T\}} \mid t \leqslant \tau_m \wedge \tau_s\right] = \int_t^T \hat{U}_s(\hat{\Gamma}^s) \cdot f(a, b; t, s) ds.$$

Observe that $\frac{d}{ds} f(a, b; t, s) = -(\lambda_D a_s + \lambda_S(1-a_s))b_s \cdot f(a, b; t, s)$. By using integration by parts, we have

$$\begin{aligned}
&\mathbb{E}^{a,b}\left[\int_t^{T \wedge \tau_m \wedge \tau_s} \phi(1 - b_s) \mid t \leqslant \tau_m \wedge \tau_s\right] \\
&= \int_t^T \left[\int_t^s \phi(1 - b_l) dl\right] \cdot (\lambda_D a_s + \lambda_S(1 - a_s))b_s \cdot f(a, b; t, s) ds \\
&\quad + \left[\int_t^T \phi(1 - b_s) ds\right] \cdot f(a, b; t, T) \\
&= \int_t^T \phi(1 - b_s) \cdot f(a, b; t, s) ds.
\end{aligned}$$

Then, by plugging the above expressions into (B.2.1), $U_t(\Gamma)$ can be rewritten as follows:

$$U_t(\Gamma) = \sup_{\tilde{b}\in\mathcal{B}_t} \int_t^T \left[R_s \cdot \lambda_D a_s \tilde{b}_s + \hat{U}_s(\hat{\Gamma}^s) \cdot \lambda_S (1-a_s)\tilde{b}_s + \phi(1-\tilde{b}_s)\right] f(a, \tilde{b}; t, s) ds. \tag{B.2.2}$$

Now we can prove Lemma B.2.1 by using the above equation. The proof is inspired by Proposition 3.2.1 in Bertsekas (1995).

Proof of Lemma B.2.1. Consider an arbitrary admissible action $\tilde{b} \in \mathcal{B}_t$. By rearranging ($\text{HJB}_{PK}$), we can derive that

$$-\dot{u}_s + (\lambda_D a_s + \lambda_S(1-a_s))\tilde{b}_s u_t \geqslant R_s \lambda_D a_s \tilde{b}_s + u_S^s \lambda_S (1-a_s)\tilde{b}_s + \phi(1-\tilde{b}_s)$$

and it is equivalent to

$$\frac{d}{ds}\left[-u_s \cdot f(a,b;t,s)\right] \geqslant \left[\left(R_s\lambda_D a_s + u_S^s \lambda_S(1-a_s)\right)\tilde{b}_s + \phi(1-\tilde{b}_s)\right] \cdot f(a,b;t,s).$$

By integrating the above inequality from t to T and using $u_T = 0$, we can derive that

$$u_t \geqslant \int_t^T \left[\left(R_s\lambda_D a_s + u_S^s \lambda_S(1-a_s)\right)\tilde{b}_s + \phi(1-\tilde{b}_s)\right] \cdot f(a,\tilde{b};t,s) ds$$

for all $\tilde{b} \in \mathcal{B}_t$.

Suppose that $b^* \in \mathcal{B}_t$ attains the maximum in the equation (HJB_{PK}) for all $0 \leqslant t \leqslant T$. Then, we have

$$\begin{aligned} u_t &= \int_t^T \left[\left(R_s\lambda_D a_s + u_S^s (1-a_s)\right)\lambda_S b_s^* + \phi(1-b_s^*)\right] \cdot f(a,b^*;t,s) ds \\ &\geqslant \int_t^T \left[\left(R_s\lambda_D a_s + u_S^s \lambda_S(1-a_s)\right)\tilde{b}_s + \phi(1-\tilde{b}_s)\right] \cdot f(a,\tilde{b};t,s) ds \end{aligned}$$

for all $\tilde{b} \in \mathcal{B}_t$. Therefore, by (B.2.2), we have $u_t = U_t(\Gamma)$. □

B.2.2 The Principal's Problem

I now consider the principal's problem. I begin by considering the incentive compatibility condition. To make a contract incentive compatible, at each point of time, the recommended effort level should coincide with the agent's choice in (HJB_{PK}), that is,

$$b \in \arg\max_{\tilde{b}\in[0,1]} \phi(1-\tilde{b}) + (R-u)\lambda_D a\tilde{b} + (u_S - u)\lambda_S(1-a)\tilde{b}. \tag{IC}$$

In addition, since (HJB_{PK}) is linear in $\tilde{b}$, it can be rewritten as follows:

$$\dot{u}_t = -\left[\phi \vee \left((R_t - u_t)\lambda_D a_t + (u_S^t - u_t)\lambda_S(1-a_t)\right)\right]. \tag{B.2.3}$$

I now explore how the principal's value function evolves. Note that $V(0) = 0$ since the agent will not participate in the contract when the continuation utility is zero. This will serve as a boundary condition. The value function $V(u_t)$ can be heuristically written as follows:

$$V(u_t) = \max_{\substack{R_t \geqslant 0,\ u_S^t \geqslant 0,\\ a_t \in\{0,1\},\ b_t\in[0,1]}} \begin{array}{l} -cdt + (\Pi - R_t)\,\lambda_D a_t b_t dt + V_S(u_S^t)\lambda_S(1-a_t)b_t dt \\ +\,(1 - \lambda_D a_t b_t dt - \lambda_S(1-a_t)b_t dt)V(u_{t+dt}) \end{array}.$$

By using $V(u_{t+dt}) = V(u_t) + V'(u_t)\dot{u}_t dt + o(dt)$, canceling $V(u_t)$ on both sides, taking the limit as $dt \to 0$ and plugging (B.2.3) in, we obtain an HJB equation:

$$0 = \max_{\substack{R\geqslant 0,\ u_S \geqslant 0,\\ a\in\{0,1\},\ b\in[0,1]}} \begin{array}{l} -c + (\Pi - R - V(u))\lambda_D ab + (V_S(u_S) - V(u))\lambda_S(1-a)b \\ -\left[\phi \vee \{(R-u)\lambda_D a + (u_S - u)\lambda_S(1-a)\}\right] V'(u) \end{array}. \tag{HJB_V}$$

Then, the principal's problem is to solve (HJB_V) subject to (IC) with the boundary condition $V(0) = 0$. The following lemma shows that the solution of the problem maximizes the principal's expected payoff subject to a promise keeping constraint $U_0(\Gamma) = u$. The proof is relegated to Appendix B.2.2.

Lemma B.2.2 (Verification Lemma)**.** *Suppose that a differentiable and concave function* $\bar{V}$ *solves* (HJB$_V$) *subject to* (IC) *with the boundary condition* $\bar{V}(0) = 0$*. Then, for any incentive-compatible contract* Γ *with* $U_0(\Gamma) = u$*,*

$$\bar{V}(u) \geqslant P_0(\Gamma).$$

Given this result, I derive the value function by using the 'guess and verify' method. I construct a differentiable and concave value function, $\bar{V} : \mathbb{R}_+ \to \mathbb{R}$, which solves (HJB$_V$) subject to (IC) for all $u \geqslant 0$. Next, for any $u \geqslant 0$, I find a deterministic and incentive-compatible contract Γ that implements $(u, \bar{V}(u))$. Then, by the above verification lemma, $\bar{V}(u)$ is the highest expected payoff among incentive-compatible contracts with $U_0(\Gamma) = u$, i.e., $\bar{V}(u) = V(u)$. In the main text, I provide intuition of how I guess the value function. The actual guess appears in Appendix B.2.4. The formal verification proof is relegated to the Online Appendix.

Proof of Lemma B.2.2

Proof of Lemma B.2.2. Consider an arbitrary (deterministic) incentive-compatible contract Γ where the agent's expected payoff is given by u_t. The payoff to the principal under Γ is

$$\begin{aligned} P_0(\Gamma) = &\int_0^T (\Pi - R_t - c \cdot t) \cdot \lambda_D a_t b_t f(a, b; 0, t) dt \\ &+ \int_0^T (V_S(u_{S,t}) - c \cdot t) \cdot \lambda_S (1 - a_t) b_t f(a, b; 0, t) dt - c \cdot T \cdot f(a, b; 0, T) \\ = &\int_0^T ((\Pi - R_t)\lambda_D a_t b_t + V_S(u_{S,t})\lambda_S(1 - a_t) b_t - c) f(a, b; 0, t) dt \end{aligned}$$

where $u_{S,t} = \hat{U}_t(\hat{\Gamma}^S)$.

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

$$\begin{aligned} 0 \geqslant &- c + (\Pi - R_t - \tilde{V}(u_t))\lambda_D a_t b_t + (V_S(u_{S,t}) - \tilde{V}(u_t))\lambda_S(1 - a_t) b_t \\ &- [\phi \vee \{(R_t - u_t)\lambda_D a_t + (u_{S,t} - u_t)\lambda_S(1 - a_t)\}] \tilde{V}'(u_t). \end{aligned}$$

By using (B.2.3), rearranging, and multiplying by $f(a,b;0,t)$, we can obtain that

$$\begin{aligned}&(\lambda_D a_t b_t + \lambda_S(1-a_t)b_t) f(a,b;0,t)\cdot \tilde{V}(u_t) - f(a,b;0,t)\cdot \tilde{V}'(u_t)\dot{u}_t \\ \geqslant& f(a,b;0,t)\left((\Pi - R_t)\lambda_D a_t b_t + V_S(u_{S,t})\lambda_S(1-a_t)b_t - c\right)\end{aligned} \tag{B.2.4}$$

Note that

$$\frac{d}{dt}\left(-f(a,b;0,t)\tilde{V}(u_t)\right) = (\lambda_D a_t b_t + \lambda_S(1-a_t)b_t) f(a,b;0,t)\cdot \tilde{V}(u_t) - f(a,b;0,t)\cdot \tilde{V}'(u_t)\dot{u}_t.$$

Then, by integrating (B.2.4) over $[0,T]$ and noting that $f(a,b;0,0)=1$, $u_T=0$ and $\tilde{V}(0)=0$, we have

$$\begin{aligned}&\tilde{V}(u_0) = \tilde{V}(u_0) - f(a,b;0,T)\tilde{V}(u_T) \\ \geqslant& \int_0^T f(a,b;0,t)\cdot\left((\Pi - R_t)\lambda_D a_t b_t + V_S(u_{S,t})\lambda_S(1-a_t)b_t - c\right) dt = P_0(\Gamma).\end{aligned}$$

Therefore, $\tilde{V}(u_0)$ is greater than or equal to any deterministic contract where the agent's expected payoff is equal to u_0. Since $\tilde{V}$ is assumed to be concave, it is greater than or equal to any randomized contract. □

B.2.3 Implementation

Proof of Proposition 3.4.1

Proof of Proposition 3.4.1. (a) Let $\Gamma_d(T)$ denote a direct-only contract with the deadline T. The agent's expected payoff is

$$\begin{aligned}U_0(\Gamma_d(T)) =& \int_0^T R_{\tau_m}\lambda_D e^{-\lambda_D\tau_m} d\tau_m \\ =& \int_0^T \phi\left[T - \tau_m + 1/\lambda_D\right]\lambda_D e^{-\lambda_D\tau_m} d\tau_m \\ =& -\phi\left(T-\tau_m\right)e^{-\lambda_D\tau_m}\Big|_0^T \\ =& \phi T.\end{aligned}$$

Therefore, $U_0(\Gamma_d(u/\phi)) = u$.

Also note that

$$P_0(\Gamma_d(T)) + U_0(\Gamma_d(T)) = \int_0^T (\Pi - c\tau_m)\,\lambda_D e^{-\lambda_D\tau_m}d\tau_m - cTe^{-\lambda_D T}$$

$$= - (\Pi - c\tau_m - c/\lambda_D)\,e^{-\lambda_D\tau_m}\Big|_0^T - cTe^{-\lambda_D T}$$

$$= - (\Pi - cT - c/\lambda_D)\,e^{-\lambda_D T} + (\Pi - c/\lambda_D) - cTe^{-\lambda_D T}$$

$$= (\Pi - c/\lambda_D)\left(1 - e^{-\lambda_D T}\right).$$

Therefore,

$$P_0(\Gamma_d(u/\phi)) = \left(\Pi - \frac{c}{\lambda_D}\right)(1 - e^{-\frac{\lambda_D}{\phi}u}) - U_0(\Gamma_d(u/\phi))$$

$$= \left(\Pi - \frac{c}{\lambda_D}\right)(1 - e^{-\frac{\lambda_D}{\phi}u}) - u = V^d(u).$$

(b) Let $\Gamma_{sd}(T_1, T)$ denote a contract with a switch from the sequential approach to the direct approach at T_1 and the deadline T. The subcontract at time $t \leqslant T_1$ is denoted by $\hat{\Gamma}_{sd}(t|T_1, T)$. Then, the agent's expected payoff for the subcontract $\hat{\Gamma}_{sd}(t|T_1, T)$ at time t is

$$U_t(\hat{\Gamma}_{sd}(t|T_1, T)) = \int_t^{T+1/\lambda_S} \phi\,(T + 1/\lambda_S - \tau_m + 1/\lambda_S)\,\lambda_S e^{-\lambda_S(\tau_m - t)}d\tau_m$$

$$= - \phi\,(T + 1/\lambda_S - \tau_m)\,e^{-\lambda_S(\tau_m - t)}\Big|_t^{T+1/\lambda_S}$$

$$= \phi\,(T + 1/\lambda_S - t)\,.$$

Also note that

$$\int_0^{T_1} U_{\tau_s}(\hat{\Gamma}_s(\tau_s|T_1, T))\lambda_S e^{-\lambda_S\tau_s}d\tau_s = \int_0^{T_1} \phi(T + 1/\lambda_S - \tau_s)\lambda_S e^{-\lambda_S\tau_s}d\tau_s$$

$$= - \phi(T - \tau_s)e^{-\lambda_S\tau_s}\Big|_0^{T_1}$$

$$= \phi T - \phi(T - T_1)e^{-\lambda_S T_1}.$$

Then, the agent's expected payoff at time 0 is

$$
\begin{aligned}
U_0(\Gamma_{sd}(T_1,T)) &= \int_0^{T_1} U_{\tau_s}(\hat{\Gamma}_s(\tau_s|T_1,T))\lambda_S e^{-\lambda_S\tau_s}d\tau_s \\
&\quad + e^{-\lambda_S T_1}\int_{T_1}^{T}\phi(T+1/\lambda_D-\tau_m)\lambda_D e^{-\lambda_D(\tau_m-T_1)}d\tau_m \\
&=\phi T-\phi(T-T_1)e^{-\lambda_S T_1}-e^{-\lambda_S T_1}\left[\phi\left(T-\tau_m\right)e^{-\lambda_D(\tau_m-T_1)}\Big|_{T_1}^{T}\right] \\
&=\phi T.
\end{aligned}
$$

Thus, $U_0(\Gamma_{sd}(T_1,u/\phi))=u$.

The sum of expected payoffs for the subcontract is

$$
\begin{aligned}
&P_t(\hat{\Gamma}_{sd}(t|T_1,T))+U_t(\hat{\Gamma}_{sd}(t|T_1,T)) \\
&=\int_t^{T+1/\lambda_S}(\Pi-c(\tau_m-t))\lambda_S e^{-\lambda_S(\tau_m-t)}d\tau_m-c\left(T+\frac{1}{\lambda_S}-t\right)e^{-\lambda_S\left(T+\frac{1}{\lambda_S}-t\right)} \\
&=-\left(\Pi-\frac{c}{\lambda_S}-c(\tau_m-t)\right)e^{-\lambda_S(\tau_m-t)}\Big|_t^{T+1/\lambda_S}-c\left(T+\frac{1}{\lambda_S}-t\right)e^{-\lambda_S\left(T+\frac{1}{\lambda_S}-t\right)} \\
&=-\left(\Pi-\frac{c}{\lambda_S}-c\left(T+\frac{1}{\lambda_S}-t\right)\right)e^{-\lambda_S\left(T+\frac{1}{\lambda_S}-t\right)}+\Pi-\frac{c}{\lambda_S} \\
&\quad -c\left(T+\frac{1}{\lambda_S}-t\right)e^{-\lambda_S\left(T+\frac{1}{\lambda_S}-t\right)} \\
&=\left(\Pi-\frac{c}{\lambda_S}\right)\left(1-e^{-\lambda_S\left(T+\frac{1}{\lambda_S}-t\right)}\right).
\end{aligned}
$$

Also note that

$$\int_0^{T_1} \left[P_{\tau_s}(\hat{\Gamma}_{sd}(\tau_s|T_1,T)) + U_{\tau_s}(\hat{\Gamma}_{sd}(\tau_s|T_1,T)) - c\tau_s\right] \lambda_S e^{-\lambda_S \tau_s} d\tau_s$$

$$= -\left(\Pi - 2c/\lambda_S - c\tau_s\right)\Big|_0^{T_1} - \left(\Pi - c/\lambda_S\right) e^{-\lambda_S(T+1/\lambda_S)}\tau_s\Big|_0^{T_1}$$

$$= \left(\Pi - \frac{2c}{\lambda_S}\right)\left(1 - e^{-\lambda_S T_1}\right) + cT_1 e^{-\lambda_S T_1} - \left(\Pi - \frac{c}{\lambda_S}\right) T_1 e^{-\lambda_S\left(T+\frac{1}{\lambda_S}\right)},$$

and

$$\int_{T_1}^{T} \left(\Pi - c(\tau_m - T_1)\right) \lambda_D e^{-\lambda_D(\tau_m - T_1)} d\tau_m - c(T-T_1)e^{-\lambda_D(T-T_1)}$$

$$=V^d((T-T_1)/\phi) + (T-T_1)/\phi = V^d(u_1) + u_1.$$

Then, we can derive that

$$P_0(\Gamma_{sd}(T_1,T)) + U_0(\Gamma_{sd}(T_1,T))$$

$$= \int_0^{T_1} \left[P_{\tau_s}(\hat{\Gamma}_s(\tau_s|T)) + U_{\tau_s}(\hat{\Gamma}_s(\tau_s|T)) - c\tau_s\right] \lambda_S e^{-\lambda_S \tau_s} d\tau_s - cT_1 e^{-\lambda_S T_1}$$

$$+ e^{-\lambda_S T_1}\left[\int_{T_1}^{T} \left(\Pi - c(\tau_m - T_1)\right) \lambda_D e^{-\lambda_D(\tau_m - T_1)} d\tau_m - c(T-T_1)e^{-\lambda_D(T-T_1)}\right]$$

$$= \left(\Pi - \frac{2c}{\lambda_S}\right)(1 - e^{-\lambda_S T_1}) - \left(\Pi - \frac{c}{\lambda_S}\right) T_1 e^{-\lambda_S\left(T+\frac{1}{\lambda_S}\right)} + e^{-\lambda_S T_1}\left(V^d(u_1) + u_1\right)$$

$$= \left(\Pi - \frac{2c}{\lambda_S}\right)\left(1 - e^{\frac{\lambda_S}{\phi}(u_1 - u)}\right) + (V^d(u_1) + u_1)e^{\frac{\lambda_S}{\phi}(u_1-u)}$$

$$- \left(\Pi - \frac{c}{\lambda_S}\right)\frac{\lambda_S}{\phi}(u - u_1)e^{-\frac{\lambda_S}{\phi}u - 1},$$

thus $P_0(\Gamma_{sd}(T_1,T)) = V^{ds}(u|u_1)$.

(c) Note that a sequential-only contract with a deadline T is equivalent to a contract with a switch from the sequential approach to the direct approach at $T_1 = T$

and a deadline T. Therefore, by the previous result, a sequential-only contract with the deadline u/ϕ implements $(V^{ds}(u|0), u)$.

$\square$

Proof of Proposition 3.5.1

Proof of Proposition 3.5.1. Let $\Gamma_{dsd}(T_1, T_2, T)$ denote a contract with two switches at T_1 and T_2 and a deadline T. Note that at time T_1 (if the project has not been successful), the remaining contract is equivalent to $\Gamma_{sd}(T_2 - T_1, T - T_1)$. Then, the agent's expected payoff at time 0 is

$$\begin{aligned}U_0(\Gamma_{dsd}(T_1, T_2, T)) = & \int_0^{T_1} \phi(T + 1/\lambda_D - \tau_m)\lambda_D e^{-\lambda_D \tau_m} d\tau_m \\ & + e^{-\lambda_S T_1} U_0(\Gamma_{sd}(T_2 - T_1, T - T_1)) \\ = & \phi T - \phi(T - T_1)e^{-\lambda_D T_1} + e^{-\lambda_D T_1}\phi(T - T_1) \\ = & \phi T.\end{aligned}$$

Thus, $U_0(\Gamma_{dsd}(T_1, T_2, u/\phi)) = u$.

Also note that

$$\begin{aligned}& P_0(\Gamma_{dsd}(T_1, T_2, T)) + U_0(\Gamma_{dsd}(T_1, T_2, T)) \\ = & \int_0^{T_1} (\Pi - c\tau_m)\lambda_D e^{-\lambda_D \tau_m} d\tau_m - cT_1 e^{-\lambda_D T_1} \\ & + e^{-\lambda_D T_1}(P_0(\Gamma_{sd}(T_2 - T_1, T - T_1)) + U_0(\Gamma_{sd}(T_2 - T_1, T - T_1))) \\ = & \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{\lambda_D T_1}\right) + \left(V^{ds}(\phi(T - T_1)|\phi(T - T_2)) + \phi(T - T_1)\right) e^{-\lambda_D T_1} \\ = & \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{\frac{\lambda_D}{\phi}(u_2 - u)}\right) + \left(V^{ds}(u_2|u_1) + u_2\right) e^{\frac{\lambda_D}{\phi}(u_2 - u)},\end{aligned}$$

thus $P_0(\Gamma_{dsd}(T_1, T_2, T)) = V^{dsd}(u|u_1, u_2) - u$. $\square$

B.2.4 Properties of Benchmark Value Functions

Proposition B.2.1. *The following equations hold:*

$$\phi V^{d\prime}(u) = -\,c + \lambda_D \left(\Pi - \frac{\phi}{\lambda_D} - u - V^d(u)\right) = \alpha_D(u), \tag{B.2.5}$$

$$\phi V^{ds\prime}(u|u_1) = -\,c + \lambda_S \left(V_S\left(u + \phi/\lambda_S\right) - V^{ds}(u|u_1)\right) = \alpha_S(u), \tag{B.2.6}$$

$$\phi V^{dsd\prime}(u|u_1, u_2) = -\,c + \lambda_D \left(\Pi - \frac{\phi}{\lambda_D} - u - V^{dsd}(u|u_1, u_2)\right) = \alpha_D(u). \tag{B.2.7}$$

Proof of Proposition B.2.1. By taking the derivative of (3.4.1) and multiplying by ϕ, we have

$$\begin{aligned}\phi V^{d\prime}(u) =& \lambda_D \left(\Pi - \frac{c}{\lambda_D}\right) e^{-\frac{\lambda_D}{\phi}u} - \phi = -c + \lambda_D \left(\Pi - \frac{\phi}{\lambda_D} - V^d(u) - u\right), \\ =& \lambda_D \Pi - \lambda_D(V^d(u) + u) - c - \phi = \alpha_D(u),\end{aligned}$$

i.e., (B.2.5) holds.

Next, by taking the derivative of (3.4.2) and multiplying by ϕ, we have

$$\begin{aligned}\phi V^{ds\prime}(u|u_1) =& \lambda_S \left(\Pi - \frac{2c}{\lambda_S}\right) e^{\frac{\lambda_S}{\phi}(u_1 - u)} - \lambda_S(V^d(u_1) + u_1) e^{\frac{\lambda_S}{\phi}(u_1 - u)} \\ & - \left(\Pi - \frac{c}{\lambda_S}\right) \lambda_S e^{-\frac{\lambda_S}{\phi}u - 1} + \lambda_S \left(\Pi - \frac{c}{\lambda_S}\right) \frac{\lambda_S}{\phi}(u - u_1) e^{-\frac{\lambda_S}{\phi}u - 1} - \phi.\end{aligned}$$

Observe that it can be rewritten as follows:

$$\begin{aligned}\phi V^{ds\prime}(u|u_1) =& -\,c + \lambda_S \left(\Pi - \frac{c}{\lambda_S}\right) \left(1 - e^{\frac{-\lambda_S}{\phi}u - 1}\right) - \lambda_S u - \lambda_S V^{ds}(u|u_1) \\ =& -\,c + \lambda_S \left(V_S(u + \phi/\lambda_S) - V^{ds}(u|u_1)\right) \\ =& \lambda_S \left(V_S(u + \phi/\lambda_S) + u + \phi/\lambda_S\right) - \lambda_S(V^{ds}(u|u_1) + u) - c - \phi = \alpha_S(u),\end{aligned}$$

i.e., (B.2.6) holds.

Last, by taking the derivative of (3.5.1) and multiplying by ϕ, we have

$$\begin{aligned}\phi V^{dsd\prime}(u|u_1,u_2) =&\lambda_D\left(\Pi-\frac{c}{\lambda_D}\right)e^{\frac{\lambda_D}{\phi}(u_2-u)}-\lambda_D\left(V^{ds}(u_2|u_1)+u_2\right)e^{\frac{\lambda_D}{\phi}(u_2-u)}-\phi\\ =&-c+\lambda_D\left(\Pi-\frac{\phi}{\lambda_D}-uV^{dsd}(u|u_1,u_2)\right),\\ =&\lambda_D(\Pi-V^{dsd}(u|u_1,u_2))-c-\phi=\alpha_D(u),\end{aligned}$$

i.e., (B.2.7) holds. □

B.3 Optimal Contracts

B.3.1 Feasibility (Π_F)

Now that we characterized the value function that solves (HJB$_V$) subject to (IC), the next step is to solve (MP) subject to $u \geqslant 0$. I begin by checking the feasibility of the project. If the maximum of the value function V is greater than 0, the principal earns positive expected payoff from the contract, thus the project is feasible. If $V'(0) > 0$, there exists $u > 0$ such that $V(u) > 0$. Thus, the project is feasible. On the other hand, if $V'(0) \leqslant 0$, the maximum of the value function is 0 at $u = 0$ since V is concave (Proposition 3.4.2 and 3.5.2). Thus, the project is infeasible. Note that from (HJB$_V$), $V'(0) > 0$ is equivalent to

$$\max\left[\lambda_D\Pi-\phi,\lambda_S V_S\left(\phi/\lambda_S\right)\right]>c, \tag{B.3.1}$$

i.e., the project is feasible if at least one of the instantaneous payoff at the deadline covers the operating cost c. Note that

$$\begin{aligned}\lambda_S V_S(\phi/\lambda_S) =&\lambda_S(\Pi-c/\lambda_S)\left(1-e^{-1}\right)-\phi\\ =&(1+\eta)(1-e^{-1})\lambda_D\Pi-(1-e^{-1})c-\phi.\end{aligned}$$

Then, we can derive that $\lambda_S V_S(\phi/\lambda_S) > c$ is equivalent to

$$\Pi>\frac{(2-e^{-1})c+\phi}{(1+\eta)(1-e^{-1})\lambda_D},$$

whereas $\lambda_D \Pi - \phi > c$ is equivalent to $\Pi > \dfrac{c+\phi}{\lambda_D}$. By simple algebra, we can show that

$$\frac{(2-e^{-1})c+\phi}{(1+\eta)(1-e^{-1})\lambda_D} > \frac{c+\phi}{\lambda_D}.$$

Therefore, $\Pi > \Pi_F \equiv \dfrac{c+\phi}{\lambda_D}$ is equivalent to (B.3.1) and the project is feasible if this condition is satisfied.

B.3.2 The Length of the Contract (Π_D)

Given that the project is feasible, there exists $\bar{u} > 0$ that maximizes $V(u)$. Since V is concave and differentiable, $\bar{u}$ is the solution of $V'(\bar{u}) = 0$. To check what type of contract would be utilized to implement $(\bar{u}, V(\bar{u}))$, we need to compare $\bar{u}$ with switch points u_1 and u_2 defined in (3.4.2) and (3.5.1).

If Π is less than or equal to $\Pi_M(\eta)$, by Proposition 3.4.2 and 3.5.2, the value function is equal to $V^d(u)$, i.e., it does not have any switch point. Therefore, it is enough to restrict attention to the case where Π is greater than $\Pi_M(\eta)$.

If Π is greater than $\Pi_M(\eta)$ and less than $\Pi_S(\eta)$, there exists a switch point $u_1 > 0$ such that $V^d(u_1) = V^{ds'}(u_1|u_1)$. Note that if Π is greater than or equal to $\Pi_S(\eta)$ and η is greater than $1/(e-1)$, the value function near $u = 0$ corresponds to $V^{ds}(u|0)$ by Proposition 3.4.2 and 3.5.2. Thus, in this case, we can consider u_1 as 0. The following lemma characterizes the threshold of Π for comparing $\bar{u}$ and u_1.

Lemma B.3.1. *Suppose that $\Pi > \Pi_M(\eta)$ and $2\lambda_D \geqslant \lambda_S > \lambda_D$ are satisfied. Then, there exists $\Pi_D(\eta) \geqslant \Pi_M(\eta)$ such that $u_1 < \bar{u}$ if and only if $\Pi > \Pi_D(\eta)$. Moreover, if $\eta \leqslant \sqrt{c/(c+\phi)}$, $\Pi_D(\eta)$ is equal to $\Pi_M(\eta)$.*

Now I compare $\bar{u}$ and the second switching point u_2. The following lemma shows that u_2 is less than $\bar{u}$ if η is sufficiently small. Thus, in this case, the optimal contract involves two switches of approaches.

Lemma B.3.2. *Suppose that* $\Pi > \Pi_M(\eta)$ *and* $\eta \leqslant c/(c+\phi)$ *are satisfied. Then,* u_2 *is less than* $\bar{u}$.

On the other hand, the following lemma shows that u_2 is greater than $\bar{u}$ if η is close to 1. Thus, in this case, the optimal contract involves at most one switch of approaches.

Lemma B.3.3. *Suppose that* $\Pi \geqslant \Pi_D(\eta)$ *and* $\eta \geqslant \sqrt{c/(c+\phi)}$ *are satisfied. Then,* u_2 *is greater than or equal to* $\bar{u}$.

The proofs for the above lemmas are relegated to Appendix B.5

B.3.3 Proofs of Theorems

Proof of Theorem 3.4.1. (a) By the argument in Appendix B.3.1, the project is infeasible when Π is less than Π_F.

(b) When $\Pi_D(1) \geqslant \Pi > \Pi_F = c/\lambda_D = \Pi_M(1)$, by Lemma B.3.1, the switching point u_1 is greater than or equal to $\bar{u}$. Then, $V(\bar{u}) = V^d(\bar{u})$ by (a) of Proposition 3.4.2. In both cases, by (a) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with the deadline $\bar{u}/\phi$.

(c) When $\Pi \in (\Pi_D(1), \Pi_S(1))$, u_1 is greater than 0 and less than $\bar{u}$ by Lemma B.3.1 and Proposition 3.4.2 (a). Then, $V(\bar{u}) = V^{ds}(\bar{u}|u_1)$ by Proposition 3.4.2. By (b) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a contract with a switch from the sequential approach to the direct approach at $(\bar{u} - u_1)/\phi$ and the deadline $\bar{u}/\phi$.

(d) When $\Pi \geqslant \Pi_S(1)$, $V(u) = V^{ds}(u|0)$ by (b) of Proposition 3.4.2. By (c) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a sequential-only contract with the deadline $\bar{u}/\phi$.

$\square$

Proof of Theorem 3.5.1. Note that u_2 is greater than $\bar{u}$ by Lemma B.3.3 since $\eta > \sqrt{c/(c+\phi)}$. When $\Pi \in [\Pi_F, \Pi_M(\eta)]$, $V(\bar{u}) = V^d(\bar{u})$ by (a) of Proposition 3.5.2. Then, by (a) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with the deadline $\bar{u}/\phi$. For other cases, the statements can be similarly proved as in Theorem 3.4.1, except that we need to use Proposition 3.5.2 instead of Proposition 3.4.2. □

Proof of Theorem 3.5.2. The proof for part (a) is same as Theorem 3.4.1, thus, it is enough to show (b) and (c).

(b) When $\Pi \in [\Pi_F, \Pi_M(\eta)]$, $V(\bar{u}) = V^d(\bar{u})$ by (a) of Proposition 3.5.2. By (a) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with the deadline $\bar{u}/\phi$.

(c) When $\Pi > \Pi_M(\eta)$, since $\eta < c/(c+\phi)$ is assumed, u_2 is less than $\bar{u}$ by Lemma B.3.2. Also note that (b) of Proposition 3.5.2 applies since $\eta < 1/(e-1)$. Therefore, $V(\bar{u})$ is equal to $V^{dsd}(u|u_1, u_2)$. By (a) of Proposition 3.4.1, $(\bar{u}, V(\bar{u}))$ is implemented by a contract with two switches at $(\bar{u}-u_2)/\phi$ and $(\bar{u}-u_1)/\phi$ and the deadline $\bar{u}/\phi$.

□

B.4 Proofs for Value Function Characterization

In this section, I provide the proofs for value function characterization. In Section B.4.1, by using the smooth pasting conditions, I identify the thresholds Π_S and Π_M which determine the number of potential switches in the value function. In Section B.4.2, I define the functions that specify deviations from the given value function and present useful lemmas. By using these results, I prove Proposition 3.4.2 and 3.5.2 in Section B.4.3.

B.4.1 Conditions for Smooth Pasting

The goal of this analysis is to construct a differentiable and concave function V that solves (HJB_V) subject to (IC) with $V(0) = 0$ and apply Lemma B.2.2. To construct such 'differentiable' function, I need to find a point where one benchmark function has the same first derivative to another benchmark function, i.e., two benchmark functions are 'smoothly pasted.' Depending on the parameter Π, such a switching point may or may not exist. In the following propositions, I formally introduce threshold of Π, Π_S and Π_M, as functions of η $(= \lambda_S/\lambda_D - 1)$ and characterize the conditions for smooth pasting by using those thresholds.

Proposition B.4.1. *Suppose that η is equal to 1. There exists $\Pi_S(1) > c/\lambda_D$ such that the following statements hold.*

(a) If $\Pi_M(1) = c/\lambda_D \leqslant \Pi \leqslant \Pi_S(1)$, there exists $u_1 \geqslant 0$ such that $V^{d\prime}(u) > V^{ds\prime}(u|u)$ for all $0 \leqslant u < u_1$, $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$, $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$, and $V^{ds\prime}(u|u_1) > V^{dsd\prime}(u|u, u_1)$ for all $u > u_1$.

(b) If $\Pi > \Pi_S(1)$, $V^{ds\prime}(0|0) > V^{d\prime}(0)$ and $V^{ds\prime}(u|0) > V^{dsd\prime}(u|0, u)$ for all $u \geqslant 0$.

Proposition B.4.2. *Suppose that η is less than 1. There exists $\Pi_S : (1/(e-1), 1) \to \mathbb{R}_+$ and $\Pi_M : (0, 1) \to \mathbb{R}_+$ such that the following statements hold.*

(a) If $c/\lambda_D < \Pi < \Pi_M(\eta)$, $V^{d\prime}(u) > V^{ds\prime}(u|u)$ for all $u \geqslant 0$.

(b) Suppose that one of the followings hold: (i) $\eta \leqslant 1/(e-1)$ and $\Pi \geqslant \Pi_M(\eta)$; (ii) $1/(e-1) < \eta < 1$ and $\Pi \in [\Pi_M(\eta), \Pi_S(\eta)]$. Then, there exist a pair (u_1, u_2) with $u_2 \geqslant u_1 \geqslant 0$ such that:

(i) $V^{d\prime}(u) > V^{ds\prime}(u|u)$ for all $0 \leqslant u < u_1$,

(ii) $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$ and $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$,

(iii) $V^{ds'}(u|u_1) > V^{dsd'}(u|u, u_1)$ *for all* $u_2 > u > u_1$,

(iv) $V^{ds'}(u_2|u_1) = V^{dsd'}(u_2|u_1, u_2)$ *and* $V^{ds''}(u_2|u_1) < V^{dsd''}(u_2|u_1, u_2)$,

(v) $V^{dsd'}(u|u_1, u_2) > \frac{1}{\phi}\left[\lambda_S(V_S(u + \phi/\lambda_S) - V^{dsd}(u|u_1, u_2)) - c\right]$ *for all* $u > u_2$.

(c) If $1/(e-1) < \eta < 1$ *and* $\Pi > \Pi_S(\eta)$, $V^{ds'}(0|0) > V^{d'}(0)$ *and there exists* $u_2 > 0$ *such that:*

(i) $V^{ds'}(u_2|0) = V^{dsd'}(u_2|0, u_2)$, $V^{ds''}(u_2|0) < V^{dsd''}(u_2|0, u_2)$ *and* $V^{ds'}(u|0) > V^{dsd'}(u|0, u)$ *for all* $u_2 > u \geqslant 0$,

(ii) $V^{dsd'}(u|0, u_2) > \frac{1}{\phi}\left[\lambda_S(V_S(u + \phi/\lambda_S) - V^{dsd}(u|0, u_2)) - c\right]$ *for all* $u > u_2$.

Useful Lemmas

Lemma B.4.1. *If* η *is less than or equal to* $1/(e-1)$, *the inequality* $V^{d'}(0) > V^{ds'}(0|0)$ *always holds. If* η *is greater than* $1/(e-1)$, $V^{d'}(0) > V^{ds'}(0|0)$ *is equivalent to* $\Pi < \Pi_S(\eta)$. *Moreover, if* $\Pi = \Pi_S(\eta)$, *then* $V^{d'}(0) = V^{ds'}(0|0)$ *and* $V^{d''}(0) < V^{ds''}(0|0)$.

Proof of Lemma B.4.1. By Proposition B.2.1, $V^{d'}(0) = \alpha_D(0)$ and $V^{ds'}(0|0) = \alpha_S(0)$. By using the arguments in Section 3.3.4, when $\eta \leqslant 1/(e-1)$, $V^{d'}(0) > V^{ds'}(0|0)$ always hold, and when $\eta > 1/(e-1)$, $V^{d'}(0) > V^{ds'}(0|0)$ is equivalent to $\Pi < \Pi_S(\eta)$.

When $\Pi = \Pi_S(\eta)$, $V^{d'}(0) = V^{ds'}(0|0)$. Also note that

$$\phi V^{d''}(0) = -\lambda_D\left(1 + V^{d'}(0)\right) = -\frac{\lambda_D}{\phi}(\lambda_D\Pi - c),$$

$$\begin{aligned}\phi V^{ds''}(0|0) &= \lambda_S V_S'(\phi/\lambda_S) - \lambda_S V^{ds'}(0|0) = \frac{\lambda_S}{\phi}\left(\phi V_S'(\phi/\lambda_S) - \phi V^{ds'}(0|0)\right)\\ &= \frac{\lambda_S}{\phi}\left(\phi V_S'(\phi/\lambda_S) - \lambda_S V_S(\phi/\lambda_S) + c\right)\\ &= \frac{\lambda_S}{\phi}\left(\lambda_S\Pi - 2\lambda_S(V_S(\phi/\lambda_S) + \phi/\lambda_S)\right).\end{aligned}$$

When $\Pi = \Pi_S(\eta)$, $\lambda_D \Pi = \lambda_S \left(V_S(\phi/\lambda_S) + \phi/\lambda_S\right)$ from $V^{d\prime}(0) = V^{ds\prime}(0|0)$. Then,

$$\begin{aligned}
\phi^2 V^{ds\prime\prime}(0|0) - \phi^2 V^{d\prime\prime}(0) &= \lambda_S(\lambda_S \Pi_S(\eta) - 2\lambda_D \Pi_S(\eta)) + \lambda_D(\lambda_D \Pi_S(\eta) - c) \\
&= \lambda_D^2 \left[\left(\frac{\lambda_S}{\lambda_D} - 1\right)^2 \Pi_S(\eta) - \frac{c}{\lambda_D}\right] \\
&= \lambda_D^2 \left[\eta^2 \Pi_S(\eta) - \frac{c}{\lambda_D}\right] \\
&= \lambda_D c \left[\frac{(e-1)\eta^2}{(e-1)\eta - 1} - 1\right].
\end{aligned}$$

Since $(e-1)x^2 > (e-1)x - 1$ for all $x > 1/(e-1)$, we can see that $V^{ds\prime\prime}(0|0) > V^{d\prime\prime}(0)$. $\square$

Lemma B.4.2. *There exists $\Pi_M(\eta) \geqslant 2c/\lambda_S$ with $\Pi_M(1) = 2c/\lambda_S = c/\lambda_D$ such that the following statements hold.*

(a) If $c/\lambda_D \leqslant \Pi < \Pi_M(\eta)$, $V^{d\prime}(u) > V^{ds\prime}(u|u)$ for all $u \geqslant 0$.

(b) Suppose that one of the following statements hold: (i) $\eta \leqslant 1/(e-1)$ and $\Pi > \Pi_M(\eta)$; (ii) $\eta > 1/(e-1)$ and $\Pi_S(\eta) \geqslant \Pi > \Pi_M(\eta)$. Then, there exists $u_1 > 0$ such that $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$, $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$ and $V^{d\prime}(u) > V^{ds\prime}(u|u)$ for all $u \in [0, u_1)$;

Proof of Lemma B.4.2. Consider a function $H_1 : \mathbb{R}_+ \to \mathbb{R}$ defined as follows:

$$H_1(u) \equiv \phi V^{ds\prime}(u|u) - \phi V^{d\prime}(u). \tag{B.4.1}$$

Then, $H_1(u)$ can be rewritten as follows:

$$\begin{aligned}
H_1(u) &= \lambda_S\left((\Pi - c/\lambda_S)(1-e^{-\frac{\lambda_S}{\phi}u-1}) - V^d(u) - u\right) - \lambda_D(\Pi - u - V^d(u)) \\
&= (\lambda_S\Pi - c)\left(1-e^{-\frac{\lambda_S}{\phi}u-1}\right) - (\lambda_S-\lambda_D)(u+V^d(u)) - \lambda_D\Pi \\
&= (\lambda_S\Pi - c)\left(1-e^{-\frac{\lambda_S}{\phi}u-1}\right) - (\lambda_S-\lambda_D)\left(\Pi - c/\lambda_D\right)(1-e^{-\frac{\lambda_D}{\phi}u}) - \lambda_D\Pi \\
&= (\lambda_S/\lambda_D - 1)(\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u} - (2-\lambda_S/\lambda_D)c - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u-1} \\
&= \eta(\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u} - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u-1} - (1-\eta)c \\
&= -\left\{(\eta+1)\lambda_D\Pi - c\right\}e^{-1}e^{-\frac{(\eta+1)\lambda_D}{\phi}u} + \eta(\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u} - (1-\eta)c
\end{aligned}$$

Define $x \equiv e^{-\frac{\lambda_D}{\phi}u}$. Then, by considering $H_1(u)$ as a function of x and Π, it can be rewritten as follows:

$$\tilde{H}_1(x;\Pi) \equiv -\left\{(\eta+1)\lambda_D\Pi - c\right\}e^{-1}x^{\eta+1} + \eta(\lambda_D\Pi - c)x - (1-\eta)c.$$

Observe that

$$\frac{\partial^2\tilde{H}_1}{\partial x^2}(x;\Pi) = -(\eta+1)\eta\left\{(\eta+1)\lambda_D\Pi - c\right\}e^{-1}x^{\eta-1},$$

thus $\tilde{H}_1$ is a strict concave function in x when $\Pi \geqslant c/\lambda_D$. Let $x^*(\Pi)$ be the solution of $\max_x H_1(x;\Pi)$ subject to $0 \leqslant x \leqslant 1$. Then, when $\Pi \geqslant c/\lambda_D$, from the first order condition, we can derive that

$$x^*(\Pi) = \left[\frac{\eta(\lambda_D\Pi - c)}{(\eta+1)\{(\eta+1)\lambda_D\Pi - c\}e^{-1}}\right]^{\frac{1}{\eta}}.[4]$$

Now define

$$h(\Pi) \equiv \tilde{H}_1(x^*(\Pi);\Pi) = K\left(\frac{\lambda_D\Pi - c}{\lambda_S\Pi - c}\right)^{\frac{1}{\eta}}(\lambda_D\Pi - c) - (1-\eta)c$$

[4] When $\Pi \leqslant \frac{(\eta+1)e^{-1}-\eta}{(\eta+1)^2e^{-1}-\eta}\cdot\frac{c}{\lambda_D}$, the solution of the maximization problem $\max_{0\leqslant x\leqslant 1} H_1(x;\Pi)$ is $x^*(\Pi) = 1$. However, we can show that $\frac{(\eta+1)e^{-1}-\eta}{(\eta+1)^2e^{-1}-\eta} < 1$ for any $0 < \eta$, which implies that we can focus on the interior solution when $\Pi \geqslant c/\lambda_D$

where $K = \frac{\eta^2}{\eta+1}\left(\frac{\eta e}{\eta+1}\right)^{\frac{1}{\eta}}$. Observe that

$$h\left(\frac{2c}{\lambda_S}\right) = (1-\eta)c\left[\frac{\eta^2}{(\eta+1)^2}\left(\frac{\eta(1-\eta)e}{(\eta+1)^2}\right)^{\frac{1}{\eta}} - 1\right] < 0$$

from $\eta < 1$ and $\eta(1-\eta)e \leqslant e/4 < 1 \leqslant (\eta+1)^2$. In addition, $\lim_{\Pi\to\infty} h(\Pi) = \infty$ and

$$h'(\Pi) = K(\lambda_D\Pi - c)^{1/\eta}(\lambda_S\Pi - c)^{-1/\eta-1}\lambda_D\lambda_S\Pi > 0.$$

Therefore, there exists a unique Π such that $h(\Pi) = 0$ and $\Pi \geqslant 2c/\lambda_S$. Let the solution of $h(\Pi) = 0$ with $\Pi \geqslant 2c/\lambda_S$ be $\Pi_M(\eta)$. Also note that when $\eta = 1$, $h(2c/\lambda_S) = 0$ thus $\Pi_M(1) = 2c/\lambda_S = c/\lambda_D$.

(a) Suppose that $c/\lambda_D \leqslant \Pi < \Pi_M(\eta)$. We have $0 > h(\Pi) = \tilde{H}_1(x^*(\Pi);\Pi) \geqslant \tilde{H}_1(x;\Pi)$ for all $0 \leqslant x \leqslant 1$. It is equivalent to $0 > H_1(u) = \phi({V^{ds}}'(u|u) - {V^{d}}'(u))$ for all $u \geqslant 0$, thus ${V^{d}}'(u) > {V^{ds}}'(u|u)$ for all $u \geqslant 0$ in this case.

(b) First, suppose that $\eta \leqslant 1/(e-1)$ and $\Pi > \Pi_M(\eta)$. Then, we have $0 < h(\Pi) = \tilde{H}_1(x^*(\Pi);\Pi)$. In addition, by Lemma B.4.1, we have $\tilde{H}_1(1;\Pi) = \phi({V^{ds}}'(0|0) - {V^{d}}'(0)) < 0$. Then, by concavity of $\tilde{H}_1$ w.r.t. x and $\frac{\partial \tilde{H}_1}{\partial x}(x^*(\Pi);\Pi) = 0$, there exists $x_1 \in (x^*(\Pi), 1]$ such that $\tilde{H}_1(x_1;\Pi) = 0$, $\frac{\partial \tilde{H}_1}{\partial x}(x_1;\Pi) < 0$ and $\tilde{H}_1(x;\Pi) < 0$ for all $x \in (x_1, 1]$. By defining $u_1 \equiv -\frac{\phi}{\lambda_D}\log x_1$, the above conditions can be translated into: ${V^{d}}'(u_1) = {V^{ds}}'(u_1|u_1)$, ${V^{d}}''(u_1) < {V^{ds}}''(u_1|u_1)$ and ${V^{d}}'(u) > {V^{ds}}'(u|u)$ for all $u \in [0, u_1)$.

Next, suppose that $\eta > 1/(e-1)$. Note that by the definition of $\Pi_S(\eta)$, if $\Pi \geqslant \Pi_S(\eta)$, $H_1(0) = \tilde{H}_1(1;\Pi) \geqslant 0$. It implies that $h(\Pi) \geqslant \tilde{H}_1(1;\Pi) \geqslant 0$ and $\Pi \geqslant \Pi_M(\eta)$. Therefore, we can see that $\Pi_S(\eta) \geqslant \Pi_M(\eta)$. If $\Pi_S(\eta) \geqslant \Pi > \Pi_M(\eta)$, we also have $\tilde{H}_1(x^*(\Pi);\Pi) > 0 > \tilde{H}_1(1;\Pi)$. By using the same arguments as above, we can show that there exists $u_1 > 0$ such that ${V^{d}}'(u_1) = {V^{ds}}'(u_1|u_1)$, ${V^{d}}''(u_1) < {V^{ds}}''(u_1|u_1)$ and ${V^{d}}'(u) > {V^{ds}}'(u|u)$ for all $u \in [0, u_1)$.

$\square$

Lemma B.4.3. *Suppose that $\Pi > c/\lambda_D$ and one of the followings hold: (i) $V^{d'}(u_1) < V^{ds'}(u_1|u_1)$; (ii) $V^{d'}(u_1) = V^{ds'}(u_1|u_1)$ and $V^{d''}(u_1) < V^{ds''}(u_1|u_1)$.*

(a) If $\eta = 1$, $V^{ds'}(u|u_1) > V^{dsd'}(u|u_1, u)$ for all $u > u_1$.

(b) If $\eta < 1$, there exists $u_2 > u_1$ such that $V^{ds'}(u_2|u_1) = V^{dsd'}(u_2|u_1, u_2)$ and $V^{ds''}(u_2|u_1) < V^{dsd''}(u_2|u_1, u_2)$ and such u_2 is unique. Moreover, $V^{ds'}(u|u_1) > V^{dsd'}(u|u_1, u)$ for all $u \in (u_1, u_2)$.

Proof of Lemma B.4.3. Define a function $H_2 : [u_1, \infty) \to \mathbb{R}$ as $H_2(u) \equiv \phi V^{dsd'}(u|u_1, u) - \phi V^{ds'}(u|u_1)$. From $\phi V^{dsd'}(u|u_1, u) = -c + \lambda_D \left(\Pi - \phi/\lambda_D - u - V^{dsd}(u|u_1, u)\right)$ (by (B.2.5)), $V^{dsd}(u|u_1, u) = V^{ds}(u|u_1)$ and (3.4.2), $H_2(u)$ can be rewritten as follows:

$$
\begin{aligned}
H_2(u) &= \lambda_D \left(\Pi - \frac{\phi}{\lambda_D} - u - V^{ds}(u|u_1)\right) - c - \phi V^{ds'}(u|u_1) \\
&= \frac{2\lambda_D - \lambda_S}{\lambda_S} c + (\lambda_S \Pi - c) e^{-1 - \frac{\lambda_S}{\phi} u_1} \left[1 + \frac{\lambda_S - \lambda_D}{\lambda_S} \frac{\lambda_S}{\phi}(u_1 - u)\right] e^{\frac{\lambda_S}{\phi}(u_1 - u)} \\
&\quad - (\lambda_S - \lambda_D) \left[\Pi - \frac{2c}{\lambda_S} - (V^d(u_1) + u_1)\right] e^{\frac{\lambda_S}{\phi}(u_1 - u)} \\
&= \frac{1-\eta}{1+\eta} c + (\lambda_S \Pi - c) e^{-1 - \frac{\lambda_S}{\phi} u_1} \left[1 + \frac{\eta}{1+\eta} \frac{\lambda_S}{\phi}(u_1 - u)\right] e^{\frac{\lambda_S}{\phi}(u_1 - u)} \\
&\quad + \eta \left[\frac{1-\eta}{1+\eta} c - (\lambda_D \Pi - c) e^{-\frac{\lambda_D}{\phi} u_1}\right] e^{\frac{\lambda_S}{\phi}(u_1 - u)}.
\end{aligned}
$$

Define $x \equiv e^{\frac{\lambda_S}{\phi}(u_1 - u)}$. Then, $H_2(u)$ can be rewritten as follows:

$$
\begin{aligned}
\tilde{H}_2(x) \equiv\ & \frac{1-\eta}{1+\eta} c + (\lambda_S \Pi - c) e^{-1 - \frac{\lambda_S}{\phi} u_1} \left[1 + \frac{\eta}{1+\eta} \log x\right] x \\
& + \eta \left[\frac{1-\eta}{1+\eta} c - (\lambda_D \Pi - c) e^{-\frac{\lambda_D}{\phi} u_1}\right] x.
\end{aligned} \tag{B.4.2}
$$

Note that $\tilde{H}_2(1) = H_2(u_1) = \phi V^{dsd\prime}(u_1|u_1, u_1) - \phi V^{ds\prime}(u_1|u_1) = \phi V^{d\prime}(u_1) - \phi V^{ds\prime}(u_1|u_1) \leqslant 0$. By differentiating $\tilde{H}$ twice, we have

$$\tilde{H}_2''(x) = \frac{\eta}{1+\eta}(\lambda_S \Pi - c)e^{-1-\frac{\lambda_S}{\phi}u_1}\frac{1}{x} > 0.$$

Since $\Pi > c/\lambda_D > c/\lambda_S$, $\tilde{H}_2$ is strictly convex in x. Also note that

$$\lim_{x\to 0} \tilde{H}_2(x) = \frac{1-\eta}{1+\eta}c,$$

and we simply denote by $\tilde{H}_2(0)$.

(a) Suppose that $\eta = 1$. Then, by the strict convexity of $\tilde{H}_2$, for all $x \in (0, 1)$,

$$\tilde{H}_2(x) < (1-x)\tilde{H}_2(0) + x\tilde{H}_2(1) \leqslant 0.$$

Therefore, for all $u > u_1$, $H_2(u) < 0$, i.e., $V^{ds\prime}(u|u_1) > V^{dsd\prime}(u|u_1, u)$.

(b) Suppose that $\eta < 1$. Then, $\tilde{H}_2(0) > 0$. If $V^{ds\prime}(u_1|u_1) < V^{d\prime}(u_1)$, $\tilde{H}_2(1) < 0$. In this case, there exists $x_2 \in (0, 1)$ such that $\tilde{H}_2(x_2) = 0$. Let $u_2 = u_1 - \frac{\phi}{\lambda_S}\log x_2$. Then, $H(u_2) = 0$, i.e., $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$.

Next, consider the case with $V^{ds\prime}(u_1|u_1) = V^{d\prime}(u_1)$ and $V^{ds\prime\prime}(u_1|u_1) > V^{d\prime\prime}(u_1)$. By differentiating (B.2.5) and (B.2.6) once, we have

$$\phi V^{d\prime\prime}(u_1) = -\lambda_D(1 + V^{d\prime}(u_1)),$$

$$\phi V^{ds\prime\prime}(u_1|u_1) = \lambda_S \left(V_S'(u + \phi/\lambda_S) - V^{ds\prime}(u|u_1)\right).$$

Then, from the above expressions and $V^{ds\prime}(u_1|u_1) = V^{d\prime}(u_1)$, $V^{ds\prime\prime}(u_1|u_1) > V^{d\prime\prime}(u_1)$ is equivalent to :

$$\lambda_S \left(V_S'(u_1 + \phi/\lambda_S) + 1\right) > (\lambda_S - \lambda_D)(1 + V^{d\prime}(u_1))$$

$$\iff \quad \frac{\lambda_S}{\phi}\lambda_S\left(\Pi - \frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}u_1 - 1} > (\lambda_S - \lambda_D)\frac{\lambda_D}{\phi}\left(\Pi - \frac{c}{\lambda_D}\right)e^{-\frac{\lambda_D}{\phi}u_1}$$

$$\iff \quad (\eta + 1)(\lambda_S \Pi - c)e^{-\frac{\lambda_S}{\phi}u_1 - 1} > \eta(\lambda_D \Pi - c)e^{-\frac{\lambda_D}{\phi}u_1}. \tag{B.4.3}$$

Note that $\tilde{H}_2(1) = (1-\eta)c + (\lambda_S\Pi - c)e^{-1-\frac{\lambda_S}{\phi}u_1} - \eta(\lambda_D\Pi - c)e^{-\frac{\lambda_S}{\phi}u_1} = 0$ from $V^{ds\prime}(u_1|u_1) = V^{d\prime}(u_1)$. Then,

$$\begin{aligned}
\tilde{H}_2'(1) &= (\lambda_S\Pi - c)e^{-1-\frac{\lambda_S}{\phi}u_1}\left[1 + \frac{\eta}{1+\eta}\right] + \eta\left[\frac{1-\eta}{1+\eta}c - (\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u_1}\right] \\
&= (\lambda_S\Pi - c)e^{-1-\frac{\lambda_S}{\phi}u_1} - \frac{\eta}{1+\eta}(\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u_1} \\
&\geqslant 0.
\end{aligned}$$

The last inequality is due to (B.4.3).

$\tilde{H}_2(1) = 0$ and $\tilde{H}_2'(1) > 0$ imply $\tilde{H}_2(1-\epsilon) < 0$ for small enough $\epsilon > 0$. Then, since $\tilde{H}_2(0) > 0$ and $\tilde{H}_2(1-\epsilon) < 0$, there exists $x_2 \in (0, 1-\epsilon)$ such that $\tilde{H}_2(x_2) = 0$ & $\tilde{H}_2'(x_2) < 0$, thus there exists u_2 such that $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$ & $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$.

Suppose that there exists another u_2' with $H_2(u_2') = 0$. Consider corresponding x_2', then $\tilde{H}_2(x_2') = 0$. If $x_2' > x_2$,

$$0 = \tilde{H}_2(x_2'|u_1) < \frac{1-x_2'}{1-x_2}\tilde{H}_2(x_2) + \frac{x_2'-x_2}{1-x_2}\tilde{H}_2(1) \leqslant 0$$

from $\tilde{H}_2(x_2) = 0$ and $\tilde{H}_2(1) \leqslant 0$. Similar logic holds for the case of $x_2' < x_2$. Therefore, there is unique u_2 satisfying $H_2(u_2) = 0$.

From $\tilde{H}_2(x_2) = 0$, $\tilde{H}_2(1) < 0$ and the strict convexity of $\tilde{H}_2$, for all $x \in (x_2, 1)$,

$$\tilde{H}_2(x) < \frac{1-x}{1-x_2}\tilde{H}_2(x_2) + \frac{x-x_2}{1-x_2}\tilde{H}_2(1) < 0.$$

Therefore, for all $u_2 > u > u_1$, $H_2(u) < 0$, i.e., $V^{ds\prime}(u|u_1) > V^{dsd\prime}(u|u_1, u)$.

$\square$

Lemma B.4.4. *Suppose that* $\Pi > c/\lambda_D$, $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$ *and* $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$. *Then,*

$$\lambda_S \left(V_S(u + \phi/\lambda_S) - V^{dsd}(u|u_1, u_2)\right) - \phi V^{dsd\prime}(u|u_1, u_2) - c < 0$$

for all $u > u_2$.

Proof of Lemma B.4.4. By (B.2.5) and (B.2.6), we have

$$\phi V^{ds\prime}(u|u_1) = \lambda_S \left(V_S \left(u + \phi/\lambda_S\right) + u + \phi/\lambda_S\right) - \lambda_S \left(V^{ds}(u|u_1) + u\right) - c - \phi,$$

$$\phi V^{dsd\prime}(u|u_1, u_2) = \lambda_D \Pi - \lambda_D \left(V^{dsd}(u|u_1, u_2) + u\right) - c - \phi.$$

By differentiating above equations, we have

$$\phi V^{ds\prime\prime}(u|u_1) = \lambda_S \left(V_S' \left(u + \phi/\lambda_S\right) + 1\right) - \lambda_S \left(V^{ds\prime}(u|u_1) + 1\right),$$

$$\phi V^{dsd\prime\prime}(u|u_1, u_2) = -\lambda_D \left(V^{dsd\prime}(u|u_1, u_2) + 1\right).$$

Then, $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$ and $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$ imply that

$$(\lambda_S - \lambda_D)(1 + V^{ds\prime}(u_2|u_1)) > \lambda_S(V_S'(u_2 + \phi/\lambda_S) + 1)$$

$$\Longleftrightarrow \qquad \eta(1 + V^{ds\prime}(u_2|u_1)) > (\eta + 1) \left(\frac{\lambda_S \Pi - c}{\phi}\right) e^{-\frac{\lambda_S}{\phi} u_2 - 1}. \qquad \text{(B.4.4)}$$

Define a function $H_3 : [u_2, \infty) \to \mathbb{R}$ as

$$
\begin{aligned}
H_3(u) &= \lambda_S \left[V_S(u + \phi/\lambda_S) - V^{dsd}(u|u_1, u_2)\right] - \phi V^{dsd\prime}(u|u_1, u_2) - c \\
&= \lambda_S \left[V_S(u + \phi/\lambda_S) + u + \phi/\lambda_S\right] - \lambda_S \left[u + V^{dsd}(u|u_1, u_2)\right] \\
&\quad - \phi(1 + V^{dsd\prime}(u|u_1, u_2)) - c \\
&= \left(\frac{\lambda_S}{\lambda_D} - 2\right) c - (\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} u - 1} \\
&\quad + \left(\frac{\lambda_S}{\lambda_D} - 1\right) \left[\lambda_D \Pi - c - \lambda_D (V^{ds}(u_2|u_1) + u_2)\right] e^{\frac{\lambda_D}{\phi}(u_2 - u)} \\
&= (\eta - 1)\, c - (\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} u_2 - 1} \cdot e^{\frac{\lambda_S}{\phi}(u_2 - u)} \\
&\quad + \eta \left[\lambda_D \Pi - c - \lambda_D \left(V^{ds}(u_2|u_1) + u_2\right)\right] e^{\frac{\lambda_D}{\phi}(u_2 - u)} \\
&= (\eta - 1)\, c - (\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} u_2 - 1} \cdot e^{\frac{\lambda_S}{\phi}(u_2 - u)} \\
&\quad + \eta \phi \left(V^{ds\prime}(u_2|u_1) + 1\right) e^{\frac{\lambda_D}{\phi}(u_2 - u)}.
\end{aligned}
$$

Also note that

$$
\begin{aligned}
H_3(u_2) &= \lambda_S \left[V_S(u_2 + \phi/\lambda_S) - V^{ds}(u_2|u_1)\right] - c - \phi V^{dsd\prime}(u_2|u_1, u_2) \\
&= \phi V^{ds\prime}(u_2|u_1) - \phi V^{dsd\prime}(u_2|u_1, u_2) = 0.
\end{aligned}
$$

Define $x \equiv e^{\frac{\lambda_D}{\phi}(u_2 - u)}$. Then, $H_3(u)$ can be rewritten as follows:

$$
\tilde{H}_3(x) = (\eta - 1)c - (\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} u_2 - 1} x^{\eta + 1} + \eta \phi \left(V^{ds\prime}(u_2|u_1) + 1\right) x
$$

and $\tilde{H}_3(1) = H_3(u_2) = 0$.

Note that

$$
\tilde{H}_3'(x) = -(\eta + 1)(\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} u_2 - 1} x^{\eta} + \eta \phi \left(V^{ds\prime}(u_2|u_1) + 1\right).
$$

By (B.4.4), we can derive that

$$\tilde{H}_3'(1) = -(\eta+1)(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2-1} + \eta\phi\left(V^{ds'}(u_2|u_1)+1\right) > 0.$$

Also note that

$$\tilde{H}_3''(x) = -(\eta+1)\eta(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2-1}x^{\eta-1} < 0.$$

Therefore, $\tilde{H}_3'(x) > 0$ for all $0 < x < 1$. Since $\tilde{H}_3(1) = 0$, $\tilde{H}_3(x) < 0$ for all $x \in (0,1)$. Thus, $\lambda_S\left(V_S(u+\phi/\lambda_S) - V^{dsd}(u|u_1,u_2)\right) - \phi V^{dsd'}(u|u_1,u_2) - c < 0$ for all $u \geqslant u_2$. $\square$

Lemma B.4.5. *Suppose that $V'(u) \geqslant -1$. The solution of* (HJB$_V$) *subject to* (IC) *involves $b = 1$.*

Proof of Lemma B.4.5. Assume that $b^* < 1$ solves (HJB$_V$) subject to (IC). Observe that $((\Pi - R - V(u))\lambda_D a + (V_S(u_S) - V(u))\lambda_S(1-a))b^* = 0$ from (HJB$_V$). This is because $(\Pi - R - V(u))\lambda_D a + (V_S(u_S) - V(u))\lambda_S(1-a) = 0$ when $b^* \in (0,1)$.

Also note that $(-\phi + (R-u)\lambda_D a + (u_S - u)\lambda_S(1-a))\, b^* = 0$ from (HJB$_{PK}$). Then, we have $\dot{u} = -\phi$. By plugging this into (HJB$_V$), we have $0 = -c - \phi V'(u)$, i.e., $V'(u) = -c/\phi < -1$. It contradicts the assumption of $V'(u) \geqslant -1$. Therefore, b should be equal to 1 for the solution of (HJB$_V$) subject to (IC). $\square$

Lemma B.4.6. *Suppose that $2\lambda_D \geqslant \lambda_S > \lambda_D$ and $\Pi > c/\lambda_D$. Then, the following statements hold:*

(a) $V^d(u) < \Pi - c/\lambda_D - u$, $V^{d'}(u) > -1$ and $V^{d''}(u) < 0$;

(b) Suppose that u_1 satisfies $V^{d'}(u_1) \leqslant V^{ds'}(u_1|u_1)$. Then, for all $u \geqslant u_1$, $V^{ds}(u|u_1) < \Pi - c/\lambda_D - u$, $V^{ds'}(u|u_1) > -1$ and $V^{ds''}(u|u_1) < 0$.

(c) *Suppose that u_1 and u_2 satisfy ${V^d}'(u_1) \leqslant {V^{ds}}'(u_1|u_1)$ and $u_2 > u_1$. Then, for all $u \geqslant u_2$, $V^{dsd}(u|u_1, u_2) < \Pi - c/\lambda_D - u$, ${V^{dsd}}'(u|u_1, u_2) > -1$ and ${V^{dsd}}''(u|u_1, u_2) < 0$.*

Proof of Lemma B.4.6. (a) By (3.4.1) and $e^{-\lambda_D u/\phi} > 0$, $V^d(u) < \Pi - c/\lambda_D - u$. By differentiating (3.4.1), we have

$$ {V^d}'(u) = \left(\Pi - \frac{c}{\lambda_D}\right) \frac{\lambda_D}{\phi} e^{-\frac{\lambda_D}{\phi} u} - 1 > -1. $$

By differentiating once again, we have

$$ {V^d}''(u) = -\left(\Pi - \frac{c}{\lambda_D}\right) \frac{\lambda_D^2}{\phi^2} e^{-\frac{\lambda_D}{\phi} u} < 0. $$

(b) By (3.4.2), $\Pi - 2c/\lambda_S \leqslant \Pi - c/\lambda_D$, $V^d(u_1) + u_1 < \Pi - c/\lambda_D$ and $u \geqslant u_1$, we can derive that

$$ V^{ds}(u|u_1) < \Pi - \frac{c}{\lambda_D} - u. $$

From (3.3.1), (B.2.5) and (B.2.6), we can derive that

$$ \begin{aligned} {V^d}'(u_1) &= -\frac{c}{\phi} - 1 + \frac{\lambda_D}{\phi}\left(\Pi - u_1 - V^d(u_1)\right), \\ {V^{ds}}'(u_1|u_1) &= -\frac{c}{\phi} - 1 + \frac{\lambda_S}{\phi}\left(\left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi} u_1 - 1}\right) - V^d(u_1) - u_1\right). \end{aligned} $$

Then, ${V^d}'(u_1) \leqslant {V^{ds}}'(u_1|u_1)$ is equivalent to

$$ u_1 + V^d(u_1) \leqslant \Pi - \frac{c}{\lambda_S - \lambda_D} - \frac{\lambda_S}{\lambda_S - \lambda_D}\left(\Pi - \frac{c}{\lambda_S}\right) e^{-\frac{\lambda_S}{\phi} u_1 - 1}. \tag{B.4.5} $$

By differentiating (3.4.2) twice, we have

$$V^{ds''}(u|u_1) = -\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_1-u)}\left[\left(\Pi-\frac{2c}{\lambda_S}\right)-\big(V^d(u_1)+u_1\big)\right]$$
$$-\left(\Pi-\frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)^2\left[-2+\frac{\lambda_S}{\phi}(u-u_1)\right]e^{-\frac{\lambda_S}{\phi}u-1}$$

$$=\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_1-u)}\cdot\big(V^d(u_1)+u_1\big)$$
$$+\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_1-u)}\cdot\left[-\left(\Pi-\frac{2c}{\lambda_S}\right)+2\left(\Pi-\frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}u_1-1}\right]$$
$$-\left(\Pi-\frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)^3(u-u_1)e^{-\frac{\lambda_S}{\phi}u-1}.$$

By using (B.4.5), we can show that

$$V^{ds''}(u|u_1)\leqslant\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(u_1-u)}\left[\frac{\lambda_S-2\lambda_D}{\lambda_S(\lambda_S-\lambda_D)}c+\frac{\lambda_S-2\lambda_D}{\lambda_S-\lambda_D}\left(\Pi-\frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}u_1-1}\right]$$
$$-\left(\Pi-\frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)^3(u-u_1)e^{-\frac{\lambda_S}{\phi}u-1}.$$

Then, from $2\lambda_D\geqslant\lambda_S>\lambda_D$ and $\Pi>c/\lambda_S$, we can derive that $V^{ds''}(u|u_1)\leqslant 0$.

Note that

$$V^{ds'}(u|u_1) =\frac{\lambda_S}{\phi}e^{\frac{\lambda_S}{\phi}(u_1-u)}\left[\left(\Pi-\frac{2c}{\lambda_S}\right)-(V^d(u_1)+u_1)\right]$$
$$-\left(\Pi-\frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)\left[1-\frac{\lambda_S}{\phi}(u-u_1)\right]e^{-\frac{\lambda_S}{\phi}u-1}-1,$$

$$\lim_{u\to\infty}V^{ds'}(u|u_1)=-1.$$

Then, by the concavity of $V^{ds}(u|u_1)$, $V^{ds'}(u|u_1)>-1$.

(c) By (3.5.1), $V^{ds}(u_2|u_1)+u_2<\Pi-c/\lambda_D$ and $u\geqslant u_2$, we can derive that

$$V^{dsd}(u|u_1,u_2)<\Pi-\frac{c}{\lambda_D}-u.$$

By differentiating (3.5.1) once, we have

$$V^{dsd'}(u|u_1,u_2) = \frac{\lambda_D}{\phi}\left(\Pi - \frac{c}{\lambda_D} - (V^{ds}(u_2|u_1,u_2)+u_2)\right)e^{\frac{\lambda_D}{\phi}(u_2-u)} - 1.$$

By the previous result ($V^{ds}(u_2|u_1,u_2) + u_2 < \Pi - c/\lambda_D$), we can derive that $V^{dsd'}(u|u_1,u_2) > -1$.

By differentiating (3.5.1) twice, we can derive that

$$V^{dsd''}(u|u_1,u_2) = -\left(\frac{\lambda_D}{\phi}\right)^2\left(\Pi - \frac{c}{\lambda_D} - (V^{ds}(u_2|u_1,u_2)+u_2)\right)e^{\frac{\lambda_D}{\phi}(u_2-u)} < 0.$$

□

Proofs for Proposition B.4.1 and B.4.2

Proof of Proposition B.4.1. (a) Observe that $\Pi_M(1) = c/\lambda_D < \Pi_S(1)$ from Lemma B.4.1 and B.4.2. Suppose that $\Pi_M(1) = c/\lambda_D < \Pi \leqslant \Pi_S(1)$. Then, by Lemma B.4.2, there exists $u_1 \geqslant 0$ such that $V^{d'}(u_1) = V^{ds'}(u_1|u_1)$, $V^{d''}(u_1) < V^{ds''}(u_1|u_1)$ and $V^{d'}(u) > V^{ds'}(u|u)$ for all $u \in [0,u_1)$. By Lemma B.4.3 (a), we have $V^{ds'}(u|u_1) > V^{dsd'}(u|u_1,u)$ for all $u > u_1$.

(b) Suppose that $\Pi > \Pi_S(1)$. By Lemma B.4.1, $V^{d'}(0) < V^{ds'}(0|0)$. Set $u_1 = 0$. Note that $V^{ds'}(0|0) > V^{d'}(0) = V^{dsd'}(0|0,0)$. Next, by Lemma B.4.3 (a), we have $V^{ds'}(u|0) > V^{dsd'}(u|0,u)$ for all $u > 0$.

□

Proof of Proposition B.4.2. (a) If $c/\lambda_D < \Pi < \Pi_M(\eta)$, by Lemma B.4.2, $V^{d'}(u) > V^{ds'}(u|u)$ for all $u \geqslant 0$.

(b) (i-ii) By Lemma B.4.2, there exists $u_1 > 0$ such that $V^{d'}(u_1) = V^{ds'}(u_1|u_1)$, $V^{d''}(u_1) < V^{ds''}(u_1|u_1)$ and $V^{d'}(u) > V^{ds'}(u|u)$ for all $u \in [0,u_1)$.

(iii-iv) By Lemma B.4.3, there exists $u_2 > u_1$ such that $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$, $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$ and $V^{ds\prime}(u|u_1) > V^{dsd\prime}(u|u_1, u)$ for all $u \in (u_1, u_2)$.

(v) By Lemma B.4.4, for all $u > u_2$, we have

$$V^{dsd\prime}(u|u_1, u_2) > \frac{1}{\phi}\left[\lambda_S\left(V_S(u + \phi/\lambda_S) - V^{dsd}(u|u_1, u_2)\right) - c\right].$$

(c) Suppose that $1/(e-1) < \eta < 1$ and $\Pi > \Pi_S(\eta)$. By Lemma B.4.1, we have $V^{d\prime}(0) < V^{ds\prime}(0|0)$. Now set $u_1 = 0$, then $V^{d\prime}(u_1) < V^{d\prime\prime}(u_1|u_1)$.

(i) By Lemma B.4.3, there exists $u_2 > u_1 = 0$ such that
$V^{ds\prime}(u_2|0) = V^{dsd\prime}(u_2|0, u_2)$, $V^{ds\prime\prime}(u_2|0) < V^{dsd\prime\prime}(u_2|0, u_2)$ and
$V^{ds\prime}(u|0) > V^{dsd\prime}(u|0, u)$ for all $u \in (0, u_2)$.

(ii) By Lemma B.4.4, for all $u > u_2$, we have

$$V^{dsd\prime}(u|u_1, u_2) > \frac{1}{\phi}\left[\lambda_S\left(V_S(u + \phi/\lambda_S) - V^{dsd}(u|u_1, u_2)\right) - c\right].$$

□

B.4.2 Functions for Deviation

In this subsection, I introduce functions that specify deviations from the given value functions. These functions will be used when I verify that the constructed value functions solve (HJB_V) subject to (IC). Then I present some properties of these functions.[5]

1. Functions for deviation given V^d

[5] This approach is inspired by the tangible first breakthrough case of Green and Taylor (2016c). 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 an approach to another approach. Thus, we need to define two functions for each case.

(a) Define

$$L_1^D(u,R) \equiv \lambda_D(\Pi - R - V^d(u)) - c - \lambda_D(R-u)V^{d\prime}(u).$$

Given u, maximizing this function with respect to $R \geqslant u + \phi/\lambda_D$ is equivalent to maximizing the right hand side of (HJB_V) under the condition that $b=1$ solves (HJB_{PK}) with $a=1$.

(b) Define

$$\begin{aligned} L_1^S(u,w) \equiv& \lambda_S(V_S(w) - V^d(u)) - c - \lambda_S(w-u)V^{d\prime}(u) \\ =& \lambda_S \Bigg[\left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi}u - \frac{\lambda_S}{\phi}(w-u)}\right) \\ & - \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{-\frac{\lambda_D}{\phi}u}\right) - (w-u)\left(\frac{\lambda_D \Pi - c}{\phi}\right)e^{-\frac{\lambda_D}{\phi}u}\Bigg] - c. \end{aligned} \tag{B.4.6}$$

Given u, maximizing this function with respect to $w \geqslant u + \phi/\lambda_S$ is equivalent to maximize the right hand side of (HJB_V) under the condition that $b=1$ solves (HJB_{PK}) with $a=0$.

2. Functions for deviation given V^{ds}

(a) Define

$$L_2^D(u,R|u_1) \equiv \lambda_D\left(\Pi - R - V^{ds}(u|u_1)\right) - c - \lambda_D(R-u)V^{ds\prime}(u|u_1).$$

(b) Define

$$L_2^S(u,w|u_1) \equiv \lambda_S\left(V_S(w) - V^{ds}(u|u_1)\right) - c - \lambda_S(w-u)V^{ds\prime}(u|u_1).$$

3. Functions for deviation given V^{dsd}

(a) Define

$$L_3^D(u,R|u_1,u_2) \equiv \lambda_D\left(\Pi - R - V^{dsd}(u|u_1,u_2)\right) - c - \lambda_D(R-u)V^{dsd\prime}(u|u_1,u_2).$$

(b) Define

$$L_3^S(u,w|u_1,u_2) \equiv \lambda_S\left(V_S(w) - V^{dsd}(u|u_1,u_2)\right) - c - \lambda_S(w-u)V^{dsd\prime}(u|u_1,u_2).$$

Useful Lemmas

Lemma B.4.7. *Suppose that $\Pi > c/\lambda_D$ and $2\lambda_D \geqslant \lambda_S > \lambda_D$ are satisfied. Then, L_1^D and L_1^S satisfy the following properties:*

(a) $L_1^D(u,R) \leqslant 0$ for all $u \geqslant 0$ and $R \geqslant u + \phi/\lambda_D$.

(b) If $\Pi \leqslant \Pi_M(\eta)$, $L_1^S(u,w) \leqslant 0$ for all $u \geqslant 0$ and $w \geqslant u + \phi/\lambda_S$.

(c) If $\Pi_M(\eta) < \Pi < \Pi_S(\eta)$, $L_1^S(u,w) \leqslant 0$ for all $u \in [0,u_1]$ and $w \geqslant u + \phi/\lambda_S$ where u_1 is the threshold defined on Proposition B.4.1 or B.4.2.

Proof of Lemma B.4.7. (a) Note that $\frac{\partial}{\partial R}L_1^D = -\lambda_D(1 + V^{d\prime}(u)) < 0$ by Lemma B.4.6. Therefore, for a fixed u, L_1^D is maximized at $R = u + \phi/\lambda_D$. Note that by the definition of V^d, $L_1^D(u, u+\phi/\lambda_D) = 0$, thus, $L_1^D(u,R) \leqslant 0$ for all $R \geqslant u + \phi/\lambda_D$.

(b) Define $x \equiv e^{-\frac{\lambda_D}{\phi}u}$ and $y \equiv w - u$. Note that $u \geqslant 0$, $w \geqslant u + \phi/\lambda_S$ and $2\lambda_D \geqslant \lambda_S > \lambda_D$ imply that $1 \geqslant x > 0$, $y \geqslant \phi/\lambda_S$ and $1 \geqslant \eta > 0$. Then, L_1^S can be rewritten as follows:

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

By $\Pi > c/\lambda_D > c/\lambda_S$, $\eta > 0$ and $x > 0$,

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

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

By differentiating $\tilde{L}_1^S$ once by x,

$$\frac{\partial\,\tilde{L}_1^S}{\partial x} = -\lambda_S\left[(\eta+1)\left(\Pi-\frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}y}x^{\eta}-\left(1-\frac{\lambda_D y}{\phi}\right)\left(\Pi-\frac{c}{\lambda_D}\right)\right].$$

If $y \geqslant \phi/\lambda_D$ and $x \geqslant 0$, $\tilde{L}_1^S$ is decreasing in x, thus, $\tilde{L}_1^S$ is maximized at $x = 0$. Then, for all $1 \geqslant x > 0$ and $y \geqslant \phi/\lambda_D$, the following inequalities hold:

$$\tilde{L}_1^S(x, y) \leqslant \tilde{L}_1^S(0, y) = (\eta - 1)c \leqslant 0,$$

thus, $L_1^S(u, w) \leqslant 0$ for all $u \geqslant 0$ and $w \geqslant u + \phi/\lambda_D$.

When $\phi/\lambda_D > y \geqslant \phi/\lambda_S$ and y is fixed, since $\tilde{L}_1^S(x, y)$ is concave in x, $\tilde{L}_1^S(\cdot, y)$ is maximized at

$$x^*(y) \equiv \left[\frac{(\lambda_D\Pi - c)\,(1-\lambda_D y/\phi)\,e^{\frac{\lambda_S}{\phi}y}}{(\lambda_S\Pi - c)}\right]^{\frac{1}{\eta}}.$$

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

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

Note that since $y \geqslant \phi/\lambda_S$ and $2\lambda_D \geqslant \lambda_S$, $g(y)$ is decreasing in y, hence, $x^*(y)$ is also decreasing in y.

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

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

Note that $(1 - \lambda_D y/\phi)e^{\lambda_D y/\phi}$ is decreasing in y,[6] thus, $\tilde{L}_1^S(x^*(y), y)$ is also decreasing in y.

If $x^*(y) \geqslant 1$, since $\dfrac{\partial \tilde{L}_1^b}{\partial x}$ is negative for all $0 \leqslant x \leqslant 1$, the maximum value of $\tilde{L}_1^S(\cdot, y)$ is

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

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

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

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

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

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

since $x^*(y)$ maximizes $\tilde{L}_1^S(x, y)$ and $\tilde{L}_1^S(x^*(y), y)$ is decreasing in y.

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

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

When $y > y^*$,

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

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

By combining the above results, we can show that

$$\max_{\substack{1 \geqslant x > 0, \\ \phi/\lambda_D > y \geqslant \phi/\lambda_S}} \tilde{L}_1^S(x,y) = \begin{cases} \tilde{L}_1^S\left(x^*\left(\phi/\lambda_S\right),\ \phi/\lambda_S\right) & \text{if } \ x^*\left(\phi/\lambda_S\right) \leqslant 1, \\ \tilde{L}_1^S\left(1,\ \phi/\lambda_S\right) & \text{if } \ x^*\left(\phi/\lambda_S\right) > 1. \end{cases}$$

Note that

$$\tilde{L}_1^S(x,\phi/\lambda_S) = L_1^S(u, u+\phi/\lambda_S) = \phi V^{ds'}(u|u) - \phi V^{d'}(u)$$

where $u = -\frac{\phi}{\lambda_D}\log x$. Then, by Lemma B.4.2, if $c/\lambda_D < \Pi \leqslant \Pi_M(\eta)$, $\tilde{L}_1^S(x,\phi/\lambda_S) \leqslant 0$ for all $x \in (0,1]$. Therefore, if $c/\lambda_D < \Pi \leqslant \Pi_M(\eta)$, for all $x \in (0,1]$, $y \geqslant \phi/\lambda_S$,

$$\tilde{L}_1^S(x,y) \leqslant \tilde{L}_1^S(x^*(\phi/\lambda_S) \wedge 1, \phi/\lambda_S) \leqslant 0,$$

thus, $L_1^S(u,w) \leqslant 0$ for all $u \geqslant 0$ and $u + \phi/\lambda_D > w \geqslant u + \phi/\lambda_S$.

(c) By Lemma B.4.1, if $\Pi_S(\eta) \geqslant \Pi$,

$$\tilde{L}_1^S(1,\phi/\lambda_S) = \phi V^{ds'}(u|u) - \phi V^{d'}(u) < 0.$$

In the previous case, we show that $\frac{\partial}{\partial y}\tilde{L}_1^S(1,y) < 0$, thus $\tilde{L}_1^S(1,y) < 0$ for all $y \geqslant \phi/\lambda_S$.

On the other hand, by the definition of u_1, $V^{d'}(u_1) = V^{ds'}(u_1|u_1)$ and $V^{d''}(u_1) < V^{ds''}(u_1|u_1)$. Let $x_1 = e^{-\frac{\lambda_S}{\phi}u_1}$. Then,

$$\tilde{L}_1^S(x_1,\phi/\lambda_S) = L_1^S(u_1,\phi/\lambda_S) = \phi V^{ds'}(u_1|u_1) - \phi V^{d'}(u_1) = 0.$$

Then, we can derive that

$$-(\lambda_S\Pi - c)e^{-1}x_1^{\eta+1} + \eta(\lambda_D\Pi - c)x_1 + (\eta-1)c = 0.$$

Note that

$$\begin{aligned}
\frac{\partial \tilde{L}_1^S}{\partial y}(x_1, y) =& \frac{\lambda_S}{\phi}\left[(\lambda_S \Pi - c)e^{-\frac{\lambda_S}{\phi}y} x_1^\eta - (\lambda_D \Pi - c)\right] x_1 \\
=& \frac{\lambda_S}{\phi}\left[(\lambda_S \Pi - c)x_1^\eta e^{-1} \cdot e^{-\frac{\lambda_S}{\phi}\left(y - \frac{\phi}{\lambda_S}\right)} - (\lambda_D \Pi - c)\right] x_1 \\
=& \frac{\lambda_S}{\phi}\left[(\eta(\lambda_D \Pi - c)x_1 + (\eta - 1)c) \cdot e^{-\frac{\lambda_S}{\phi}\left(y - \frac{\phi}{\lambda_S}\right)} - (\lambda_D \Pi - c)x_1\right] \\
=& \frac{\lambda_S}{\phi}\left[(\lambda_D \Pi - c) \cdot \left(\eta e^{-\frac{\lambda_S}{\phi}\left(y - \frac{\phi}{\lambda_S}\right)} - 1\right) x_1 + (\eta - 1)c \cdot e^{-\frac{\lambda_S}{\phi}\left(y - \frac{\phi}{\lambda_S}\right)}\right].
\end{aligned}$$

Since $\eta \leqslant 1$, $y \geqslant \phi/\lambda_S$ and $\lambda_D \Pi > c$, $\dfrac{\partial \tilde{L}_1^S}{\partial y}(x_1, y) < 0$. Therefore, $\tilde{L}_1^S(x_1, y) < 0$ for all $y \geqslant \phi/\lambda_S$.

In the previous case, we show that $\tilde{L}_1^S$ is strictly concave, thus, for all $x_1 \leqslant x \leqslant 1$ and $y \geqslant \phi/\lambda_S$,

$$\tilde{L}_1^S(x, y) \leqslant \frac{x - x_1}{1 - x_1}\, \tilde{L}_1^S(1, y) + \frac{1 - x}{1 - x_1}\, \tilde{L}_1^S(x_1, y) \leqslant 0.$$

Therefore, $L_1^S(u, w) \leqslant 0$ for all $u \in [0, u_1]$ and $w \geqslant u + \phi/\lambda_S$.

$\square$

Lemma B.4.8. *Suppose that* $2\lambda_D \geqslant \lambda_S > \lambda_D$, $\Pi > \Pi_M(\eta)$ *and* $V^{d\prime}(u_1) < V^{ds\prime}(u_1|u_1)$ *or* $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$ *&* $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$ *are satisfied. Then,* L_2^D *and* L_2^S *satisfy the following properties:*

(a) $L_2^S(u, w|u_1) \leqslant 0$ *for all* $u \geqslant u_1$ *and* $w \geqslant u + \phi/\lambda_S$.

(b) When $\eta = 1$, *for all* $u > u_1$ *and* $R \geqslant u + \phi/\lambda_D$, $L_2^D(u, R|u_1) < 0$ *is satisfied.*

(c) When $\eta < 1$, $L_2^D(u_2, u_2 + \phi/\lambda_D|u_1) = 0$, $L_2^D(u, R|u_1) < 0$ *for all* $u_1 < u < u_2$ *and* $R \geqslant u + \phi/\lambda_D$, *where* u_2 *is the threshold defined on B.4.2.*

Proof of Lemma B.4.8. (a) Note that $\frac{\partial}{\partial w}L_2^S = \lambda_S(V_S'(w) - V^{ds\prime}(u|u_1))$ and $\frac{\partial^2}{\partial w^2}L_2^S = \lambda_S V_S''(w) < 0$. By differentiating (B.2.6) once, we can derive that

$$\phi V^{ds\prime\prime}(u|u_1) = \lambda_S \left(V_S'\left(u + \frac{\phi}{\lambda_S}\right) - V^{ds\prime}(u|u_1)\right).$$

By (b) of Lemma B.4.6, $V^{ds\prime\prime}(u|u_1) < 0$. Thus, the following inequality holds:

$$0 > \frac{\partial}{\partial w}L_2^S(u, u+\phi/\lambda_S|u_1) = \lambda_S \left(V_S'\left(u + \frac{\phi}{\lambda_S}\right) - V^{ds\prime}(u|u_1)\right).$$

Then, $L_2^S(u, w|u_1)$ subject to $w \geqslant u + \phi/\lambda_S$ is maximized at $w = u + \phi/\lambda_S$ for a given u. Also note that $L_2^S(u, u+\phi/\lambda_S|u_1) = 0$ holds by (B.2.6). Therefore, $L_2^S(u, w|u_1) \leqslant 0$ for all $u \geqslant u_1$ and $w \geqslant u + \phi/\lambda_S$.

(b) Note that if $V^{d\prime}(u_1) \leqslant V^{ds\prime}(u_1|u_1)$, $\frac{\partial}{\partial R}L_2^D = -\lambda_D(1 + V^{ds\prime}(u|u_1)) < 0$ by Lemma B.4.6. Therefore, for all $u > u_1$ and $R \geqslant u + \phi/\lambda_D$,

$$L_2^D(u, R|u_1) \leqslant L_2^D(u, u+\phi/\lambda_D|u_1).$$

Note that

$$L_2^D(u, u+\phi/\lambda_D|u_1) = \phi V^{dsd\prime}(u|u_1, u) - \phi V^{ds\prime}(u|u_1)$$

By (a) of Lemma B.4.3, when $\eta = 1$, $V^{dsd\prime}(u|u_1, u) < V^{ds\prime}(u|u_1)$ for all $u > u_1$, thus, $L_2^D(u, R|u_1) < 0$ for all $u \geqslant u_1$ and $R \geqslant u + \phi/\lambda_D$.

(c) By the definition of u_2, $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$ and $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$, thus $L_2^D(u_2, u_2+\phi/\lambda_D|u_1) = \phi V^{dsd\prime}(u_2|u_1, u) - \phi V^{ds\prime}(u_2|u_1) = 0$. Moreover, $V^{dsd\prime}(u|u_1, u) < V^{ds\prime}(u|u_1)$ for all $u \in (u_1, u_2)$, thus, $L_2^D(u, R|u_1) < 0$ for all $u \in (u_1, u_2)$ and $R \geqslant u + \phi/\lambda_D$.

$\square$

Lemma B.4.9. *Suppose that $2\lambda_D \geqslant \lambda_S > \lambda_D$, $\Pi > \Pi_M(\eta)$, $V^{d\prime}(u_1) < V^{ds\prime}(u_1|u_1)$ or $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$ & $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$, $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$ and $V^{ds\prime\prime}(u_2|u_1) < V^{dsd\prime\prime}(u_2|u_1, u_2)$ are satisfied. Then, L_3^D and L_3^S satisfy the following properties:*

(a) $L_3^D(u, R|u_1, u_2) \leqslant 0$ for all $u \geqslant u_2$ and $R \geqslant u + \phi/\lambda_D$.

(b) $L_3^S(u, w|u_1, u_2) \leqslant 0$ for all $u \geqslant u_2$ and $w \geqslant u + \phi/\lambda_S$.

Proof of Lemma B.4.9. (a) Note that $\frac{\partial}{\partial R} L_3^D(u, R) = -\lambda_D \left(1 + V^{dsd\prime}(u|u_1, u_2)\right)$. By (c) of Lemma B.4.6, $V^{dsd\prime}(u|u_1, u_2) > -1$, thus $\frac{\partial}{\partial R} L_3^D(u, R)$ and $L_3^D(u, R) \leqslant L_3^D(u, u + \phi/\lambda_D)$ for all $u \geqslant u_2$ and $R \geqslant u + \phi/\lambda_D$.

By (B.2.5),

$$
\begin{aligned}
L_3^D\left(u, u + \frac{\phi}{\lambda_D} \mid u_1, u_2\right) =& \lambda_D(\Pi - u - \phi/\lambda_D - V^{dsd}(u|u_1, u_2)) \\
& - c - \phi V^{dsd\prime}(u|u_1, u_2) \\
=& 0.
\end{aligned}
$$

Therefore, $L_3^D(u, R) \leqslant 0$ for all $u \geqslant u_2$ and $R \geqslant u + \phi/\lambda_D$.

(b) By differentiating L_3^S by w, we have

$$
\begin{aligned}
& \frac{\partial}{\partial w} L_3^S(u, w|u_1, u_2) = \lambda_S \left[V_S'(w) - V^{dsd\prime}(u|u_1, u_2)\right] \\
& = \lambda_S \left[\frac{\lambda_S}{\phi}\left(\Pi - \frac{c}{\lambda_S}\right) e^{-\frac{\lambda_S}{\phi} w} - \frac{1}{\phi}\left(-c + \lambda_D(\Pi - u - V^{dsd}(u|u_1, u_2))\right)\right] \\
& = \frac{\lambda_S}{\phi}\left[(\lambda_S \Pi - c) e^{-\frac{\lambda_S}{\phi} w} - \lambda_D\left(\Pi - \frac{c}{\lambda_D} - (V^{ds}(u_2|u_1) + u_2)\right) e^{\frac{\lambda_D}{\phi}(u_2 - u)}\right].
\end{aligned}
$$

By $w \geqslant u + \phi/\lambda_S$,

$$\begin{aligned}
&\frac{\partial}{\partial w} L_3^S(u, w|u_1, u_2) \\
<&\frac{\lambda_S}{\phi}\left[(\lambda_S \Pi - c)e^{-\frac{\lambda_S}{\phi}u - 1} - \lambda_D\left(\Pi - \frac{c}{\lambda_D} - (V^{ds}(u_2|u_1) + u_2)\right)e^{\frac{\lambda_D}{\phi}(u_2 - u)}\right] \\
=& \frac{\lambda_S}{\phi}e^{\frac{\lambda_D}{\phi}(u_2-u)}\left[\begin{array}{l} (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2 - 1}e^{\frac{\lambda_S - \lambda_D}{\phi}(u_2 - u)} \\ -\lambda_D\left\{\Pi - \dfrac{c}{\lambda_D} - (V^{ds}(u_2|u_1) + u_2)\right\}\end{array}\right]
\end{aligned}$$

Since $u \geqslant u_2$ and $\lambda_S > \lambda_D$, we have

$$\frac{\partial}{\partial w} L_3^S(u, w|u_1, u_2) < \frac{\lambda_S}{\phi}e^{\frac{\lambda_D}{\phi}(u_2-u)}\left[\begin{array}{l} (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2 - 1} \\ -\lambda_D\left\{\Pi - \dfrac{c}{\lambda_D} - (V^{ds}(u_2|u_1) + u_2)\right\}\end{array}\right].$$

By (B.2.5), (B.2.6), $V^{ds}(u_2|u_1) = V^{dsd}(u_2|u_1, u_2)$ and $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$, we can derive that

$$(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2 - 1} = (\lambda_S - \lambda_D)(\Pi - (V^{ds}(u_2|u_1) + u_2)) - c.$$

By plugging this into the above inequality, we have

$$\frac{\partial}{\partial w} L_3^S(u, w|u_1, u_2) < \frac{\lambda_S}{\phi}e^{\frac{\lambda_D}{\phi}(u_2-u)}(\lambda_S - 2\lambda_D)\left(\Pi - (V^{ds}(u_2|u_1) + u_2)\right),$$

thus, since $2\lambda_D \geqslant \lambda_S$ and $\Pi - c/\lambda_D > V^{ds}(u_2|u_1) + u_2$ ((b) of Lemma B.4.6), $L_3^S(u, w|u_1, u_2)$ is decreasing in w. Therefore,

$$L_3^S(u, w|u_1, u_2) \leqslant L_3^S(u, u + \phi/\lambda_S \mid u_1, u_2)$$

for all $w \geqslant u + \phi/\lambda_S$.

Note that $L_3^S(u, u+\phi/\lambda_S|u_1, u_2) = H_3(u)$ and $H_3(u) \leqslant 0$ for all $u \geqslant u_2$ by Lemma B.4.4. Therefore, $L_3^S(u, w|u_1, u_2) \leqslant 0$ for all $u \geqslant u_2$ and $w \geqslant u + \phi/\lambda_S$.

□

B.4.3 Proof of Proposition 3.4.2 and 3.5.2

Proof of Proposition 3.4.2. (a) From Proposition B.4.1 (a), there exists $u_1 \geqslant 0$ such that $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$ and $V^{d\prime\prime}(u_1) < V^{ds\prime\prime}(u_1|u_1)$. Therefore, the function defined in (3.4.3) is differentiable.

First, consider the case where $u \in [0, u_1]$. Then, by Lemma B.4.6 (a), we have $V^{d\prime}(u) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (c) of Lemma B.4.7, $L_1^S(u, u_S) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (a) of Lemma B.4.7, $0 \geqslant L_1^D(u, R)$ for all $R \geqslant u + \phi/\lambda_D$. In addition, from the definition of V^d, we have $L_1^D(u, u + \phi/\lambda_D) = 0$. Therefore, when $u \in [0, u_1]$, $V^d(u)$ solves (HJB$_V$) subject to (IC).

Second, consider the case where $u > u_1$. Then, by Lemma B.4.6 (b), we have $V^{ds\prime}(u|u_1) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (b) of Lemma B.4.8, $0 \geqslant L_2^D(u, R|u_1)$ for all $R \geqslant u + \phi/\lambda_D$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (a) of Lemma B.4.8, $L_2^S(u, u_S|u_1) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{ds}, we have $L_2^S(u, u + \phi/\lambda_D|u_1) = 0$. Thus, for all $u > u_1$, $V^{ds}(u|u_1)$ solves (HJB$_V$) subject to (IC). Therefore, the value function specified in (3.4.3) solves (HJB$_V$) subject to (IC).

Next, by (a) of Lemma B.4.6, $V^{d\prime\prime}(u) < 0$ for all $0 < u < u_1$. By (b) of Lemma B.4.6 and $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$, we have $V^{ds\prime\prime}(u|u_1) < 0$ for all $u > u_1$. It remains to show that the function defined in (3.4.3) is concave at u_1. Observe that

$$V^d(u) < V^d(u_1) - V^{d\prime}(u_1)(u_1 - u) \tag{B.4.7}$$

for all $0 < u < u_1$, and

$$V^{ds}(u|u_1) < V^{ds}(u_1|u_1) + V^{ds\prime}(u_1|u_1)(u - u_1) \qquad \text{(B.4.8)}$$

for all $u_1 < u$. Therefore, the function defined in (3.4.3) is concave for all $u > 0$.

By Lemma B.2.2, any incentive compatible contract delivers the principal's expected payoff less than or equal to $V^d(u)$ for $u \leqslant u_1$ and $V^{ds}(u|u_1)$ for $u > u_1$. In addition, by Proposition 3.4.1 (a) and (b), these values can be implemented by some incentive compatible contracts, i.e., they are maximized expected payoff of the principal given the agent's promised utility. Thus, the function defined in (3.4.3) is the principal's value function.

(b) By Proposition B.4.1 (b), we have $V^{ds\prime}(0|0) > V^{d\prime}(0)$ ($\Pi > \Pi_S(1)$) or $V^{ds\prime}(0|0) = V^{d\prime}(0)$ & $V^{ds\prime\prime}(0|0) > V^{d\prime\prime}(0)$ ($\Pi = \Pi_S(1)$). Consider any $u \geqslant 0$. By using $V^{ds\prime}(0|0) \geqslant V^{d\prime}(0)$ and Lemma B.4.6 (b), we have $V^{ds\prime}(u|0) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (b) of Lemma B.4.8, $0 \geqslant L_2^D(u, R|0)$ for all $R \geqslant u + \phi/\lambda_D$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (a) of Lemma B.4.8, $L_2^S(u, u_S|0) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{ds}, we have $L_2^S(u, u + \phi/\lambda_D|0) = 0$. Therefore, the value function $V^{ds}(u|0)$ solves (HJB_V) subject to (IC).

By (b) of Lemma B.4.6 and $V^{d\prime}(0) < V^{ds\prime}(0|0)$, we have $V^{ds\prime\prime}(u|0) < 0$ for all $u > 0$, i.e., $V^d(\cdot|0)$ is concave. By Lemma B.2.2 and Proposition 3.4.1 (c), $V^{ds}(u|0)$ is the principal's value function.

□

Proof of Proposition 3.5.2. (a) By Lemma B.4.6 (a), we have $V^{d\prime}(u) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 0$, (IC) with $b = 1$ is equivalent to

$u_S \geqslant u + \phi/\lambda_S$. By (b) of Lemma B.4.7, $L_1^S(u, u_S) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (a) of Lemma B.4.7, $0 \geqslant L_1^D(u, R)$ for all $R \geqslant u + \phi/\lambda_D$. In addition, from the definition of V^d, we have $L_1^D(u, u + \phi/\lambda_D) = 0$. Therefore, $V^d(u)$ solves (HJB$_V$) subject to (IC). By (a) of Lemma B.4.6, V^d is concave. By Lemma B.2.2 and Proposition 3.4.1 (a), V^d is the principal's value function.

(b) Set the thresholds u_1 and u_2 as in Proposition B.4.2 (b).

First, when $u \in [0, u_1]$, by following the same steps as in Proposition 3.4.2 (a), we can show that $V^d(u)$ solves (HJB$_V$) subject to (IC).

Second, consider the case where $u_1 < u \leqslant u_2$. Then, by Lemma B.4.6 (b), we have $V^{ds\prime}(u|u_1) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (c) of Lemma B.4.8, $0 \geqslant L_2^D(u, R|u_1)$ for all $R \geqslant u + \phi/\lambda_D$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (a) of Lemma B.4.8, $L_2^S(u, u_S|u_1) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{ds}, we have $L_2^S(u, u + \phi/\lambda_D|u_1) = 0$. Thus, for all $u > u_1$, $V^{ds}(u|u_1)$ solves (HJB$_V$) subject to (IC).

Last, consider the case where $u > u_2$. Then, by Lemma B.4.6 (c), we have $V^{dsd\prime}(u|u_1, u_2) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (b) of Lemma B.4.9, $L_3^S(u, u_S|u_1, u_2) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (a) of Lemma B.4.9, $L_3^S(u, u_S|u_1, u_2) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{dsd}, we have $L_3^D(u, u + \phi/\lambda_D|u_1, u_2) = 0$. Therefore, the value function specified in (3.5.2) solves (HJB$_V$) subject to (IC).

By using Lemma B.4.6, $V^{d\prime}(u_1) = V^{ds\prime}(u_1|u_1)$ and $V^{ds\prime}(u_2|u_1) = V^{dsd\prime}(u_2|u_1, u_2)$, the function defined in (3.5.2) is concave.

(c) Set the thresholds $u_1 = 0$ and $u_2 > 0$ as in Proposition B.4.2 (c).

First, consider the case where $u \in [0, u_2]$, so $V(u) = V^{ds}(u|0)$. By using $V^{ds\prime}(0|0) > V^{d\prime}(0)$ and Lemma B.4.6 (b), we have $V'(u) = V^{ds\prime}(u|u_1) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (b) of Lemma B.4.8, $0 \geqslant L_2^D(u, R|0)$ for all $R \geqslant u + \phi/\lambda_D$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (a) of Lemma B.4.8, $L_2^S(u, u_S|0) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{ds}, we have $L_2^S(u, u + \phi/\lambda_D|0) = 0$. Therefore, for all $u \in [0, u_2]$, the value function $V^{ds}(u|0)$ solves (HJB$_V$) subject to (IC).

Next, consider the case where $u > u_2$, so $V(u) = V^{dsd}(u|0, u_2)$. By using $V^{ds\prime}(0|0) > V^{d\prime}(0)$ and Lemma B.4.6 (c), we have $V'(u) = V^{dsd\prime}(u|0, u_2) > -1$. By Lemma B.4.5, we can set $b = 1$. When $a = 0$, (IC) with $b = 1$ is equivalent to $u_S \geqslant u + \phi/\lambda_S$. By (b) of Lemma B.4.9, $L_3^S(u, u_S|0, u_2) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. When $a = 1$, (IC) with $b = 1$ is equivalent to $R \geqslant u + \phi/\lambda_D$. By (a) of Lemma B.4.9, $L_3^S(u, u_S|0, u_2) \leqslant 0$ for all $u_S \geqslant u + \phi/\lambda_S$. In addition, from the definition of V^{dsd}, we have $L_3^D(u, u + \phi/\lambda_D|0, u_2) = 0$. Thus, for all $u > u_2$, the value function $V^{dsd}(u|0, u_2)$ solves (HJB$_V$) subject to (IC). Therefore, the value function specified in (3.5.3) solves (HJB$_V$) subject to (IC).

By using Lemma B.4.6, $V^{d\prime}(0) < V^{ds\prime}(0|0)$ and $V^{ds\prime}(u_2|0) = V^{dsd\prime}(u_2|0, u_2)$, we can show that V is concave.

□

B.5 Proofs of Lemmas in Section B.3

B.5.1 Proof of Lemma B.3.1

Proof of Lemma B.3.1. If Π is greater than or equal to $\Pi_S(\eta)$ and η is greater than $1/(e-1)$, u_1 is equal to zero by Proposition B.4.1 and B.4.2. Note that

$$\Pi_S(\eta) = \frac{e-1}{(e-1)\eta - 1} \cdot \frac{c}{\lambda_D} \geqslant \frac{e-1}{e-2} \cdot \frac{c}{\lambda_D} > \frac{2c}{\lambda_D} \geqslant \frac{c+\phi}{\lambda_D} = \Pi_F.$$

Then, the project is feasible and $\bar{u}$ is greater than 0, thus, $\bar{u}$ is always greater than u_1.

Now suppose that $\Pi_D(\eta) < \Pi < \Pi_S(\eta)$. Since V is strictly concave, $u_1 < \bar{u}$ is equivalent to $0 < V'(u_1) = V^{d'}(u_1)$. Note that $0 < V^{d'}(u_1)$ is equivalent to:

$$\frac{\phi}{\lambda_D \Pi - c} < e^{-\frac{\lambda_D}{\phi} u_1}.$$

Recall that u_1 is the solution of

$$-H_1(u) = ((\eta+1)\lambda_D \Pi - c)e^{-1} e^{-\frac{(\eta+1)\lambda_D}{\phi} u} - \eta(\lambda_D \Pi - c) e^{-\frac{\lambda_D}{\phi} u} + (1-\eta)c = 0.$$ [7]

Define $x_1 \equiv e^{-\frac{\lambda_D}{\phi} u_1}$. Then, x_1 is the solution of

$$\tilde{H}_1(x) = ((\eta+1)\lambda_D \Pi - c)e^{-1} x^{\eta+1} - \eta(\lambda_D \Pi - c)x + (1-\eta)c = 0,$$

and we need to identify a condition for $x_1 > \phi/(\lambda_D \Pi - c)$.

Note that

$$\frac{\partial^2 \tilde{H}_1}{\partial x^2}(x) = (\eta+1)\eta((\eta+1)\lambda_D \Pi - c)e^{-1} x^{\eta-1} > 0,$$

thus there exists $\underline{x} \in (0,1)$ that minimizes $\tilde{H}_1$. Also note that $\Pi_S(\eta) > \Pi > \Pi_M(\eta)$ imply that $\tilde{H}_1(1) > 0 > \tilde{H}_1(\underline{x})$ and $x_1 \in (\underline{x}, 1)$.

[7] See the proof of Lemma B.4.2 for the definition of H_1.

There are two possible cases that satisfy $x_1 > \phi/(\lambda_D \Pi - c)$: (i) $\underline{x} \geqslant \phi/(\lambda_D \Pi - c)$; (ii) $\phi/(\lambda_D \Pi - c) > \underline{x}$ and $\tilde{H}_1(\phi/(\lambda_D \Pi - c)) < 0$.

The first case is equivalent to $\tilde{H}_1'(\phi/(\lambda_D \Pi - c)) < 0$. By algebra, we can show that it is equivalent to

$$\frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \frac{\eta e}{\eta+1}\phi^{-\eta}. \tag{B.5.1}$$

The second case is equivalent to $\tilde{H}_1'(\phi/(\lambda_D \Pi - c)) \geqslant 0$ and $\tilde{H}_1(\phi/(\lambda_D \Pi - c)) < 0$. By algebra, we can show that it is equivalent to

$$\frac{\eta e}{\eta+1}\phi^{-\eta} \leqslant \frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \left(\eta(c+\phi) - c\right) e \phi^{-\eta-1}. \tag{B.5.2}$$

Last, by the proof of Lemma B.4.2, we can show that $\Pi > \Pi_M(\eta)$ is equivalent to

$$\frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \left(\frac{\eta^2}{1-\eta^2}\right)^{\eta} \frac{\eta e}{1+\eta} c^{-\eta}. \tag{B.5.3}$$

Now I compare the above three conditions. When $\eta > \sqrt{c/(c+\phi)}$, by simple algebra, we can show that

$$\frac{\eta e}{\eta+1}\phi^{-\eta} < (\eta(c+\phi) - c) e \phi^{-\eta-1} < \left(\frac{\eta^2}{1-\eta^2}\right)^{\eta} \frac{\eta e}{1+\eta} c^{-\eta}.$$

Therefore, the inequality

$$\frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \left(\eta(c+\phi) - c\right) e \phi^{-\eta-1}$$

imply that (B.5.1), (B.5.2) and (B.5.3). Define $\Pi_D(\eta)$ be the value of Π that makes both sides of the above inequality equal. Then, $\Pi_D(\eta) > \Pi_M(\eta)$ since $\Pi < \Pi_M(\eta)$ implies $\Pi < \Pi_D(\eta)$. Therefore, there exists $\Pi_D(\eta) > \Pi_M(\eta)$ such that $u_1 < \bar{u}$ if and only if $\Pi > \Pi_D(\eta)$.

When $\eta \leqslant \sqrt{c/(c+\phi)}$, by simple algebra, we can show that

$$\frac{\eta e}{\eta+1}\phi^{-\eta} \geqslant (\eta(c+\phi)-c)e\phi^{-\eta-1} \quad \& \quad \frac{\eta e}{\eta+1}\phi^{-\eta} \geqslant \left(\frac{\eta^2}{1-\eta^2}\right)^{\eta}\frac{\eta e}{1+\eta}c^{-\eta}.$$

Therefore, (B.5.2) cannot hold in this case and (B.5.3) implies (B.5.1). It means that $\Pi > \Pi_M(\eta)$ implies $\tilde{H}_1'(\phi/(\lambda_D\Pi - c)) < 0$. Moreover, $\Pi > \Pi_M(\eta)$ is necessary for the existence of u_1. Hence, $u_1 < \bar{u}$ holds if and only if $\Pi > \Pi_M(\eta)$. □

B.5.2 Proof of Lemma B.3.2

Proof of Lemma B.3.2. Since V is strictly concave, $u_2 < \bar{u}$ is equivalent to $0 < V'(u_2)$. By (b) of Proposition 3.5.2, we have $V'(u_2) = {V^{ds}}'(u_2|u_1) = {V^{dsd}}'(u_2|u_1,u_2)$. By (B.2.5) and $V^{dsd}(u_2|u_1,u_2) = V^{ds}(u_2|u_1)$, $0 < {V^{dsd}}'(u_2|u_1,u_2)$ is equivalent to:

$$\lambda_D(u_2 + V^{ds}(u_2|u_1)) < \lambda_D\Pi - c - \phi. \tag{B.5.4}$$

Also note that ${V^{dsd}}'(u_2|u_1,u_2) = \phi {V^{ds}}'(u_2|u_1)$ and $V^{dsd}(u_2|u_1,u_2) = V^{ds}(u_2|u_1)$ imply that

$$\lambda_D(\Pi - u_2 - V^{ds}(u_2|u_1)) = \lambda_S\left(V_S\left(u_2 + \frac{\phi}{\lambda_S}\right) + u_2 + \frac{\phi}{\lambda_S}\right) - \lambda_S\left(V^{ds}(u_2|u_1) + u_2\right)$$

by (B.2.5) and (B.2.6). By plugging (3.3.1) into the above equation, we can derive that

$$\begin{aligned}
(\lambda_S - \lambda_D)(V^{ds}(u_2|u_1) + u_2) =& \lambda_S\left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi}u_2 - 1}\right) - \lambda_D\Pi \\
\iff \qquad \eta\lambda_D(V^{ds}(u_2|u_1) + u_2) =& \eta\lambda_D\Pi - c - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2 - 1}.
\end{aligned}$$

Then, by plugging this into (B.5.4), $0 < {V^{dsd}}'(u_2|u_1,u_2)$ is equivalent to

$$(\eta - 1)c + \eta\phi < (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_2 - 1}.$$

Since $\Pi > c/\lambda_D > c/\lambda_S$, the right hand side of the above inequality is always greater than 0. Since it is assumed that $\eta > \dfrac{c}{c+\phi}$, the left hand side of the above inequality is always less than 0. Therefore, the above inequality always holds, i.e., u_2 is less than $\bar{u}$.

□

B.5.3 Proof of Lemma B.3.3

Proof of Lemma B.3.3. By following the proof of Lemma B.3.2, $u_2 \geqslant \bar{u}$ is equivalent to

$$\hat{y} \equiv \frac{(\eta-1)c+\eta\phi}{(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}u_1 - 1}} \geqslant e^{\frac{\lambda_S}{\phi}(u_1-u_2)} \tag{B.5.5}$$

By the proof of Lemma B.4.3, $x_2 \equiv e^{\frac{\lambda_S}{\phi}(u_1-u_2)}$ is the solution, which is not equal to 1, of $\tilde{H}_2(x) = 0$.[8]. Since $u_2 \geqslant u_1$, if $\hat{y} \geqslant 1$, the above inequality holds, thus, I restrict attention to the case of $\hat{y} < 1$. Observe that the inequality $\tilde{H}_2(\hat{y}) \leqslant 0$ implies (B.5.5) because $\tilde{H}_2$ is strictly convex in x and $\tilde{H}_2(1) \leqslant 0$.

Note that $\tilde{H}_2(x)$ can be rewritten as follows:

$$\tilde{H}_2(x) = \frac{1-\eta}{1+\eta}c - H_1(u_1)x + \left[-\frac{1-\eta}{1+\eta}c + \frac{\eta}{1+\eta}(\lambda_S\Pi - c)e^{-\frac{\lambda_D}{\phi}u_1 - 1}\log x\right]x$$

where H_1 is a function defined in (B.4.1). Also note that $H_1(u_1) = \phi V^{ds''}(u_1|u_1) - \phi V^{d'}(u_1) \geqslant 0$.

By plugging the definition of $\hat{y}$ into the above equation, we can derive that

$$\tilde{H}_2(\hat{y}) = \frac{1-\eta}{1+\eta}c(1-\hat{y}) - H_1(u_1)\hat{y} + \frac{\eta}{1+\eta}\left((\eta-1)c+\eta\phi\right)\log\hat{y}.$$

Now define a new function G as follows:

$$G(y) \equiv \frac{1-\eta}{1+\eta}c(1-y) - H_1(u_1)y + \frac{\eta}{1+\eta}\left((\eta-1)c+\eta\phi\right)\log y,$$

[8] The function $\tilde{H}_2$ is defined in (B.4.2)

and it is enough to show that $G(y) \leqslant 0$ for all $y < 1$.

Note that

$$G''(y) = -\frac{\eta}{1+\eta}\left(\frac{(\eta-1)c+\eta\phi}{y^2}\right) < 0$$

from $\eta \geqslant \sqrt{c/(c+\phi)} > c/(c+\phi)$. Also note that

$$G'(1) = -H_1(u_1) + \frac{1}{1+\eta}\left((\eta^2-1)c+\eta^2\phi\right) < 0.$$

from $\eta \geqslant \sqrt{c/(c+\phi)}$ and $H_1(u_1) \geqslant 0$. Lastly, note that $G(1) = -H_1(u_1) \leqslant 0$. Therefore, for all $y < 1$, $G(y) \leqslant G(1) + G'(1)(1-y) \leqslant 0$. Therefore, $\tilde{H}_2(\hat{y}) \leqslant 0$ and $u_2 \geqslant \bar{u}$. □

I also explore three variations of the model that reflect relevant features in some economic applications. First, I analyze the case of asymmetric arrival rates of subprojects and ask whether this can change the form of the optimal contract (Section B.6). Next, I consider outside options for both players that emerge after the completion of the subproject under the sequential approach and investigate how they affect the optimal contract (Section B.7). Last, I introduce discounting and discuss how it affects the desirability of each approach (Section B.8).

B.6 Asymmetric Arrival Rates for Subprojects

Here I investigate a setting where the arrival rates for the subprojects in the sequential approach are no longer the same. Let $\lambda_{S,1}$ and $\lambda_{S,2}$ denote the arrival rates for the first and second subprojects. The ratio between these two arrival rates, $\lambda_{S,2}/\lambda_{S,1}$, is denoted by κ. In Theorem B.6.1, I show that the form of the optimal contract depends crucially on κ.

To simplify the discussion, I restrict attention to the case where there is no efficiency loss from monitoring. I begin by considering the first-best scenario to derive the condition that fulfills this restriction. Under the assumption that the agent's allocation of effort is observable to the principal, the principal's expected profit from the sequential-only contract without a deadline is derived as follows:

$$\int_0^\infty \int_{\tau_s}^\infty \left(\Pi - \int_0^{\tau_m} c\ dt\right) \lambda_{S,2} e^{-\lambda_{S,2}(\tau_m - \tau_s)} d\tau_m\ \lambda_{S,1} e^{-\lambda_{S,1}\tau_s} d\tau_s = \Pi - \frac{c}{\lambda_{S,1}} - \frac{c}{\lambda_{S,2}}.$$

Note that the principal's expected profit from the direct-only contract without a deadline is $\Pi - c/\lambda_D$. By using similar steps to those in Section 3.2.2, we can show that there is no efficiency loss from monitoring if and only if $1/\lambda_D = 1/\lambda_{S,1} + 1/\lambda_{S,2}$.

Under this condition, by using the definition of κ, we can derive that $\lambda_{S,1} = (1 + 1/\kappa)\,\lambda_D$ and $\lambda_{S,2} = (1 + \kappa)\lambda_D$. From these equations, we observe that higher κ

implies that it is harder to achieve the first subproject (lower $\lambda_{S,1}$) but it is easier to complete the second subproject (higher $\lambda_{S,2}$). Note that the completion of the first subproject is monitored by the principal. If the first subproject becomes more difficult and the second subproject becomes easier, we can think of monitoring as relatively more effective. Thus, κ can be interpreted as a parameter that measures the *effectiveness* of monitoring. The following theorem characterizes the optimal contract under an arbitrary κ.

Theorem B.6.1. *Suppose that there is no efficiency loss from monitoring and the arrival rates of the subprojects are not necessarily symmetric. There is a threshold $\kappa^* > 0$ such that the optimal contract is implemented as follows.*

(a) If $\kappa < \kappa^*$, there exist $\Pi_S^A(\kappa) > \Pi_D^A(\kappa) > \Pi_F^A(\kappa) = \Pi_F = (c+\phi)/\lambda_D$ such that the optimal contract is determined as follows:

(i) when $\Pi \leqslant \Pi_F^A(\kappa) = \Pi_F$, the project is infeasible;

(ii) when $\Pi_D^A(\kappa) \geqslant \Pi > \Pi_F^A = \Pi_F$, $(\bar{u}, V(\bar{u}))$ is implemented by a direct-only contract with a deadline $\bar{u}/\phi$;

(iii) when $\Pi_S^A(\kappa) > \Pi > \Pi_D^A(\kappa)$, there exists $u_1 \in (0, \bar{u})$ such that $(\bar{u}, V(\bar{u}))$ is implemented by a contract with a switch from the sequential approach to the direct approach at $(\bar{u} - u_1)/\phi$ and a deadline $\bar{u}/\phi$;

(iv) when $\Pi \geqslant \Pi_S^A(\kappa)$, $(\bar{u}, V(\bar{u}))$ is implemented by a sequential-only contract with a deadline $\bar{u}/\phi$.

(b) If $\kappa \geqslant \kappa^*$, there exists $\Pi_F^A(\kappa) \leqslant \Pi_F$ such that the optimal contract is determined as follows:

(i) when $\Pi \leqslant \Pi_F^A(\kappa)$, the project is infeasible;

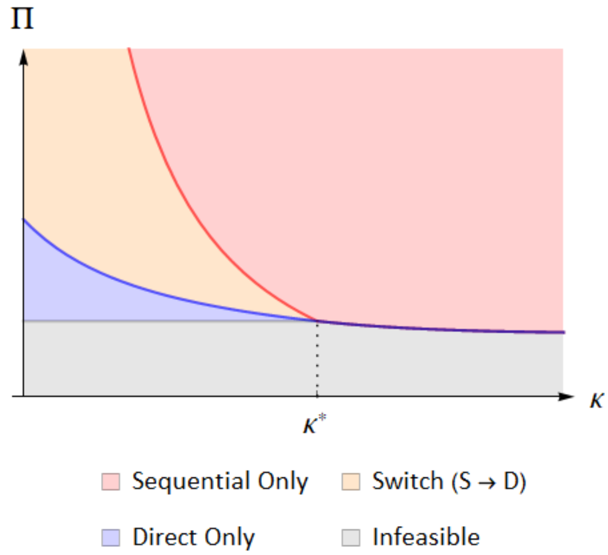


FIGURE B.1: Optimal contracts under asymmetric arrival rates and no efficiency loss

(ii) when $\Pi > \Pi_F^A(\kappa)$, $(\bar{u}, V(\bar{u}))$ is implemented by a sequential-only contract with a deadline $\bar{u}/\phi$.

This theorem is illustrated in Figure B.1. When monitoring is relatively ineffective, i.e., κ is below κ^*, optimal contracts are characterized in the same way as in Theorem 3.4.1: (i) a sequential-only contract when Π is high; (ii) a contract with one switch when Π is intermediate; (iii) a direct-only contract when Π is relatively low; (iv) infeasible when Π is very low. However, when monitoring is relatively effective, i.e., κ is above κ^*, optimal contracts are characterized in a different manner. In this case, the optimal contract is in the form of a sequential-only contract when Π is not low, and the project is infeasible when Π is low.

From this exercise, we can conclude that the form of the optimal contract is determined not only by the return of the project (Π) but also by the effectiveness of monitoring (κ). In particular, when monitoring is effective enough, monitoring is a more important factor than time pressure. Thus, even when the return of the project is low, the principal always chooses the sequential approach so long as the project is feasible.

The rest of this section is devoted to the derivation of Theorem B.6.1.

B.6.1 Benchmark Value Functions

Under the assumption that the arrival rates for the subprojects are no longer symmetric, the agent's HJB equation (HJB_{PK}) can be rewritten as follows:

$$0 = \max_{\substack{a_t \in \{0,1\}, \\ b_t \in [0,1]}} \dot{u}_t + \phi(1 - b_t) + (R_t - u_t)\lambda_D a_t b_t + (u_S^t - u_t)\lambda_{S,1}(1 - a_t)b_t. \qquad (\text{PK}_A)$$

Then, we can observe that the sequential approach can be induced with a minimum incentive by setting $u_S^t = u_t + \phi/\lambda_{S,1}$.

Also note that the principal's value function given that the subproject, V_S^A, is already completed can be written as follows:

$$V_S^A(u_S) = \left(\Pi - \frac{c}{\lambda_{S,2}} - u_S\right) - \left(\Pi - \frac{c}{\lambda_{S,2}}\right) e^{-\frac{\lambda_{S,2}}{\phi} u_S}. \qquad (\text{B.6.1})$$

Let the principal's value function prior to the completion of the subproject be V^A. The HJB equation for $V^A(u)$ is

$$0 = \max_{\substack{R \geqslant 0,\ u_S \geqslant 0, \\ a \in \{0,1\},\ b \in [0,1]}} -c + (\Pi - R - V^A(u))\lambda_D ab + (V_S^A(u_S) - V^A(u))\lambda_{S,1}(1-a)b + V^{A\prime}(u)\, \dot{u}. \qquad (\text{HJB}_A)$$

Then, the principal's problem is to solve (HJB_A) subject to (PK_A). We can derive two benchmark value functions as follows:[9]

1. When the agent's promised utility is lower than a switching point u_1, the principal would recommend the direct approach and the benchmark value function,

[9] Since I focus on the case of no efficiency loss from splitting the project, a benchmark value function involving two switches such as (3.5.1) is not needed.

V_A^d, would be the same as (3.4.1):

$$V_A^d(u) = \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{-\frac{\lambda_D}{\phi}u}\right) - u.$$

2. When the agent's promised utility is higher than a switching point u_1, the principal would recommend the sequential approach. By following the similar steps in Section B.2.4, the benchmark value function, V_A^{ds}, is derived as follows:

$$V_A^{ds}(u|u_1) = \left(\Pi - \frac{c}{\lambda_{S,1}} - \frac{c}{\lambda_{S,2}}\right)\left(1 - e^{\frac{\lambda_{S,1}}{\phi}(u_1-u)}\right) + (V_A^d(u_1) + u_1)e^{\frac{\lambda_{S,1}}{\phi}(u_1-u)}$$
$$- \left(\Pi - \frac{c}{\lambda_{S,2}}\right)\frac{e^{\frac{\lambda_{S,2}}{\phi}(u_1-u)} - e^{\frac{\lambda_{S,1}}{\phi}(u_1-u)}}{\kappa - 1}e^{-\frac{\lambda_{S,1}}{\phi}u_1 - \kappa} - u.\text{[10]} \tag{B.6.2}$$

B.6.2 Value Function Derivation

To derive the value function, we need to characterize a threshold of Π that determines the recommended approach at the deadline as in Lemma B.4.1.

Lemma B.6.1. *Suppose that there is no efficiency loss from splitting the project and $\kappa = \lambda_{S,2}/\lambda_{S,1}$. The inequality $V_A^{d\prime}(0) > V_A^{ds\prime}(0|0)$ holds if and only if*

$$\Pi < \Pi_S^A(\kappa) \equiv \frac{1 - e^{-\kappa}}{1 - (\kappa + 1)e^{-\kappa}} \cdot \frac{c}{\lambda_D}.\text{[11]} \tag{B.6.3}$$

In addition, $\Pi_S^A(\kappa)$ is decreasing in κ.

[10] Note that

$$\lim_{\kappa\to 1}\frac{e^{\frac{\lambda_{S,2}}{\phi}(u_1-u)} - e^{\frac{\lambda_{S,1}}{\phi}(u_1-u)}}{\kappa - 1} = \frac{(u - u_1)e^{\frac{\lambda_{S,2}}{\phi}(u_1-u)}}{\phi}.$$

Hence, when $\kappa = 1$, i.e., $\lambda_{S,1} = \lambda_{S,2}$, (B.6.2) is equivalent to (3.4.2).

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

Proof of Lemma B.6.1. Note that

$$V_A^{d\prime}(0) = \frac{1}{\phi}\left(\lambda_D \Pi - \phi - c\right),$$

$$\begin{aligned} V_A^{ds\prime}(0|0) &= \frac{1}{\phi}\left(\lambda_{S,1} V_S^A\left(\frac{\phi}{\lambda_{S,1}}\right) - c\right) \\ &= \frac{1}{\phi}\left(\left(1 + \frac{1}{\kappa}\right)(1 - e^{-\kappa})\lambda_D \Pi - \frac{1}{\kappa}(1 - e^{-\kappa})c - \phi - c\right). \end{aligned}$$

Therefore, $V_A^{d\prime}(0) > V_A^{ds\prime}(0|0)$ is equivalent to

$$(1 - e^{-\kappa})c > (1 - (\kappa + 1)e^{-\kappa})\lambda_D \Pi.$$

Since $1-(\kappa+1)e^{-\kappa} > 0$ for all $\kappa > 0$, the above inequality is equivalent to $\Pi_S^A(\kappa) > \Pi$.

By differentiating Π_S^A with respect to κ, we have

$$\Pi_S^{A\prime}(\kappa) = -\frac{e^{\kappa}\left(e^{-\kappa} + \kappa - 1\right)}{\left(1 - (\kappa + 1)e^{-\kappa}\right)^2} \cdot \frac{c}{\lambda_D}.$$

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

Proposition B.6.1. *Assume that both technologies are equally efficient and $\lambda_{S,2}/\lambda_{S,1}$ is equal to κ. Then, the following statements hold:*

(a) When $\Pi \in (c/\lambda_D, \Pi_S^A(\kappa))$, for some $u_1 > 0$, $V_A^{d\prime}(u_1) = V_A^{ds\prime}(u_1|u_1)$ and

$$V^A(u) = \begin{cases} V_A^d(u) & \text{if } 0 \leqslant u \leqslant u_1, \\ V_A^{ds}(u|u_1) & \text{if } u_1 < u \end{cases} \tag{B.6.4}$$

solves (HJB_A) *subject to* (PK_A).

(b) When $\Pi \geqslant \Pi_S^A(\kappa)$, $V^A(u) = V_A^{ds}(u|0)$ solves (HJB_A) *subject to* (PK_A).

B.6.3 Thresholds

Feasibility (Π_F^A)

Now I derive a threshold that determines the feasibility of the project. By using the same logic in Appendix B.3.1, the project is feasible if and only if $V'(0) > 0$, and it is equivalent to

$$\max\left[\lambda_D \Pi - \phi,\ \lambda_{S,1} V_S^A(\phi/\lambda_{S,1})\right] > c.$$

Note that

$$\begin{aligned}\lambda_{S,1} V_S^A(\phi/\lambda_{S,1}) =& \lambda_{S,1}\left(\Pi - \frac{c}{\lambda_{S,2}}\right)(1-e^{-\kappa}) - \phi \\ =& \left(1+\frac{1}{\kappa}\right)\lambda_D(1-e^{-\kappa})\Pi - \frac{1-e^{-\kappa}}{\kappa}c - \phi.\end{aligned}$$

Therefore, $\lambda_{S,1} V_S^A(\phi/\lambda_{S,1}) > c$ is equivalent to

$$\Pi > \frac{1}{(\kappa+1)\lambda_D}\left[c + \frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right].$$

Then, the project is feasible if and only if

$$\Pi > \min\left[\frac{c+\phi}{\lambda_D}, \frac{1}{(\kappa+1)\lambda_D}\left[c + \frac{\kappa}{1-e^{-\kappa}}(c+\phi)\right]\right] \equiv \Pi_F^A(\kappa).$$

The Length of the Contract (Π_D^A)

Next, I derive another threshold that determines whether there is a switch in the optimal contract. Let $\bar{u}$ denote the value that maximizes $V^A(u)$. Then, we need to compare $\bar{u}$ with a switch point u_1. The following lemma characterizes the threshold Π_D^A.

Lemma B.6.2. *Suppose that there is no efficiency loss from splitting the project and $\kappa = \lambda_{S,2}/\lambda_{S,1}$. Then, there exists $\Pi_D^A(\kappa)$ such that $u_1 < \bar{u}$ if and only if $\Pi > \Pi_D^A(\kappa)$.*

Proof of Lemma B.6.2. I begin by deriving the closed form of u_1. Define a function H_1^A as in (B.4.1):

$$H_1^A(u) \equiv \phi V_A^{ds\prime}(u|u) - \phi V_A^{d\prime}(u). \tag{B.6.5}$$

Then, $H_1^A(u)$ can be rewritten as follows:

$$\begin{aligned}
H_1^A(u) &= \lambda_{S,1}\left((\Pi - c/\lambda_{S,2})(1 - e^{-\frac{\lambda_{S,2}}{\phi}u-\kappa}) - V_A^d(u) - u\right) - \lambda_D(\Pi - u - V_A^d(u)) \\
&= (\lambda_{S,1}\Pi - c/\kappa)\left(1 - e^{-\frac{\lambda_{S,2}}{\phi}u-\kappa}\right) - (\lambda_{S,1} - \lambda_D)(u + V_A^d(u)) - \lambda_D\Pi \\
&= (\lambda_{S,1}\Pi - c/\kappa)\left(1 - e^{-\frac{\lambda_{S,2}}{\phi}u-\kappa}\right) - (\lambda_{S,1} - \lambda_D)\left(\Pi - c/\lambda_D\right)(1 - e^{-\frac{\lambda_D}{\phi}u}) - \lambda_D\Pi \\
&= (\lambda_{S,1}/\lambda_D - 1)(\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u} - (1/\kappa - \lambda_{S,1}/\lambda_D + 1)c \\
&\quad - (\lambda_{S,1}\Pi - c/\kappa)e^{-\frac{\lambda_{S,1}}{\phi}u-\kappa} \\
&= (\lambda_D\Pi - c)e^{-\frac{\lambda_D}{\phi}u}/\kappa - (\lambda_{S,2}\Pi - c)e^{-\frac{\lambda_{S,1}}{\phi}u-\kappa}/\kappa.
\end{aligned}$$

By the definition of u_1, it is the solution of $H_1^A(u) = 0$. Then, by using the above equation, we can derive that

$$u_1 = \frac{\phi}{\lambda_D}\left[\frac{1}{\kappa}\log\left(\frac{\lambda_{S,2}\Pi - c}{\lambda_D\Pi - c}\right) - 1\right].$$

The next step is to compare u_1 with $\bar{u}$. Since V^A is strictly concave, $u_1 < \bar{u}$ is equivalent to $0 < V^{A\prime}(u_1) = V_A^{d\prime}(u_1)$. By using the above equation, we can derive that $u_1 = \bar{u}$ is equivalent to

$$\frac{(\lambda_D\Pi - c)^{1+\kappa}}{(1+\kappa)\lambda_D\Pi - c}\left(\frac{e}{\phi}\right)^{\kappa} = 1. \tag{B.6.6}$$

Note that the left hand side is increasing in Π, is equal to zero when Π is equal to c/λ_D, and diverges as Π goes to infinity. Therefore, there exists a unique solution that satisfies (B.6.6) and I denote the solution as $\Pi_D^A(\kappa)$. Then, $\Pi > \Pi_D^A(\kappa)$ is equivalent to $\bar{u} > u_1$. $\square$

Comparison of Thresholds

Lemma B.6.3. *Let κ^* be the positive solution of the equation*

$$0 = \phi + (c+\phi)\kappa - \phi e^{\kappa}. \tag{B.6.7}$$

Then,

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

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

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

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

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

$$\Pi_F^A(\kappa) = \frac{c+\phi}{\lambda_D} < \Pi_D^A(\kappa) < \Pi_S^A(\kappa);$$

4. *as $\kappa \to 0$,*

$$\lim_{\kappa\to 0}\Pi_D^A(\kappa) = \frac{c + \phi\cdot\psi\left(c/\phi\right)}{\lambda_D} \quad \textit{and} \quad \lim_{\kappa\to 0}\Pi_S^A(\kappa) = \infty,$$

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

Proof of Lemma B.6.3. Note that for all $\kappa > 0$ and $\lambda_D > 0$,

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

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,

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

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

$$L(\Pi) = (\kappa+1)\lambda_D\Pi - c,$$

$$R_F(\Pi|\kappa) = \begin{cases} \dfrac{\kappa}{1-e^{-\kappa}}(c+\phi) & \text{if } \ \kappa \geqslant \kappa^* \\ \kappa(c+\phi)+\phi & \text{if } \ \kappa \leqslant \kappa^* \end{cases},^{12}$$

$$R_W(\Pi|\kappa) = \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_D\Pi - c}{\phi}\right)^{\kappa+1},$$

$$R_S(\Pi|\kappa) = \phi \cdot e^{\kappa} \cdot \left(\frac{\lambda_D\Pi - c}{\phi}\right).$$

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

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

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

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

and it implies that $\Pi_j(\kappa)$ is lower 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 $\Pi_F^A(\kappa) > \max\left[\Pi_D^A(\kappa), \Pi_S^A(\kappa)\right]$, it is enough to show that $R_F(\Pi_F^A(\kappa)|\kappa) < R_W(\Pi_F^A(\kappa)|\kappa)$ and $R_F(\Pi_F^A(\kappa)|\kappa) < R_S(\Pi_F^A(\kappa)|\kappa)$.

   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

$$\frac{\lambda_D \Pi_F^A(\kappa) - c}{\phi} = \frac{\kappa}{\kappa+1}\left(\frac{c + e^\kappa \phi}{e^\kappa - 1}\right) = \frac{x(\kappa)+\kappa}{\kappa+1}.$$

   By using the definition of $x(\kappa)$ and the above equation, we can see that

$$\begin{aligned} R_F(\Pi_F^A(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot x(\kappa), \\ R_W(\Pi_F^A(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right)^{\kappa+1}, \\ R_S(\Pi_F^A(\kappa)|\kappa) &= \phi \cdot e^\kappa \cdot \left(\frac{x(\kappa)+\kappa}{\kappa+1}\right). \end{aligned} \tag{B.6.8}$$

   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, $R_W(\Pi_F^A(\kappa)|\kappa) > R_F(\Pi_F^A(\kappa)|\kappa)$. Also, we can easily see that $\dfrac{x+\kappa}{\kappa+1} > x$ is equivalent to $x < 1$, i.e., $R_S(\Pi_F^A(\kappa)|\kappa) > R_F(\Pi_F^A(\kappa)|\kappa)$.

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

Note that $x(\kappa^*) = 1$. Hence, by (B.6.8), $R_F(\Pi_F^A(\kappa)|\kappa) = R_W(\Pi_F^A(\kappa)|\kappa) = R_S(\Pi_F^A(\kappa)|\kappa)$.

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

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

   Also note that

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

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

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

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

   Define $y(\kappa) \equiv \left(\lambda_D \Pi_D^A(\kappa) - c\right)/\phi > 0$. Then, from (B.6.6), $y(\kappa)$ 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

$$\lim_{\kappa \to 0} \Pi_D^A(\kappa) = \frac{c + \phi \cdot y(0)}{\lambda_D} = \frac{c + \phi \cdot \psi\left(\frac{c}{\phi}\right)}{\lambda_D}.$$

□

B.6.4 Proof of Theorem B.6.1

Proof of Theorem B.6.1. (a) By Lemma B.6.3, when $\kappa < \kappa^*$, $\Pi_S^A(\kappa) > \Pi_D^A(\kappa) > \Pi_F^A(\kappa) = (c + \phi)/\lambda_D$.

(i) By the argument in Section B.6.3, the project is infeasible if $\Pi \leqslant \Pi_F^A(\kappa)$.

(ii) When $\Pi_D^A(\kappa) \geqslant \Pi > \Pi_F^A(\kappa)$, u_1 is greater than or equal to $\bar{u}$ by Lemma B.6.2. By Proposition B.6.1, the value function is $V^A(u) = V_A^g(u)$ for $u \leqslant \bar{u} \leqslant u_1$. Thus, the optimal contract is to execute the direct approach for all $u \leqslant \bar{u}$. Therefore, the direct-only contract with $T = \bar{u}/\phi$ implements $(\bar{u}, V(\bar{u}))$.

(iii) When $\Pi_S^A(\kappa) > \Pi > \Pi_D^A(\kappa)$, u_1 is less than $\bar{u}$ by Lemma B.6.2 and greater than zero by Lemma B.6.1. By Proposition B.6.1, the value function is $V(u) = V_A^{ds}(u|u_1)$ for $u_1 < u < \bar{u}$ and $V(u) = V_A^g(u)$ for $0 \leqslant u \leqslant u_1$. Thus, the optimal contract is to execute the sequential approach for $u_1 < u \leqslant \bar{u}$ and the direct approach for $0 \leqslant u \leqslant u_1$. Therefore, the contract with a switch from the sequential approach to the direct approach at $(\bar{u} - u_1)/\phi$ and a deadline $\bar{u}/\phi$.

(iv) When $\Pi \geqslant \Pi_S^A(\kappa)$, $V^A(u) = V_A^{ds}(u|0)$ by Proposition B.6.1. Thus, the optimal contract is to execute the sequential approach for $0 \leqslant u \leqslant \bar{u}$. Therefore, the sequential-only contract with $T = \bar{u}/\phi$ implements $(\bar{u}, V(\bar{u}))$.

(b) By Lemma B.6.3, when $\kappa \geqslant \kappa^*$, $\Pi_F \geqslant \Pi_F^A(\kappa) \geqslant \Pi_S^A(\kappa)$.

(i) By the argument in Section B.6.3, the project is infeasible if $\Pi \leqslant \Pi_F^A(\kappa)$.

(ii) When $\Pi > \Pi_F^A(\kappa)$, Π is greater than $\Pi_S^A(\kappa)$. Thus, $V^A(u) = V_A^{ds}(u|0)$ by Proposition B.6.1 and the sequential-only contract with $T = \bar{u}/\phi$ implements $(\bar{u}, V(\bar{u}))$.

□

B.7 Independent Values from a Subproject

In many relevant applications, both players may benefit from the completion of the subproject. In particular, their outside options will differ before and after the intermediate breakthrough. For example, by adding experience to his resume, the agent can get a better offer from other firms. For the principal, even if the agent does not finish the main project, she may be able to license out the current progress to another company.

The goal of this section is to investigate how these outside options affect the optimal contract. I address this question by exploring a numerical example. I fix parameter values as follows: $\Pi = 5$, $c = 1$, $\phi = .5$, $\lambda_D = 1$, $\lambda_S = 2$. Note that λ_S is equal to $2\lambda_D$ which means that two approaches are equally efficient. Moreover, Π is greater than $\Pi_S(1)$ implying that the optimal contract takes a form of a sequential-only contract with a deadline when there are no intermediate values from the subproject.

Let $B_P \geqslant 0$ and $B_A \geqslant 0$ denote the outside options for the principal and the agent when the subproject is completed. We can interpret a high B_P involving an active buyout market for projects and a high B_A involving an active labor market for experienced workers. The goal is to provide comparative statics on the deadlines and the principal's expected payoffs from the optimal contracts with respect to B_A and B_P. To derive the optimal contracts, we need to reformulate the principal's problem by adding B_A and B_P.

I begin by rewriting the promise-keeping constraint (HJB_{PK}). Once I introduce the outside option B_A for the agent, the effective promised utility for the agent at time t would be $u_S^{B,t} \equiv u_S^t - B_A$ where u_S^t is the agent's promised utility for the success of the subproject at t. Then, (HJB_{PK}) can be rewritten as follows:

$$0 = \max_{\substack{a_t \in \{0,1\}, \\ b_t \in [0,1]}} \dot{u}_t + \phi(1-b_t) + (R_t - u_t)\lambda_D a_t b_t + (u_S^{B,t} - u_t + B_A)\lambda_S(1-a_t)b_t. \quad (\text{HJB}_{PK}^B)$$

From this equation, we can infer that to induce the sequential approach, $u_S^{B,t}$ has to be greater than or equal to $u_t + \phi/\lambda_S - B_A$. To simplify discussion, I focus on the case where the principal "extends" the deadline after the completion of the subproject, i.e., $B_A < \phi/\lambda_S = .25$.

Next, we need to reconsider the principal's expected payoff after the completion of the subproject. Note that the additional value for the second subproject would be $\Pi - B_P$ since the principal has the outside option of B_P. In addition, the principal needs to work with the agent's effective promised utility rather than the original promised utility. Then, given the subproject completion, the principal's value function can be written as in (3.3.1):

$$V_S^B(u_S^B) = \left(\Pi - B_P - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi}u_S^B}\right) - u_S^B. \quad (\text{B.7.1})$$

This value stands for the principal's expected payoff additional to the outside option

B_P when the agent's effective promised utility is u_S^B. Thus, $V_S^B(u_S^B) + B_P$ is the principal's expected payoff after the completion of the subproject.

Note that the introduction of B_A and B_P does not affect the direct approach, thus V^d remains the same. Since (HJB$_{PK}$) is replaced by (HJB$^B_{PK}$), when the sequential approach is chosen, we need to substitute $u_S^B = u + \phi/\lambda_S - B_A$ for u_S. The principal's value function after the intermediate breakthrough ($V_S(u_S)$) also needs to be replaced by $V_S^B(u_S^B) + B_P$. Then, the HJB equation (B.2.6) can be rewritten as follows:

$$0 = -c + \lambda_S \left(V_S^B \left(u + \phi/\lambda_S - B_A\right) + B_P - V_B^{ds}(u|u_1) \right) - \phi V_B^{ds\prime}(u|u_1) \tag{B.7.2}$$

where V_B^{ds} is the new value function.

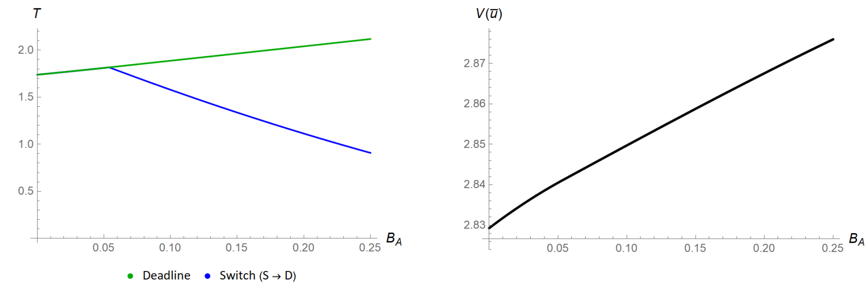
By plugging (B.7.1) into (B.7.2) and following the same steps in Appendix B.2.4, we can derive the closed form of V_B^{ds} as follows:

$$V_B^{ds}(u|u_1) = \left(\Pi - \frac{2c}{\lambda_S} + B_A\right)\left(1 - e^{\frac{\lambda_S}{\phi}(u_1 - u)}\right) + (V^d(u_1) + u_1)e^{\frac{\lambda_S}{\phi}(u_1 - u)}$$
$$- \left(\Pi - B_P - \frac{c}{\lambda_S}\right)\frac{\lambda_S}{\phi}(u - u_1)e^{\frac{\lambda_S}{\phi}(B_A - u) - 1} - u.$$

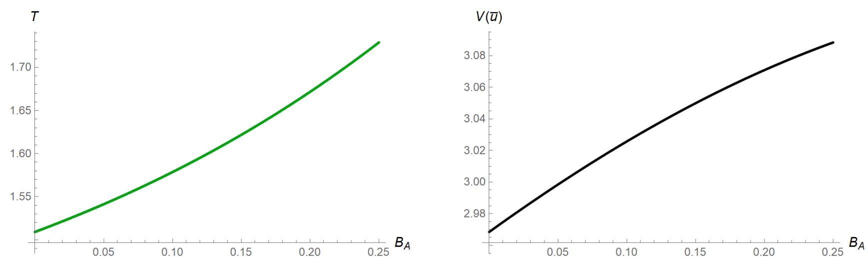
If there is no switching point, $V_B^{ds}(u|0)$ would serve as the value function. If there exists a switching point, we can pin down the switching point by following the same steps in Appendix B.4.1. When the switching point is u_1, the value function for $u \leqslant u_1$ is $V^d(u)$ and that for $u > u_1$ is $V^{ds}(u|u_1)$.

In Figure B.2, I illustrate the comparative statics on the optimal deadlines and the principal's expected payoffs under four different scenarios. In Figure B.2a and B.2b, I demonstrate the comparative statics with respect to B_A when B_P is 0 or 3. In Figure B.2c and B.2d, I display the comparative statics with respect to B_P when B_A is 0 or .2. In the left panels, the green curves demonstrate the optimal deadlines and the blue curves display the switching time from the sequential approach to the direct approach. If the blue curves are not present, it means that the sequential

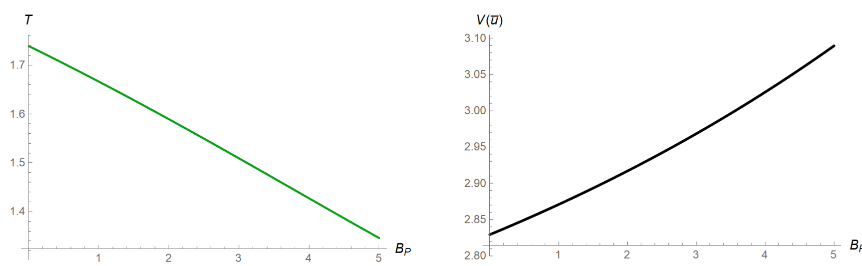
(a) Deadlines and the principal expected payoffs when $B_P = 0$



(b) Deadlines and the principal expected payoffs when $B_P = 3$



(c) Deadlines and the principal expected payoffs when $B_A = 0$



(d) Deadlines and the principal expected payoffs when $B_A = .2$

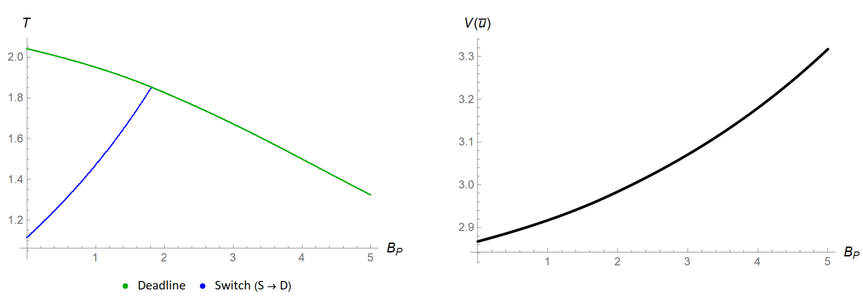


FIGURE B.2: Comparative Statics with respect to B_A and B_P

approach is chosen until the deadline. The principal's expected payoffs (before the intermediate breakthrough) are shown in the right panels.

First, we can observe that the principal's expected payoffs are increasing in B_A or B_P in every panel in Figure B.2. This means that the principal benefits from both the active buyout market for projects and the labor market for experienced workers. Since the active buyout market allows the principal to liquidate the project easier, it directly helps the principal. When the labor market for experienced workers becomes more active, it is easier for the principal to incentivize the agent to split the project. Given the intermediate breakthrough, while the principal needs to extend the deadline by $1/\lambda_S$ without the labor market, she only needs to extend the deadline by $1/\lambda_S - B_A/\phi$ with the labor market. This numerical example shows that the principal benefits from this decreased incentive under the restriction that B_A is less than ϕ/λ_S.[13]

Next, in the left panels of Figure B.2a and B.2b, we can see that the optimal deadline increases as B_A increases. The principal expects that the deadline extension would be shortened as B_A increases, thus she preemptively prolonds the deadline. Another interesting feature is that switching to the direct approach occurs when B_A is large and B_P is small as in the left panel of Figure B.2a. As B_A rises, the sequential approach becomes less appealing to the principal since the probability of success is lowered by the shortened deadline extension. Therefore, it may be more profitable to switch to the direct approach near the deadline. When B_P is large, even though the sequential approach becomes less desirable, it may still be better than the direct approach. Thus, there may not exist a switch at all as in B.2b.

Last, the left panels of Figure B.2c and B.2d show that the optimal deadline

[13] The case where B_A is greater than ϕ/λ_S is problematic for the following reason. In this case, if the principal provides the minimum incentive not to shirk, the deadline would be shortened after the completion of the subproject. This lowers the probability for the success of the project implying that it may not be optimal to employ "the minimum incentive contract" in the first place. This possibility significantly complicates the analysis of the model even numerically.

decreases as B_P rises. This is simply because an "early buyout" becomes more attractive as the buyout market grows. Moreover, as B_P increases, the sequential approach becomes more desirable. The left panel of Figure B.2d shows that the direct approach is optimal when B_P is small, but the sequential approach eventually dominates as B_P grows.

B.8 The Role of the Discount Rate

Finally, I introduce discounting to the model and investigate how it distorts the efficiency of each approach. Let $r \geqslant 0$ denote the common discount rate for the principal and the agent. The expected payoffs for indefinitely employing the direct approach and the sequential approach can be rewritten respectively as follows:

$$\begin{aligned}
F_D(r) \equiv & \int_0^\infty \left(e^{-r\tau_m}\Pi - \int_0^{\tau_m} e^{-rt} c dt \right) \lambda_D e^{-\lambda_D \tau_m} d\tau_m \\
= & \int_0^\infty \left(\frac{\lambda_D}{\lambda_D + r} \cdot \left(\Pi + \frac{c}{r}\right) \cdot (\lambda_D + r) e^{-(\lambda_D + r)\tau_m} - \frac{c}{r} \cdot \lambda_D e^{-\lambda_D \tau_m} \right) d\tau_m \\
= & \frac{\lambda_D}{\lambda_D + r} \left(\Pi - \frac{c}{\lambda_D} \right) = \frac{\lambda_D}{\lambda_D + r} \Pi - \frac{1}{\lambda_D + r} c,
\end{aligned}$$

$$\begin{aligned}
F_S(r) \equiv & \int_0^\infty \int_{\tau_s}^\infty \left(e^{-r\tau_m}\Pi - \int_0^{\tau_m} e^{-rt} c dt \right) \lambda_S e^{-\lambda_S(\tau_m - \tau_s)} d\tau_m \; \lambda_S e^{-\lambda_S \tau_s} d\tau_s \\
= & \int_0^\infty \left(\frac{\lambda_S}{\lambda_S + r} \cdot \left(\Pi + \frac{c}{r}\right) \cdot e^{-r\tau_s} - \frac{c}{r} \right) \lambda_S e^{-\lambda_S \tau_s} d\tau_s \\
= & \frac{\lambda_S^2}{(\lambda_S + r)^2} \left(\Pi + \frac{c}{r} \right) - \frac{c}{r} = \frac{\lambda_S^2}{(\lambda_S + r)^2} \Pi - \frac{2\lambda_S + r}{(\lambda_S + r)^2} c.
\end{aligned}$$

From these expressions, when $r > 0$, we observe that the efficiency relationship depends not only on the arrival rates but also on the return of the project (Π) and the flow cost (c). This is significantly different from the no-discounting case. When $r = 0$, the efficiency relationship is simply determined by comparing the expected durations under the two approaches ($1/\lambda_D$ and $2/\lambda_S$). Thus, the presence of the discount factor complicates the analysis.

To simplify the argument, I focus on the case where there is no efficiency loss from monitoring; i.e., $2\lambda_D = \lambda_S$. The following proposition shows how the discount rate distorts the efficiency of each approach.

Proposition B.8.1. *Suppose that $2\lambda_D = \lambda_S$ and $\Pi > c/\lambda_D$. Then, for all $r > 0$, the following inequality holds: $F_D(0) = F_S(0) > F_D(r) > F_S(r)$.*

Proof of Proposition B.8.1. Note that

$$F_D(0) = \Pi - \frac{c}{\lambda_D} = \Pi - \frac{2c}{\lambda_S} = F_S(0),$$

and

$$F_D(0) = \frac{\lambda_D \Pi - c}{\lambda_D} > \frac{\lambda_D \Pi - c}{\lambda_D + r} = F_D(r)$$

since $\Pi > c/\lambda_D$ and $r > 0$. Also note that

$$F_D(r) - F_S(r) = \frac{\lambda_D r}{(\lambda_D + r)(\lambda_S + r)^2}(r\Pi + c) > 0,$$

thus $F_D(r) > F_S(r)$. $\square$

This proposition says that the introduction of the discount rate harms the efficiency of the sequential approach more than that of the direct approach. In other words, if players begin to discount the future, the sequential approach becomes less appealing in terms of efficiency. Recall that the sequential approach is less advantageous in the short run. Thus, a high discount rate distorts the efficiency of the sequential approach more than that of the direct approach.

Appendix C

Appendix for Chapter 4

C.1 Proofs for Public Information Setting

C.1.1 Transformation

It is useful to write conditions in terms of the rate μ. The following lemma summarizes this transformation.

Lemma C.1.1. *Let* $\overline{\mu} \equiv \frac{2\lambda_L\lambda_H}{\lambda_H-\lambda_L}$ *and* $\underline{\mu} \equiv \frac{\lambda_L(\lambda_H+\lambda_L)}{\lambda_H-\lambda_L}$. *Then,* $\eta \geqslant \overline{\eta}(\delta)$ *is equivalent to* $\mu \geqslant \overline{\mu}$ *and* $\eta \leqslant \underline{\eta}(\delta)$ *is equivalent to* $\mu \leqslant \underline{\mu}$.[1]

Proof of Lemma C.1.1. First, we have:

$$\eta - \overline{\eta}(\delta) = \eta - (1+\delta) = \frac{\mu\lambda_H}{\lambda_L(\lambda_H+\mu)} - \frac{\lambda_H}{\lambda_H+\mu} - 1$$
$$= \frac{\mu(\lambda_H-\lambda_L) - 2\lambda_L\lambda_H}{\lambda_L(\lambda_H+\mu)} = \frac{\lambda_H-\lambda_L}{\lambda_L(\lambda_H+\mu)}(\mu-\overline{\mu}).$$

[1] Although $\overline{\mu}$ and $\underline{\mu}$ are the functions of λ_L and λ_H, we suppress them to simplify.

Therefore, $\eta \geqslant \overline{\eta}(\delta)$ is equivalent to $\mu \geqslant \overline{\mu}$. Next, we have

$$\begin{aligned}
\left(\underline{\eta}(\delta) - \eta\right)\left(\eta - \frac{1-\sqrt{1+4\delta(1-\delta)}}{2}\right) &= -\,\eta^2 + \eta + (1-\delta)\delta \\
&= -\left(\frac{\mu\lambda_H}{\lambda_L(\lambda_H+\mu)}\right)^2 + \frac{\mu\lambda_H}{\lambda_L(\lambda_H+\mu)} + \frac{\mu\lambda_H}{(\lambda_H+\mu)^2} \\
&= \frac{\mu\lambda_H(\lambda_H-\lambda_L)}{\lambda_L^2(\lambda_H+\mu)^2}\left(\underline{\mu} - \mu\right).
\end{aligned}$$

Note that $\eta > 1$ implies $\eta - \left\{1 - \sqrt{1+4\delta(1-\delta)}\right\}/2 > 0$. Thus, $\eta \leqslant \underline{\eta}(\delta)$ is equivalent to $\mu \leqslant \underline{\mu}$. $\square$

C.1.2 Proof of Proposition 4.3.1

Proposition 4.3.1 characterizes the MPE of the setting with observable technologies. A Markov strategy for player i is a mapping $\sigma^i : \Omega \to [0,1]$, where the state Ω represents the set of firms that possess the new technology. A MPE consists of a profile of Markov strategies such that each of the players is best responding to the strategy of their opponent.[2]

The existence of a best response that is Markov to a Markov opponent strategy is supported by the stationarity of the problem from the firm's perspective. Thus, the best Markov response must also be a best response over all strategies. This means that we need only consider Markov deviations to determine the Markov Perfect Equilibria.

Given a Markov strategy profile, we can compute the expected value for each player at every state. Let U^j_ω denote Firm j's continuation value at the state ω. Notice that for every strategy, a positive transition rate from state ω to ω' implies that

[2] Existance of these equilibria for a larger class of continuous-time stochastic games with finite states and actions has been studied in Neyman (2017).

$\omega \subseteq \omega'$. This allows us to obtain the equilibrium Markov strategies and continuation value at any MPE via *backward induction.*

We begin with the state $\omega = \{A, B\}$. For any firm that possesses the new technology, it is optimal to develop with it. Thus, in any MPE, $\sigma^A(\{A,B\}) = \sigma^B(\{A,B\}) = 1$ and the continuation value is

$$U^A_{\{A,B\}} = U^B_{\{A,B\}} = \frac{1}{2}\left(\Pi - \frac{c}{\lambda_H}\right) = V_C \tag{C.1.1}$$

Next, we consider jointly the states $\omega = \{A\}$ and $\omega = \{B\}$, i.e. the cases in which only one of the firms possesses the new technology. The firm with the new technology will trivially find it optimal to develop with it. Thus, $\sigma^A(\{A\}) = \sigma^B(\{B\}) = 1$ in any MPE. The following lemma characterizes $\sigma^i(\{j\})$, i.e. the opponent firm's optimal action.

Lemma C.1.2. *If $\eta < \overline{\eta}(\delta)$ then, in any MPE, $\sigma^A(\{B\}) = \sigma^B(\{A\}) = 0$. When $\eta > \overline{\eta}(\delta)$ then, in any MPE, $\sigma^A(\{B\}) = \sigma^B(\{A\}) = 1$.*

Proof of Lemma C.1.2. Let $\omega = \{i\}$, i.e., only firm i possesses the new technology. Firm i will develop with the new technology. For firm j, the problem is to choose σ to maximize his continuation value:

$$U^j_{\{i\}} = \max_{\sigma\in[0,1]} \frac{\sigma\mu V_C + (1-\sigma)\lambda_L \Pi - c}{\sigma\mu + (1-\sigma)\lambda_L + \lambda_H}$$

Taking the derivative of the previous objective function with respect to the choice variable σ, and using (C.1.1) we obtain

$$\frac{(\lambda_H \Pi + c)\cdot[\mu(\lambda_H - \lambda_L) - 2\lambda_H\lambda_L]}{2\lambda_H(\lambda_H + \lambda_L + \sigma(\mu - \lambda_L))^2} = \frac{(\lambda_H \Pi + c)(\lambda_H - \lambda_L)(\mu - \overline{\mu})}{2\lambda_H(\lambda_H + \lambda_L + \sigma(\mu - \lambda_L))^2}$$

Note that the above equation is positive if and only if $\mu > \bar{\mu}$, or equivalently by Lemma C.1.1, $\eta > \overline{\eta}(\delta)$. Therefore, the best response is $\sigma^i(\{j\}) = 1$ if $\eta > \overline{\eta}(\delta)$, and

$\sigma^i(\{j\}) = 0$ if $\eta < \overline{\eta}$. In addition, we have

$$U^j_{\{i\}} = \begin{cases} \dfrac{\mu V_C - c}{\mu + \lambda_H}, & \text{if } \eta \geqslant \overline{\eta}(\delta), \\ \dfrac{\lambda_L \Pi - c}{\lambda_L + \lambda_H}, & \text{if } \eta < \overline{\eta}(\delta). \end{cases}$$

□

Finally, it remains to analyze the state $\omega = \{\varnothing\}$. We divide the analysis into the following three lemmas.

Lemma C.1.3. *Assume $\eta < \underline{\eta}(\delta)$. Then the unique MPE involves $\sigma^A(\varnothing) = \sigma^B(\varnothing) = 0$.*

Proof of Lemma C.1.3. Notice that $\eta < \underline{\eta}(\delta) \leqslant \overline{\eta}(\delta)$. Thus, By Lemma C.1.2, once a firm finds the new technology, i.e., $\omega = \{A\}$ or $\{B\}$, the other one switches to developing with the incumbent technology since. In these cases, the values of the firms are given as follows:

$$U^A_{\{A\}} = U^B_{\{B\}} = \frac{\lambda_H \Pi - c}{\lambda_L + \lambda_H} \quad \text{and} \quad U^B_{\{A\}} = U^A_{\{B\}} = \frac{\lambda_L \Pi - c}{\lambda_L + \lambda_H}. \tag{C.1.2}$$

Back to the state $\varnothing$, let $\pi_i(\sigma^i_\varnothing, \sigma^j_\varnothing)$ be the expected continuation payoff of Firm i when firms play the actions $\sigma^i_\varnothing$ and $\sigma^j_\varnothing$ as long as neither firm obtains the new technology, and continue with the optimal actions thereafter.

Given that firm j plays $\hat{\sigma}$ at state $\omega = \varnothing$, the best response of firm i is:

$$\max_{\sigma \in [0,1]} \frac{\sigma \mu U^i_{\{i\}} + (1-\sigma)\lambda_L \Pi + \hat{\sigma}\mu U^i_{\{j\}} - c}{\sigma\mu + (1-\sigma)\lambda_L + \hat{\sigma}\mu + (1-\hat{\sigma})\lambda_L} \tag{C.1.3}$$

Taking the derivative of the previous objective function with respect to σ and using eq. C.1.2, we get

$$\frac{[\mu(\lambda_H - \lambda_L) - \lambda_L(\lambda_H + \lambda_L)] \cdot [c + \Pi(\lambda_L + \hat{\sigma}(\mu - \lambda_L)]}{(\lambda_H + \lambda_L)(\lambda_L(2 - \sigma - \hat{\sigma}) + \mu(\sigma + \hat{\sigma}))^2} \tag{C.1.4}$$

This is negative since $\eta < \underline{\eta}(\delta)$ implies, by Lemma C.1.1, $\mu < \underline{\mu} = \frac{\lambda_L(\lambda_L+\lambda_H)}{\lambda_H-\lambda_L}$. Therefore, the best action independently of the action of the opponent at state $\varnothing$ is to use the incumbent technology, given optimal continuation. $\square$

Lemma C.1.4. *Assume $\eta \in (\underline{\eta}(\delta), \overline{\eta}(\delta))$. Then the unique MPE involves $\sigma^A(\varnothing) = \sigma^B(\varnothing) = 1$.*

Proof of Lemma C.1.4. The problem of firm i at state $\varnothing$ is, as before,

$$\max_{\sigma\in[0,1]} \quad \frac{\sigma\mu U^i_{\{i\}} + (1-\sigma)\lambda_L\Pi + \hat{\sigma}\mu U^i_{\{j\}} - c}{\sigma\mu + (1-\sigma)\lambda_L + \hat{\sigma}\mu + (1-\hat{\sigma})\lambda_L}$$

By the assumption $\eta < \overline{\eta}(\delta)$ and Lemma C.1.2, the equations (C.1.2) also hold for this case. Thus, the derivative is the same as we obtained in eq. C.1.4. The difference is that now, since $\eta > \underline{\eta}(\delta)$), the derivative changes sign. Thus, independently of the action chosen by the opponent in state $\varnothing$, it is optimal to do research (given optimal continuation).

Thus, the strategy profile that both firms play fall-back strategies constitutes the unique equilibrium. $\square$

Lemma C.1.5. *Assume $\eta > \overline{\eta}(\delta)$. Then the unique MPE involves $\sigma^A(\varnothing) = \sigma^B(\varnothing) = 1$.*

Proof of Lemma C.1.5. By Lemma C.1.2, in any MPE the firms do research once the opponent obtains the new technology. Thus,

$$U^i_{\{i\}} = \frac{\lambda_H\Pi + \mu V_C - c}{\mu + \lambda_H} \qquad \text{and} \qquad U^i_{\{j\}} = \frac{\mu V_C - c}{\mu + \lambda_H}.$$

Using these in the problem of firm i in state $\varnothing$ (eq. C.1.3) and taking derivatives with respect to the choice variable σ we obtain:

$$\frac{\lambda_H\Pi(\mu(\lambda_H-\lambda_L)-\lambda_H\lambda_L)[\lambda_L+\hat{\sigma}(\mu-\lambda_L)]+c(\lambda_H^2(\mu-\lambda_L)+\mu^2(\lambda_H-\lambda_L)-3\lambda_L\lambda_H\mu)}{\lambda_H(\lambda_H+\mu)(\lambda_L(2-\sigma-\hat{\sigma})+\mu(\sigma+\hat{\sigma}))^2}$$

Note that the denominator is positive. The numerator can be rewritten as

$$\{\mu(\lambda_H-\lambda_L)-\lambda_H\lambda_L\}\cdot\{\lambda_H\Pi\cdot(\lambda_L+\hat{\sigma}(\mu-\lambda_L))+c\cdot(\mu+\lambda_H)\}$$

Which is positive since $\mu > \frac{2\lambda_H\lambda_L}{\lambda_H-\lambda_L} > \frac{\lambda_H\lambda_L}{\lambda_H-\lambda_L}$. Thus, the best response at state $\varnothing$, for every $\hat{\sigma}$, is to choose $\sigma = 1$. □

C.2 Private Information: Evolution of Beliefs

Proof of Lemma 4.4.1. Suppose that Firm j allocates σ_t^j attention to research conditional on not having found the good project so far. Finally, let $\Sigma_t = \int_0^t \sigma_s^j \, ds$.

Let p_t^i be the belief of Firm i that Firm j obtained the new technology by time t.

$$\begin{aligned}\frac{1-p_t}{p_t} &= \frac{\Pr(\text{no success in research and development})}{\Pr(\text{success in research but no success in development})} \\ &= \frac{e^{-\mu\Sigma_t}\cdot e^{-\lambda_L(t-\Sigma_t)}}{\int_0^t \mu\cdot\sigma_s\cdot e^{-\mu\Sigma_s}\cdot e^{-\lambda_L(s-\Sigma_s)}\cdot e^{-\lambda_H(t-s)}\,ds}\end{aligned} \tag{C.2.1}$$

Let U_t and R_t be the numerator and the denominator of the right-hand side of (C.2.1). Note that

$$\frac{\partial U_t}{\partial t} = -U_t\cdot\big((\mu-\lambda_L)\sigma_t^j+\lambda_L\big)$$

$$\frac{\partial R_t}{\partial t} = \mu\cdot\sigma_t^j\cdot U_t - \lambda_H\cdot R_t$$

By Differentiating (C.2.1) and multiplying R_t/U_t,

$$\begin{aligned}-\frac{\dot{p}_t^i}{(1-p_t^i)p_t^i} &= \frac{-U_t\cdot\big((\mu-\lambda_L)\sigma_t^j+\lambda_L\big)\cdot R_t-\mu\cdot\sigma_t^j\cdot U_t^2+\lambda_H\cdot R_t\cdot U_t}{R_t^2}\cdot\frac{R_t}{U_t} \\ &= -\big((\mu-\lambda_L)\sigma_t^j+\lambda_L\big)+\lambda_H-\mu\cdot\sigma_t^j(1-p_t^i)/p_t^i \\ &= \big\{\lambda_H-\lambda_L(1-\sigma_t^j)\big\}-\mu\cdot\sigma_t^j/p_t^i.\end{aligned}$$

By multiplying $-(1-p_t^i)p_t^i$, we have (4.4.1). $\square$

Proof of Lemma 4.4.2. By plugging $\sigma_t = 1$ to (4.4.1), we have $\dot{p}_t = (\mu - \lambda_H p_t)(1-p_t)$. By rearranging the differential equation, we can derive that

$$\lambda_H - \mu = (\lambda_H - \mu)\frac{\lambda_H \dot{p}_t}{(\mu - \lambda_H p_t)(\lambda_H - \lambda_H p_t)} = \frac{d}{dt}\log\left(\frac{\lambda_H - \lambda_H p_t}{\mu - \lambda_H p_t}\right).$$

Then, from $p_0 = 0$, we can derive that

$$\frac{\lambda_H(1-p_T)}{\mu - \lambda_H p_T} = \frac{\lambda_H}{\mu}e^{(\lambda_H - \mu)T}.$$

By rearranging the above equation, we have (4.4.2).

Observe that

$$\dot{q}_T = \frac{\mu(\lambda_H - \mu)^2 e^{(\lambda_H + \mu)T}}{(\lambda_H e^{\lambda_H T} - \mu e^{\mu T})^2} > 0.$$

Thus, q is increasing in T.

When $\mu > \lambda_H$,

$$\lim_{T\to\infty} q_T = \lim_{T\to\infty} \frac{\frac{1}{\lambda_H}\left(e^{(\lambda_H - \mu)T} - 1\right)}{\frac{1}{\mu}e^{(\lambda_H-\mu)T} - \frac{1}{\lambda_H}} = 1.$$

When $\mu < \lambda_H$,

$$\lim_{T\to\infty} q_T = \lim_{T\to\infty} \frac{\frac{1}{\lambda_H}\left(1 - e^{(\mu - \lambda_H)T}\right)}{\frac{1}{\mu} - \frac{1}{\lambda_H}e^{(\mu-\lambda_H)T}} = \frac{\mu}{\lambda_H}.$$

$\square$

C.3 Private Information: Equilibria

C.3.1 Cutoff Structure

Lemmas

Lemma C.3.1. *If the belief process $\{p_t\}$ is derived from a symmetric Markov strategy σ, then $\dot{p}_t \geqslant 0$ for all $t \geqslant 0$.*

Proof of Lemma C.3.1. Suppose that $\dot{p}_t < 0$ for some $t \geqslant 0$. Note that the belief is nonnegative. Then, $p_{t-\eta} > p_t \geqslant 0$ for a small $\eta > 0$. Also note that $p_0 = 0$ and $\dot{p}_0 = \mu \cdot \sigma_0 \geqslant 0$ by (4.4.1). By the continuity of p, there exists $t' < t$ such that $p_{t'} = p_t$ and $\dot{p}_{t'} \geqslant 0$. However, since the strategy is Markov, $\sigma_{t'} = \sigma_t$, which gives $\dot{p}_{t'} = \dot{p}_t$ and contradicts $\dot{p}_{t'} \geqslant 0 > \dot{p}_t$. Therefore, $\dot{p}_t \geqslant 0$ for all $t \geqslant 0$. □

Lemma C.3.2. *If σ constitutes a symmetric Markov equilibrium, $\sigma_t = 0$ for all $t \geqslant 0$ or $\sigma_t > 0$ for all $t \geqslant 0$.*

Proof of Lemma C.3.2. Consider the case with $p_t > 0$. If $\sigma_t = 0$, by (4.4.1), $\dot{p}_t = -(\lambda_H - \lambda_L)p_t(1-p_t)$. Since p_t cannot be equal to 1, $\dot{p}_t < 0$, which contradicts the previous result. Therefore, $\sigma_t > 0$ whenever $p_t > 0$. If $\sigma_0 = 0$, then $\dot{p}_0 = 0$ and the belief stays at 0. By the Markov property, $\sigma_t = 0$ for all $t \geqslant 0$. If $\sigma_0 > 0$, then $\dot{p}_0 > 0$ and $p_t > 0$ for a small enough $t > 0$. Then, $\sigma_t = 0$ will never be chosen, i.e., $\sigma_t > 0$ for all $t \geqslant 0$. □

Lemma C.3.3. *If σ constitutes a symmetric Markov equilibrium and $\sigma_S \in (0,1)$ for some $S \geqslant 0$, then $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$ for all $t \geqslant S$.*

Proof of Lemma C.3.3. By Lemma C.3.2, if $\sigma_S > 0$ for some $S \geqslant 0$, $\sigma_t > 0$ for all $t > 0$, thus, $\mu(V_t^1 - V_t^0) \geqslant \lambda_L(\Pi - V_t^0)$ for all $t \geqslant 0$.

Assume the contrary. Then, we can properly define $T \equiv \inf\{t > S \mid \mu(V_t^1 - V_t^0) > \lambda_L(\Pi - V_t^0)\}$. Then, $\mu(V_s^1 - V_s^0) = \lambda_L(\Pi - V_s^0)$ holds for $S \leqslant s \leqslant T$, and for some $\delta > 0$, $\mu(V_s^1 - V_s^0) > \lambda_L(\Pi - V_s^0)$ for all $T < s < T + \delta$.

When $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$, we show that $\mu(\dot{V}_t^1 - \dot{V}_t^0) \leqslant -\lambda_L \dot{V}_t^0$ if and only if

$$\lambda_H p_t + \lambda_L(1 - p_t)(1 - \sigma_t) \leqslant \frac{\mu(\lambda_H - \lambda_L)(\Pi - V_t^1) - \lambda_L c}{\lambda_L \Pi}.$$

First, by (HJB_1) and (HJB_0), we have

$$\mu \dot{V}_t^1 = \mu(\lambda_H + X_t)V_t^1 - \mu(\lambda_H \Pi - c), \tag{C.3.1}$$

$$(\mu - \lambda_L)\dot{V}_t^0 = (\mu - \lambda_L)X_t V_t^0 - (\mu - \lambda_L)(\mu(V_t^1 - V_t^0) - c), \tag{C.3.2}$$

where $X_t = \lambda_H p_t + \lambda_L(1 - p_t)(1 - \sigma_t)$. Also note that $(\mu - \lambda_L)(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^1)$. Then, $\mu(\dot{V}_t^1 - \dot{V}_t^0) \leqslant -\lambda_L \dot{V}_t^0$ is equivalent to:

$$\begin{aligned}
\lambda_L \Pi \cdot X_t &= \left\{\mu V_t^1 - (\mu - \lambda_L)V_t^0\right\} X_t \\
&\leqslant \mu(\lambda_H \Pi - c) - \mu \lambda_H V_t^1 + (\mu - \lambda_L)c - \mu(\mu - \lambda_L)(V_t^1 - V_t^0) \\
&= \mu(\lambda_H - \lambda_L)(\Pi - V_t^1) - \lambda_L c.
\end{aligned}$$

Let $\sigma_{T^-} := \lim_{t \to T^-} \sigma_t$ and $\sigma_{T^+} := \lim_{t \to T^+} \sigma_t$. Note that $\sigma_{T^+} = 1$. By the continuity of p and V^1, we have $p_{T^-} = p_{T^+} = p_T$ and $V_{T^-}^1 = V_{T^+}^1 = V_T^1$.

First, consider the case with $\sigma_{T^-} < 1$.[3] In this case, we have

$$\begin{aligned}
X_{T^+} =& \lambda_H p_T + \lambda_L(1 - p_T)(1 - \sigma_{T^+}) \\
&< \lambda_H p_T + \lambda_L(1 - p_T)(1 - \sigma_{T^-}) = \frac{\mu(\lambda_H - \lambda_L)(\Pi - V_T^1) - \lambda_L c}{\lambda_L \Pi}.
\end{aligned}$$

Then, $\mu(\dot{V}_{T^+}^1 - \dot{V}_{T^+}^0) < -\lambda_L \dot{V}_{T^+}^0$. Since $\mu(V_T^1 - V_T^0) = \lambda_L(\Pi - V_T^0)$, for small enough $\eta > 0$, $\lambda_L(\Pi - V_{T+\eta}^0) > \mu(V_{T+\eta}^1 - V_{T+\eta}^0)$ which contradicts $\sigma_{T^+} = 1$.

[3] It is possible that σ_{T^-} is not properly defined. Even in this case, the following argument still holds by considering a converging subsequence of $\{\sigma_t\}$.

Next, consider the case with $\sigma_{T-} = 1$. Note that $\mu(V_T^1 - V_T^0) = \lambda_L(\Pi - V_T^0)$ and $\mu(\dot{V}_T^1 - \dot{V}_T^0) = -\lambda_L \dot{V}_T^0$. If we show that $\mu(\ddot{V}_{T+}^1 - \ddot{V}_{T+}^0) < -\lambda_L \ddot{V}_{T+}^0$, it contradicts $\sigma_{T+} = 1$. Observe that $\sigma_{T+} = 1$ and $\dot{\sigma}_{T+} = 0$, thus, $\dot{X}_{T+} = \lambda_H \dot{p}_T$. By taking derivatives for (C.3.1) and (C.3.2), we can derive that

$$\begin{aligned}\mu\ddot{V}_{T+}^1 =& \mu\lambda_H\left[(1+p_T)\dot{V}_T^1 + \dot{p}_T V_T^1\right],\\ (\mu-\lambda_L)\ddot{V}_{T+}^1 =& (\mu-\lambda_L)\left[\lambda_H p_T \dot{V}_T^0 + \lambda_H \dot{p}_T V_T^0 - \mu(\dot{V}_T^1 - \dot{V}_T^0)\right].\end{aligned}$$

Then, we have

$$\begin{aligned}\mu\ddot{V}_{T+}^1 - (\mu-\lambda_L)\ddot{V}_{T+}^0 =& \lambda_H p_T\left\{\mu\dot{V}_T^1 - (\mu-\lambda_L)\dot{V}_T^0\right\} + \mu\lambda_H \dot{V}_T^1\\ &+ \lambda_H\dot{p}_T\left\{\mu V_T^1 - (\mu-\lambda_L)V_T^0\right\} + \mu(\mu-\lambda_L)(\dot{V}_T^1 - \dot{V}_T^0)\\ =& \mu(\lambda_H - \lambda_L)\dot{V}_T^1.\end{aligned} \tag{C.3.3}$$

Note that $(\lambda_H\Pi - c)/(\lambda_H + \lambda_H p_T)$ is the expected payoff of the firm (with the new technology) at time T if the opponent shuts down when it does not possess the new technology at time T. Then, in equilibrium, the expected payoff cannot exceed this value: $V_T^1 < (\lambda_H\Pi - c)/(\lambda_H + \lambda_H p_T)$. From (C.3.1), we have $\dot{V}_T^1 < 0$. By (C.3.3), $\mu(\ddot{V}_{T+}^1 - \ddot{V}_{T+}^0) < -\lambda_L\ddot{V}_{T+}^0$, which contradicts $\sigma_{T+} = 1$. $\square$

Lemma C.3.4. *Suppose that there exists $0 \leqslant T < \infty$ such that σ is a symmetric Markov equilibrium with $\sigma_t = 1$ for all $t < T$ and $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$ for all $t \geqslant T$. Then, for all $t \geqslant T$, $\sigma_t = \sigma^*$, $p_t = p^*$, $V_t^1 = V_1^*$, and $V_t^0 = V_0^*$, i.e., σ is a stationary fall-back equilibrium. In addition, $T = \frac{1}{\lambda_H - \mu}\log\left(\frac{\mu(1-p^*)}{\mu - \lambda_H p^*}\right)$.*

Proof of Lemma C.3.4. To have $0 < \sigma_t < 1$ for all $t \geqslant T$,

$$\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0). \tag{C.3.4}$$

By taking a derivative, we also have

$$\mu\dot{V}_t^1 = (\mu - \lambda_L)\dot{V}_t^0. \tag{C.3.5}$$

Define $X(p_t, \sigma_t) \equiv \lambda_H p_t + \lambda_L(1-p_t)(1-\sigma_t)$. By ($\mathrm{HJB}_1$), we have

$$\mu\dot{V}_t^1 = X(p_t,\sigma_t)\mu \cdot V_t^1 - \mu\lambda_H(\Pi - V_t^1) + \mu c. \tag{C.3.6}$$

By (HJB_0) and (C.3.4), we also have

$$\begin{aligned}(\mu-\lambda_L)\dot{V}_t^0 =& X(p_t,\sigma_t)(\mu-\lambda_L)V_t^0 - (\mu-\lambda_L)\mu(V_t^1 - V_t^0) + (\mu-\lambda_L)c \\ =& X(p_t,\sigma_t)(\mu V_t^1 - \lambda_L\Pi) - \mu\lambda_L(\Pi - V_t^1) + (\mu-\lambda_L)c.\end{aligned} \tag{C.3.7}$$

By using (C.3.5), (C.3.6) and (C.3.7), we have

$$X(p_t,\sigma_t) = \frac{\mu(\lambda_H-\lambda_L)(\Pi - V_t^1) - \lambda_L c}{\lambda_L\Pi}. \tag{C.3.8}$$

By plugging (C.3.8) into (HJB_1), we can derive that

$$\begin{aligned}0 =& \dot{V}_t^1 - \frac{1}{\lambda_L\Pi}\left\{\mu(\lambda_H-\lambda_L)(\Pi-V_t^1) - \lambda_L c\right\}V_t^1 + \lambda_H(\Pi - V_t^1) - c \\ =& \dot{V}_t^1 - \frac{\mu(\lambda_H-\lambda_L)}{\lambda_L\Pi}\left\{V_t^1 - \frac{\lambda_L(\lambda_H\Pi - c)}{\mu(\lambda_H-\lambda_L)}\right\}(\Pi - V_t^1) \\ =& \dot{V}_t^1 - \frac{\alpha}{\Pi-\beta}(V_t^1-\beta)(\Pi - V_t^1)\end{aligned} \tag{C.3.9}$$

where $\alpha \equiv \frac{\mu(\lambda_H-\lambda_L)(\Pi-\beta)}{\lambda_L\Pi}$ and $\beta \equiv \frac{\lambda_L(\lambda_H\Pi-c)}{\mu(\lambda_H-\lambda_L)}$. Note that

$$\Pi - \beta = \frac{\lambda_L\lambda_H}{\lambda_H-\lambda_L}\left[\left(\frac{1}{\lambda_L} - \frac{1}{\lambda_H} - \frac{1}{\mu}\right)\Pi + \frac{c}{\mu}\right] > 0.$$

Thus, α and β are strictly positive.

Also note that $\Pi - c/\lambda_H > V_t^1$ for all $t \geqslant 0$ since the firm's expected profit under the competition cannot exceed that without the competition. If $V_T^1 > \beta$, $V_t^1 > \beta$ for all $t \geqslant T$. If not, there exists $t > T$ such that $V_t > \beta$ and $\dot{V}_t = 0$, which contradict

(C.3.9). Likewise, if $V_T^1 < \beta$, $V_t^1 < \beta$ for all $t \geqslant T$. Now suppose that $V_T^1 \neq \beta$. By (C.3.9), we have

$$\alpha = \frac{(\Pi - \beta)}{(V_t^1 - \beta)(\Pi - V_t^1)}\dot{V}_t^1 = \frac{d}{dt}\log\left(\frac{|\beta - V_t^1|}{\Pi - V_t^1}\right)$$

By integrating the above equation side-by-side from T to t, we have

$$\alpha(t-T) = \log\left(\frac{|\beta - V_t^1|}{\Pi - V_t^1}\right) - \log\left(\frac{|\beta - V_T^1|}{\Pi - V_T^1}\right)$$

$$\iff \qquad \frac{|\beta - V_T^1|}{\Pi - V_T^1}e^{\alpha(t-T)} = \frac{|\beta - V_t^1|}{\Pi - V_t^1}.$$

Notice that the right-hand-side is bounded above and below since $V_t^1 < \Pi - c/\lambda_H$. The left-hand-side diverges to positive or negative infinite. Thus, it must be that $V_T^1 = \beta$.

In addition, by solving (C.3.9) with the initial condition $V_T^1 = \beta$, we have $V_t^1 = \beta = V_1^*$ for all $t \geqslant T$. By plugging $V_t^1 = \beta$ into (C.3.4), we also have

$$V_t^0 = \frac{\mu\beta - \lambda_L\Pi}{\mu - \lambda_L} = \frac{\lambda_L}{\mu - \lambda_L}\left(-\Pi + \frac{\lambda_H\Pi - c}{\lambda_H - \lambda_L}\right) = \frac{\lambda_L}{\mu - \lambda_L}\cdot\frac{\lambda_L\Pi - c}{\lambda_H - \lambda_L} = V_0^*$$

Next, by plugging $V_t^1 = \beta$ into (C.3.8), we have

$$\lambda_H p_t + \lambda_L(1-p_t)(1-\sigma_t) = \frac{\mu(\lambda_H - \lambda_L)(\Pi - \beta) - \lambda_L c}{\lambda_L\Pi} = \frac{\lambda_H - \lambda_L}{\lambda_L}\mu - \lambda_H,$$

or equivalently,

$$1 - \sigma_t = \frac{\left(\frac{\lambda_H - \lambda_L}{\lambda_L}\right)\mu - \lambda_H(1 + p_t)}{\lambda_L(1-p_t)}. \tag{C.3.10}$$

From (4.4.1), we have

$$\begin{aligned}
\dot{p}_t =&(1-p_t)\left[\mu-\lambda_H p_t-(1-\sigma_t)(\mu-\lambda_L p_t)\right] \\
=&(1-p_t)(\mu-\lambda_H p_t)-\left\{\left(\frac{\lambda_H-\lambda_L}{\lambda_L}-\lambda_H(1+p_t)\right)\mu\right\}\left(\frac{\mu}{\lambda_L}-p_t\right) \\
=&\frac{\mu(\lambda_H-\lambda_L)}{\lambda_L^2}(\underline{\mu}-\mu)+\left(\frac{\lambda_H-\lambda_L}{\lambda_L}\right)(2\mu-\overline{\mu})p_t \\
=&\frac{(\lambda_H-\lambda_L)(2\mu-\overline{\mu})}{\lambda_L}(p_t-p^*).
\end{aligned}$$

If $p_T \neq p^*$, then the solution of the above differential equation diverges and contradicts $0 \leqslant p_t \leqslant 1$ for all $t \geqslant T$. Therefore, $p_T = p^*$ and it also gives $p_t = p^*$ for all $t \geqslant T$. Also note that

$$1-p^*=\frac{(\overline{\mu}-\mu)(\mu-\lambda_L)}{\lambda_L(2\mu-\overline{\mu})}.$$

By plugging this into (C.3.10), for all $t \geqslant T$, we have

$$\begin{aligned}
\sigma_t =&1-\frac{\left(\frac{\lambda_H-\lambda_L}{\lambda_L}\right)\mu-2\lambda_H+\lambda_H(1-p^*)}{\lambda_L(1-p^*)} \\
=&1-\frac{\frac{\lambda_H-\lambda_L}{\lambda_L}(\mu-\overline{\mu})+\lambda_H\frac{(\overline{\mu}-\mu)(\mu-\lambda_L)}{\lambda_L(2\mu-\overline{\mu})}}{\frac{(\overline{\mu}-\mu)(\mu-\lambda_L)}{(2\mu-\overline{\mu})}} \\
=&1-\frac{-(\lambda_H-\lambda_L)(2\mu-\overline{\mu})+\lambda_H(\mu-\lambda_L)}{\lambda_L(\mu-\lambda_L)} \\
=&1-\frac{-(\lambda_H-\lambda_L)(\mu-\lambda_L)+\lambda_L(\mu+\lambda_L)}{\lambda_L(\mu-\lambda_L)}=\sigma^*.
\end{aligned}$$

Last, since $\sigma_t = 1$ for all $0 \leqslant t \leqslant T$, the belief that the opponent has the new technology at time t is $p^* = p_T = q_T$ where q is defined as in (4.4.2). By the definition of q, we can derive that $e^{(\lambda_H-\mu)T} = \frac{\mu(1-p^*)}{\mu-\lambda_H p^*}$, or equivalently, $T = \frac{1}{\lambda_H-\mu}\log\left(\frac{\mu(1-p^*)}{\mu-\lambda_H p^*}\right)$. □

Proof of Proposition 4.4.1 (a)

Proof of Proposition 4.4.1 (a). By Lemma C.3.1, σ satisfies $\sigma_t = 0$ for all $t \geqslant 0$ or $\sigma_t > 0$ for all $t \geqslant 0$. In the former case, we can set $T^* = 0$ and $\sigma^* = 0$, then we have $\sigma_t = \sigma^*$ for all $t > T^* = 0$. Now consider the latter case. If $\{t \geqslant 0 | \sigma_t \in (0,1)\}$ is empty, set $T^* = \infty$. Then, $\sigma_t = 1$ for all $t < T^*$. If $\{t \geqslant 0 | \sigma_t \in (0,1)\}$ is nonempty, we can properly define $T^* \equiv \inf\{t \geqslant 0 | \sigma_t \in (0,1)\} < \infty$. Then, $\sigma_t = 1$ for all $t < T^*$. In addition, by Lemma C.3.3, $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$ for all $t \geqslant T^*$. By Lemma C.3.4, $\sigma_t = \sigma^*$ for all $t \geqslant T^*$. □

C.3.2 Equilibrium Characterization

The Equilibrium with Incumbent Strategies

Lemma C.3.5. *Suppose that σ is the incumbent strategy, i.e., $\sigma_t = 0$ for all $t \geqslant 0$. Then, $\sigma^A = \sigma^B = \sigma$ constitutes a symmetric Markov equilibrium if and only if $\mu \leqslant \underline{\mu}$.*

Proof of Lemma C.3.5. Suppose that the incumbent strategy constitutes an equilibrium. Since neither firm conducts research, the belief that the other firm possesses the new technology is 0, i.e., $p_t = 0$ for all $t \geqslant 0$. Observe that $V_t^0 = \frac{\lambda_L \Pi - c}{2\lambda_L}$ since both firms develop with the incumbent technology. If a firm happens to have the new technology and the other firm sticks with the strategy, the expected payoff is $\frac{\lambda_H \Pi - c}{\lambda_H + \lambda_L}$, i.e., $V_t^1 = \frac{\lambda_H \Pi - c}{\lambda_H + \lambda_L}$ for all $t \geqslant 0$. To support this equilibrium, from (HJB$_0$), $\mu(V_t^1 - V_t^0) \leqslant \lambda_L(\Pi - V_t^0)$ needs to hold. By plugging V_t^1 and V_t^0 in, we have

$$\mu \left(\frac{\lambda_H \Pi - c}{\lambda_H + \lambda_L} - \frac{\lambda_L \Pi - c}{2\lambda_L} \right) \leqslant \lambda_L \left(\Pi - \frac{\lambda_L \Pi - c}{2\lambda_L} \right)$$

$$\iff \qquad \frac{\mu(\lambda_L \Pi + c)(\lambda_H - \lambda_L)}{2\lambda_L(\lambda_H + \lambda_L)} \leqslant \frac{\lambda_L \Pi + c}{2}$$

$$\iff \qquad \mu \leqslant \frac{\lambda_L(\lambda_H + \lambda_L)}{\mu(\lambda_H - \lambda_L)} = \underline{\mu}.$$

Now suppose that $\mu \leqslant \underline{\mu}$. By the above inequality, the strategy profile with $\sigma_t = 0$ for all $t \geqslant 0$ constitutes an equilibrium, i.e., the incumbent equilibrium exists. $\square$

The Equilibrium with Research Strategies

Lemma C.3.6. *Suppose that for some T, both firms play $\sigma_t = 1$ for all $0 \leqslant t \leqslant T$. Then, there exist $C_0, C_1 \in \mathbb{R}$ such that the expected payoffs of the firm with and without the new technology at time $t \in [0, T]$ is given as follows:*

$$V_t^1 = \overline{V}_1(q_t) + C_1 \cdot (1-q_t) \cdot \left(\frac{\mu - \lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}, \tag{C.3.11}$$

$$V_t^0 = \overline{V}_0(q_t) + \left(C_0\left(\frac{\mu}{\lambda_H} - q_t\right) - C_1\frac{\mu}{\lambda_H}\right) \cdot \left(\frac{\mu - \lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}.[4] \tag{C.3.12}$$

Moreover, if both firms play the research strategy ($T = \infty$), $C_1 = C_0 = 0$, i.e., $V_t^1 = \overline{V}_1(q_t)$ and $V_t^0 = \overline{V}_0(q_t)$.

Proof of Lemma C.3.6. Consider V_t^n as a value function with respect to the belief process q_t defined as in (4.4.2): $V_t^n = V_n(q_t)$. Note that $\dot{V}_t^n = V_n'(q_t)\dot{q}_t = V_n'(q_t)(\mu - \lambda_H q_t)(1-q_t)$. By plugging this into (HJB$_1$), we have

$$0 = V_1'(q)(\mu - \lambda_H q)(1-q) - \lambda_H(1+q)V_1(q) + \lambda_H \Pi - c. \tag{C.3.13}$$

By multiplying $(\mu - \lambda_H q)^{-\frac{2\mu}{\mu-\lambda_H}}(1-q)^{\frac{3\lambda_H-\mu}{\mu-\lambda_H}}$ and rearranging the equation, for all $0 = q_0 \leqslant q \leqslant q_T$, we can derive that

$$0 = \frac{d}{dq}\left[\frac{(1-q)^{\frac{2\lambda_H}{\mu-\lambda_H}}}{(\mu-\lambda_H q)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}}\left\{V_1(q) - \overline{V}_1(q)\right\}\right]. \tag{C.3.14}$$

[4] If $\mu = \lambda_H$, we need to replace $\left(\frac{\mu - \lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}$ to $e^{\frac{2}{1-q_t}}$.

Therefore, for all $0 = q_0 \leqslant q \leqslant q_T$, we have

$$V_1(q) = \overline{V}_1(q) + C_1 \cdot (1-q) \cdot \left(\frac{\mu - \lambda_H q}{1-q}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} \tag{C.3.15}$$

for some $C_1 \in \mathbb{R}$. By $V_t^1 = V_1(q_t)$ for all $0 \leqslant t \leqslant T$, (C.3.11) holds.

Next, plug $\dot{V}_t^0 = V_0'(q_t)(\mu - \lambda_H q_t)(1-q_t)$ into (HJB_0):

$$\begin{aligned} 0 =& V_0'(q)(\mu - \lambda_H q)(1-q) - \lambda_H q V_0(q) - c + \mu(V_1(q) - V_0(q)) \\ =& V_0'(q)(\mu - \lambda_H q)(1-q) - V_0(q)(\lambda_H q + \mu) - c \\ &+ \mu\left(\Pi - \frac{c}{\lambda_H}\right)\left(\frac{1}{2} + \frac{\lambda_H(1-q)}{2(\lambda_H+\mu)}\right) + \mu C_1 \cdot (1-q) \cdot \left(\frac{\mu - \lambda_H q}{1-q}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}. \end{aligned} \tag{C.3.16}$$

By multiplying $(1-q)^{\frac{2\lambda_H}{\mu-\lambda_H}}(\mu - \lambda_H q)^{-\frac{3\mu-\lambda_H}{\mu-\lambda_H}}$ and rearranging the equation, $0 \leqslant q \leqslant q_T$, we have

$$0 = \frac{d}{dq}\left[\frac{(1-q)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}}{(\mu-\lambda_H q)^{\frac{2\mu}{\mu-\lambda_H}}}\left\{V_0(q) - \overline{V}_0(q) + C_1 \cdot \frac{\mu}{\lambda_H} \cdot \left(\frac{\mu-\lambda_H q}{1-q}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}\right\}\right].$$

Therefore, we have

$$V_0(q) = \overline{V}_0(q) + \left(C_0\left(\frac{\mu}{\lambda_H} - q\right) - C_1\frac{\mu}{\lambda_H}\right) \cdot \left(\frac{\mu - \lambda_H q}{1-q}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}}. \tag{C.3.17}$$

for some $C_0 \in \mathbb{R}$. By $V_t^0 = V_0(q_t)$ for all $0 \leqslant t \leqslant T$, (C.3.12) holds.

Now suppose that both firms play research-first strategy. Then, (C.3.11) and (C.3.12) hold for all $t \geqslant 0$. When $\mu > \lambda_H$, by Lemma 4.4.2, $\lim_{t\to\infty} q_t = 1$. Since $\lim_{t\to\infty}(1-q_t)\left(\frac{\mu-\lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} = \infty$ and $\lim_{t\to\infty}\left(\frac{\mu-\lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} = \infty$, to make the value functions converge, $C_1 = C_0 = 0$. When $\mu < \lambda_H$, by Lemma 4.4.2, $\lim_{t\to\infty} q_t = \mu/\lambda_H$, which also implies $\lim_{t\to\infty}(1-q_t)\left(\frac{\mu-\lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} = \infty$ and $\lim_{t\to\infty}\left(\frac{\mu-\lambda_H q_t}{1-q_t}\right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} = \infty$. Likewise, we also have $C_1 = C_0 = 0$ in this case to make the value functions converge. □

Lemma C.3.7. *Suppose that σ is the research strategy, i.e., $\sigma_t = 1$ for all $t \geqslant 0$. Then, $\sigma^A = \sigma^B = \sigma$ constitutes a symmetric Markov equilibrium if and only if $\mu \geqslant \min\{\overline{\mu}, \hat{\mu}\}$.*

Proof of Lemma C.3.7. Suppose that both firms play the research-first strategy. By Lemma C.3.6, the expected payoffs at time t with and without the new technology are $V_t^1 = \overline{V}_1(q_t)$ and $V_t^0 = \overline{V}_0(q_t)$.

Both firms playing the research strategy constitutes an equilibrium if and only if $\mu(\overline{V}_1(q_t) - \overline{V}_0(q_t)) \geqslant \lambda_L(\Pi - \overline{V}_0(q_t))$ for all $t \geqslant 0$. Note that

$$\frac{d}{dq}\left[\mu(\overline{V}_1(q) - \overline{V}_0(q)) - \lambda_L(\Pi - \overline{V}_0(q))\right] = -\frac{\left(\Pi - \frac{c}{\mu} - \frac{c}{\lambda_H}\right) + \frac{c}{\lambda_L}}{2\left(\frac{\lambda_H + \mu}{\lambda_H \lambda_L}\right)} < 0.$$

Therefore, it is enough to check whether the following inequality holds:

$$\lim_{t\to\infty}\left[\mu(\overline{V}_1(q_t) - \overline{V}_0(q_t)) - \lambda_L(\Pi - \overline{V}_0(q_t))\right] \geqslant 0. \tag{C.3.18}$$

When $\mu \geqslant \lambda_H$, by $\lim_{t\to\infty} q_t = 1$, (C.3.18) is equivalent to

$$\mu(\overline{V}_1(1) - \overline{V}_0(1)) - \lambda_L(\Pi - \overline{V}_0(1)) = \frac{(\lambda_H\Pi + c)\mu\lambda_L\left(\mu - \overline{\mu}\right)}{2(\lambda_H + \mu)} \geqslant 0. \tag{C.3.19}$$

When $\lambda_H > \mu$, by $\lim_{t\to\infty} q_t = \mu/\lambda_H$, (C.3.18) is equivalent to

$$\begin{aligned} &\mu\left(\overline{V}_1\left(\mu/\lambda_H\right) - \overline{V}_0\left(\mu/\lambda_H\right)\right) - \lambda_L\left(\Pi - \overline{V}_0\left(\mu/\lambda_H\right)\right) \\ &= \frac{(\mu\Pi + c)\lambda_L\lambda_H\left((\lambda_H - 2\lambda_L)\mu - \lambda_L\lambda_H\right)}{2\mu(\lambda_H + \mu)} \geqslant 0. \end{aligned} \tag{C.3.20}$$

Observe that when $\lambda_H < 3\lambda_L$, $\lambda_H < \overline{\mu} < \hat{\mu}$. In this case, by (C.3.19), (C.3.18) holds iff $\mu \geqslant \overline{\mu} = \min\{\mu, \hat{\mu}\}$. When $\lambda_H \geqslant 3\lambda_L$, note that $\lambda_H \geqslant \overline{\mu} \geqslant \hat{\mu}$. If $\mu \geqslant \lambda_H \geqslant \overline{\mu}$, (C.3.18) holds by (C.3.19). If $\lambda_H > \mu \geqslant \hat{\mu}$, (C.3.18) holds by (C.3.20). Therefore, (C.3.18) holds iff $\mu \geqslant \hat{\mu} = \min\{\mu, \hat{\mu}\}$. $\square$

Lemma C.3.8. *Suppose that σ is a stationary fall-back strategy, i.e., for some $T \geqslant 0$ and $\sigma^* \in [0, 1)$, $\sigma_t = 1$ for all $t < T$ and $\sigma_t = \sigma^*$ for all $t > T$. If $\sigma^A = \sigma^B = \sigma$ constitutes a symmetric Markov equilibrium, then $\underline{\eta}(\delta) < \eta < \min\{1 + \delta, 2 - \delta\}$. Conversely, if $\underline{\eta}(\delta) < \eta < \min\{1 + \delta, 2 - \delta\}$, there exists a unique symmetric Markov equilibrium and it is a stationary fall-back strategy.*

Proof of Lemma C.3.8. Suppose that $\sigma^A = \sigma^B = \sigma$ constitutes an equilibrium. By Lemma 4.4.2 and C.3.4, $p^* = p_T = q_T > 0$. Since $2\mu > \overline{\mu}$, to have $p^* > 0$, $\mu > \underline{\mu}$ has to hold.

Next, we show that $\mu < \min\{\hat{\mu}, \overline{\mu}\}$. When $\lambda_H < \mu$,

$$\lim_{\tilde{T} \to \infty} q_{\tilde{T}} = 1 > q_T = p^*.$$

By using the definition of p^*, $\overline{\mu} = \underline{\mu} + \lambda_L$, $1 > p^*$ is equivalent to $(\overline{\mu} - \mu)(\mu - \lambda_L) > 0$, thus, $\overline{\mu} > \mu$. Then, $\overline{\mu} > \mu > \lambda_H$ is equivalent to $3\lambda_L > \lambda_H$, which implies $\hat{\mu} > \overline{\mu}$. Therefore, $\min\{\hat{\mu}, \overline{\mu}\} > \mu$ in this case. Consider the case with $\lambda_H > \mu$ and $3\lambda_L \geqslant \lambda_H$. In this case, $\hat{\mu} > \overline{\mu} \geqslant \lambda_H > \mu$, thus, $\min\{\hat{\mu}, \overline{\mu}\} > \mu$. Last, consider the case with $\lambda_H > \mu$ and $3\lambda_L < \lambda_H$. Then, we have $\overline{\mu} > \hat{\mu}$ and

$$\lim_{\tilde{T} \to \infty} q_{\tilde{T}} = \frac{\mu}{\lambda_H} > q_T = p^*.$$

By rearranging the inequality, we have $\lambda_L \lambda_H = \lambda_H \underline{\mu} - \lambda_L \overline{\mu} > (\lambda_H - 2\lambda_L)\mu$, which is equivalent to $\min\{\hat{\mu}, \overline{\mu}\} = \hat{\mu} > \mu$.

Now we assume that $\underline{\mu} < \mu < \min\{\overline{\mu}, \hat{\mu}\}$ and show that the stationary fall-back strategy defined in Lemma C.3.4 constitutes an equilibrium. By the construction of the strategy, for all $t \geqslant T$, $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$, which supports $\sigma_t \in (0, 1)$. Next, we need to show that $\mu(V_t^1 - V_t^0) \geqslant \lambda_L(\Pi - V_t^0)$ for all $0 \leqslant t < T$ to support $\sigma_t = 1$. Assume the contrary: $\mu(V_s^1 - V_s^0) < \lambda_L(\Pi - V_s^0)$ for some $0 \leqslant s < T$. Since

$\mu(V_T^1 - V_T^0) = \lambda_L(\Pi - V_T^0)$, there exists $s < t \leqslant T$ such that $\mu(V_t^1 - V_t^0) = \lambda_L(\Pi - V_t^0)$ and $\mu(\dot{V}_t^1 - \dot{V}_t^0) > -\lambda_L \dot{V}_t^0$, or equivalently,

$$\lambda_L \dot{V}_t^1 > (\mu - \lambda_L)(\dot{V}_t^0 - \dot{V}_t^1). \tag{C.3.21}$$

As a first step, we show that there exists $C_1 < 0$ such that V_t^1 is given as (C.3.11) in Lemma C.3.6 for all $0 \leqslant t < T$. By $V_T^1 = V_1^*$ and $q_T = p^*$, we have

$$C_1 = \frac{1}{(1-p^*)} \left(\frac{1-p^*}{\mu - \lambda_H p^*} \right)^{\frac{\mu+\lambda_H}{\mu-\lambda_H}} \left(V_1^* - \overline{V}_1(p^*) \right)$$

where $\overline{V}_1$ is defined as in (4.4.3). With some algebra and $\min\{\overline{\mu}, \hat{\mu}\} > \mu$, we can derive that

$$\overline{V}_1(p^*) - V_1^* = \left(\Pi - \frac{c}{\lambda_H} \right) \cdot \frac{(\overline{\mu} - \mu)^2 \left(\lambda_L \lambda_H - (\lambda_H - 2\lambda_L)\mu \right)}{2\lambda_L \lambda_H (\lambda_H + \lambda_L)(2\mu - \overline{\mu})} > 0. \tag{C.3.22}$$

Therefore, $C_1 < 0$. Then, for all $0 \leqslant t < T$, we have

$$\dot{V}_t^1 = \dot{q}_t \left[-\left(\Pi - \frac{c}{\lambda_H} \right) \frac{\lambda_H}{2(\lambda_H + \mu)} + C_1 \cdot \frac{\lambda_H (1+q_t)}{1-q_t} \left(\frac{1-q_t}{\mu - \lambda_H q_t} \right)^{\frac{2\lambda_H}{\lambda_H - \mu}} \right] < 0. \tag{C.3.23}$$

By (HJB_1) and (HJB_0), we have

$$\begin{aligned} \dot{V}_t^1 =& \lambda_H (1+q_t) V_t^1 + c - \lambda_H \Pi \\ \dot{V}_t^0 =& \lambda_H q_t V_t^0 + c - \mu(V_t^1 - V_t^0). \end{aligned}$$

By using $\lambda_L(\Pi - V_t^0) = \mu(V_t^1 - V_t^0)$, we can derive that

$$\begin{aligned} \dot{V}_t^0 - \dot{V}_t^1 =& \lambda_H (1+q_t)(V_t^0 - V_t^1) + \mu(V_t^0 - V_t^1) + \lambda_H(\Pi - V_t^0) \\ =& \left[(\lambda_H - \lambda_L)\mu - \lambda_H \lambda_L (1+q_t) \right] \left(\frac{\Pi - V_t^0}{\mu} \right). \end{aligned}$$

Note that $\Pi > V_t^0$ since the expected payoff cannot exceed the rent Π. By using $\Pi > V_t^0$, $p^* \geqslant q_t$ and $\min\{\overline{\mu}, \hat{\mu}\} > \mu$, we can derive that

$$\begin{aligned}\dot{V}_t^0 - \dot{V}_t^1 \geqslant & \left[(\lambda_H - \lambda_L)\mu - \lambda_H\lambda_L(1+p^*)\right]\left(\frac{\Pi - V_t^0}{\mu}\right) \\ = & \frac{(\overline{\mu} - \mu)\left(\lambda_L\lambda_H - (\lambda_H - 2\lambda_L)\mu\right)}{2\mu - \overline{\mu}}\left(\frac{\Pi - V_t^0}{\mu}\right) > 0.\end{aligned} \tag{C.3.24}$$

Then, (C.3.23) and (C.3.24) contradict (C.3.21). Therefore, $\mu(V_t^1 - V_t^0) \geqslant \lambda_L(\Pi - V_t^0)$ for all $0 \leqslant t < T$, and the stationary fall-back strategy constitutes an equilibrium.

□

Bibliography

Akcigit, U. and Liu, Q. (2016), "The role of information in innovation and competition," *Journal of the European Economic Association*, 14, 828–870.

Anton, J. J. and Yao, D. A. (2004), "Little patents and big secrets: managing intellectual property," *RAND Journal of Economics*, pp. 1–22.

Athey, S. and Levin, J. (2017), "The value of information in monotone decision problems," *Research in Economics*.

Bertsekas, D. P. (1995), *Dynamic programming and optimal control*, vol. 1, Athena scientific Belmont, MA.

Bhattacharya, S. and Guriev, S. (2006), "Patents vs. trade secrets: Knowledge licensing and spillover," *Journal of the European Economic Association*, 4, 1112–1147.

Bhattacharya, S., Glazer, J., and Sappington, D. E. (1992), "Licensing and the sharing of knowledge in research joint ventures," *Journal of Economic Theory*, 56, 43–69.

Biais, B., Mariotti, T., Rochet, J.-C., and Villeneuve, S. (2010), "Large risks, limited liability, and dynamic moral hazard," *Econometrica*, 78, 73–118.

Blackwell, D. (1951), "Comparison of Experiments," in *Second Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 93–102.

Blackwell, D. (1953), "Equivalent comparisons of experiments," *The annals of mathematical statistics*, 24, 265–272.

Bobtcheff, C., Bolte, J., and Mariotti, T. (2017), "Researcher's dilemma," *The Review of Economic Studies*, 84, 969–1014.

Bonatti, A. and Hörner, J. (2011), "Collaborating," *American Economic Review*, 101, 632–63.

Bonatti, A. and Hörner, J. (2017), "Career concerns with exponential learning," *Theoretical Economics*, 12, 425–475.

Bryan, K. A. and Lemus, J. (2017), "The direction of innovation," *Journal of Economic Theory*, 172, 247–272.

Bush, V. (1945), "Science: The Endless Frontier," *Transactions of the Kansas Academy of Science (1903-)*, 48, 231–264.

Cabrales, A., Gossner, O., and Serrano, R. (2013), "Entropy and the value of information for investors," *The American economic review*, 103, 360–377.

Capponi, A. and Frei, C. (2015), "Dynamic contracting: Accidents lead to nonlinear contracts," *SIAM Journal on Financial Mathematics*, 6, 959–983.

Carnehl, C. and Schneider, J. (2021), "On Risk and Time Pressure: When to Think and When to Do," *mimeo*.

Carnehl, C. and Schneider, J. (2022), "on Risk and Time Pressure: When to Think and When to Do," *Journal of the European Economic Association*.

Chen, M., Sun, P., and Xiao, Y. (2020), "Optimal monitoring schedule in dynamic contracts," *Operations Research*, 68, 1285–1314.

Chi, C. K. (2015), "The Value of Information and Dispersion," .

Coase, R. (1960), "The Problem of Social Cost," *The Journal of Law & Economics*, 3, 1–44.

Das, K. and Klein, N. (2020), "Do Stronger Patents Lead To Faster Innovation? The Effect Of Duplicative Search," .

d'Aspremont, C., Bhattacharya, S., and Gerard-Varet, L.-A. (2000), "Bargaining and sharing innovative knowledge," *The Review of Economic Studies*, 67, 255–271.

Denicolò, V. (2000), "Two-stage patent races and patent policy," *the RAND Journal of Economics*, pp. 488–501.

Denicolò, V. and Franzoni, L. A. (2004), "Patents, Secrets, and the First-Inventor Defense," *Journal of Economics & Management Strategy*, 13, 517–538.

Denicolò, V. and Franzoni, L. A. (2010), "On the winner-take-all principle in innovation races," *Journal of the European Economic Association*, 8, 1133–1158.

Dewatripont, M., Jewitt, I., and Tirole, J. (2000), "Multitask agency problems: Focus and task clustering," *European economic review*, 44, 869–877.

Di Tillio, A., Ottaviani, M., and Sørensen, P. N. (2021), "Strategic sample selection," *Econometrica*, 89, 911–953.

Dilmé, F. and Garrett, D. F. (2019), "Residual deterrence," *Journal of the European Economic Association*, 17, 1654–1686.

Durrett, R. (2019), *Probability: theory and examples*, vol. 49, Cambridge university press.

Gallini, N. T. (1992), "Patent policy and costly imitation," *The RAND Journal of Economics*, pp. 52–63.

Ganuza, J.-J. and Penalva, J. S. (2010), "Signal orderings based on dispersion and the supply of private information in auctions," *Econometrica*, 78, 1007–1030.

Gilbert, R. J. and Katz, M. L. (2011), "Efficient division of profits from complementary innovations," *International Journal of Industrial Organization*, 29, 443–454.

Golany, B. and Shtub, A. (2001), "Work breakdown structure," *Handbook of Industrial Engineering: Technology and Operations Management*, pp. 1263–1280.

Gossner, O. (2000), "Comparison of information structures," *Games and Economic Behavior*, 30, 44–63.

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

Green, B. and Taylor, C. R. (2016b), "Breakthroughs, deadlines, and self-reported progress: Contracting for multistage projects," *American Economic Review*, 106, 3660–99.

Green, B. and Taylor, C. R. (2016c), "On Breakthroughs, Deadlines, and the Nature of Progress: Contracting for Multistage Projects," Working Paper.

Green, J. R. and Scotchmer, S. (1995), "On the division of profit in sequential innovation," *The Rand journal of economics*, pp. 20–33.

Guesnerie, R. and Laffont, J.-J. (1984), "A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm," *Journal of Public Economics*, 25, 329–369.

Holmstrom, B. and Milgrom, P. (1991), "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design," *Journal of Law, Economics, & Organization*, 7, 24.

Hopenhayn, H. A. and Squintani, F. (2016), "Patent rights and innovation disclosure," *The Review of Economic Studies*, 83, 199–230.

Horstmann, I., MacDonald, G. M., and Slivinski, A. (1985), "Patents as information transfer mechanisms: To patent or (maybe) not to patent," *Journal of Political Economy*, 93, 837–858.

Hu, Z. (2014), "Financing Innovation with Unobserved Progress," .

Jewitt, I. (2007), "Information order in decision and agency problems," Nuffield College.

Karlin, S. and Rubin, H. (1956), "The theory of decision procedures for distributions with monotone likelihood ratio," *The Annals of Mathematical Statistics*, pp. 272–299.

Keller, G., Rady, S., and Cripps, M. (2005), "Strategic experimentation with exponential bandits," *Econometrica*, 73, 39–68.

Kim, S. K. (1995), "Efficiency of an Information System in an Agency Model," *Econometrica*, 63, 89–102.

Kim, Y. (2022), "Managing a Project by Splitting it into Pieces," Available at SSRN: https://ssrn.com/abstract=3450802.

Kultti, K., Takalo, T., and Toikka, J. (2007), "Secrecy versus patenting," *The RAND Journal of Economics*, 38, 22–42.

Kwon, I. (2012), "Patent races with secrecy," *The Journal of Industrial Economics*, 60, 499–516.

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

Lee, T. and Wilde, L. L. (1980), "Market structure and innovation: A reformulation," *The Quarterly Journal of Economics*, 94, 429–436.

Lehmann, E. (1988), "Comparing Location Experiments," *The Annals of Statistics*, pp. 521–533.

Levin, J. (2001), "Information and the Market for Lemons," *RAND Journal of Economics*, 32, 657–66.

Li, J. and Zhou, J. (2016), "Blackwell's informativeness ranking with uncertainty-averse preferences," *Games and Economic Behavior*, 96, 18–29.

Li, J. and Zhou, J. (2020), "Information order in monotone decision problems under uncertainty," *Journal of Economic Theory*, 187, 105012.

Lobel, O. (2013), "Filing for a patent versus keeping your invention a trade secret," *Harvard Business Review*, 21.

Loury, G. C. (1979), "Market structure and innovation," *The quarterly journal of economics*, pp. 395–410.

Manso, G. (2011), "Motivating innovation," *The Journal of Finance*, 66, 1823–1860.

Marinovic, I. and Szydlowski, M. (2019), "Monitoring with Career Concerns," *Available at SSRN 3450629*.

Marinovic, I. and Szydlowski, M. (2020), "Monitor reputation and transparency," *Available at SSRN 3703870*.

Maskin, E. and Riley, J. (1984), "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15, 171–196.

Mekonnen, T. and Leal Vizcaíno, R. (2021), "Bayesian Comparative Statics," *Available at SSRN 3227239*.

Milgrom, P. and Shannon, C. (1994), "Monotone comparative statics," *Econometrica*, 62, 157–180.

Milgrom, P. R. (1981), "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

Moroni, S. (2022), "Experimentation in organizations," *Theoretical Economics*, 17, 1403–1450.

Myerson, R. B. (2015), "Moral hazard in high office and the dynamics of aristocracy," *Econometrica*, 83, 2083–2126.

Neyman, A. (2017), "Continuous-time stochastic games," *Games and Economic Behavior*, pp. 92–130.

Organ, C. and Bottorff, C. (2022), "Work Breakdown Structure (WBS) in Project Management," *Forbes*.

Orlov, D. (2015), "Frequent monitoring in dynamic contracts," *Available at SSRN 2524628*.

Ottaviani, M. and Prat, A. (2001), "The Value of Public Information in Monopoly," *Econometrica*, 69, 1673–1683.

Persico, N. (2000), "Information acquisition in auctions," *Econometrica*, 68, 135–148.

Piskorski, T. and Westerfield, M. M. (2016), "Optimal dynamic contracts with moral hazard and costly monitoring," *Journal of Economic Theory*, 166, 242–281.

Poggi, F. (2021), "The Timing of Complementary Innovations," Working Paper.

Project Management Institute (2006), *Practice Standard for Work Breakdown Structures*.

Project Management Institute (2017), *A guide to the Project Management Body of Knowledge (PMBOK guide)*, Project Management Institute, 6th edn.

Quah, J. K.-H. and Strulovici, B. (2007), "Comparative Statics with the Interval Dominance Order: Some Extensions." .

Quah, J. K.-H. and Strulovici, B. (2009), "Comparative statics, informativeness, and the interval dominance order," *Econometrica*, 77, 1949–1992.

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

Roesler, A.-K. (2015), "Information Disclosure in Markets: Auctions, Contests, and Matching Markets," Tech. rep.

Scotchmer, S. and Green, J. (1990), "Novelty and disclosure in patent law," *The RAND Journal of Economics*, pp. 131–146.

Song, Y. and Zhao, M. (2021), "Dynamic R&D competition under uncertainty and strategic disclosure," *Journal of Economic Behavior & Organization*, 181, 169–210.

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

Spiegel, Y. (2008), "Licensing interim R&D knowledge," Tech. rep.

Sun, P. and Tian, F. (2017), "Optimal contract to induce continued effort," *Management Science*, 64, 4193–4217.

Szydlowski, M. (2019), "Incentives, project choice, and dynamic multitasking," *Theoretical Economics*, 14, 813–847.

Takalo, T. (1998), "Innovation and imitation under imperfect patent protection," *Journal of Economics*, 67, 229–241.

Templeton, B. (2019), "Elon Musk's war on LIDAR: who is right and why do they think that," *Forbes https://www. forbes. com/sites/bradtempleton/2019/05/06/elon-musks-war-on-lidar-who-is-right-and-why-do-they-think-that/7fe42c4f2a3b*.

Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton University Press.

Varas, F. (2017), "Managerial short-termism, turnover policy, and the dynamics of incentives," *The Review of Financial Studies*, 31, 3409–3451.

Varas, F., Marinovic, I., and Skrzypacz, A. (2020), "Random inspections and periodic reviews: Optimal dynamic monitoring," *The Review of Economic Studies*, 87, 2893–2937.

Wolf, C. (2018), "Informative milestones in experimentation," *University of Mannheim, Working Paper*.

Zhang, T. (2012), "Patenting in the shadow of independent discoveries by rivals," *International Journal of Industrial Organization*, 30, 41–49.