Managing a Project by Splitting it into Pieces

Yonggyun Kim*

Florida State University

ykim22@fsu.edu

August 28, 2024

Abstract

I study a dynamic principal-agent problem where there are two routes to complete 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.

JEL Classification: J41, L14, D86

Keywords: Dynamic Agency, Multiple Approaches, Poisson Arrival, Basic and Applied Research

*I am indebted to Fei Li and Curtis Taylor for their guidance and support in this project. I am also grateful to Attila Ambrus, Arjada Bardhi, Christoph Carnehl, Yeon-Koo Che, Jinwoo Kim, Pino Lopomo, Francisco Poggi, Ludvig Sinander, Huseyin Yildirim, anonymous referees and seminar participants at Duke, Seoul National University, and the Southern Economic Association conference for comments and suggestions. All remaining errors are my own.

1 Introduction

In project management, a work breakdown structure (WBS)—a step-by-step approach to completing projects—is widely used (Organ and Bottorff, 2022).[1] The applications of 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 to employing the WBS (e.g., clarifying goals, improving communication, etc.), a fundamental benefit is the ability to monitor progress. As a project is split into smaller components, a manager can better audit a subordinate's progress, potentially reducing the issue of moral hazard. However, decomposing a project into too small pieces can make the project rigid (Golany and Shtub, 2001). This rigidity may lead the manager to micromanage the project, which can slow down progress and generate inefficiencies. Thus, when a manager splits a project, she faces a trade-off between *monitoring* and *efficiency*.

In this article, I introduce a stylized model to study this tension between efficiency and monitoring in splitting a project. 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), which requires a breakthrough that 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, each arriving at a higher rate. 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, the reward to be paid upon success, and a termination policy. If the agent accepts the contract, then at each point in time, he chooses whether to work on the specified approach or shirk

[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.

for private benefit. The principal does not observe the agent's effort, which creates a moral hazard problem.

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, which allows the contract to be extended upon subproject completion. This feature provides a monitoring advantage, as it enables the principal to better oversee progress. The second assumption is that the direct approach is more efficient than the sequential one. Thus, while the sequential approach has the advantage of better monitoring the agent, it may have a disadvantage in efficiency compared to the direct approach.

In addition to these elements, a third important economic factor affects the choice of methodology: the *deadline effect*. The principal needs to impose a deadline because, without one, the agent could (and likely would) shirk indefinitely, never completing the project. Therefore, the deadline is essential to overcoming moral hazard. Furthermore, the deadline is imposed by the principal, i.e., it is endogenously determined. Consequently, the optimal contract is shaped by the interplay of these three factors: monitoring, efficiency, and endogenous deadlines.

I first set aside the efficiency concern—considering the case where both approaches are equally efficient—to focus on the interaction between monitoring and deadline effects. In this context, I show that the optimal contract involves at most one switch between the approaches. Specifically, the principal initially employs the sequential approach up to an intermediate deadline. If the subproject is completed by then, the principal extends the deadline; otherwise, the principal switches to the direct approach until the final deadline.

Intuitively, near the deadline, the direct approach has a comparative advantage because it requires only one breakthrough, while the sequential approach requires two. In contrast, when the deadline is distant, the sequential approach is more appealing due to the possibility of extending the deadline, which arises from monitoring. This suggests that the principal would prefer the sequential approach when the deadline is far off and the direct approach when it is imminent. However, the monitoring advantage might be so significant that the principal could opt for the sequential approach even close to the deadline: the intermediate deadline equals to the final deadline. Conversely, the optimal deadline might be short enough that the principal would prefer the direct approach right from the outset: the intermediate deadline is set to 0.

I show that the optimal contract crucially depends on the project's return—the gross value to the principal from completing the project, given its operating cost. When the project return is low, the optimal deadline is short, and the principal opts for the direct approach. When the project return is high, monitoring is highly advantageous, leading the principal to choose the sequential approach. In intermediate cases, there is a switch from the sequential approach to the direct approach (Theorem 2).

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 (Theorem 3). 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 (Theorem 4).

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).[3] 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.[4]

The remainder of this article is organized as follows. Related literature is discussed below. Section 2 introduces the basic setup of the model and analyzes the first-best case. Section 3 provides a planner's problem with deadlines. Then, Section 4 and 5 characterize the optimal contracts for the cases with and without the efficiency losses from splitting the project. Section 6 concludes. 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, 2016b). 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

[3] 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.

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

approaches. Moreover, this choice problem arises at every point in time. Therefore, the principal's problem becomes more complex from a dynamic perspective.

An article that has a similar flavor is Carnehl and Schneider (2023). 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.[5]

This article is also 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); Wong (2023). 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

[5] However, the economic forces that drive these results 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.

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. The most relevant paper in this line is Kim and Poggi (2024), which introduces an innovation race model with two R&D routes: one requiring a single breakthrough (direct development) and the other requiring two breakthroughs (research and development). However, to my knowledge, most studies in this literature focus on the problems involving competing firms or a single decision maker, whereas this article addresses an agency problem.

[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.

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).

2 Model

Preliminaries 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 $\Pi > 0$, referred to as 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; that is, 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. Both players have outside options of zero.

Paths toward project completion There are two routes to completing the project. The first is to tackle the project directly, namely the *direct approach*. The second is to break the main project into two subprojects, which I term the *sequential approach*. Completing the first subproject does not provide any independent value for the principal or the agent. 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.

Contracts and arrival rates At time 0, the principal offers the agent a contract consisting of (i) the deadlines at which the project is terminated; (ii) the reward schedules upon project completion; (iii) the approaches to be taken;

[7] See Remark 1 for further discussion of limited liability.

and (iv) the agent's recommended effort. The principal can fully commit to these contractual terms. See Section OA.1 for the formal definition of the contract.

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 ($1 - \tilde{b}_t$).[8] 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 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 $\lambda_S \tilde{b}_t$.

No-deadline benchmark As a first benchmark, I consider the problem of a social planner who is able to observe the agent's effort, directs which approach to take, and faces no exogenous deadline. This scenario represents the first-best situation where the planner has perfect information, control over the agent's actions, and no deadline constraints. Since the benefit from shirking is less than the flow cost, it is optimal for the planner to make the agent work in this case.

To determine which approach is more efficient, I compare the expected surpluses of each approach. Since there is no deadline and the agent never shirks, the project will surely be completed, and the planner will receive Π (recall that neither the principal nor the agent discounts the future). If the planner

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

adopts the direct approach indefinitely, i.e., until the project is completed, the expected duration of the project is $\frac{1}{\lambda_D}$, resulting in the expected cost is $\frac{c}{\lambda_D}$. Similarly, if the planner employs the sequential approach indefinitely, the expected duration of each breakthrough is $\frac{1}{\lambda_S}$, thus, the total expected duration is $\frac{2}{\lambda_S}$, and the expected cost is $\frac{2c}{\lambda_S}$. Therefore, the expected surpluses of employing each approach are derived as follows.

$$\text{Direct approach}: \ \Pi - \frac{c}{\lambda_D}, \qquad \text{Sequential approach}: \ \Pi - \frac{2c}{\lambda_S}.$$

Parametric assumptions I begin the analysis by focusing on the case where there is no efficiency loss from splitting the project: $2\lambda_D = \lambda_S$. Then, in Section 5, I consider the case where splitting the project harms the efficiency: $2\lambda_D > \lambda_S$. Additionally, I assume that the project is profitable enough: $\Pi > \frac{2c}{\lambda_S} \geq \frac{c}{\lambda_D}$.

Remark 1. 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 $\frac{\phi}{\lambda_D} + \epsilon$ to the agent for some small $\epsilon > 0$. Observe that the agent will always work because the instantaneous payoff from working is $\lambda_D(\frac{\phi}{\lambda_D} + \epsilon) - \phi$ whereas that from shirking is zero. Then, the agent's expected payoff is ϵ, and the principal's expected payoff is $\Pi - \frac{c}{\lambda_D} - \epsilon$. By sending ϵ to zero, the principal is able to achieve the first-best outcome.

3 Planner's Problem with Deadlines

As an intermediate step toward characterizing the optimal contract, I consider a problem where a planner faces exogenous deadlines. These deadlines generate inefficiencies, as Π is assumed to be large enough that continuing the project is more beneficial in expectation than terminating it.

Assume that the project is terminated when time passes a deadline T. Additionally, the deadline is extended by $\Delta \geq 0$ when the subproject is completed.[9] In other words, the planner faces two exogenous deadlines: the original deadline T under no subproject completion, and the extended deadline $T+\Delta$ under a subproject completion. Given these deadlines, the planner chooses which approach to take at each point in time.

I begin by introducing a benchmark policy where the planner initially employs the sequential approach and later switches to the direct approach.

Definition 3.1. A policy is called a *one-switch policy* if there exists an intermediate deadline $S \in [0, T]$ such that (i) the sequential approach is employed up to S, (ii) if the subproject is completed before S, the planner continues working on the remaining subproject until the extended deadline $T + \Delta$, and (iii) if the subproject is not completed by S, the planner switches to the direct approach until the deadline T.

This class of policies includes two extreme cases. Observe that a one-switch policy with $S = 0$ does not involve any sequential approach, and is referred to as the *direct-only policy.* Conversely, a one-switch policy with $S = T$ does not involve any direct approach, and is referred to as the *sequential-only policy.*

The following theorem shows that the optimal policy takes the form of a one-switch policy.

Theorem 1. *Suppose that there is no efficiency loss from splitting the project ($\lambda_S = 2\lambda_D$). When the planner faces the deadline T and the extension Δ resulting from the subproject completion, the optimal policy is characterized as follows:*

(a) ***(Long extension)*** *if $\Delta \geq \bar{\Delta} \equiv \frac{1}{\lambda_S} \log\left[\frac{\lambda_S \Pi - c}{(\lambda_S - \lambda_D)\Pi - c}\right]$, the sequential-only policy is optimal;*

(b) ***(Short extension)*** *if $\Delta < \bar{\Delta}$, there exists $\hat{T} > 0$ such that*

[9] It is possible to consider the case where the deadline is shortened after the subproject completion. However, in the context of the analysis focusing on the agency problem, I specifically examine the scenario of deadline extension because a shortened deadline would disincentivize the agent from working.

(i) when $T < \hat{T}$, the direct-only policy is optimal;

(ii) when $T > \hat{T}$, the one-switch policy with the intermediate deadline $T - \hat{T}$ is optimal.

This theorem implies that the direct approach is never employed when the deadline extension is sufficiently long. On the other hand, when the deadline extension is relatively short, there exists a time $\hat{T}$ such that the direct approach begins to be employed when fewer than $\hat{T}$ units of time remain.

Direct-only vs. sequential-only policies To provide an intuition for this theorem, I compare the expected surpluses of the project under the direct-only policy (d) and the sequential-only policy (s). Note that the expected surplus, $\mathcal{W}^i(T, \Delta)$ for $i \in \{d, s\}$, can be expressed as follows:

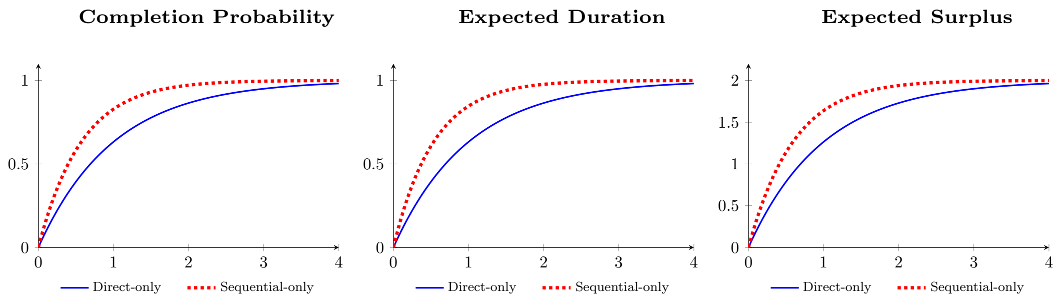
$$\mathcal{W}^i(T, \Delta) = \Pi \cdot \mathcal{P}^i(T, \Delta) - c \cdot \mathcal{D}^i(T, \Delta),$$

where $\mathcal{P}^i(T, \Delta)$ is the probability that the project is completed by the deadline—namely, the completion probability—and $\mathcal{D}^i_{T,\Delta}$ is the expected duration of the project. The formal definitions of each term in the above equation are presented in Appendix A.1.

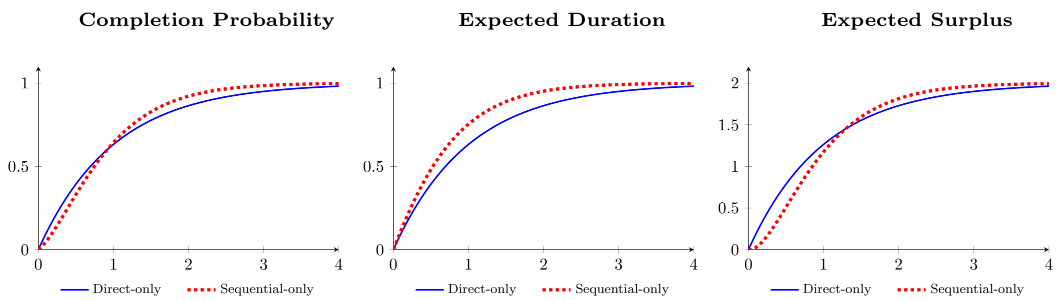
I begin by comparing the completion probabilities across policies when there is no deadline extension ($\Delta = 0$). The following proposition shows that $\mathcal{P}^d(T, 0)$ and $\mathcal{P}^s(T, 0)$ cross once as T increases. The proof is in Appendix A.2.

Proposition 3.1. *Suppose that $\Delta = 0$. There exists $\check{T}$ such that $\mathcal{P}^d(T, 0) > \mathcal{P}^s(T, 0)$ for all $T < \check{T}$ and $\mathcal{P}^s(T, 0) > \mathcal{P}^d(T, 0)$ for all $T > \check{T}$.*

Intuitively, under the direct-only policy, the planner needs only one breakthrough, whereas the sequential-only policy requires two breakthroughs, which is challenging within a short timeframe. Therefore, when the deadline is short, the completion probability under the direct-only policy is higher than that under the sequential-only policy. On the other hand, when the deadline is relatively long, achieving two faster breakthroughs can be easier than achieving



(a) Long Deadline Extension: $\Delta = 1$, $\lambda_D = 1$, $\lambda_S = 2$, $\Pi = 3$, $c = 1$



(b) Short Deadline Extension: $\Delta = .1$, $\lambda_D = 1$, $\lambda_S = 2$, $\Pi = 3$, $c = 1$

Figure 1: Completion Probabilities, Expected Durations, and Expected Surpluses

one slower breakthrough. In this case, the sequential-only policy may have a higher chance of project completion than the direct-only policy. I call this dynamic the 'deadline effect.'

Next, observe that the deadline extension provides additional benefits to the sequential-only policy: $\mathcal{P}^s(T, \Delta)$ increases as Δ increases. I refer to this as the 'monitoring effect' because this increase in the completion probability occurs due to the planner's ability to observe intermediate progress under the sequential approach.

Based on these deadline and monitoring effects, I can infer that the direct-only policy has a higher completion probability than the sequential-only policy when both the original deadline and the deadline extension are short. The two leftmost figures in Figure 1 illustrate this. The horizontal axis represents the deadline T. With the long deadline extension, as depicted in Figure 1a, the sequential-only policy always have higher completion probabilities—the monitoring effect outweighs the deadline effect. When the extension is short, as depicted in Figure 1b, the direct-only policy has higher completion probabilities with short deadlines, whereas the sequential-only policy has higher completion probabilities with long deadlines.

These effects play crucial roles in comparing expected surpluses between the two policies. When computing the expected surpluses, the expected duration—illustrated in the middle figures of Figure 1—needs to be considered: the sequential-only policy always has a longer expected duration due to the possibility of an extension. Nevertheless, the rightmost figures in Figure 1 demonstrate that the expected surpluses share a key characteristic with the completion probabilities: the direct-only policy can achieve higher expected surpluses only when both the deadline and the extension are short.

In light of this intuition, we can guess that the one-switch policy—employing (i) the sequential approach when the deadline is distant and (ii) the direct approach when the deadline is near—may be optimal. Theorem 1 verifies that it indeed is.

Comparative statics of the extension cutoff Theorem 1 shows that there exists a cutoff for the deadline extension, $\overline{\Delta}$, such that the sequential-only policy is optimal if the extension is longer than the cutoff; otherwise, the direct approach is employed near the deadline. The following proposition demonstrates how this cutoff changes with respect to the project payoff Π.

Proposition 3.2. *Suppose that $\lambda_S = 2\lambda_D$ and $\Pi > \frac{c}{\lambda_D}$. The extension cutoff $\overline{\Delta}$ decreases in the project payoff Π.*

Proof of Proposition 3.2. Note that

$$\frac{\partial\overline{\Delta}}{\partial\Pi} = -\frac{\lambda_D c}{\lambda_S(\lambda_S\Pi - c)((\lambda_S - \lambda_D)\Pi - c)}.$$

From $\lambda_S = 2\lambda_D$ and $\Pi > \frac{c}{\lambda_D}$, the above equation is negative. $\square$

Intuitively, as Π increases, the completion probability becomes a more influential factor than the expected duration in deriving the optimal policy. For example, in Figure 1a, the sequential-only policy has a higher completion probability than the direction-only policy, but it also has a longer expected duration, resulting in higher costs. In the figure, Π is chosen to be large enough that the expected surplus under the sequential-only policy is higher than that under the direct-only policy, even near the deadline. However, if Π were relatively low, the higher expected duration could cause the expected surplus under the sequential-only policy to be lower than that under the direct-only policy near the deadline.

4 Optimal Contract Derivation

In this section, I characterize the optimal contract in the case where there is no efficiency loss from splitting the project. As in the tangible progress case in Green and Taylor (2016b), the optimal contract can be implemented with three key properties: (i) the contract is terminated after a deadline; (i) the reward for the project completion, R_t, linearly diminishes over time; and (ii) the deadline is extended by $\frac{1}{\lambda_S}$ upon subproject completion.[10]

A distinctive feature of this model is that the principal must decide 'which path to take.' Since the contract involves a deadline and an extension of it, based on the result of the previous section, we can naturally infer that the optimal choice of approaches over time may involve one switch from the sequential approach to the direct approach—or no switch at all. In light of these insights, I define contracts involving the above characteristics.

Definition 4.1. A contract is called a *one-switch contract* with a final deadline T and an intermediate deadline $S \in (0, T)$ if

(i) the sequential approach is employed up to S,

[10] The details of these properties will be addressed in Section 4.2.

(ii) when the subproject is completed before S, the contract is extended by $\frac{1}{\lambda_S}$ and the reward upon project completion at time t is $R_t^S \equiv \phi(T - t + \frac{2}{\lambda_S})$,

(iii) if the subproject is not completed by S, the direct approach is employed up to T and the reward upon project completion at time t is $R_t^D \equiv \phi(T - t + \frac{1}{\lambda_D})$, and

(iv) the contract is terminated if the project is not completed by T.

When $S = T$, we call the contract a *sequential-only contract*, and when $S = 0$, we call the contract a *direct-only contract*.

The following theorem shows that the optimal contract takes one of the above forms.

Theorem 2. *Suppose that there is no efficiency loss from splitting the project ($\lambda_S = 2\lambda_D$). There exist thresholds Π_F, Π_S and Π_D such that $\Pi_S > \Pi_D > \Pi_F \equiv \frac{c+\phi}{\lambda_D}$ and the optimal contract can be implemented as follows:*

(a) when $\Pi > \Pi_S$, a sequential-only contract is optimal;

(b) when $\Pi_S > \Pi > \Pi_D$, there exists a one-switch contract that is optimal;

(c) when $\Pi_D > \Pi > \Pi_F$, a direct-only contract is optimal; and

(d) when $\Pi < \Pi_F$, the project is infeasible.

As discussed in Proposition 3.2, the direct approach is preferred when Π is lower and the sequential approach is preferred when Π is higher. The above theorem aligns with that intuition. In the subsequent subsections, I provide the details of the derivation of this result.

4.1 Promised Utility and Incentive Compatibility

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 (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 payoff $P_0(\Gamma)$ subject to the promise-keeping constraint $U_0(\Gamma) = u$ and the incentive compatibility condition, which will be demonstrated later in this subsection. 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

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

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 $V(u)$.

Promised utility upon subproject completion I begin by considering the principal's problem, given that the subproject is completed at time t. Let u_S^t denote the agent's promised utility, which will be considered as a state variable. Since this case only requires one more breakthrough, it is identical to the single-stage benchmark of Green and Taylor (2016a). They show that the optimal contract is to impose a deadline $t + \frac{u_S^t}{\phi}$ and a linearly diminishing reward schedule $\{\check{R}_s^t\}_{t \leq s \leq t + u_S^t/\phi}$ where

$$\check{R}_s^t = u_S^t + \frac{\phi}{\lambda_S} - \phi(s - t) \tag{4.1}$$

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

Incentive compatibility conditions 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. From this, we can derive the incentive compatibility constraint under the direct approach ($a_t = 1$):

$$R_t \geq u_t + \frac{\phi}{\lambda_D}. \tag{IC$_1$}$$

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, the incentive compatibility constraint under the sequential approach ($a_t = 0$) is

$$u_S^t \geq u_t + \frac{\phi}{\lambda_S}. \tag{IC$_0$}$$

4.2 Value Function Characterization

In this subsection, I characterize the value function of the principal. A natural conjecture is that the principal's expected payoff is maximized when incentive compatibility conditions bind. I outline key properties of contracts make IC conditions bind, and then characterize the value function.

Deadline and extension With binding IC conditions, the agent's promised utilities should be consumed at the same rate with the benefit from shirking: $\frac{du}{dt} = \dot{u}_t = -\phi$, or equivalently, $u_t = u_0 - \phi t$. If the completion has not been made by $\frac{u_0}{\phi}$, 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 $\frac{u_0}{\phi}$.

When the sequential approach is chosen, to make (IC$_0$) bind, we have

$$t + \frac{u_S^t}{\phi} = t + \frac{u_t}{\phi} + \frac{1}{\lambda_S} = \frac{u_0}{\phi} + \frac{1}{\lambda_S}.$$

It implies that upon the subproject completion, the updated deadline $t + \frac{u_S^t}{\phi}$ extends the original deadline $\frac{u_0}{\phi}$ by $\frac{1}{\lambda_S}$.

Linearly diminishing rewards Let T denote the deadline $\frac{u_0}{\phi}$. By using $u_t = u_0 - \phi t = \phi(T-t)$, to make (IC$_1$) bind, the reward of completion at time t via the direct approach is

$$R_t = u_t + \frac{\phi}{\lambda_D} = \phi\left(T - t + \frac{1}{\lambda_D}\right),$$

which corresponds to R_t^D in Definition 4.1.

Next, when the subproject is completed at $\check{t}$, to make (IC$_0$) bind, I have $u_S^{\check{t}} = u_{\check{t}} + \frac{\phi}{\lambda_S}$. Then, by (4.1), the reward of completion at time $t \in [\check{t}, T + \frac{1}{\lambda_S}]$ via the sequential approach is

$$\check{R}_t^{\check{t}} = u_{\check{t}} + \frac{\phi}{\lambda_S} + \frac{\phi}{\lambda_S} - \phi(t - \check{t}) = \phi\left(T - t + \frac{2}{\lambda_S}\right),$$

which corresponds to R_t^S in Definition 4.1.

Value Function Based on the above observations, I infer that the value function is linked to the planner's problem with a deadline and its extension, as explored in the previous section. Let $\mathcal{W}^*(T,\Delta)$ denote the optimal expected surplus under the deadline T and extension Δ, derived from the optimal policy in Theorem 1. Then, when the agent's promised utility is u, a conjecture for the principal's value function is the expected surplus from the deadline $\frac{u}{\phi}$ and the extension $\frac{1}{\lambda_S}$, $\mathcal{W}^*(\frac{u}{\phi}, \frac{1}{\lambda_S})$, net of u.

The following proposition verifies this conjecture, with the proof provided in Online Appendix OA.2.

Proposition 4.1. *The principal's value function $\mathcal{V}$ is characterized as follows:*

$$\mathcal{V}(u) = \mathcal{W}^* \left(\tfrac{u}{\phi}, \tfrac{1}{\lambda_S} \right) - u. \tag{4.2}$$

Moreover, $\mathcal{V}$ is concave.

A key step in proving this proposition is finding a contract implementing the pair of the agent's promised utility, u, and the principal's expected payoff, $V(u)$. In terms of choosing which approach to take, the principal's incentives are perfectly aligned with those of the planner—who faces the same deadlines as the principal—in that both want to maximize the expected surplus. Since the planner's policy with at most one switch is optimal, I show that the principal can implement the pair using a contract that involves at most one switch with linearly diminishing rewards (Online Appendix OA.2.2).

4.3 Proof of Theorem 2

Now that I have characterized the value function, the next step is to pin down the optimal initial promised utility level, $\bar{u}$, which is the solution to (MP). This will establish the starting point of the contract in Figure 2 and determine the deadline length, $\frac{\bar{u}}{\phi}$. This can be interpreted as the principal endogenously imposing the deadline to overcome the agent's moral hazard.

Recall that the direct approach is never employed when the extension is greater than $\overline{\Delta}$ (Theorem 1 (a)) and the cutoff is decreasing in Π (Proposition 3.2). Let Π_S be the solution of $\overline{\Delta} = \frac{1}{\lambda_S}$. Then, for all $\Pi > \Pi_S$, the direct approach will not be employed, even near the deadline. This establishes Theorem 2 (a) and is illustrated in Figure 2 (c).

When $\Pi < \Pi_S$, Theorem 1 (b) indicates that the optimal approach is switched from the sequential approach to the direct approach when $\hat{T}$ units of time remain. In terms of the promised utility, the switch happens at $\hat{u}_1 \equiv \phi\hat{T}$. Then, the form of the optimal contract depends on whether $\bar{u}$ is greater than $\hat{u}_1$ or not. For example, the value functions in Figure 2a and 2b both involve a switching point $\hat{u}_1$, however, $\bar{u}$ is greater than $\hat{u}_1$ in Figure 2a and less than

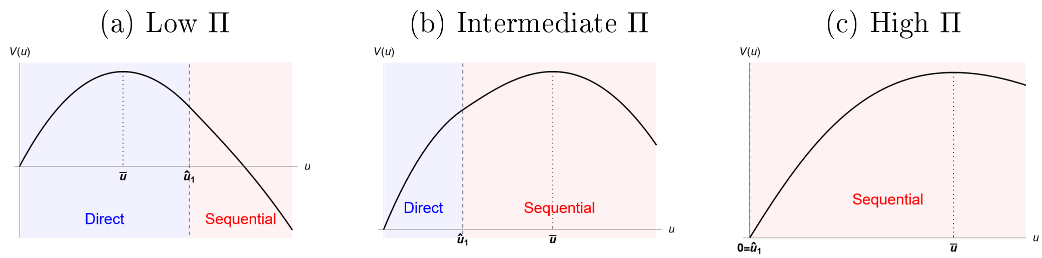


Figure 2: Value functions when there is no efficiency loss

$\hat{u}_1$ in Figure 2b. Thus, the optimal contracts are a direct-only contract in Figure 2a and a contract with a switch from the sequential approach to the direct approach in Figure 2b. By choosing Π_D such that $\overline{u}$ and $\hat{u}_1$ same, we see that the optimal contract involves a switch when $\Pi \in (\Pi_D, \Pi_S)$, establishing Theorem 2 (b). Conversely, when $\Pi < \Pi_D$, the direct approach will only appear in the optimal contract.

Last, the feasibility of the project depends on whether $\bar{u}$ is greater than 0 or not. When $\bar{u}$ is equal to zero, the principal's expected payoff is maximized at $u = 0$, meaning it is optimal for the principal not to initiate the contract—the project is infeasible. This occurs when the principal's flow profit is negative near the deadline T. Since the promised utility u is close to zero near the deadline, the reward R is approximately $\frac{\phi}{\lambda_D}$. Then, the principal's flow profit in $[T - dt, T]$ is approximately

$$\lambda_D dt \cdot \left(\Pi - \frac{\phi}{\lambda_D}\right) - cdt = \lambda \left(\Pi - \frac{\phi + c}{\lambda_D}\right) dt.$$

Therefore, the project is feasible if and only if Π is greater than $\Pi_F \equiv \frac{c+\phi}{\lambda_D}$. I show that $\Pi_D \in (\Pi_F, \Pi_S)$ (Lemma OA.2.7). Then, when $\Pi \in (\Pi_F, \Pi_D)$, the direct-only contract is optimal (Theorem 2 (c)); and when $\Pi < \Pi_F$, the project becomes infeasible (Theorem 2 (d)).

Remark 2. 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 upon the one characterized above. This follows because the value function $\mathcal{V}$ is concave (Lemma OA.2.5 (c)).

Consider a set of contracts $\{\Gamma_i\}_{1\leq i\leq n}$ where the agent's expected utility under Γ_i is u_i, and the weight is w_i with $\sum_{i=1}^n w_i = 1$ and $\sum_{i=1}^n w_i \cdot u_i = u$. The principal's expected payoff from this mixture is $\sum_{i=1}^n w_i \cdot P_0(\Gamma_i)$ and the agent's expected utility is u. By concavity, I have $V(u) \geq \sum_{i=1}^n w_i V(u_i)$. Additionally, $V(u_i) \geq P_0(\Gamma_i)$ holds for all $1 \leq i \leq n$ 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 upon the characterized contract.

5 The Optimal Contract under Efficiency Loss

I now consider the case where splitting the project generates an efficiency loss, i.e., $\lambda_S < 2\lambda_D$. This introduces efficiency as another economic force, alongside monitoring and deadline effects, that determines the optimal contract.

I define a parameter $\eta \equiv \frac{\lambda_S}{\lambda_D} - 1$, which measures the efficiency of the sequential approach. Note that $0 < \eta < 1$, and the efficiency loss increases as η decreases. In this section, I characterize the optimal contracts for two cases: (i) when the efficiency loss is small ($\eta > \overline{\eta} \equiv \max\{\sqrt{\frac{c}{c+\phi}}, \frac{1}{e-1}\}$); and (ii) when the efficiency loss is large ($\eta < \underline{\eta} \equiv \min\{\frac{c}{c+\phi}, \frac{1}{e-1}\}$).[11]

In Figures 3 and 4, I illustrate the principal's value functions when there are efficiency losses from splitting the project. A key characteristic of these value functions is that the direct approach is employed when the promised utility is high, indicating that the deadline is far off. To understand this dynamics, I compare the direct-only and sequential-only contracts again. As time horizons become longer, the sums of expected payoffs for both players from

[11] These do not cover cases where the efficiency loss is intermediate. In such cases, the form of the optimal contract depends heavily on the parameter values η and Π, resulting in many subcases to analyze. Thus, I focus on the extreme cases to provide results with clear economic implications.

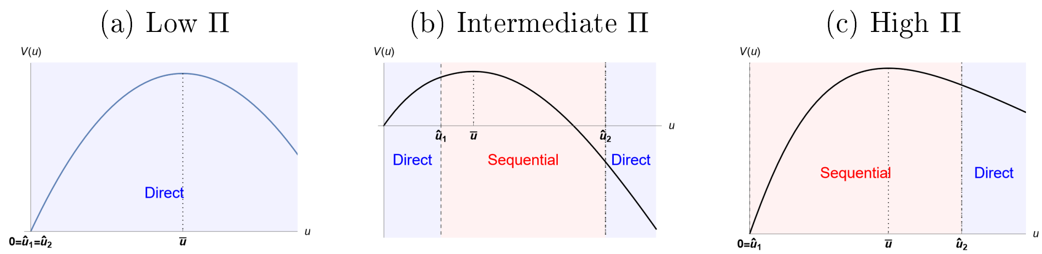


Figure 3: Value functions when the efficiency loss is small

these contracts converge to the expected surpluses of the no-deadline benchmark: $\Pi - \frac{c}{\lambda_D}$ for the direct-only contract and $\Pi - \frac{2c}{\lambda_S}$ for the sequential-only contract. Therefore, 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 chose the direct approach when the deadline is distant.

This observation, combined with the effects of monitoring and deadline discussed in the previous sections, leads us to conjecture that there will be two switching points $\hat{u}_1$ and $\hat{u}_2$ in determining the value function. The direct approach is chosen when $u > \hat{u}_2$ or $u < \hat{u}_1$, and the sequential approach is chosen when $u \in (\hat{u}_1, \hat{u}_2)$.

When the efficiency loss is relatively small, I show that $\hat{u}_2$ is always greater than the optimal initial promised utility level $\overline{u}$ (Lemma OA.3.9). It implies that the switch occurs at most once in the optimal contract. Therefore, a result similar to the no-efficiency-loss case holds. In other words, Theorem 2 is robust to small efficiency losses.

Theorem 3. *Suppose that $\eta \in (\overline{\eta}, 1)$, i.e., the efficiency loss from splitting the project is small. There exist thresholds $\tilde{\Pi}_D(\eta)$ and $\tilde{\Pi}_S(\eta)$ with $\tilde{\Pi}_S(\eta) > \tilde{\Pi}_D(\eta) > \Pi_F$ such that the optimal contract is determined as follows:*

(a) when $\Pi > \tilde{\Pi}_S(\eta)$, a sequential-only contract is optimal;

(b) when $\tilde{\Pi}_S(\eta) > \Pi > \tilde{\Pi}_D(\eta)$, there exists a one-switch contract that is optimal;

(c) when $\tilde{\Pi}_D(\eta) > \Pi > \Pi_F$, a direct-only contract is optimal.

(a) Low Π (b) High Π

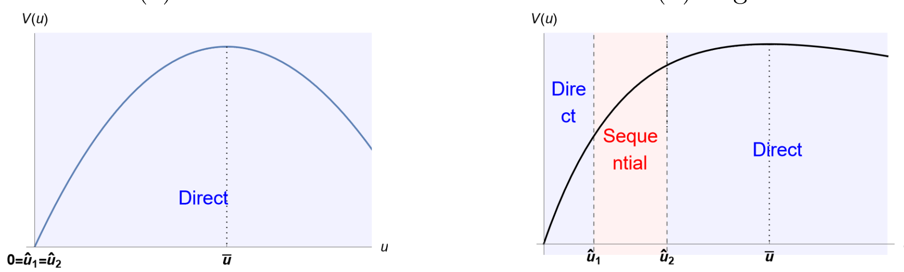


Figure 4: Value functions when the efficiency loss is large

Now suppose that the efficiency loss is large. Figure 4 illustrates that the sequential approach is either not employed at all (for small Π) or is employed in the middle of the contract (for large Π). Intuitively, as Π increases, the monitoring effect becomes more significant. However, the direct approach is preferred at the beginning of the contract due to its efficiency and at the end of the contract due to the deadline effect. Therefore, if the sequential approach is ever employed, the contract will involve two switches. I formally define the two-switch contract and state the theorem for the case of large efficiency loss.

Definition 5.1. A contract is called a *two-switch contract* with a final deadline T and two intermediate deadlines $0 < S_1 < S_2 < T$ if

(i) the direct approach is employed up to S_1 and the reward upon project completion at time t is R_t^D,

(ii) if the project is not completed by S_1, the sequential approach is employed up to S_2, and if the subproject is completed before S_2, the contract is extended by $\frac{1}{\lambda_S}$ with the reward upon project completion at time t being R_t^S,

(iii) if the subproject is not completed by S_2, the direct approach is employed up to T and the reward upon project completion at time t is R_t^D, and

(iv) the contract is terminated if the project is not completed by T.[12]

[12] The rewards R_t^D and R_t^S are defined in the same way as in the one-switch contract.

Theorem 4. *Suppose that η is less than $\underline{\eta}$, i.e., the efficiency loss from splitting the project is large. There exists a threshold $\tilde{\Pi}_M(\eta)$ with $\tilde{\Pi}_M(\eta) > \Pi_F$ such that the optimal contract is determined as follows:*

(a) when $\Pi > \tilde{\Pi}_M(\eta)$, there exists a two-switch contract that is optimal;

(b) when $\tilde{\Pi}_M(\eta) > \Pi > \Pi_F$, a direct-only contract is optimal.

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 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).

6 Conclusion

In this article, 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.[13] I leave these intriguing questions–and others–for future work.

Appendix

A Proofs for Section 3

A.1 Direct-only vs. Sequential-only Policies

In this section, I provide formal representations of the completion probabilities, expected durations and expected social surpluses under the direct-only and the sequential-only policies.

[13] This possibility contrasts with the setting considered by Carnehl and Schneider (2023), where learning causes the expected arrival rate to fall.

Suppose that the planner adopts the direct-only policy. Note that the direct-only policy is not affected by the deadline extension Δ, so we can omit it. Let τ_m denote the date of the project's completion. The completion probability is

$$\mathcal{P}^d(T) \equiv \int_0^T \lambda_D \cdot e^{-\lambda_D \tau_m} d\tau_m = 1 - e^{-\lambda_D T}, \tag{A.1}$$

and the expected duration of the project is

$$\mathcal{D}^d(T) \equiv \int_0^T \tau_m \cdot \lambda_D \cdot e^{-\lambda_D \tau_m} d\tau_m + T \cdot e^{-\lambda_D T} = \frac{1}{\lambda_D}(1 - e^{-\lambda_D T}). \tag{A.2}$$

Then, the expected social surplus of the direct-only policy is

$$\mathcal{W}^d(T) \equiv \Pi \cdot \mathcal{P}^d(T) - c \cdot \mathcal{D}^d(T) = \left(\Pi - \frac{c}{\lambda_D}\right) \cdot (1 - e^{-\lambda_D T}). \tag{A.3}$$

Now suppose that the planner employs the sequential-only policy. Let τ_s denote the date of the first subproject's completion. The probability of project completion can be derived as follows:

$$\begin{aligned} \mathcal{P}^s(T, \Delta) &\equiv \int_0^T \left[\int_{\tau_s}^{T+\Delta} \lambda_S e^{-\lambda_S(\tau_m - \tau_s)} d\tau_m\right] \cdot \lambda_S e^{-\lambda_S \tau_s} d\tau_s \\ &= 1 - (1 + \lambda_S \cdot T \cdot e^{-\lambda_S \Delta}) \cdot e^{-\lambda_S T}. \end{aligned} \tag{A.4}$$

Next, conditional on the first subproject being completed at τ_s, the expected duration is

$$\mathcal{D}^1(T, \Delta, \tau_s) \equiv \int_{\tau_s}^{T+\Delta} \tau_m \cdot \lambda_S e^{-\lambda_S(\tau_m - \tau_s)} d\tau_m + (T + \Delta) \cdot e^{-\lambda_S(T+\Delta-\tau_s)}.$$

Then, the expected duration of the project can also be derived as follows:

$$\begin{aligned} \mathcal{D}^s(T, \Delta) &\equiv \int_0^T \mathcal{D}^1(T, \Delta, \tau_s) \cdot \lambda_S e^{-\lambda_S \tau_s} d\tau_s + T \cdot e^{-\lambda_S T} \\ &= \frac{2}{\lambda_S}(1 - e^{-\lambda_S T}) - T \cdot e^{-\lambda_S(T+\Delta)}. \end{aligned} \tag{A.5}$$

Then, the expected social surplus of the sequential-only policy is

$$\begin{aligned}\mathcal{W}^s(T,\Delta) &\equiv \Pi\cdot\mathcal{P}^s(T,\Delta) - c\cdot\mathcal{D}^s(T,\Delta) \\ &= \left(\Pi - \frac{2c}{\lambda_S}\right)\cdot(1-e^{-\lambda_S T}) - \lambda_S\left(\Pi - \frac{c}{\lambda_S}\right)\cdot T\cdot e^{-\lambda_S(T+\Delta)}.\end{aligned} \tag{A.6}$$

A.2 Proof of Proposition 3.1

Proof of Proposition 3.1. Observe that $\mathcal{P}^d(T,0) \geq \mathcal{P}^s(T,0)$ is equivalent to:

$$(1+\lambda_S T) \geq e^{(\lambda_S-\lambda_D)T}.$$

Note that the equality holds at $T=0$. While the left hand side linearly increases with the slope λ_S, he right-hand side exponentially increases and the slope at $T=0$ is $\lambda_S-\lambda_D$, which is lower than λ_S. Therefore, for small enough T, $\mathcal{P}^d(T,0) > \mathcal{P}^s(T,0)$, but there exists $\check{T}>0$, which makes the two sides equal. Then, we have $\mathcal{P}^d(T,0) > \mathcal{P}^s(T,0)$ for all $T<\check{T}$ and $\mathcal{P}^d(T,0) < \mathcal{P}^s(T,0)$ for all $T>\check{T}$. □

A.3 Proof of Theorem 1

The Expected Surplus upon the Subproject Completion Let W_x^1 represent the expected surplus when the first subproject is completed and the remaining time is x. By following steps similar to those used in the derivation of (A.3), we have

$$W_x^1 \equiv \left(\Pi - \frac{c}{\lambda_S}\right)\cdot(1-e^{-\lambda_S x}). \tag{A.7}$$

Suppose that the subproject is completed at calendar time $T-z$, meaning that z units of time remain until the original deadline. Then, the subproject completion extends the deadline by Δ, giving the planner $z+\Delta$ units of time to complete the project. Therefore, the expected surplus in this situation is $W_{z+\Delta}^1$.

The Expected Surplus without the Subproject Completion Now we consider the situation that neither the subproject nor the main project has been completed by calendar time $T - z$. Then, the (optimal) expected surplus $W_z^{0,\Delta}$ can heuristically be written as follows:

$$W_z^{0,\Delta} = \max_{a_z \in \{0,1\}} \begin{array}{l} \Pi \cdot \lambda_D a_z \cdot dz + W_{z+\Delta}^1 \cdot \lambda_S (1 - a_z) \cdot dz - c\, dz \\ + \{1 - \lambda_D a_z \cdot dz - \lambda_S (1 - a_z) \cdot dz\} \cdot W_{z-dz}^{0,\Delta}. \end{array}$$

By using a Taylor expansion, $W_{z-dz}^{0,\Delta} = W_z^{0,\Delta} - \dot{W}_z^0 dz$, canceling out $W_z^{0,\Delta}$ on both sides, and taking the limit as $dz \to 0$, we obtain a Hamilton-Jacobi-Bellman (HJB) equation:

$$\dot{W}_z^{0,\Delta} = \max_{a_z \in \{0,1\}} \left[\lambda_D a_z \cdot (\Pi - W_z^{0,\Delta}) + \lambda_S (1 - a_z) \cdot (W_{z+\Delta}^1 - W_z^{0,\Delta}) - c \right]. \tag{HJB$_W$}$$

Since the project is terminated at the deadline, $W_0^{0,\Delta} = 0$. Then, by using standard verification arguments (e.g., Proposition 3.2.1 in Bertsekas (1995)), if a function $w : [0, T] \to \mathbb{R}$ is differentiable and satisfies (HJB$_W$) and $w(0) = 0$, then, $w(z) = W_z^{0,\Delta}$.

The Expected Surplus of the One-switch Policy I now derive the expected surplus of the one-switch policy with an intermediate deadline S and a deadline T. Denote $Z \equiv T - S$. The one-switch policy implies that $a_z = 1$ for all $0 \leq z < Z$, and $a_z = 0$ for all $Z \leq z \leq T$. Let $\mathcal{W}_Z^{sd}(z, \Delta)$ denote the expected surplus of this policy when the remaining time is z. The following differential equations then hold, where $\dot{W}_Z^{sd} = \frac{\partial W_Z^{sd}(z,\Delta)}{\partial z}$:

$$\dot{\mathcal{W}}_Z^{sd}(z, \Delta) = \lambda_S \cdot \left(W_{z+\Delta}^1 - \mathcal{W}_Z^{sd}(z, \Delta) \right) - c \quad \text{for } z \geq Z, \tag{A.8}$$

$$\dot{\mathcal{W}}_Z^{sd}(z, \Delta) = \lambda_D \cdot \left(\Pi - \mathcal{W}_Z^{sd}(z, \Delta) \right) - c \quad \text{for } z < Z. \tag{A.9}$$

By solving this with $\mathcal{W}^{sd}_Z(0,\Delta)=0$, I derive

$$\mathcal{W}^{sd}_Z(z,\Delta)=\begin{cases}\mathcal{W}^d(z), & \text{if } z\le Z,\\ \left(\Pi-\dfrac{2c}{\lambda_S}\right)\cdot\left(1-e^{-\lambda_S(z-Z)}\right)+\mathcal{W}^d(Z)\cdot e^{-\lambda_S(z-Z)} & \\ \qquad -\lambda_S\left(\Pi-\dfrac{c}{\lambda_S}\right)\cdot(z-Z)\cdot e^{-\lambda_S(z+\Delta)}, & \text{if } z>Z.\end{cases} \tag{A.10}$$

Also note that $\mathcal{W}^{sd}_0(z,\Delta)=\mathcal{W}^s(z,\Delta)$ is the expected surplus of the sequential-only policy and $\mathcal{W}^{sd}_T(z,\Delta)=\mathcal{W}^d(z)$ is the expected surplus of the direct-only policy.

We will prove the theorem by showing that there exists $Z\in[0,T]$ such that $\mathcal{W}^{sd}_Z(z,\Delta)$ solves (HJB$_W$).

The Optimal Approach at the Deadline Note that $W^1_\Delta=\left(\Pi-\frac{c}{\lambda_S}\right)\cdot(1-e^{-\lambda_S\Delta})$ and $W^{0,\Delta}_0=0$. Then, at the deadline, the sequential approach is preferred over the direct approach if and only if

$$\begin{aligned} & \lambda_D(\Pi-W^{0,\Delta}_0)\le\lambda_S(W^1_\Delta-W^{0,\Delta}_0)\\ \iff\quad & \lambda_D\Pi\le(\lambda_S\Pi-c)\cdot(1-e^{-\lambda_S\Delta}). \end{aligned} \tag{A.11}$$

With simple algebra, we can derive that (A.11) is equivalent to $\Delta\ge\bar{\Delta}$.

Optimal Policy Derivation I introduce two crucial lemmas, then complete the proof of Theorem 1. The proof of the lemmas are in the following subsection.

Lemma A.1. *Suppose that $\lambda_S=2\lambda_D$ and $\Delta<\bar{\Delta}$. Then, there exists $\hat{T}$ such that (i) $\lambda_D(\Pi-\mathcal{W}^d(z))>\lambda_S(W^1_{z+\Delta}-\mathcal{W}^d(z))$ for all $z<\hat{T}$; and (ii) $\lambda_D(\Pi-\mathcal{W}^d(\hat{T}))=\lambda_S(W^1_{\hat{T}+\Delta}-\mathcal{W}^d(\hat{T}))$.*

Lemma A.2. *Suppose that $\lambda_S=2\lambda_D$ and $\lambda_S(W^1_{Z+\Delta}-\mathcal{W}^d(Z))\ge\lambda_D(\Pi-\mathcal{W}^d(Z))$ for some $Z\ge 0$. Then, $\lambda_S(W^1_{z+\Delta}-\mathcal{W}^{sd}_Z(z,\Delta))\ge\lambda_D(\Pi-\mathcal{W}^{sd}_Z(z,\Delta))$ for all $z>Z$.*

Proof of Theorem 1. (a) Suppose that $\Delta \geq \bar{\Delta}$. From (A.11) and $\mathcal{W}^d(0) = 0$, we have $\lambda_S(W^1_\Delta - \mathcal{W}^d(0)) \geq \lambda_D(\Pi - \mathcal{W}^d(0))$. Then, by Lemma A.2, $\lambda_S(W^1_{z+\Delta} - \mathcal{W}^{sd}_0(z,\Delta)) \geq \lambda_D(\Pi - \mathcal{W}^{sd}_0(z,\Delta))$ for all $z > 0$. Then, by (A.8), $\mathcal{W}^s_{z,\Delta} = \mathcal{W}^{sd}_0(z,\Delta)$ solves (HJB_W) for all $z \in \mathbb{R}_+$, i.e., the sequential-only policy is optimal.

(b) Suppose that $\Delta < \bar{\Delta}$. Let $\hat{T}$ be the time defined in Lemma A.1. If $T < \hat{T}$, $\lambda_D(\Pi - \mathcal{W}^d(z)) > \lambda_S(W^1_{z+\Delta} - \mathcal{W}^d(z))$ for all $z \in [0,T]$. Then, by (A.9), $\mathcal{W}^d(z)$ solves (HJB_W) for all $z \in [0,T]$, i.e., the direct-only policy is optimal.

Now consider the case with $T \geq \hat{T}$. Note that $\mathcal{W}^{sd}_{\hat{T}}(z,\Delta) = W^d_z$ for all $z \in [0,\hat{T}]$. From Lemma A.1, we have $\lambda_D(\Pi - \mathcal{W}^{sd}_{\hat{T}}(z,\Delta)) > \lambda_S(W^1_{z+\Delta} - \mathcal{W}^{sd}_{\hat{T}}(z,\Delta))$ for all $z \in (\hat{T},T]$. Then, by (A.9), $\mathcal{W}^{sd}_{\hat{T}}(z,\Delta)$ solves (HJB_W) for all $z > \hat{T}$. In addition, we have $\lambda_D(\Pi - \mathcal{W}^{sd}_{\hat{T}}(\hat{T},\Delta)) = \lambda_S(W^1_{\hat{T}+\Delta} - \mathcal{W}^{sd}_{\hat{T}}(\hat{T},\Delta))$. By applying Lemma A.2 for $Z = \hat{T}$, we have $\lambda_S(W^1_{z+\Delta} - \mathcal{W}^{sd}_{\hat{T}}(z,\Delta)) \geq \lambda_D(\Pi - \mathcal{W}^{sd}_{\hat{T}}(z,\Delta))$ for all $z > \hat{T}$. Then, by (A.8), $\mathcal{W}^{sd}_{\hat{T}}(z,\Delta)$ solves (HJB_W) for all $z > \hat{T}$. Therefore, $\mathcal{W}^{sd}_{\hat{T}}(z,\Delta)$ solves (HJB_W) for all $z \in [0,T]$, i.e., the one-switch policy with $T - \hat{T}$ is optimal.

□

A.3.1 Proof of Lemmas

Proof of Lemma A.1. Define $H^1_z \equiv \lambda_S\left(W^1_{z+\Delta} - \mathcal{W}^d(z)\right) - \lambda_D\left(\Pi - \mathcal{W}^d(z)\right)$ and $x \equiv e^{-\lambda_D z}$. Then, with some algebra, H^1_z is equivalent to

$$H_1(x) \equiv (\lambda_S\Pi - c)\cdot(1 - e^{-\lambda_S\Delta}\cdot x^{\frac{\lambda_S}{\lambda_D}}) - \lambda_D\Pi - (\lambda_S - \lambda_D)\cdot\left(\Pi - \frac{c}{\lambda_D}\right)\cdot(1-x). \tag{A.12}$$

By using $\lambda_S = 2\lambda_D$, with some algebra, $H_1(x)$ can be rewritten as follows:

$$H_1(x) = (\lambda_D\Pi - c)\cdot x - (\lambda_S\Pi - c)\cdot e^{-\lambda_S\Delta}\cdot x^2$$

Define

$$\hat{x} \equiv \frac{\lambda_D \Pi - c}{\lambda_S \Pi - c} \cdot e^{\lambda_S \Delta}. \tag{A.13}$$

Note that $\hat{x} < 1$ when $\Delta < \overline{\Delta}$. Additionally, observe that $H_1(\hat{x}) = 0$ and $H_1(x) < 0$ for all $\hat{x} < x \leq 1$.

Now set $\hat{T} \equiv -\frac{\log(\hat{x})}{\lambda_D}$. Then, $\hat{T} > z$ is equivalent to $x > \hat{x}$, which implies $H_1(x) < 0$. Therefore, for all $z < \hat{T}$, $\lambda_D(\Pi - \mathcal{W}^d(z)) > \lambda_S(W^1_{z+\Delta} - \mathcal{W}^d(z))$. In addition, $H_1(\hat{x}) = 0$ implies $\lambda_D(\Pi - \mathcal{W}^d(\hat{T})) = \lambda_S(W^1_{\hat{T}+\Delta} - \mathcal{W}^d(\hat{T}))$. □

Proof of Lemma A.2. Define $H^2_z \equiv \lambda_S(W^1_{z+\Delta} - \mathcal{W}^{sd}_Z(z,\Delta)) - \lambda_D(\Pi - \mathcal{W}^{sd}_Z(z,\Delta))$ and $y \equiv e^{-\lambda_S(z-Z)}$. Note that, for any $z > Z$,

$$\begin{aligned} \mathcal{W}^{sd}_Z(z,\Delta) =& \mathcal{W}^d(Z) + \left(\Pi - \frac{2c}{\lambda_S} - \mathcal{W}^d(Z)\right) \cdot (1-y) \\ &+ \left(\Pi - \frac{c}{\lambda_S}\right) \cdot e^{-\lambda_S(Z+\Delta)} \cdot \log(y) \cdot y \end{aligned}$$

and

$$W^1_{z+\Delta} = W^1_{Z+\Delta} + \left(\Pi - \frac{c}{\lambda_S} - W^1_{Z+\Delta}\right) \cdot (1-y).$$

Then, with some algebra, H^2_z can be rewritten as follows:

$$H_2(y) \equiv H^2_Z + h_1 \cdot (1-y) + h_2 \cdot \log(y) \cdot y \tag{A.14}$$

where

$$\begin{aligned} h_1 \equiv& \lambda_S \cdot \left(\Pi - \frac{c}{\lambda_S} - W^1_{Z+\Delta}\right) - (\lambda_S - \lambda_D) \cdot \left(\Pi - \frac{2c}{\lambda_S} - \mathcal{W}^d(Z)\right), \\ h_2 \equiv& \left(\Pi - \frac{c}{\lambda_S}\right) \cdot e^{-\lambda_S(Z+\Delta)} > 0. \end{aligned}$$

Observe that

$$H_2''(y) = -\frac{h_2}{y} < 0,$$

i.e., H_2 is strictly concave. By the assumption, we have $H_2(1) = H^2_S \geq 0$. In addition, we have

$$\lim_{y\to 0} H_2(y) = \lambda_S\left(\Pi - \frac{c}{\lambda_S}\right) - \lambda_D\Pi - (\lambda_S - \lambda_D)\left(\Pi - \frac{2c}{\lambda_S}\right) = \frac{\lambda_S - 2\lambda_D}{\lambda_S}c = 0. \tag{A.15}$$

Then, by using the strict concavity of H_2, $H_2(1) \geq 0$ and $\lim_{y\to 0} H_2(y) = 0$, we have $H_2(y) \geq 0$ for all $y \in (0,1)$. Therefore, $\lambda_S(W^1_{z+\Delta} - \mathcal{W}^{sd}_Z(z,\Delta)) \geq \lambda_D(\Pi - \mathcal{W}^{sd}_Z(z,\Delta))$ for all $z > Z$. $\square$

References

Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA, 1995.

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

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

Kevin A Bryan and Jorge Lemus. The direction of innovation. *Journal of Economic Theory*, 172:247–272, 2017.

Vannevar Bush. Science: The endless frontier. *Transactions of the Kansas Academy of Science (1903-)*, 48(3):231–264, 1945. URL https://www.jstor.org/stable/3625196.

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

Christoph Carnehl and Johannes Schneider. On risk and time pressure: When to think and when to do. *Journal of the European Economic Association*, 21(1):1–47, 2023.

Mingliu Chen, Peng Sun, and Yongbo Xiao. Optimal monitoring schedule in dynamic contracts. *Operations Research*, 68(5):1285–1314, 2020.

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

Francesc Dilmé and Daniel F Garrett. Residual deterrence. *Journal of the European Economic Association*, 17(5):1654–1686, 2019.

Richard J Gilbert and Michael L Katz. Efficient division of profits from complementary innovations. *International Journal of Industrial Organization*, 29(4):443–454, 2011.

Boaz Golany and Avraham Shtub. Work breakdown structure. *Handbook of Industrial Engineering: Technology and Operations Management*, pages 1263–1280, 2001.

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

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

Jerry R Green and Suzanne Scotchmer. On the division of profit in sequential innovation. *The Rand journal of economics*, pages 20–33, 1995.

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

Zehao Hu. Financing innovation with unobserved progress. 2014.

Yonggyun Kim and Francisco Poggi. Research or development in innovation races. *Available at SSRN 4638677*, 2024.

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

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

Iván Marinovic and Martin Szydlowski. Monitoring with career concerns. *Available at SSRN 3450629*, 2019.

Iván Marinovic and Martin Szydlowski. Monitor reputation and transparency. *Available at SSRN 3703870*, 2020.

Sofia Moroni. Experimentation in organizations. *Theoretical Economics*, 17 (3):1403–1450, 2022.

Roger B Myerson. Moral hazard in high office and the dynamics of aristocracy. *Econometrica*, 83(6):2083–2126, 2015.

Christine Organ and Cassie Bottorff. Work breakdown structure (wbs) in project management. *Forbes*, March 2022. URL https://www.forbes.com/advisor/business/what-is-work-breakdown-structure/.

Dmitry Orlov. Frequent monitoring in dynamic contracts. *Available at SSRN 2524628*, 2015.

Tomasz Piskorski and Mark M Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory*, 166: 242–281, 2016.

Francisco Poggi. The timing of complementary innovations. 2021.

Project Management Institute. *Practice Standard for Work Breakdown Structures*. 2006.

Project Management Institute. *A guide to the Project Management Body of Knowledge (PMBOK guide)*. Project Management Institute, 6th edition, 2017.

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

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

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

Martin Szydlowski. Incentives, project choice, and dynamic multitasking. *Theoretical Economics*, 14(3):813–847, 2019. doi: 10.3982/TE2858. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/TE2858.

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

Felipe Varas, Iván Marinovic, and Andrzej Skrzypacz. Random inspections and periodic reviews: Optimal dynamic monitoring. *The Review of Economic Studies*, 87(6):2893–2937, 2020.

Christoph Wolf. Informative milestones in experimentation. *University of Mannheim, Working Paper*, 2018. URL https://carnehl.github.io/IME.pdf.

Yu Fu Wong. Dynamic monitoring design. *Available at SSRN 4466562*, 2023.

Online Appendix for
"Managing a Project by Splitting it into Pieces"

Yonggyun Kim

OA.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 $\Gamma \equiv \left\{T, \{a_t, b_t, R_t, \hat{\Gamma}^t\}_{0\le t\le T}\right\}$, where each variable is defined as follows at the calendar time t:[14]

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;

4. $R_t \ge 0$: the monetary payment from the principal to the agent for the success of the main project at t;[15]

5. $\hat{\Gamma}^t \equiv \{T^t, \{b^t_s, R^t_s\}_{t\le s\le T^t}\}$: an updated contract when the subproject is completed at t;

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

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

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

[14] See Remark 2 for discussion on deterministic and mixed contracts.

[15] 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.

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 $\hat{\mathcal{B}}^t \equiv \{\{\check{b}_s\}_{t\le s\le T^t} \ : \ \check{b}_s \in [0,1]\}$. The agent's action profile $\tilde{b}^t \equiv \{\tilde{b}^t_s\}_{t\le s\le T^t} \in \hat{\mathcal{B}}^t$ 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

$$\hat{P}^t(\hat{\Gamma}^t) = \mathbb{E}^{b^t}\left[\left(\Pi - R^t_{\tau_m}\right)\cdot \mathbf{1}_{\{t\le\tau_m\le T^t\}} - \int_t^{T^t\wedge\tau_m} c\ ds\right],$$ [16]

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

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

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 $\mathcal{B} \equiv \{\{\check{b}_t\}_{0\le t\le T} \ : \ \check{b}_t \in [0,1]\}$. 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

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

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

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

where the first term is the payoff from the main project completion, the second term is the

[16] 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.

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

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

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

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

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

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., $U_0(\Gamma)\ge 0$. Designate such a contract as an *optimal contract*.

OA.2 Proofs for Section 4

In this section, I provide the proof for the value function characterization when there is no efficiency loss from splitting the project (Proposition 4.1). I begin by recursively formulating the agent's and the principal's problems. Then, I explore some properties of the value function candidate suggested in Proposition 4.1, including that this value function can be implemented by a one-switch contract. Based on these properties, I prove Proposition 4.1 in Section OA.2.3. I also provide some properties of the thresholds for Π, which are necessary in the proof of Theorem 2, in Section OA.2.4. Proofs of the lemmas are relegated to Section OA.2.5.

OA.2.1 Recursive Formulation

OA.2.1.1 The Agent's Problem

In this subsection, I formally derive the agent's continuation utility and the Hamilton-Jacobi-Bellman (HJB) equation of it.

I begin by specifying the probability distribution functions for possible events given a profile of approaches $a = \{a_s\}_{t \le s \le T}$ 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

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

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

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

$$\mathbb{E}^{a,b}\left[\int_t^{T \wedge \tau_m \wedge \tau_s} \phi(1 - b_s) \mid t \le \tau_m \wedge \tau_s\right] = \int_t^T \phi(1 - b_s) \cdot f(a, b; t, s) ds.$$

Given a contract Γ, the agent's maximized continuation utility at time t, $U_t(\Gamma)$, can be derived as follows:

$$U_t(\Gamma) \equiv \sup_{\tilde{b} \in \mathcal{B}_t} \int_t^T \left[R_s \cdot \lambda_D a_s \tilde{b}_s + \hat{U}_s(\hat{\Gamma}^s) \cdot \lambda_S (1 - a_s) \tilde{b}_s + \phi(1 - \tilde{b}_s)\right] f(a, \tilde{b}; t, s) ds, \quad \text{(OA.2.1)}$$

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

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 $U_t(\Gamma)$.

Lemma OA.2.1. *Given a contract Γ, suppose that a continuous and differentiable process*

$\{u_t\}_{0\le t\le T}$ *satisfies* $u_T = 0$

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

where $\dot{u}_t \equiv \frac{du_t}{dt}$. *Then,* $u_t = U_t(\Gamma)$.

OA.2.1.2 The Principal's Problem

The Value Function upon the Subproject Completion Since this case only requires one more breakthrough, it is identical to the single-stage benchmark of Green and Taylor (2016a). They show that the principal's value function V_S is characterized as follows:

$$V_S(u_S^t) = W^1_{u_S^t/\phi} - u_S^t = \left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi}u_S^t}\right) - u_S^t. \qquad (\text{OA.2.2})$$

The Value Function without the Subproject Completion I now consider the principal's problem without the subproject completion. 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,

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

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

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

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:

$$V(u_t) = \max_{R_t, u_S^t, a_t, b_t} \begin{array}{c} -\,cdt + (\Pi - R_t)\,\lambda_D a_t b_t dt + V_S(u_S^t)\lambda_S(1-a_t)b_t dt \\ +\,(1 - \lambda_D a_t b_t dt - \lambda_S(1-a_t)b_t dt)V(u_{t+dt}) \end{array}.$$

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 (OA.2.3) in, we obtain an HJB equation:

$$0 = \max_{R, u_S, a, b} \mathcal{J}(V(\cdot), R, u_S, a, b). \qquad (\text{HJB}_V)$$

where

$$\mathcal{J}(V(\cdot), R, u_S, a, b) \equiv -c + (\Pi - R - V(u))\lambda_D ab + (V_S(u_S) - V(u))\lambda_S(1-a)b \\ - [\phi \vee \{(R-u)\lambda_D a + (u_S - u)\lambda_S(1-a)\}]\, V'(u) \qquad \text{(OA.2.4)}$$

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 $U_0(\Gamma) = u$.

Lemma OA.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 $U_0(\Gamma) = u$,*

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

Lemma OA.2.3. *Suppose that $V'(u) \geq -1$. The solution of (HJB_V) subject to (IC) involves $b = 1$.*

OA.2.2 Value Function Candidates and Implementation

Lemma OA.2.4. *The following statements hold.*

(a) A direct-only contract with the deadline $\frac{u}{\phi}$ implements a pair of expected payoffs of the principal and the agent $(V^d(u), u)$ where

$$V^d(u) \equiv \mathcal{W}^d\left(\frac{u}{\phi}\right) - u. \qquad \text{(OA.2.5)}$$

(b) When $0 < u_1 < u$, a one-switch contract with the intermediate deadline $\frac{u-u_1}{\phi}$ and the final deadline $\frac{u}{\phi}$ implements $(V^{sd}(u|u_1), u)$ where

$$V^{sd}(u|u_1) \equiv \mathcal{W}^{sd}_{u_1/\phi}\left(\frac{u}{\phi}, \frac{1}{\lambda_S}\right) - u. \qquad \text{(OA.2.6)}$$

(c) A sequential-only contract with the deadline $\frac{u}{\phi}$ implements $(V^{sd}(u|0), u)$.

(d) *The following differential equations hold:*

$$\phi V^{d'}(u) = \lambda_D \left(\Pi - \frac{\phi}{\lambda_D} - u - V^d(u)\right) - c, \tag{OA.2.7}$$

$$\phi V^{sd'}(u|u_1) = \lambda_S \left(V_S(u + \tfrac{\phi}{\lambda_S}) - V^{sd}(u|u_1)\right) - c. \tag{OA.2.8}$$

Together with Theorem 1, this lemma implies that $(\mathcal{V}(u), u)$—defined in (4.2)—can be implemented by one of the above three contracts. Moreover, there exists $\hat{u}_1 \geq 0$ such that $\mathcal{V}$ can be rewritten as follows:

$$\mathcal{V}(u) = \begin{cases} V^d(u), & \text{if } u < \hat{u}_1, \\ V^{sd}(u|\hat{u}_1), & \text{if } u \geq \hat{u}_1. \end{cases} \tag{OA.2.9}$$

Specifically, $\hat{u}_1$ is chosen to be equal to $\phi\hat{T}$ if $\frac{1}{\lambda_S} < \overline{\Delta}$, and 0 if $\frac{1}{\lambda_S} \geq \overline{\Delta}$. The following lemma provides useful properties of $\mathcal{V}$ and $\hat{u}_1$.

Lemma OA.2.5. *Suppose that $\lambda_S = 2\lambda_D$. The following statements hold.*

(a) *if $\hat{u}_1 > 0$, $V^{sd'}(\hat{u}_1|\hat{u}_1) = V^{d'}(\hat{u}_1)$ and $V^{sd'}(u|u) < V^{d'}(u)$ for all $u < u_1$, and if $\hat{u}_1 = 0$, $V^{sd'}(0|0) \geq V^{d'}(0)$.*

(b) *$\mathcal{V}'(u) \geq -1$ for all $u \geq 0$.*

(c) *$\mathcal{V}$ is concave.*

OA.2.3 Value Function Verification (Proposition 4.1)

The goal of this subsection is to prove Proposition 4.1. Specifically, I show that the value function defined in the previous section solves (HJB_V) subject to (IC). To achieve this, I introduce functions that specify potential deviations and then establish useful properties as a lemma, followed by the proof for Proposition 4.1.

First, define

$$L^D(u, R) \equiv \mathcal{J}(\mathcal{V}(\cdot), R, \cdot, 1, 1) = \lambda_D(\Pi - R - \mathcal{V}(u)) - c - \lambda_D(R - u)\mathcal{V}'(u). \tag{OA.2.10}$$

Given u, maximizing this function with respect to $R \geq u + \frac{\phi}{\lambda_D}$ is equivalent to maximizing the right hand side of (HJB_V) under the condition that $b = 1$ solves (HJB_{PK}) with $a = 1$.

Similarly, define

$$L^S(u,w) \equiv \mathcal{J}(\mathcal{V}(\cdot),\cdot,u_S,0,1) = \lambda_S(V_S(w)-\mathcal{V}(u)) - c - \lambda_S(w-u)\mathcal{V}'(u).$$

Given u, maximizing this function with respect to $w \geq u + \frac{\phi}{\lambda_S}$ is equivalent to maximize the right hand side of (HJB$_V$) under the condition that $b=1$ solves (HJB$_{PK}$) with $a=0$.

Lemma OA.2.6. *Suppose that $\Pi > \frac{c}{\lambda_D}$ and $\lambda_S = 2\lambda_D$. Then, for any $u \geq 0$, $L^D(u,R) \leq 0$ for all $R \geq u + \frac{\phi}{\lambda_D}$, and $L^S(u,w) \leq 0$ for all $w \geq u + \frac{\phi}{\lambda_S}$.*

Proof of Proposition 4.1. I begin by showing that $\mathcal{J}$ becomes zero when the value function $\mathcal{V}$ is utilized alongside contractual terms with binding ICs.

When $u \geq \hat{u}_1$, $\mathcal{V}(u) = V^{sd}(u|\hat{u}_1)$. Then, by (OA.2.8),

$$\mathcal{J}\left(V^{sd}(u|\hat{u}_1),\ \cdot,\ u+\tfrac{\phi}{\lambda_S},\ 0,\ 1\right) = 0.$$

Likewise, when $u < \hat{u}_1$, $\mathcal{V}(u) = V^d(u|)$, and by (OA.2.7),

$$\mathcal{J}\left(V^d(u),\ u+\tfrac{\phi}{\lambda_D},\ \cdot,,\ 1,\ 1\right) = 0.$$

Next, by Lemma OA.2.5 (b), we have $\mathcal{V}'(u) \geq -1$, which guarantees $b=1$ by Lemma OA.2.3. Additionally, Lemma OA.2.6 shows that $\mathcal{J}$ is nonpositive for any feasible deviations. Therefore, $\mathcal{V}$ solves (HJB$_V$) subject to (IC).

The concavity of $\mathcal{V}$ is shown in Lemma OA.2.5 (c). If $\hat{u}_1 = 0$, $\mathcal{V}(u) = V^{sd}(u|0)$ is differentiable for all $u \geq 0$. If $\hat{u}_1 > 0$, $V^d(u)$ is differentiable for all $u < \hat{u}_1$ and $V^{sd}(u|\hat{u}_1)$ is differentiable for all $u > \hat{u}$. By Lemma OA.2.5 (a), $\mathcal{V}$ is differentiable at $\hat{u}_1$ as well. Also note that $\mathcal{V}(0) = 0$. Therefore, by Lemma OA.2.2, for any incentive compatible contract promising the agent u units of utility, the principal's expected payoff is lower than or equal to $\mathcal{V}(u)$.

Last, by Lemma OA.2.4, there exists a contract implementing $(\mathcal{V}(u), u)$. Therefore, $\mathcal{V}(u)$ is the principal's maximized expected payoff subject to the promise-keeping constraint $U_0(\Gamma) = u$ and the incentive compatibility constraints. □

OA.2.4 Thresholds

In this section, I explain how to pin down the thresholds Π_D and Π_S and provide some properties of them.

First, recall that Π_S is the solution of $\frac{1}{\lambda_S} = \overline{\Delta}$. This gives us

$$\frac{1}{\lambda_S} = \frac{1}{\lambda_S} \log \left[\frac{\lambda_S \Pi_S - c}{(\lambda_S - \lambda_D)\Pi_S - c} \right] \qquad \Leftrightarrow \qquad \frac{\lambda_D \Pi_S}{c} = \frac{e-1}{e-2} \approx 2.392. \tag{OA.2.11}$$

Also recall that the threshold is relevant to whether the switching point $\hat{u}_1$ is greater than $\overline{u}$ or not. Since $\mathcal{V}$ is concave, $\hat{u}_1 \leq \overline{u}$ if and only if $\mathcal{V}'(\hat{u}_1) \geq 0$. Observe that by using the formula of V^d and (A.13), we can derive that

$$\mathcal{V}'(\hat{u}_1) = V^{d\prime}(\hat{u}_1) = \frac{(\lambda_D \Pi - c)^2}{\phi \cdot (\lambda_S \Pi - c)} e - 1.$$

By solving the equation making the above equal to zero, it follows that $\mathcal{V}'(\hat{u}_1) \geq 0$ if and only if

$$\Pi \leq \Pi_D \equiv \frac{c}{\lambda_D} \cdot \frac{c \cdot e + \phi + \sqrt{\phi(c \cdot e + \phi)}}{c \cdot e}. \tag{OA.2.12}$$

I conclude the section by showing that Π_D lies between Π_F and Π_S.

Lemma OA.2.7. *When $\lambda_S = 2\lambda_D$, $\Pi_D \in (\Pi_F, \Pi_S)$.*

Proof of Lemma OA.2.7. From $\phi \leq c$, we have

$$\frac{\lambda_D \Pi_D}{c} \leq \frac{(e+1) + \sqrt{e+1}}{e} \approx 2.077.$$

Therefore, by (OA.2.11), $\Pi_D < \Pi_S$.

Next, observe that

$$\frac{\lambda_D(\Pi_D - \Pi_F)}{c} = \frac{-\phi(e-1) + \sqrt{\phi(c \cdot e + \phi)}}{c \cdot e} \geq \frac{\phi(\sqrt{e+1} - (e-1))}{c \cdot e}.$$

Since $\sqrt{e+1} - (e-1) \approx .21 > 0$, $\Pi_D > \Pi_F$. ☐

OA.2.5 Proof of Lemmas

Proof of Lemma OA.2.1. The proof is inspired by Proposition 3.2.1 in Bertsekas (1995). Consider an arbitrary admissible action $\tilde{b} \in \mathcal{B}_t$. By rearranging (HJB$_{PK}$), we can derive that

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

and it is equivalent to

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

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

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

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 $0 \leq t \leq T$. Then, we have

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

for all $\tilde{b} \in \mathcal{B}_t$. Therefore, by (OA.2.1), we have $u_t = U_t(\Gamma)$. □

Proof of Lemma OA.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

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

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

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

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

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

$$
\begin{aligned}
&(\lambda_D a_t b_t + \lambda_S (1-a_t) b_t) f(a,b;0,t) \cdot \tilde{V}(u_t) - f(a,b;0,t) \cdot \tilde{V}'(u_t)\dot{u}_t \\
\geq & f(a,b;0,t) \left((\Pi - R_t)\lambda_D a_t b_t + V_S(u_{S,t})\lambda_S(1-a_t)b_t - c \right)
\end{aligned}
\tag{OA.2.13}
$$

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 (OA.2.13) over $[0, T]$ and noting that $f(a, b; 0, 0) = 1$, $u_T = 0$ and $\tilde{V}(0) = 0$, we have

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

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. □

Proof of Lemma OA.2.3. 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., $V'(u) = -\frac{c}{\phi} < -1$. It contradicts the assumption of $V'(u) \geq -1$. Therefore, b should be equal to 1 for the solution of (HJB_V) subject to (IC). □

Proof of Lemma OA.2.4. (a) Let $\Gamma_d(T)$ denote a direct-only contract with the deadline T.

The agent's expected payoff is

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

Therefore, $U_0(\Gamma_d(\frac{u}{\phi})) = u$.

Also note that the sum of the expected payoffs of the principal and the agent should equal to the expected surplus from the direct-only policy with a deadline of T:

$$P_0(\Gamma_d(T)) + U_0(\Gamma_d(T)) = \mathcal{W}^d(T).$$

Therefore,

$$P_0\left(\Gamma_d\left(\frac{u}{\phi}\right)\right) = \mathcal{W}^d\left(\frac{u}{\phi}\right) - u = V^d(u).$$

(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 $t \leq T_1$ 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

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

Also note that

$$\begin{aligned}
\int_0^{T_1} U_{\tau_s}(\hat{\Gamma}_{sd}(\tau_s|T_1, T))\lambda_S e^{-\lambda_S \tau_s} d\tau_s &= \int_0^{T_1} \phi(T + \frac{1}{\lambda_S} - \tau_s)\lambda_S e^{-\lambda_S \tau_s} d\tau_s \\
&= -\phi(T - \tau_s)e^{-\lambda_S \tau_s}\Big|_0^{T_1} \\
&= \phi T - \phi(T - T_1)e^{-\lambda_S T_1}.
\end{aligned}$$

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

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

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

As in the previous case, the sum of the expected payoffs of the principal and the agent is equal to the one-switch policy with the intermediate deadline T_1, the deadline T, and the extension $\frac{1}{\lambda_S}$:

$$P_0(\Gamma_{sd}(T_1, T)) + U_0(\Gamma_{sd}(T_1, T)) = \mathcal{W}^{sd}_{T-T_1}(T, \tfrac{1}{\lambda_S}).$$

By plugging in $T = \frac{u}{\phi}$ and $T_1 = \frac{u-u_1}{\phi}$, (OA.2.6) holds.

(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 $\frac{u}{\phi}$ implements $(V^{sd}(u|0), u)$.

(d) By the construction of $\mathcal{W}^d$,

$$\dot{\mathcal{W}}^d(T) = \lambda_D(\Pi - \mathcal{W}^d(T)) - c, \tag{OA.2.14}$$

for all $T \geq 0$. Similarly,

$$\dot{\mathcal{W}}^{sd}_{\hat{T}}(T, \tfrac{1}{\lambda_S}) = \lambda_S(W^1_{T+\frac{1}{\lambda_S}} - \mathcal{W}^{sd}_{\hat{T}}(T, \tfrac{1}{\lambda_S})) - c, \tag{OA.2.15}$$

for all $T \geq \hat{T}$.

Using the definitions of V^d, V^{sd} and V_S, (OA.2.7) and (OA.2.8) follow.

□

Proof of Lemma OA.2.5. (a) Suppose that $\hat{u}_1 > 0$, which implies that $\frac{1}{\lambda_S} < \overline{\Delta}$. Now set $\Delta = \frac{1}{\lambda_S}$. In Lemma A.1, $\hat{T}$ is chosen to satisfy $\lambda_D(\Pi - \mathcal{W}^d(\hat{T})) = \lambda_S(W^1_{\hat{T}+\frac{1}{\lambda_S}} - \mathcal{W}^d(\hat{T}))$ and $\lambda_D(\Pi - \mathcal{W}^d(z)) > \lambda_S(W^1_{z+\frac{1}{\lambda_S}} - \mathcal{W}^d(z))$ for all $z < \hat{T}$.

Using Lemma OA.2.4 (d) and $\mathcal{W}^d(\hat{T}) = \mathcal{W}^{sd}_{\hat{T}}(\hat{T}, \frac{1}{\lambda_S})$, we can derive that $V^{sd\prime}(\hat{u}_1|\hat{u}_1) = V^{d\prime}(\hat{u}_1)$ and $V^{sd\prime}(u|u) < V^{d\prime}(u)$ for all $u < \hat{u}_1$.

When $\hat{u}_1 = 0$, and thereby $\frac{1}{\lambda_S} \geq \overline{\Delta}$, it follows from (A.11) that $V^{sd\prime}(\hat{u}_1|\hat{u}_1) \geq V^{d\prime}(\hat{u}_1)$.

(b) Since a longer the deadline increases the expected social surplus, $\mathcal{W}^*(T, \Delta)$ is increasing in T. Therefore,

$$\mathcal{V}'(u) = \frac{\dot{\mathcal{W}}^*(\frac{u}{\phi}, \frac{1}{\lambda_S})}{\phi} - 1 \geq -1.$$

(c) If $u \le \hat{u}_1$,

$$\mathcal{V}''(u) = V^{d''}(u) = -\left(\Pi - \frac{c}{\lambda_D}\right)\frac{\lambda_D^2}{\phi^2}e^{-\frac{\lambda_D}{\phi}u} < 0.$$

Now assume that $u > \hat{u}_1$. By using (OA.2.14), (OA.2.15) and $V^{d'}(\hat{u}_1) \le V^{sd'}(\hat{u}_1|\hat{u}_1)$, we can derive that

$$\mathcal{W}^d(\tfrac{\hat{u}_1}{\phi}) \le \Pi - \frac{c}{\lambda_S - \lambda_D} - \frac{\lambda_S}{\lambda_S - \lambda_D}\left(\Pi - \frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}\hat{u}_1 - 1}. \tag{OA.2.16}$$

Using (A.10), we can derive the followings though some algebra:

$$\begin{aligned} V^{sd''}(u|\hat{u}_1) = &\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(\hat{u}_1 - u)}\left[\mathcal{W}^d(\tfrac{\hat{u}_1}{\phi}) - \left(\Pi - \frac{2c}{\lambda_S}\right) + 2\left(\Pi - \frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}\hat{u}_1 - 1}\right] \\ &- \left(\Pi - \frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)^3 (u - \hat{u}_1)e^{-\frac{\lambda_S}{\phi}(u + \frac{1}{\lambda_S})} \end{aligned}$$

By plugging (OA.2.16) in, we have

$$\begin{aligned} V^{sd''}(u|\hat{u}_1) \le &\left(\frac{\lambda_S}{\phi}\right)^2 e^{\frac{\lambda_S}{\phi}(\hat{u}_1 - u)}\left[\frac{\lambda_S - 2\lambda_D}{\lambda_S(\lambda_S - \lambda_D)}c + \frac{\lambda_S - 2\lambda_D}{\lambda_S - \lambda_D}\left(\Pi - \frac{c}{\lambda_S}\right)e^{-\frac{\lambda_S}{\phi}\hat{u}_1 - 1}\right] \\ &- \left(\Pi - \frac{c}{\lambda_S}\right)\left(\frac{\lambda_S}{\phi}\right)^3 (u - \hat{u}_1)e^{-\frac{\lambda_S}{\phi}(u + \frac{1}{\lambda_S})}. \end{aligned} \tag{OA.2.17}$$

Then, from $\lambda_S = 2\lambda_D$, $\mathcal{V}''(u) = V^{sd''}(u|\hat{u}_1) \le 0$ for all $u \ge \hat{u}_1$.

□

Proof of Lemma OA.2.6. I begin by showing $L^D(u, R) \le 0$ for all $R \ge u + \frac{\phi}{\lambda_D}$. Observe that

$$\frac{\partial L^D}{\partial R} = -\lambda_D(1 + \mathcal{V}'(u)) \le 0.$$

from Lemma OA.2.5 (b). Also note that

$$\begin{aligned} L^D(u, u + \tfrac{\phi}{\lambda_D}) = &\lambda_D(\Pi - u - \mathcal{V}(u)) - c - \phi(\mathcal{V}'(u) + 1) \\ = &\lambda_D(\Pi - \mathcal{W}^*(\tfrac{u}{\phi}, \tfrac{1}{\lambda_S})) - c - \dot{\mathcal{W}}^*(\tfrac{u}{\phi}, \tfrac{1}{\lambda_S}) \le 0. \end{aligned}$$

The last inequality is due to (HJB$_W$). Therefore, $L^D(u, R) \le 0$ for all $R \ge u + \frac{\phi}{\lambda_D}$.

For L^S, observe that

$$\frac{\partial L^S}{\partial u} = -\lambda_S(w - u)\mathcal{V}''(u) \ge 0$$

by the concavity of $\mathcal{V}$. Therefore, it is sufficient to check whether $L^S(u, u + \frac{\phi}{\lambda_S}) \leq 0$ for all $u \geq 0$.

Note that for all $u \geq \hat{u}_1$, $L^S(u, u + \frac{\phi}{\lambda_S}) = 0$ holds by (OA.2.8). Now, suppose that $u < \hat{u}_1$, thereby $\mathcal{V}(u) = V^d(u)$. Using (OA.2.8) and Lemma OA.2.5 (a), we have $L^S(u, u + \frac{\phi}{\lambda_S}) = \phi(V^{sd'}(u|u) - V^{d'}(u)) < 0$ for all $u < \hat{u}_1$. □

OA.3 Proofs for Section 5

In this section, I provide the proofs for the value function and optimal contract characterizations when there is an efficiency loss from splitting the project (Proposition OA.3.1, Theorem 3, and Theorem 4). Some results in Section OA.2 can still be utilized (e.g., Lemma OA.2.1, Lemma OA.2.2, Lemma OA.2.3, and Lemma OA.2.4) as they do not use the parametric assumption $\lambda_S = 2\lambda_D$.

OA.3.1 Value Function Characterization

I begin by specifying a value function that can be implemented by a two-switch contract defined in Definition 5.1.

Lemma OA.3.1. *The following statements hold.*

(a) When $0 < u_1 < u_2 < u$, a two-switch contract with the intermediate deadlines $\frac{u-u_2}{\phi}$, $\frac{u-u_1}{\phi}$ and the final deadline $\frac{u}{\phi}$ implements $(V^{dsd}(u|u_1, u_2), u)$ where

$$V^{dsd}(u|u_1, u_2) \equiv \left(\Pi - \frac{c}{\lambda_D}\right)\left(1 - e^{\frac{\lambda_D}{\phi}(u_2 - u)}\right) + (V^{sd}(u_2|u_1) + u_2)e^{\frac{\lambda_D}{\phi}(u_2 - u)} - u. \tag{OA.3.1}$$

(b) The following differential equation holds:

$$\phi V^{dsd'}(u|u_1, u_2) = \lambda_D\left(\Pi - \frac{\phi}{\lambda_D} - u - V^{dsd}(u|u_1, u_2)\right) - c. \tag{OA.3.2}$$

Based on the intuition presented in the main text, I conjecture the value function defined as follows.

$$\mathcal{V}(u) = \begin{cases} V^d(u), & \text{if } 0 \le u \le \hat{u}_1, \\ V^{sd}(u|\hat{u}_1), & \text{if } \hat{u}_1 < u \le \hat{u}_2, \\ V^{dsd}(u|\hat{u}_1, \hat{u}_2), & \text{if } \hat{u}_2 < u. \end{cases} \tag{OA.3.3}$$

The following proposition shows that there exist $\hat{u}_1$ and $\hat{u}_2$ such that the above three value functions are smoothly pasted, and the resulting function is the principal's value function.

Proposition OA.3.1. *Suppose that η is less than 1.*

(a) ***(Smooth Pasting)*** *There exist $\hat{u}_2 \ge \hat{u}_1 \ge 0$ such that*

i. $V^{d\prime}(u) > V^{sd\prime}(u|u)$ *for all* $0 \le u < \hat{u}_1$;

ii. $V^{sd\prime}(\hat{u}_1|\hat{u}_1) \ge V^{d\prime}(\hat{u}_1)$, *and if the equality holds,* $V^{sd\prime\prime}(\hat{u}_1|\hat{u}_1) > V^{d\prime\prime}(\hat{u}_1)$;

iii. $V^{sd\prime}(u|\hat{u}_1) > V^{dsd\prime}(u|\hat{u}_1, u)$ *for all* $\hat{u}_1 < u < \hat{u}_2$;

iv. $V^{dsd\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2) = V^{sd\prime}(\hat{u}_2|\hat{u}_1)$ *and* $V^{dsd\prime\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2) > V^{sd\prime\prime}(\hat{u}_2|\hat{u}_1)$;

v. $V^{dsd\prime}(u|\hat{u}_1, \hat{u}_2) > \frac{1}{\phi}\left[\lambda_S(V_S(u + \frac{\phi}{\lambda_S}) - V^{dsd}(u|\hat{u}_1, \hat{u}_2)) - c\right]$ *for all* $u > \hat{u}_2$.

(b) ***(Large Loss)*** *If* $\eta \le \frac{1}{e-1}$, *there exists* $\tilde{\Pi}_M(\eta)$ *such that*

i. $\hat{u}_2 = \hat{u}_1 = 0$ *if* $\tilde{\Pi}_M(\eta) \ge \Pi > \frac{c}{\lambda_D}$;

ii. $\hat{u}_2 > \hat{u}_1 > 0$ *if* $\Pi > \tilde{\Pi}_M(\eta)$.

(c) ***(Small Loss)*** *If* $\frac{1}{e-1} < \eta < 1$, *there exist* $\tilde{\Pi}_S(\eta) > \tilde{\Pi}_M(\eta)$ *such that*

i. $\hat{u}_2 = \hat{u}_1 = 0$ *if* $\tilde{\Pi}_M(\eta) \ge \Pi > \frac{c}{\lambda_D}$;

ii. $\hat{u}_2 > \hat{u}_1 > 0$ *if* $\tilde{\Pi}_S(\eta) \ge \Pi > \tilde{\Pi}_M(\eta)$;

iii. $\hat{u}_2 > \hat{u}_1 = 0$ *if* $\Pi \ge \tilde{\Pi}_S(\eta)$.

(d) The function $\mathcal{V}$ defined in (OA.3.3), with $\hat{u}_1$ and $\hat{u}_2$ derived in (a), serves as the principal's value function.

OA.3.1.1 Proof of Proposition OA.3.1

I begin by identifying which approach will be chosen at the deadline. Note that the direct approach is chosen at the deadline if and only if $V^{d'}(0) > V^{sd'}(0|0)$. The following lemma provides the parametric condition for this.

Lemma OA.3.2. *If $\eta \leq \frac{1}{e-1}$, the inequality $V^{d'}(0) > V^{sd'}(0|0)$ always holds. If $\eta > \frac{1}{e-1}$, $V^{d'}(0) > V^{sd'}(0|0)$ is equivalent to*

$$\Pi < \tilde{\Pi}_S(\eta) \equiv \frac{e-1}{\eta(e-1)-1} \cdot \frac{c}{\lambda_D}.$$

Moreover, if $\Pi = \tilde{\Pi}_S(\eta)$, then $V^{d'}(0) = V^{sd'}(0|0)$ and $V^{d''}(0) < V^{sd''}(0|0)$.

Next, I establish a condition under which the direct approach is always employed. When this condition does not hold, a switch from the direct to the sequential approach occurs. I show the existence of the switching point $\hat{u}_1$.

Lemma OA.3.3. *There exists $\tilde{\Pi}_M(\eta) \geq \frac{2c}{\lambda_S}$ with $\tilde{\Pi}_M(1) = \frac{2c}{\lambda_S} = \frac{c}{\lambda_D}$ such that the following statements hold.*

(a) If $\frac{c}{\lambda_D} \leq \Pi < \tilde{\Pi}_M(\eta)$, $V^{d'}(u) > V^{sd'}(u|u)$ for all $u \geq 0$.

(b) Suppose that one of the following statements hold: (i) $\eta \leq \frac{1}{e-1}$ and $\Pi > \tilde{\Pi}_M(\eta)$; (ii) $\eta > \frac{1}{e-1}$ and $\tilde{\Pi}_S(\eta) \geq \Pi > \tilde{\Pi}_M(\eta)$. Then, there exists $\hat{u}_1 > 0$ such that $V^{d'}(\hat{u}_1) = V^{sd'}(\hat{u}_1|\hat{u}_1)$, $V^{d''}(\hat{u}_1) < V^{sd''}(\hat{u}_1|\hat{u}_1)$ and $V^{d'}(u) > V^{sd'}(u|u)$ for all $u \in [0, \hat{u}_1)$;

The following lemma shows that when there is an efficiency loss from splitting the project and the sequential approach is employed, there will be an additional switching point, $\hat{u}_2$.

Lemma OA.3.4. *Suppose that $\eta < 1$, $\Pi > \frac{c}{\lambda_D}$ and one of the followings hold: (i) $V^{d'}(\hat{u}_1) < V^{sd'}(\hat{u}_1|\hat{u}_1)$; (ii) $V^{d'}(\hat{u}_1) = V^{sd'}(\hat{u}_1|\hat{u}_1)$ and $V^{d''}(\hat{u}_1) < V^{sd''}(\hat{u}_1|\hat{u}_1)$. Then, there exists $\hat{u}_2 > \hat{u}_1$ such that $V^{sd'}(\hat{u}_2|\hat{u}_1) = V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ and $V^{sd''}(\hat{u}_2|\hat{u}_1) < V^{dsd''}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ and such $\hat{u}_2$ is unique. Moreover, $V^{sd'}(u|\hat{u}_1) > V^{dsd'}(u|\hat{u}_1, u)$ for all $u \in (\hat{u}_1, \hat{u}_2)$.*

Next, when there is a switching point $\hat{u}_2$, the following lemma shows that the direct approach is employed for all $u > \hat{u}_2$.

Lemma OA.3.5. *Suppose that $\Pi > \frac{c}{\lambda_D}$, $V^{sd'}(\hat{u}_2|\hat{u}_1) = V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ and $V^{sd''}(\hat{u}_2|\hat{u}_1) < V^{dsd''}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$. Then, $\lambda_S\left(V_S(u + \frac{\phi}{\lambda_S}) - V^{dsd}(u|\hat{u}_1, \hat{u}_2)\right) - \phi V^{dsd'}(u|\hat{u}_1, \hat{u}_2) - c < 0$ for all $u > \hat{u}_2$.*

Lastly, I show that the resulting value function is concave, and that L^D and L^S—the functions specifying deviations, defined in (OA.2.10) and (OA.2.3)—are nonpositive.

Lemma OA.3.6. *Suppose that $\Pi > \frac{c}{\lambda_D}$ and $\eta < 1$.*

(a) $\mathcal{V}$ is concave;

(b) for any $u \geq 0$, $L^D(u, R) \leq 0$ for all $R \geq u + \frac{\phi}{\lambda_D}$, and $L^S(u, w) \leq 0$ for all $w \geq u + \frac{\phi}{\lambda_S}$.

Proof of Proposition OA.3.1. I start by showing that, for each condition in (b) and (c), the switching points are as stated and the conditions in (a) also hold.

(b-i & c-i) By Lemma OA.3.3 (a), $V^{d\prime}(u) > V^{sd\prime}(u|u)$ for all $u > 0$. Note that $V^d(u) = V^{dsd}(u|0,0)$ and

$$\phi V^{sd\prime}(u|u) = \lambda_S(V_S(u + \tfrac{\phi}{\lambda_S}) - V^{dsd}(u|0,0)) - c$$

by (OA.2.8) and $V^{sd}(u|u) = V^d(u)$. Therefore, with $\hat{u}_1 = \hat{u}_2 = 0$, the conditions (a-i)–(a-iv) hold trivially, and the condition (a-v) holds as demonstrated above.

(b-ii & c-ii) By Lemma OA.3.3 (b), there exists $\hat{u}_1 > 0$ such that the conditions (a-i) and (a-ii) hold. Next, by Lemma OA.3.4, there exists $\hat{u}_2 > \hat{u}_1$ such that the conditions (a-iii) and (a-iv) hold. By Lemma OA.3.5, the condition (a-v) holds.

(c-iii) By Lemma OA.3.2, $V^{sd\prime}(0|0) > V^d(0)$. By setting $\hat{u}_1 = 0$, the conditions (a-i) and (a-ii) hold trivially. Next, by Lemma OA.3.4, there exists $\hat{u}_2 > 0$ such that the conditions (a-iii) and (a-iv) hold. By Lemma OA.3.5, the condition (a-v) holds.

(d) Following the same steps of the proof of Proposition 4.1, $\mathcal{V}$ solves (HJB$_V$) subject to (IC). □

OA.3.2 Proofs of Theorem 3 and Theorem 4

Lemma OA.3.7. *Suppose that $\Pi > \tilde{\Pi}_M(\eta)$ and $\eta \leq \frac{c}{c+\phi}$. Then, $\hat{u}_2$ is less than $\bar{u}$.*

Lemma OA.3.8. *Suppose that $\eta > \overline{\eta} = \max\{\frac{1}{e-1}, \sqrt{\frac{c}{c+\phi}}\}$. There exists $\tilde{\Pi}_D(\eta) \in (\tilde{\Pi}_M(\eta), \tilde{\Pi}_S(\eta))$ such that $\hat{u}_1 < \overline{u}$ if and only if $\Pi > \tilde{\Pi}_D(\eta)$.*

Lemma OA.3.9. *Suppose that* $\eta > \overline{\eta} = \sqrt{\frac{c}{c+\phi}}$*. Then,* $\hat{u}_2 \geq \overline{u}$*.*

Proof of Theorem 3. Note that $\hat{u}_2$ is always greater than $\bar{u}$ by Lemma OA.3.9 since $\eta > \sqrt{\frac{c}{c+\phi}}$. Additionally, using Lemma OA.3.2, Lemma OA.3.3 and Lemma OA.3.8, we have

$$\mathcal{V}(\overline{u}) = \begin{cases} V^d(\overline{u}), & \text{if } \Pi_F < \Pi < \tilde{\Pi}_D(\eta), \\ V^{sd}(\overline{u}|\hat{u}_1), & \text{if } \tilde{\Pi}_D(\eta) < \Pi < \tilde{\Pi}_S(\eta), \\ V^{sd}(\overline{u}|0), & \text{if } \tilde{\Pi}_S(\eta) < \Pi. \end{cases}$$

As in Theorem 2, the value functions above can be implemented by direct-only, one-switch and sequential-only contracts, respectively. □

Proof of Theorem 4. By Proposition OA.3.1 (b-i), when $\Pi \leq \tilde{\Pi}_M(\eta)$, $\hat{u}_1 = \hat{u}_2 = 0$. By Proposition OA.3.1 (b-ii) and Lemma OA.3.7, when $\Pi > \tilde{\Pi}_M(\eta)$, $\overline{u} > \hat{u}_2 > \hat{u}_2 > 0$. Therefore,

$$\mathcal{V}(\overline{u}) = \begin{cases} V^{dsd}(\overline{u}|0,0) = V^d(\overline{u}), & \text{if } \Pi_F < \Pi \leq \tilde{\Pi}_M(\eta), \\ V^{dsd}(\overline{u}|\hat{u}_1, \hat{u}_2), & \text{if } \tilde{\Pi}_M(\eta). \end{cases}$$

By Lemma OA.3.1, $(V^{dsd}(\overline{u}|\hat{u}_1, \hat{u}_2), \overline{u})$ can be implemented by a two-switch contract. Therefore, when $\Pi \in (\Pi_F, \tilde{\Pi}_M(\eta)]$, the direct-only contract is optimal, and when $\Pi > \tilde{\Pi}_M(\eta)$, there exists a two-switch contract that is optimal. □

OA.3.3 Proof of Lemmas

Proof of Lemma OA.3.1. (a) 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 $\Gamma_{sd}(T_2 - T_1, T - T_1)$. Recall that $U_0(\Gamma_{sd}(T_2 - T_1, T - T_1)) = \phi(T - T_1)$. Then, the agent's expected payoff at time 0 is

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

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

Also note that

$$
\begin{aligned}
&P_0(\Gamma_{dsd}(T_1,T_2,T))+U_0(\Gamma_{dsd}(T_1,T_2,T))\\
=&\int_0^{T_1}(\Pi-c\tau_m)\lambda_D e^{-\lambda_D\tau_m}d\tau_m-cT_1e^{-\lambda_D T_1}\\
&+e^{-\lambda_D T_1}(P_0(\Gamma_{sd}(T_2-T_1,T-T_1))+U_0(\Gamma_{sd}(T_2-T_1,T-T_1))).
\end{aligned}
$$

Recall that $U_0(\Gamma_{sd}(T_2',T_1'))+P_0(\Gamma_{sd}(T_2',T_1'))=V^{sd}(\phi T_1'|\phi(T_1'-T_2'))+\phi(T_1'-T_2')$. By plugging in $T_1'=T-T_1$, $T_2'=T_2-T_1$, $\phi T_1=u-u_2$ and $\phi T_2=u-u_1$, the right hand side of the above equation becomes $V^{dsd}(u)+u$, thus, $P_0(\Gamma_{dsd}(T_1,T_2,T))=V^{dsd}(u|u_1,u_2)-u$.

(b) Last, by taking the derivative of (OA.3.1) and multiplying by ϕ, we have

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

thus, (OA.3.2) holds.

□

Proof of Lemma OA.3.2. By $V^d(0)=V^{sd}(0|0)=0$ and Lemma OA.2.4 (d), we have

$$
\begin{aligned}
\phi V^{d\prime}(0)&=\lambda_D\Pi-\phi-c,\\
\phi V^{sd\prime}(0|0)&=\lambda_S V_S(\frac{\phi}{\lambda_S})-c=\lambda_S\left(\Pi-\frac{c}{\lambda_S}\right)(1-e^{-1})-\phi-c.
\end{aligned}
$$

Therefore, $V^{d\prime}(0)>V^{sd\prime}(0|0)$ is equivalent to:

$$(\eta(e-1)-1)\,\lambda_D\Pi<c(e-1).$$

Therefore, when $\eta\le\frac{1}{e-1}$, $V^{d\prime}(0)>V^{sd\prime}(0|0)$ always holds, and when $\eta>\frac{1}{e-1}$, $V^{d\prime}(0)>V^{sd\prime}(0|0)$ is equivalent to $\Pi<\tilde{\Pi}_S(\eta)$.

Next, assume that $\eta>\frac{1}{e-1}$ and $\Pi=\tilde{\Pi}_S(\eta)$. With some algebra, it follows that

$$\phi^2V^{sd\prime\prime}(0|0)-\phi^2V^{d\prime\prime}(0)=\lambda_D c\left[\frac{(e-1)\eta^2}{(e-1)\eta-1}\right].$$

The right hand side is positive from $\eta > \frac{1}{e-1}$, thus, $V^{sd''}(0|0) > V^{d''}(0)$. $\square$

Proof of Lemma OA.3.3. Define $x \equiv e^{-\frac{\lambda_D}{\phi}u}$. Using (OA.2.7) and (OA.2.8), $\phi V^{sd'}(u|u) - \phi V^{d'}(u)$ can be expressed in the form of $H_1(x)$ with $\Delta = \frac{1}{\lambda_S}$, as defined in (A.12) in the proof of Lemma A.1. I also consider this as a function of Π. With the definition of η, it can be rewritten as follows:

$$\tilde{H}_1(x;\Pi) \equiv -\left\{(\eta+1)\lambda_D\Pi - c\right\} e^{-1}x^{\eta+1} + \eta(\lambda_D\Pi - c)x - (1-\eta)c. \tag{OA.3.4}$$

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 $\Pi \geq \frac{c}{\lambda_D}$. Let $x^*(\Pi)$ be the solution of $\max_x H_1(x;\Pi)$ subject to $0 \leq x \leq 1$. Then, when $\Pi \geq \frac{c}{\lambda_D}$, from the first order condition, we can derive that

$$x^*(\Pi) = \left[\frac{\eta(\lambda_D\Pi - c)}{(\eta+1)\{(\eta+1)\lambda_D\Pi - c\}e^{-1}}\right]^{\frac{1}{\eta}}. \text{[17]} \tag{OA.3.5}$$

Now define

$$h(\Pi) \equiv \tilde{H}_1(x^*(\Pi);\Pi) = K\left(\frac{\lambda_D\Pi - c}{\lambda_S\Pi - c}\right)^{\frac{1}{\eta}}(\lambda_D\Pi - c) - (1-\eta)c$$

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 $\eta(1-\eta)e \leq \frac{e}{4} < 1 \leq (\eta+1)^2$. In addition, $\lim_{\Pi\to\infty} h(\Pi) = \infty$ and

$$h'(\Pi) = K(\lambda_D\Pi - c)^{\frac{1}{\eta}}(\lambda_S\Pi - c)^{-\frac{1}{\eta}-1}\lambda_D\lambda_S\Pi > 0.$$

Therefore, there exists a unique Π such that $h(\Pi) = 0$ and $\Pi \geq \frac{2c}{\lambda_S}$. Let the solution of $h(\Pi) = 0$ with $\Pi \geq \frac{2c}{\lambda_S}$ be $\tilde{\Pi}_M(\eta)$. Also note that when $\eta = 1$, $h(\frac{2c}{\lambda_S}) = 0$ thus $\tilde{\Pi}_M(1) = \frac{2c}{\lambda_S} =$

[17] When $\Pi \leq \frac{(\eta+1)e^{-1}-\eta}{(\eta+1)^2e^{-1}-\eta} \cdot \frac{c}{\lambda_D}$, the solution of the maximization problem $\max_{0\leq x\leq 1} H_1(x;\Pi)$ 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 $\Pi \geq \frac{c}{\lambda_D}$

$\frac{c}{\lambda_D}$.

(a) Suppose that $\frac{c}{\lambda_D} \leq \Pi < \tilde{\Pi}_M(\eta)$. We have $0 > h(\Pi) = \tilde{H}_1(x^*(\Pi);\Pi) \geq \tilde{H}_1(x;\Pi)$ for all $0 \leq x \leq 1$. It is equivalent to $V^{d'}(u) > V^{sd'}(u|u)$ for all $u \geq 0$ in this case.

(b) First, suppose that $\eta \leq \frac{1}{e-1}$ and $\Pi > \tilde{\Pi}_M(\eta)$. Then, we have $0 < h(\Pi) = \tilde{H}_1(x^*(\Pi);\Pi)$. In addition, by Lemma OA.3.2, we have $\tilde{H}_1(1;\Pi) = \phi(V^{sd'}(0|0) - V^{d'}(0)) < 0$. 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 $\hat{u}_1 \equiv -\frac{\phi}{\lambda_D}\log x_1$, the above conditions can be translated into: $V^{d'}(\hat{u}_1) = V^{sd'}(\hat{u}_1|\hat{u}_1)$, $V^{d''}(\hat{u}_1) < V^{sd''}(\hat{u}_1|\hat{u}_1)$ and $V^{d'}(u) > V^{sd'}(u|u)$ for all $u \in [0, \hat{u}_1)$.

   Next, suppose that $\eta > \frac{1}{e-1}$. Note that by the definition of $\tilde{\Pi}_S(\eta)$, if $\Pi \geq \tilde{\Pi}_S(\eta)$, $\tilde{H}_1(1;\Pi) \geq 0$. It implies that $h(\Pi) \geq \tilde{H}_1(1;\Pi) \geq 0$ and $\Pi \geq \tilde{\Pi}_M(\eta)$. Therefore, we can see that $\tilde{\Pi}_S(\eta) \geq \tilde{\Pi}_M(\eta)$. If $\tilde{\Pi}_S(\eta) \geq \Pi > \tilde{\Pi}_M(\eta)$, 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 $\hat{u}_1 > 0$ such that $V^{d'}(\hat{u}_1) = V^{sd'}(\hat{u}_1|\hat{u}_1)$, $V^{d''}(\hat{u}_1) < V^{sd''}(\hat{u}_1|\hat{u}_1)$ and $V^{d'}(u) > V^{sd'}(u|u)$ for all $u \in [0, \hat{u}_1)$.

□

Proof of Lemma OA.3.4. Using (OA.3.2), (OA.2.8) and $V^{dsd}(u|\hat{u}_1, u) = V^{sd}(u|\hat{u}_1)$, $\phi V^{dsd'}(u|\hat{u}_1, u) - \phi V^{sd'}(u|\hat{u}_1)$ can be rewritten as follows:

$$\lambda_D \Pi - \lambda_S \left\{ V_S(u + \frac{\phi}{\lambda_S}) + u + \frac{\phi}{\lambda_S} \right\} + (\lambda_S - \lambda_D)(V^{sd}(u|\hat{u}_1) + u).$$

By performing a similar derivation as in (A.14) and using $\eta = \frac{\lambda_S}{\lambda_D} - 1$ and $y \equiv e^{-\frac{\lambda_S}{\phi}(u-\hat{u}_1)}$, the above expression can be further rewritten as follows:

$$\begin{aligned} \tilde{H}_2(y) \;\equiv\; & \frac{1-\eta}{1+\eta}c + (\lambda_S \Pi - c)e^{-1-\frac{\lambda_S}{\phi}\hat{u}_1}\left[1 + \frac{\eta}{1+\eta}\log y\right] y \\ & +\eta\left[\frac{1-\eta}{1+\eta}c - (\lambda_D \Pi - c)e^{-\frac{\lambda_D}{\phi}\hat{u}_1}\right] y. \end{aligned} \tag{OA.3.6}$$

Note that $\tilde{H}_2(1) = \phi V^{dsd'}(\hat{u}_1|\hat{u}_1, \hat{u}_1) - \phi V^{sd'}(\hat{u}_1|\hat{u}_1) = \phi V^{d'}(\hat{u}_1) - \phi V^{sd'}(\hat{u}_1|\hat{u}_1) \leq 0$ by assumption. By differentiating $\tilde{H}_2$ twice, we have

$$\tilde{H}_2''(y) = \frac{\eta}{1+\eta}(\lambda_S \Pi - c)e^{-1-\frac{\lambda_S}{\phi}\hat{u}_1}\frac{1}{y} > 0.$$

Since $\Pi > \frac{c}{\lambda_D} > \frac{c}{\lambda_S}$, $\tilde{H}_2$ is strictly convex in y. Also note that

$$\lim_{y\to 0} \tilde{H}_2(y) = \frac{1-\eta}{1+\eta}c > 0.$$

By the convexity of $\tilde{H}_2$, there exists $y_2 \in (0,1)$ such that (i) $\tilde{H}_2(y) < 0$ for all $y \in (y_2, 1)$, (ii) $\tilde{H}_2(y_2) = 0$, and (iii) $\tilde{H}_2'(y_2) < 0$. Let $\hat{u}_2 = \hat{u}_1 - \frac{\phi}{\lambda_S}\log y_2$. Then, from (i) and (ii), we have $V^{sd'}(u|\hat{u}_1) > V^{dsd'}(u|\hat{u}_1, u)$ for all $u \in (\hat{u}_1, \hat{u}_2)$ and $V^{sd'}(\hat{u}_2|\hat{u}_1) = V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$. Additionally, since y is decreasing in u, $\tilde{H}_2'(y_2) < 0$ implies that $V^{sd''}(\hat{u}_2|\hat{u}_1) < V^{dsd''}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$. $\square$

Proof of Lemma OA.3.5. By differentiating (OA.2.7) and (OA.2.8), we have

$$\phi V^{sd''}(u|\hat{u}_1) = \lambda_S \left(V_S'\left(u + \tfrac{\phi}{\lambda_S}\right) + 1\right) - \lambda_S\left(V^{sd'}(u|\hat{u}_1) + 1\right),$$

$$\phi V^{dsd''}(u|\hat{u}_1, \hat{u}_2) = -\lambda_D\left(V^{dsd'}(u|\hat{u}_1, \hat{u}_2) + 1\right).$$

Then, $V^{sd'}(\hat{u}_2|\hat{u}_1) = V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ and $V^{sd''}(\hat{u}_2|\hat{u}_1) < V^{dsd''}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ imply that

$$(\lambda_S - \lambda_D)(1 + V^{sd'}(\hat{u}_2|\hat{u}_1)) > \lambda_S(V_S'(\hat{u}_2 + \tfrac{\phi}{\lambda_S}) + 1)$$

$$\iff \qquad \eta(1 + V^{sd'}(\hat{u}_2|\hat{u}_1)) > (\eta+1)\left(\frac{\lambda_S\Pi - c}{\phi}\right)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}. \tag{OA.3.7}$$

Define a function $H_3 : [\hat{u}_2, \infty) \to \mathbb{R}$ as

$$H_3(u) \equiv \lambda_S\left[V_S(u + \tfrac{\phi}{\lambda_S}) - V^{dsd}(u|\hat{u}_1, \hat{u}_2)\right] - \phi V^{dsd'}(u|\hat{u}_1, \hat{u}_2) - c.$$

With some algebra, we can derive that

$$H_3(u) = (\eta - 1)\,c - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}\cdot e^{\frac{\lambda_S}{\phi}(\hat{u}_2 - u)} + \eta\phi\left(V^{sd'}(\hat{u}_2|\hat{u}_1) + 1\right)e^{\frac{\lambda_D}{\phi}(\hat{u}_2 - u)}.$$

Also note that

$$\begin{aligned} H_3(\hat{u}_2) &= \lambda_S\left[V_S(\hat{u}_2 + \tfrac{\phi}{\lambda_S}) - V^{sd}(\hat{u}_2|\hat{u}_1)\right] - c - \phi V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2) \\ &= \phi V^{sd'}(\hat{u}_2|\hat{u}_1) - \phi V^{dsd'}(\hat{u}_2|\hat{u}_1, \hat{u}_2) = 0. \end{aligned}$$

Define $x \equiv e^{\frac{\lambda_D}{\phi}(\hat{u}_2 - u)}$. Then, $H_3(u)$ can be rewritten as follows:

$$\tilde{H}_3(x) = (\eta-1)c - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}x^{\eta+1} + \eta\phi\left(V^{sd'}(\hat{u}_2|\hat{u}_1) + 1\right)x$$

and $\tilde{H}_3(1) = H_3(\hat{u}_2) = 0$.

Note that

$$\tilde{H}_3'(x) = -(\eta+1)(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}x^\eta + \eta\phi\left(V^{sd'}(\hat{u}_2|\hat{u}_1) + 1\right).$$

By (OA.3.7), we can derive that

$$\tilde{H}_3'(1) = -(\eta+1)(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1} + \eta\phi\left(V^{sd'}(\hat{u}_2|\hat{u}_1) + 1\right) > 0.$$

Also note that

$$\tilde{H}_3''(x) = -(\eta+1)\eta(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}x^{\eta-1} < 0.$$

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, $\lambda_S\left(V_S(u + \frac{\phi}{\lambda_S}) - V^{dsd}(u|\hat{u}_1,\hat{u}_2)\right) - \phi V^{dsd'}(u|\hat{u}_1,\hat{u}_2) - c < 0$ for all $u \geq \hat{u}_2$. □

Proof of Lemma OA.3.6. (a) When $u < \hat{u}_1$, $\mathcal{V}''(u) = V^{d''}(u) < 0$ from Lemma OA.2.5 (c).

When $\hat{u}_1 < u < \hat{u}_2$, the inequality (OA.2.17) is still applicable, and from $\lambda_S < 2\lambda_D$, we have $\mathcal{V}''(u) = V^{sd''}(u|\hat{u}_1) \leq 0$.

When $u > \hat{u}_2$, by differentiating (OA.3.1) twice, we have

$$V^{dsd''}(u|\hat{u}_1,\hat{u}_2) = -\left(\frac{\lambda_D}{\phi}\right)^2 \cdot \left(\Pi - \frac{c}{\lambda_D} - (V^{sd}(\hat{u}_2|\hat{u}_1,\hat{u}_2) + \hat{u}_2)\right)e^{-\frac{\lambda_D}{\phi}(u-\hat{u}_2)}.$$

Note that $V^{sd}(\hat{u}_2|\hat{u}_1,\hat{u}_2) + \hat{u}_2$ cannot exceed the first-best expected surplus $\Pi - \frac{c}{\lambda_D}$, thus, the above expression is negative.

Since these component functions are smoothly pasted at $\hat{u}_1$ and $\hat{u}_2$, the entire value function is concave.

(b) As in the no efficiency loss case, $\mathcal{V}(u) + u$ is increasing in u, thus, $\mathcal{V}'(u) \geq -1$ and it gives $\frac{\partial L^D}{\partial R} \leq 0$. Thus, it is sufficient to show that $L^D(u, u + \frac{\phi}{\lambda_D}) \leq 0$ for all $u \geq 0$. Observe that from (OA.2.7), (OA.2.8) and (OA.3.2), we have

$$L^D(u, u + \tfrac{\phi}{\lambda_D}) = \begin{cases} 0, & \text{if } u \leq \hat{u}_1 \text{ or } u \geq \hat{u}_2, \\ \phi V^{dsd'}(u|\hat{u}_1, u) - \phi V^{sd'}(u|\hat{u}_1), & \text{if } u \in (\hat{u}_1, \hat{u}_2). \end{cases}$$

Since $\hat{u}_1$ and $\hat{u}_2$ are chosen to satisfy $V^{dsd'}(u|\hat{u}_1, u) < V^{sd'}(u|\hat{u}_1)$ for all $u \in (\hat{u}_1, \hat{u}_2)$, L^D is always nonpositive.

Likewise, from the concavity of $\mathcal{V}$, $\frac{\partial L^S}{\partial u} \geq 0$. Thus, it is sufficient to show that $L^S(u, u+\frac{\phi}{\lambda_S}) \leq 0$ for all $u \geq 0$. Then, we have

$$L^S(u, u+\tfrac{\phi}{\lambda_S}) = \begin{cases} \phi V^{sd\prime}(u|u) - \phi V^{d\prime}(u), & \text{if } u \leq \hat{u}_1, \\ 0, & \text{if } u \in (\hat{u}_1, \hat{u}_2), \\ \lambda_S(V_S(u + \frac{\phi}{\lambda_S}) - V^{dsd}(u|\hat{u}_1, \hat{u}_2)) - c - \phi V^{dsd\prime}(u|\hat{u}_1, \hat{u}_2) & \text{if } u \geq \hat{u}_2 \end{cases}$$

By Lemma OA.3.3 and Lemma OA.3.5, L^S is always nonpositive.

□

Proof of Lemma OA.3.7. Since $\mathcal{V}$ is strictly concave, $\hat{u}_2 < \bar{u}$ is equivalent to $0 < \mathcal{V}'(\hat{u}_2) = V^{sd\prime}(\hat{u}_2|\hat{u}_1) = V^{dsd\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$. Then, $0 < V^{dsd\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ is equivalent to:

$$\lambda_D(\hat{u}_2 + V^{ds}(\hat{u}_2|\hat{u}_1)) < \lambda_D\Pi - c - \phi. \tag{OA.3.8}$$

Also note that $V^{dsd\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2) = V^{ds\prime}(\hat{u}_2|\hat{u}_1)$ and $V^{dsd}(\hat{u}_2|\hat{u}_1, \hat{u}_2) = V^{sd}(\hat{u}_2|\hat{u}_1)$ imply that

$$\lambda_D(\Pi - \hat{u}_2 - V^{sd}(\hat{u}_2|\hat{u}_1)) = \lambda_S \left(V_S\left(\hat{u}_2 + \frac{\phi}{\lambda_S}\right) + \hat{u}_2 + \frac{\phi}{\lambda_S}\right) - \lambda_S \left(V^{ds}(\hat{u}_2|\hat{u}_1) + \hat{u}_2\right)$$

by (OA.2.7) and (OA.2.8). By plugging (OA.2.2) into the above equation, we can derive that

$$\begin{aligned} (\lambda_S - \lambda_D)(V^{ds}(\hat{u}_2|\hat{u}_1) + \hat{u}_2) =& \lambda_S \left(\Pi - \frac{c}{\lambda_S}\right)\left(1 - e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}\right) - \lambda_D\Pi \\ \iff \qquad \eta\lambda_D(V^{ds}(\hat{u}_2|\hat{u}_1) + \hat{u}_2) =& \eta\lambda_D\Pi - c - (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}. \end{aligned}$$

Then, by plugging this into (OA.3.8), $0 < V^{dsd\prime}(\hat{u}_2|\hat{u}_1, \hat{u}_2)$ is equivalent to

$$\eta(c + \phi) - c < (\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_2 - 1}.$$

Since $\Pi > \frac{c}{\lambda_D} > \frac{c}{\lambda_S}$, 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., $\hat{u}_2$ is less than $\bar{u}$. □

Proof of Lemma OA.3.8. From $\eta > \frac{1}{e-1}$, $\tilde{\Pi}_S(\eta)$ exists. Suppose that $\Pi \geq \tilde{\Pi}_S(\eta)$. By Lemma

OA.3.2, $\hat{u}_1 = 0$. Note that

$$\tilde{\Pi}_S(\eta) = \frac{e-1}{(e-1)\eta - 1} \cdot \frac{c}{\lambda_D} \geq \frac{e-1}{e-2} \cdot \frac{c}{\lambda_D} > \frac{2c}{\lambda_D} \geq \frac{c+\phi}{\lambda_D} = \Pi_F.$$

Then, the project is feasible and $\bar{u}$ is greater than 0, i.e., $\bar{u} > \hat{u}_1$.

Now suppose that $\Pi_M(\eta) < \Pi < \Pi_S(\eta)$. Since V is strictly concave, $\hat{u}_1 < \bar{u}$ is equivalent to $0 < \mathcal{V}'(\hat{u}_1) = V^{d'}(\hat{u}_1)$. Note that $0 < V^{d'}(\hat{u}_1)$ is equivalent to:

$$\frac{\phi}{\lambda_D \Pi - c} < e^{-\frac{\lambda_D}{\phi}\hat{u}_1} = \hat{x}_1. \tag{OA.3.9}$$

Recall that $\hat{x}_1$ is a solution where $\tilde{H}_1(x)$, as defined in (OA.3.4), equals zero. Additionally, $\Pi_M(\eta) < \Pi < \Pi_S(\eta)$ implies that $\tilde{H}_1(1) < 0 < \tilde{H}_1(x^*)$ where x^* is defined in (OA.3.5).[18]

There are two possible cases that satisfy (OA.3.9): (i) $x^* \geq \frac{\phi}{\lambda_D \Pi - c}$; (ii) $\frac{\phi}{\lambda_D \Pi - c} > x^*$ and $\tilde{H}_1(\frac{\phi}{\lambda_D \Pi - c}) < 0$.

The first case is equivalent to $\tilde{H}_1'(\frac{\phi}{\lambda_D \Pi - c}) < 0$. By algebra, we can show that it is equivalent to

$$\frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \frac{\eta e}{\eta+1}\phi^{-\eta}. \tag{OA.3.10}$$

The second case is equivalent to $\tilde{H}_1'(\frac{\phi}{\lambda_D \Pi - c}) \geq 0$ and $\tilde{H}_1(\frac{\phi}{\lambda_D \Pi - c}) < 0$. By algebra, we can show that it is equivalent to

$$\frac{\eta e}{\eta+1}\phi^{-\eta} \leq \frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < (\eta(c+\phi) - c)\, e\phi^{-\eta-1}. \tag{OA.3.11}$$

Last, by the proof of Lemma OA.3.3, we can show that $\Pi > \Pi_M(\eta)$ is equivalent to

$$\frac{(\eta+1)\lambda_D \Pi - c}{(\lambda_D \Pi - c)^{\eta+1}} < \left(\frac{\eta^2}{1-\eta^2}\right)^{\eta} \frac{\eta e}{1+\eta} c^{-\eta}. \tag{OA.3.12}$$

Now I compare the above three conditions. Using $\eta > \sqrt{c/(c+\phi)}$, 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

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

[18] For simplicity, Π is omitted from the definition of $\tilde{H}_1$ and x^*.

imply that (OA.3.10), (OA.3.11) and (OA.3.12). 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 $\hat{u}_1 < \bar{u}$ if and only if $\Pi > \Pi_D(\eta)$. $\square$

Proof of Lemma OA.3.9. By following the proof of Lemma OA.3.7, $\hat{u}_2 \geq \bar{u}$ is equivalent to

$$\overline{y} \equiv \frac{(\eta-1)c+\eta\phi}{(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_1 - 1}} \geq e^{\frac{\lambda_S}{\phi}(\hat{u}_1-\hat{u}_2)} = \hat{y}_2 \tag{OA.3.13}$$

By the proof of Lemma OA.3.4, $\hat{y}_2$ is the solution, which is not equal to 1, of $\tilde{H}_2(y) = 0$.[19]. Since $\hat{u}_2 \geq \hat{u}_1$, if $\overline{y} \geq 1$, the above inequality holds, thus, I restrict attention to the case of $\overline{y} < 1$. Observe that the inequality $\tilde{H}_2(\overline{y}) \leq 0$ implies (OA.3.13) because $\tilde{H}_2$ is strictly convex in y and $\tilde{H}_2(1) \leq 0$.

Using the definition of $\tilde{H}_1$ in (OA.3.4) and $\hat{x}_1 \equiv e^{-\frac{\lambda_D}{\phi}\hat{u}_1}$, $\tilde{H}_2(y)$ can be rewritten as follows:

$$\tilde{H}_2(y) = \frac{1-\eta}{1+\eta}c - \tilde{H}_1(\hat{x}_1)y + \left[-\frac{1-\eta}{1+\eta}c + \frac{\eta}{1+\eta}(\lambda_S\Pi - c)e^{-\frac{\lambda_S}{\phi}\hat{u}_1 - 1}\log y\right] y.$$

Also note that $\hat{x}_1$ is chosen to satisfy $\tilde{H}_1(\hat{x}_1)$ being greater than equal to zero.

By plugging the definition of $\overline{y}$ into the above equation, we can derive that

$$\tilde{H}_2(\overline{y}) = \frac{1-\eta}{1+\eta}c(1-\overline{y}) - \tilde{H}_1(\hat{x}_1)\overline{y} + \frac{\eta}{1+\eta}\left((\eta-1)c+\eta\phi\right)\log\overline{y}.$$

Now define a new function G as follows:

$$G(y) \equiv \frac{1-\eta}{1+\eta}c(1-y) - \tilde{H}_1(\hat{x}_1)y + \frac{\eta}{1+\eta}\left((\eta-1)c+\eta\phi\right)\log y,$$

and it is enough to show that $G(y) \leq 0$ for all $y < 1$.

Note that

$$G''(y) = -\frac{\eta}{1+\eta}\left(\frac{(\eta-1)c+\eta\phi}{y^2}\right) < 0$$

from $\eta \geq \sqrt{\frac{c}{c+\phi}} > \frac{c}{c+\phi}$. Also note that

$$G'(1) = -\tilde{H}_1(\hat{x}_1) + \frac{1}{1+\eta}\left((\eta^2-1)c+\eta^2\phi\right) < 0.$$

[19]The function $\tilde{H}_2$ is defined in (OA.3.6)

from $\eta \geq \sqrt{\frac{c}{c+\phi}}$ and $\tilde{H}_1(\hat{x}_1) \geq 0$. Lastly, note that $G(1) = -\tilde{H}_1(\hat{x}_1) \leq 0$. Therefore, for all $y < 1$, $G(y) \leq G(1) + G'(1)(1-y) \leq 0$. Therefore, $\tilde{H}_2(\overline{y}) \leq 0$ and $u_2 \geq \bar{u}$. $\square$