Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects*

Brett Green
UC Berkeley (Haas)
greenb@berkeley.edu
faculty.haas.berkeley.edu/bgreen

Curtis R. Taylor
Duke University
crtaylor@duke.edu
people.duke.edu/∼crtaylor

June 3, 2016

Abstract

We study the optimal incentive scheme for a multistage project in which the agent privately observes intermediate progress. The optimal contract involves a *soft deadline* wherein the principal guarantees funding up to a certain date – if the agent reports progress at that date, then the principal gives him a relatively short *hard deadline* to complete the project – if progress is not reported at that date, then a probationary phase begins in which the project is randomly terminated at a constant rate until progress is reported. Self-reported progress plays a crucial (but non-stationary) role in implementation. We explore several variants of the model with implications for optimal project design. In particular, we show that the principal benefits by imposing a small cost on the agent in order to submit a progress report or by making the first stage of the project somewhat "harder" than the second. On the other hand, the principal does strictly worse by impairing the agent's ability to observe intermediate breakthroughs.

JEL Classification: J41, L14, M55, D82

Keywords: Dynamic Contracting, Project Management, Moral Hazard, Private Information, Progress Reports, Strategic Communication, Milestones, R&D

*We are grateful to our editor and three anonymous referees for useful comments and suggestions. We thank Bruno Biais, Simon Board, Alessandro Bonatti, George Georgiadis, Barney Hartman-Glaser, R. Vijay Krishna, John Morgan, Sofia Moroni, Steve Tadelis, Andy Skrzypacz, Toby Stuart and Felipe Varas for their feedback. We are also grateful to conference participants at the IO Theory Conference, ESSFM Gersenzee, FTG, NBER, SITE, Olin/CFAR, and seminar participants at ASU, HEC, HKU, HKUST, Humboldt, MIT, NTU, TSE, Bonn, Chicago, Miami, and UNC. Andy Schwartz and Andrew Steck provided excellent research assistance.

In this paper we analyze the following situation. A principal contracts with an agent to complete a project. The project requires the successful completion of two stages or *breakthroughs* in order to realize its benefits. The arrival rate of breakthroughs depends on the agent's hidden action. In particular, the agent can surreptitiously divert funds for private benefit, thereby reducing the rate of a breakthrough. The second breakthrough (i.e., the date at which the project is completed) is publicly observed, but the first breakthrough is privately observed by the agent and cannot be verified by the principal. Both players are risk neutral and the agent is protected by limited liability. We are interested in how the principal should optimally design the incentive scheme and to what extent it relies on communication from the agent to the principal about whether progress has been made.

Our investigation of this setting is motivated by three key ingredients often encountered in complex real-world projects involving research or implementation of novel technical designs. First, such projects are typically organized under an agency relationship because they require both a high level of technical expertise as well as substantial capital. Thus, there is generally some degree of separation between the expert individuals responsible for execution of a venture and the entity that provides its financial backing. Moreover, the preferences of the entrepreneur, contractor, or researcher (the "agent") regarding the timing, intensity and direction of investment are unlikely to be perfectly aligned with the preferences of the financier, end user, or institution (the "principal"). For example, Tirole (2006) suggests several reasons why the relationship between researchers and their funding sources "is fraught with moral hazard."

Second, complex projects often require completion of multiple sequential stages before their benefits can be realized. For example, developing a new drug requires identifying specific chemical compounds or molecules in vitro (i.e., in test tubes) and then demonstrating their efficacy in pre-clinical (i.e., animal) trials. After this phase is successfully completed, clinical (i.e., human) trials begin, progressing through small scale and then large-scale. Each of these stages must be successfully completed before a drug can be submitted to the FDA for approval, and only after approval can the drug be brought to market, generating revenue for its developer and health benefits for society. In a similar vein, large-scale software engineering projects are often purported to follow the celebrated "waterfall model" of development involving six sequential phases: conception, initiation, analysis, design, construction, testing, and implementation (Royce, 1970). Indeed, in 1985 the United States Department of Defense codified six phases in their standards for working with software development contractors (DOD-STD-2167A). Other examples of multistage projects are ubiquitous. For example, most large-scale procurement projects (e.g., construction, defense) involve multiple sequential stages. In venture capital, entrepreneurs generally progress through numerous stages (e.g., patenting, prototype development, and manufacturing) before realizing profits. Basic research, almost by definition, requires successive advancements before its societal benefits are realized.

Third, when and whether the agent makes progress on the project – that is, when and whether it transitions from one phase to the next – is often difficult, even impossible, for the principal to ascertain directly. This could obtain either because the principal lacks the technical proficiency to evaluate progress or because it is not possible to substantiate progress at a reasonable cost. In either case, project sponsors are often forced to rely on unverifiable progress reports made by the very individuals responsible for moving the project forward. Detecting fraudulent claims of progress in complex environments is, not surprisingly, notoriously difficult. Take, for example, the respective cases of extensive data falsification by Diederik Stapel, formerly a professor of social psychology at Tilburg University, and Anil Potti, formerly a cancer researcher at Duke University. The misconduct of both academics went undetected by their universities and funding agencies for a number of years. Similarly, in a 2006 survey of software practitioners, 86% of respondents reported having encountered false reports (Glass et al., 2008). The most common occurrences were in estimation and status reporting. "Respondents said that when lying happens, developers at the bottom level of the management hierarchy are most aware of the lying; they often know it's happening even when their management doesn't."

These three common ingredients: agency, multiple stages, and intangible progress, give rise to the question at the heart of this paper. Specifically, if the principal cannot observe evolution of the project herself, can she nevertheless use unverifiable reports from the agent to monitor his progress and provide incentives for project advancement? Given the prevalence of misreporting, one might easily imagine that the answer to this question is 'no.' In fact, we find that self-reported progress is an essential component of an optimally designed incentive scheme for promoting project development. To ensure the veracity of progress reports, the principal must, however, use them judiciously.

To build intuition and a baseline of comparison for our main results, we start by analyzing a single-stage project in which only one breakthrough is needed to complete the project. In this case, the optimal incentive scheme can be implemented with a *simple contract*, which involves a single deadline T_1^* and a reward that depends only on the date at which the breakthrough arrives. If the agent realizes the breakthrough at $t \leq T_1^*$, he collects the reward that is decreasing in t. If a breakthrough is not realized by the deadline, the principal terminates the project. While the first-best policy never involves project termination, the use of a deadline plays a crucial role in the provision of incentives in our (second-best) setting. In choosing the optimal deadline the principal faces a trade off; a longer deadline yields a higher probability of project completion but also requires giving the agent more rents in order to prevent shirking.

Having established this single-stage benchmark, we analyze the situation of interest by adding a second stage to the project. Thus, there is a preliminary (e.g., research) stage, which must be successfully completed before moving on to the ultimate (e.g., development) stage. A breakthrough in the preliminary stage does not generate any direct benefits, is privately observed by the agent, and is unverifiable. On the other hand, project completion (i.e., when the second breakthrough has been made) is verifiable. In this setting, simple contracts are no longer optimal. If the principal uses a simple contract with deadline T then as the deadline approaches, an agent who has not yet made a breakthrough will "run out of steam" at some point $t_0 < T$ and begin shirking. Thus, a simple contract can be improved upon.[1] If the principal could observe the project stage directly, then she would simply fire the agent for lack of progress at t_0, thereby averting the cost of shirking *ex post* and enhancing incentives *ex ante*. When progress is intangible, however, the principal cannot threaten to fire the agent for lack of reported progress at t_0 without inducing him to make false reports.

Instead, the optimal contract involves a first-stage deadline that is not deterministic; funding for the project is guaranteed up to some *soft deadline*. If the agent has not reported progress by the soft deadline, then a probationary phase ensues in which the principal randomizes over whether or not to terminate the project.[2] The soft deadline provides incentives for the agent not to shirk or falsely report progress in the first stage while not further reducing the probability of success in the second stage conditional on making a breakthrough. Moreover, an agent who has made progress strictly prefers to report truthfully at the soft deadline rather than risk being terminated while under probation. In short, the lottery over termination dates serves as a mechanism to induce both effort and truthful reporting while simultaneously maximizing the probability of project success.

Self-reported progress from the agent to the principal regarding the status of the project, therefore, plays an essential role in the optimal incentive scheme. Interestingly, the required communication is non-stationary over the duration of the project. Early in the life of the project, communication is unnecessary. If the project is completed before the soft deadline, then the agent's compensation depends solely on the date of completion. If the project is not completed by the soft deadline, communication is required. More specifically, at the soft deadline the principal asks the agent to report whether he has made a breakthrough. If he

[1] This is analogous to why "review strategies" are suboptimal in Radner (1985) with discounting. Under a review strategy, the agent's continuation value depends only on whether his performance exceeds a threshold during the review period. Hence, near the end of a review period, the agent will shirk if his performance is sufficiently below (or above) the threshold.

[2] The principal need not randomize over terminating the project at every instant. Rather, she may randomly and secretly draw a fixed date at which she will end probation and fire the agent for lack of progress.

answers "yes" then the principal gives him a relatively short hard deadline to complete the project—if the project is not completed by the hard deadline then it is terminated. If the agent answers "no" then the probationary phase begins. During the probationary phase, the agent is asked to report any progress *immediately*, at which point he is given the same hard deadline to complete the project as if he had reported "yes" at the soft deadline.

We derive a closed-form expression for the principal's value function, the downward-sloping portion of which corresponds to the Pareto frontier. We show that each pair of payoffs on this frontier can be implemented with a contract that differs only in the length of the soft deadline: the principal's highest payoff corresponding to the shortest soft deadline and the agent's highest payoff corresponding to the longest one. An implication of this result is that when human capital is in short supply relative to financial and physical capital (i.e., where experts have relatively more bargaining power), we should expect to see contracts that involve longer horizons and correspondingly higher completion rates.

We also explore three extensions of the model with implications for optimal project design. First, we ask whether there is scope for making communication costly. We show that the principal can indeed benefit by imposing a small cost on the agent of reporting progress. Therefore, formal channels of communication that require time and effort (e.g., paperwork) can be useful even if the same information could be communicated at no cost. Second, we consider a project with asymmetric stages and show that *ceteris paribus* the principal's payoff is higher when the first stage is moderately more difficult than the second stage. However, the principal does strictly worse by making the first stage too difficult relative to the second. Finally we consider a setting in which progress is unobservable to both players in order to see whether the principal can benefit from suppressing the agent's access to information about the status of the project. We find that information suppression is suboptimal; the principal does better under the optimal contract with progress reports from the agent than in a setting where neither party observes the time of the first breakthrough. An alternative interpretation of this result is that it is better to divide up a project into two phases even if the agent privately observes when the first is completed.

We believe our findings are not only of theoretical interest, but also have practical relevance. For example, making continued funding contingent on periodic progress reports appears to be a common arrangement. The following text that appears on the website of the Amyotrophic Lateral Sclerosis (ALS) Association is broadly representative.

> The ALS Association financial officer makes grant award payments to the Principal Investigator institution for disbursement for the project. Payments are made on a specified quarterly schedule and in the case of multi-year grants, after the first

year, are contingent upon the receipt by ALS Association of satisfactory annual progress reports and documentation of research funds expended.

Regarding the use of probation, funding entities do not explicitly specify random deadlines (at least not to our knowledge). However, grant policies are often quite vague about the consequences for delays in reporting progress. For instance, the National Institute of Health's (NIH) web site states "If your [progress] report is extremely late, *you risk losing funding* for the period of time between the end of the current budget period and when we finish processing your report."[3] Similarly, the National Science Foundations (NSF) Grant Policy Manual states, "NSF *reserves the right, ... to withhold future payments* after a specified date if the recipient fails to comply with the conditions of an NSF grant, including the reporting requirements."[4] While these policies do not specify random termination as such, it certainly seems more appropriate to view the indefinite penalties for late reporting as involving *soft* deadlines rather than *hard* ones.

The soft deadline and ensuing probationary phase of our optimal incentive scheme also resembles how "project slippage," (i.e., missing a deadline) is handled in project management.[5] In practice, "slippage time" (i.e., the amount of time the project is in a state of slippage) generally extends the deadline for completing the entire project. That is, contractors and employees are not generally required to recover slippage time in one phase of a project by shortening subsequent phases. Of course, the greater the slippage time, the more likely it is that the sponsor will cancel the project. Our results suggest that some degree of slippage should, however, be tolerated on complex projects where sponsors are unable to assess progress directly.

Missing an intermediate deadline is also akin to the concepts of "slack" or "float" formalized in the widely applied Project Evaluation and Review Technique (PERT) and the Critical Path Method (CPM) both of which were developed in the late 1950s for managing complex, multistage projects.[6] Under PERT and CPM, once a project has used all its free float on earlier stages, the time remaining to complete the project is constrained by the duration of the so-called critical path. This resembles our finding that the agent should be given a relatively short hard deadline if he reports progress at or after the soft deadline.

The remainder of the paper proceeds as follows. In the next section, we discuss related literature. In Section 2, we present a single-stage version of the model as a benchmark. We introduce a second-stage to the model and present our main results in Section 3. Section 4

[3] `http://www.niaid.nih.gov/researchfunding/qa/pages/pp.aspx#late`: emphasis added.
[4] `http://www.nsf.gov/pubs/manuals/gpm05_131/gpm4.jsp`: emphasis added.
[5] See e.g., Ewusi-Mensah and Przasnyski (1991).
[6] See e.g., Malcolm et al. (1959) and Kelley (1961).

explores the role of costly reporting, projects with asymmetric stages, and the possibility of restricting the agent's access to information. Concluding remarks appear in Section 5. In Appendix A, we show how to derive the optimal contract and payoff frontier. Formal proofs are located in Appendix B and an online appendix.

1 Related Literature

There is a large and growing literature studying the optimal provision of incentives in dynamic environments.[7] Our benchmark single-stage model is similar to Shavell and Weiss (1979) and Hopenhayn and Nicolini (1997),[8] who look at providing incentives to search for employment while simultaneously providing unemployment insurance. It is also similar to Mason and Välimäki (2015) who consider a dynamic moral hazard setting in which the principal must provide incentives to the agent to complete a project.[9] A novel aspect of our model is that we explicitly consider environments with multiple sequential stages in which the agent privately observes progress. Thus, we "sandwich" a hidden-information problem between two stages of a hidden-action problem.

Our multistage environment is related to Biais et al. (2010), who analyze a model in which large (observable) losses may arrive via a Poisson process, and an agent must exert unobservable effort in order to minimize the likelihood of their arrival. They allow for investment and characterize firm dynamics as well as asymptotic properties. Our model differs in that it (i) features only a finite number of arrivals, (ii) the arrival of a breakthrough is "good news", and (iii) the first arrival is privately observed by the agent. Several other recent papers that involve observable Poisson arrivals include Hoffmann and Pfeil (2010), Piskorski and Tchistyi (2011), DeMarzo et al. (2014). Given the multistage setting, a key difference in this paper is that the agent's continuation utility is not a sufficient state variable. Toxvaerd (2006) considers a setting in which a finite number of (observable) arrivals are needed in order to complete a project. In his setting, the agent is risk averse and the optimal contract trades off optimal risk-sharing for incentive provision, but does not involve deadlines or inefficient termination. Optimal dynamic mechanisms are explored in other settings by

[7] A non-exhaustive list includes Green (1987), Spear and Srivastava (1987), Phelan and Townsend (1991), Quadrini (2004), Clementi and Hopenhayn (2006), DeMarzo and Sannikov (2006), DeMarzo and Fishman (2007) and Sannikov (2008).

[8] See also Lewis (2012) who considers a delegated search model in which the optimal contract includes a deadline and a bonus for early completion.

[9] Though the problems are similar in spirit, the optimal contract in our one-stage benchmark looks quite different from Mason and Välimäki (2015). In our model, it is optimal to use termination deadlines to provide incentives to the agent, whereas in their setting, the use of termination is strictly suboptimal. This difference arises due to the nature of the moral hazard problem we consider (e.g., private benefit from shirking) rather than their costly effort model. See Remark 1 for further discussion.

Board (2007), Eso and Szentes (2007), Bergemann and Valimaki (2010) and Pavan et al. (2014) among others.

In our model, the agent has access to private information that is persistent. Dynamic contracting with persistent private information has been studied in discrete type settings by Fernandes and Phelan (2000), Battaglini (2005), Tchistyi (2013), and with a continuum of types using a first order approach by Williams (2011) and Edmans et al. (2012). Our approach is most similar to Zhang (2009) and Guo and Hörner (2015). From a theoretical perspective, our work differs from this literature along several dimensions. One key difference in our environment is the presence of public (contractible) information (i.e., the ultimate success of the project), which can be used to screen the agent's underlying private information. Another difference is that the transition probabilities across states are endogenously determined in our setting by the agent's action.

In a working paper, Hu (2014) considers a discrete-time setting that is otherwise similar to ours. However, he restricts attention to deterministic deadlines and does not allow for communication between the principal and the agent, which leads him to conclude that the optimal contract involves a single deadline in which an agent who has not made progress shirks as the deadline approaches. Our results show these restrictions are not without loss of generality and lead to substantively different findings. In particular, we show that without these restrictions, the optimal mechanism requires communication, involves random termination, and does not involve shirking.

There is a rich existing literature exploring settings where parties *learn* about the value of a project over time.[10] In these environments, lack of success typically indicates that the project is *bad*, and it is socially efficient to discontinue investment at some point. A common finding within this literature is that agency considerations may cause the principal to terminate the project earlier than socially optimal. By contrast, we focus on a setting in which the project is commonly known to be *good* from the outset (i.e., there is no learning and the project is always funded until completion under the first-best contract) in order to isolate the extent to which progress can be used to provide stronger incentives. Bonatti and Hörner (2011) study experimentation in teams for a project that requires a single breakthrough, which introduces a free-riding problem. They show that the equilibria of the game involve inefficient delays in effort provision and that deadlines, which terminate the project prior to the socially efficient time, are useful in mitigating delays despite forfeiting value when the deadline is reached. The free-riding problem also arises in Moroni (2015), who studies experimentation with multiple agents for a project that requires several observable breakthroughs.

[10] See e.g., Levitt and Snyder (1997), Bergemann and Hege (1998, 2005), Inderst and Mueller (2010), Manso (2011), Hörner and Samuelson (2014), Halac et al. (2015).

Holmstrom and Milgrom (1991) and Laux (2001) look at settings with simultaneous tasks whereas in our setting the project involves stages that must be completed sequentially. Varas (2015) studies a dynamic multi-tasking model in which an agent can complete the project faster by reducing its quality. Under certain conditions, he finds that a random termination policy is optimal. However, the underlying mechanism is somewhat different from ours. In his model, stochastic termination is used to prevent multi-tasking and arises as part of the optimal contract only when the agent is sufficiently more impatient than the principal. Moreover the use of a hard deadline is suboptimal in his model unlike in our multistage setting where the principal uses a combination of both soft and hard deadlines.

2 Single-Stage Benchmark

In this section we present a benchmark version of the model in which only a single breakthrough is required to attain project benefits. While the single-stage model is of some independent interest, its primary role is to build intuition and facilitate discussion of the multistage model.

2.1 The Setting

A principal (she) contracts with an agent (he) to undertake a project. Time is continuous and the project can be operated over a potentially infinite horizon. The project requires the successful completion of a stage, also referred to as a *breakthrough*, in order for its benefits to be realized. Operating the project requires resources, which we model as a flow cost of c per unit time that the project is in operation. The principal has unlimited resources to fund the project. The agent has no funds and is protected by limited liability, but has the skills necessary to run the project. Both parties are risk-neutral and do not discount the future.[11] The principal can terminate the project (i.e., discontinue paying the flow cost) at any point in time. Project termination is irreversible; if the principal terminates the project prior to the breakthrough, the project delivers no benefit and the game ends.

While the project is in operation, the agent chooses an action $a_t \in \{0, 1\}$, where $a_t = 1$ indicates that he appropriately invests or "works," and $a_t = 0$ indicates that he diverts funds or "shirks" for private benefit. The arrival rate of a breakthrough is then given by λa_t. Thus, if the agent works over an interval of length dt, then the probability of a breakthrough in the interval is λdt. If the agent shirks, then the arrival rate of a breakthrough is zero but he receives a private flow benefit of ϕdt, where $\phi > 0$ measures the severity of the agency problem. The agent receives no intrinsic benefit from project success; he benefits solely from

[11] It is fairly straightforward to incorporate a common discount rate into the model. However, with discounting, closed-form solutions are no longer available in the multistage setting and therefore our method of proof for some results does not generalize directly.

the compensation delivered by the principal and any private benefits from shirking.

Remark 1. *For many relevant applications, the most natural interpretation of the moral hazard problem is that the agent can divert the principal's investment for private benefit. For example, an entrepreneur can use venture capital funding for private consumption, or a scientist may fund a pet project not authorized under his current grant. Nevertheless, we will adopt the standard "shirk"/"work" terminology (e.g., Tirole, 2006).*[12]

The success of the project is publicly observed and contractible. Upon the arrival of a breakthrough the principal realizes a payoff $\Pi > 0$, makes any outstanding contractual payments to the agent and the game ends. Let τ denote the random variable representing the date of project success. Throughout our analysis, we employ the following assumptions.

Assumption 1. *The expected value of the project (absent agency costs) is strictly positive*

$$\lambda\Pi - c > 0.$$

Assumption 2. *Shirking is non-trivial and inefficient*

$$0 < \phi \leq c.$$

Remark 2. *Assumption 1 and 2 imply that under the first-best policy, the agent never shirks and the project is never terminated prior to success.*

2.2 The Principal's Problem

At $t = 0$, the principal offers the agent a contract. If the agent rejects the proposed contract, then both parties receive their outside options normalized to zero. We assume throughout that the principal can fully commit to all contractual terms.

A contract is denoted by a triple, $\Gamma \equiv \{a, R, T\}$, where a_t is the recommended action to the agent at time t, R_t is a monetary payment made to the agent for success at t, and T is the date at which the project is terminated absent a prior breakthrough.[13] An action process, a, induces a probability distribution $\mathbb{P}^a$ over τ. Let $\mathbb{E}^a$ denote the corresponding expectation

[12] In the analog of our model where the agent incurs a cost of effort rather than a benefit from diversion (i.e., arrival rate of breakthroughs is zero without effort), the limited liability constraint has no bite and the first-best outcome is attainable. In a costly-effort model with discounting and a strictly positive arrival rate even when the agent shirks, it is possible to obtain results similar to those presented here under certain parametric restrictions.

[13] As is typical of most principal-agent models, making a payment to the agent upon "failure" (i.e., termination prior to project success) is suboptimal. Moreover, because both parties have linear utility and are equally patient, it is without loss of generality to backload all monetary payments to the agent (see e.g., Ray, 2002).

operator. If the agent adheres to the recommended action, the principal's (ex-ante) expected utility is given by

$$P_0(\Gamma) = \mathbb{E}^a \left[(\Pi - R_\tau) \cdot \mathbb{1}_{\{\tau \leq T\}} - \int_0^{T \wedge \tau} c \, dt \right], \tag{1}$$

and the agent's expected utility is given by

$$U_0(\Gamma) = \mathbb{E}^a \left[R_\tau \cdot \mathbb{1}_{\{\tau \leq T\}} + \int_0^{T \wedge \tau} \left(\phi(1 - a_t) \right) \, dt \right]. \tag{2}$$

The contract Γ is said to be incentive compatible if a maximizes the agent's expected utility given (R, T). The principal's problem is to find an incentive compatible contract that maximizes $P_0(\Gamma)$ subject to delivering an expected utility to the agent weakly larger than his outside option. We refer to a solution to the principal's problem as an *optimal contract*. That is, an optimal contract is one that the principal would offer if she was endowed with all the bargaining power.

2.3 The Optimal Single-Stage Contract

Following what is now common practice (see e.g., Spear and Srivastava, 1987), we solve the principal's problem by formulating it recursively, where the state variable is the promised utility to the agent, denoted by $u \in \mathbb{R}_+$. Let $V(u)$ denote the principal's value function, which represents her expected utility under an optimal contract subject to the additional "promise keeping" constraint that $U_0(\Gamma) = u$.[14]

Proposition 1. *For a one-stage project, the principal's value function is given by*

$$V(u) = \bar{V}(u) \equiv \left(\Pi - \frac{c}{\lambda} \right) \left(1 - e^{-\frac{\lambda u}{\phi}} \right) - u. \tag{3}$$

This payoff can be attained under the contract with a deadline $T_1(u) = \frac{u}{\phi}$, and a reward payment for success that is decreasing over time according to

$$R_1(\tau) = \phi \left(\tfrac{1}{\lambda} + T_1(u) - \tau \right), \quad \forall \tau \in [0, T_1(u)]. \tag{4}$$

To help with intuition, observe that inducing the agent to work at time t requires

$$\lambda(R_t - u) \geq \phi,$$

where the left side corresponds to the *net* expected benefit to the agent from working and

[14] Note that $\{(u, v) \in \mathbb{R}_+^2 : v \leq V(u)\}$ is the set of feasible, individually rational payoffs.

the right is his instantaneous return from shirking. Moreover, this incentive constraint binds with equality under any optimal contract. To see why, note that any contract in which the agent shirks over some interval of time can be improved upon with a direct payment and no shirking. Further, paying the agent more than necessary to induce effort can be improved upon with a lower reward and a longer deadline.

Thus, Proposition 1 says that the agent is given a deadline corresponding to the date at which his continuation utility will hit zero absent a breakthrough. If he innovates prior to that date, then he receives a reward compensating him for the instantaneous incentive to shirk ϕ/λ plus the opportunity cost of shirking for the remaining time on the clock, $\phi(T_1(u) - \tau)$. In other words, the sooner the agent makes a breakthrough, the larger is his share of the concomitant benefit. In this light, the principal's value function is simply first-best surplus scaled by the probability that the agent innovates prior to the deadline net of his promised continuation utility.

The only thing left to pin down is the initial utility level for the agent, which is equivalent to the optimal deadline. We call a project *feasible* if there is an optimal contract that does not involve immediate termination.

Corollary 1. *A single-stage project is feasible if and only if*

$$\lambda\Pi - c > \phi. \tag{C.1}$$

If (C.1) *holds, then the optimal contract has a deadline* $T_1^* \equiv \frac{1}{\lambda}\ln\left(\frac{\lambda\Pi - c}{\phi}\right)$.

The (principal's) optimal deadline balances the probability that the project ultimately succeeds, $(1 - e^{-\lambda T_1^*})$, against the rents extracted by the agent, $u_1 = \phi T_1^*$. The feasibility condition ensures that $T_1^* > 0$. Intuitively, the principal must give the agent rents of ϕ per unit time in order to prevent shirking. The condition therefore ensures that the principal's expected flow benefit, $\lambda\Pi$, outweighs her total flow cost of operating the project and inducing the agent to invest $(c + \phi)$. Notice that this inequality is stronger than would be needed in a first-best situation where the principal ran the project herself, namely $\lambda\Pi - c > 0$, in which case, as noted in Remark 2, the first-best policy is to invest indefinitely until the innovation arrives. Also notice that the optimal deadline increases as the agency problem becomes less severe, with $\lim_{\phi\to 0} T_1^* = \infty$. That is, the (second-best) outcome and principal payoffs converge to first-best as the agency conflict goes to zero. Hence, the deadline exists only to mitigate agency costs.

3 Two-Stage Projects

We now introduce a second stage to the project. Henceforth, the agent must make two breakthroughs in order for the principal to realize the project benefits. To simplify exposition, we assume that the parameters (ϕ, c, λ) are the same for both stages.[15] Analogous to Assumption 1, we assume that the expected value of the two-stage project is strictly positive; i.e., $\Pi - 2c/\lambda > 0$. Note that this assumption does *not* imply that the (two-stage) project is feasible. Indeed part of our interest is in characterizing the conditions under which the principal can profitably undertake a multistage project.

We let τ_1 and τ_2 denote the (random) times at which the first and second breakthrough occur. We distinguish between the *first stage* of the project, $t \in [0, \tau_1)$, and the *second stage*, $t \in [\tau_1, \tau_2)$. As before, the date at which the project ultimately succeeds, now denoted τ_2, is publicly observed and can be directly contracted upon.[16] For the majority of our analysis, we assume that intermediate progress (i.e., τ_1) is privately observed by the agent and cannot be verified by the principal.[17] For this reason, we will refer to intermediate progress as being *intangible*.

Intangible progress can only be contracted upon indirectly via unverifiable "progress reports" from the agent. *A priori*, it is not clear that such reports have value to the principal. Recall that knowing the stage of the project is irrelevant for the first-best (i.e., socially optimal) investment policy, hence progress reports are only beneficial to the principal if can be used to reduce agency rents. Of course, in order to induce truthful reporting the agent must be given appropriate incentives. So, the question becomes whether the savings in rents associated with inducing more efficient investment (i.e., preventing shirking) justify the payment of the rents necessary to induce truthful reports. For convenience, we will sometimes refer to the agent who has made a breakthrough as the "high" type and agent who has not yet made a breakthrough as the "low" type.

3.1 Simple Contracts

In a single-stage project, the principal can implement the optimal contract with a deterministic deadline and reward scheme that depends only on the completion date. We refer to such contracts as *simple contracts*. Notice that simple contracts effectively preclude any meaningful communication with the agent. This leads to the undesirable feature that the low-type agent

[15] We analyze projects with asymmetric stages in Section 4.2.

[16] This assumption is stronger than necessary. Our results continue to hold if τ_2 is privately observed by the agent but can be verified by the principal because, under the optimal contract, the agent prefers to reveal the second breakthrough as soon as it occurs.

[17] In Section 4, we consider the case in which intermediate progress is unobserved by both the agent and principal. In a working paper, we consider the case in which intermediate progress is publicly observable.

will begin shirking as the deadline approaches.

Proposition 2. *For any two-stage project and any simple contract with a bounded reward scheme and deadline T, there exists a $\Delta > 0$ such that if the agent has not made the first breakthrough by $T - \Delta$, he will shirk at all $t > T - \Delta$.*

Proof. Suppose the agent has not made the first breakthrough at time $T - \Delta$. The probability the agent can complete the project by working until time T is

$$\Pr(\tau_2 \leq T | \tau_1 > T - \Delta) = 1 - e^{\lambda \Delta} (1 + \lambda \Delta) . \tag{5}$$

As $\Delta \to 0$, (5) converges to zero at a rate proportional to Δ^2, whereas the benefit of shirking during the time remaining is proportional to Δ, implying that for t close enough to T, the agent will prefer to shirk unless the reward for success is arbitrarily large. □

In essence, an agent who has not yet made progress "runs out of steam" as the deadline approaches. A corollary of Proposition 2 is that simple contracts are not the most efficient way to provide incentives for multistage projects. If τ_1 was publicly observable and contractible, the principal could improve upon a simple contract by incorporating a first-stage deadline whereby the project is terminated if the agent has not made a breakthrough by $T - \Delta$.[18] However, because progress is intangible, the principal cannot employ a first-stage deadline and expect the agent to report truthfully. Doing so would induce the low-type agent to make false claims of progress (just prior to $T - \Delta$) and then shirk for the remaining allotted time.

Nevertheless, a simple contract can be improved upon. To see how, suppose that after reaching $T - \Delta$, the principal asks the agent "have you made a breakthrough yet?" If the agent reports "no", then the project is terminated but the principal gives him a severance payment of $\phi\Delta$, which is exactly what he could obtain by diverting cash flows until T. If the agent answers "yes" then the principal continues funding the project until T with the same reward scheme. This arrangement induces truth-telling for an agent who has not made a breakthrough (since he is just indifferent). It also induces truthful reporting for an agent who has made a breakthrough provided that following a breakthrough working is incentive compatible under the original contract. Moreover, this modification does not change the probability of project success (since the agent without a breakthrough would have shirked anyway) and saves the principal $(c - \phi)\Delta \geq 0$ conditional on a "no" report. Thus, communication with the agent can improve the principal's expected payoff.

[18] In Green and Taylor (2015), we solve for the optimal contract when progress is publicly observable and show that it closely resembles this multi-deadline arrangement.

3.2 The Optimal Contract (with Minimal Communication)

The optimal contract can be implemented using an arrangement similar to the one just described save for one major difference. Instead of deterministic termination and a severance payment following a negative report at $T - \Delta$, the principal begins a probationary phase in which she effectively "freezes" time (and the agent's promised utility), but starts randomly terminating the project at a constant rate. By doing this, the principal is still able to induce truthful reporting by giving the low-type agent continuation utility of $\phi\Delta$ while on probation, but also induces him to continue working, thereby increasing the probability that the project ultimately succeeds relative to a contract with deterministic termination and a severance payment at $T - \Delta$.

In order to implement the optimal contract, communication between the principal and agent is necessary. However, this communication is not required until the *soft deadline* is reached. That is, early on in the life of the project the principal need not communicate with the agent regarding progress. Instead, at $t = 0$ the principal can simply give the agent a future date T_s (i.e., the soft deadline), at which a progress report is required. If the ultimate success of the project is realized prior to T_s then there is no need for the agent to make a report. The principal simply compensates him based on τ_2. On the other hand, at all $t \geq T_s$, the agent must report a breakthrough *as soon as it arrives* in order to avoid suboptimal termination. Let $\hat{\tau}_1$ denote the time at which the agent reports the first breakthrough. The optimal contract can be implemented as follows.

Theorem 1. *There is a pair $(T_s, u_s) \in \mathbb{R}^2_+$ such that the optimal contract can be implemented by use of a soft deadline T_s, a long clock $T_s + u_s/\phi$, a termination rate $\sigma = \phi/u_s$, and a reward function, $R_s(\hat{\tau}_1, \tau_2)$, such that:*

- *If the project is not completed prior to the soft deadline T_s, the principal asks the agent for a progress report at $t = T_s$.*
  - *If the agent reports that he has made the first breakthrough ($\hat{\tau}_1 \leq T_s$), then he is given the remaining time on the long clock, u_s/ϕ, to complete the project.*
  - *If the agent reports that he has not yet made the first breakthrough, then the principal stops the long clock and initiates a probationary phase in which the project is terminated at constant rate σ.*
  - *If the agent reports a breakthrough during the probationary phase, then he is given the remaining time on the long clock, u_s/ϕ, to complete the project.*

- *The agent gets the reward* $R_s(\hat{\tau}_1, \tau_2)$ *only if the project succeeds prior to being terminated, where*

$$R_s(\hat{\tau}_1,\tau_2) = \begin{cases} \phi\left(\frac{2}{\lambda}+\frac{1}{\sigma}+T_s-\tau_2\right), & \textit{if } \tau_2 \le T_s \\ \phi\left(1+\frac{\sigma}{\lambda}\right)\left(\frac{1}{\lambda}+\frac{1}{\sigma}+\max\{T_s,\hat{\tau}_1\}-\tau_2\right) & \textit{if } 0<\tau_2-\max\{T_s,\hat{\tau}_1\}\le\frac{1}{\sigma} \\ 0 & \textit{otherwise.} \end{cases}$$

Even though the principal cannot substantiate the agent's progress reports, such reports are an essential aspect of the optimal contract because they are used to govern the continuation contract: a report of "no" resulting in probation and one of "yes" resulting in a relatively short time to complete the project.[19] Indeed, permitting the agent to "remain silent" regarding his progress would make the principal worse off as was demonstrated in Proposition 2. To develop additional intuition for the structure of the optimal contract, it is useful to explain why it is optimal for the principal to randomize during the probationary phase, and why it is incentive compatible for the agent to report truthfully at (and after) the soft deadline.

Why is it optimal for the principal to randomize during the probationary phase? Prior to T_s, the low-type agent's continuation value decreases at a constant rate. Once the soft deadline is reached, his continuation value is relatively small, which inherently limits the remaining time the principal can grant the agent to complete the project.[20] Further, while the agent's continuation utility increases by ϕ/λ when he truthfully reports the first breakthrough (see Figure 1), the principal cannot convert this increase in agent utility into additional time to complete the project without inducing false reports.

Thus, if the agent reports "no" at the soft deadline then the principal can only benefit if the agent makes two breakthroughs in a relatively short amount of time. Moreover, the probability of making two breakthroughs in Δ units of time is convex when Δ is small (see equation (5)). Therefore, after a negative report at T_s, rather than let the long clock run down further, the principal does better by randomizing between terminating the project and extending it, which allows her to "pause" the long clock and preserve time to make the second breakthrough conditional on making the first one before termination occurs.

Why does the agent report truthfully? By construction the low-type agent is indifferent between honestly reporting his lack of progress (thereby entering the probation phase) and falsely reporting progress (and then optimally shirking over the time left on the long clock). The high-type agent strictly prefers reporting his true status at this point – that is, he

[19] Lemma B.2 shows that the amount of time the agent has to complete the project after reporting "yes" is less than the optimal deadline for a one-stage project (i.e., $u_s/\phi \le T_1^*$).

[20] In particular, if the agent's continuation value is u, then the expected time to termination must be weakly less than u/ϕ.

strictly prefers having a short period of time to complete the project to facing probation. Intuitively, the low-type weakly prefers a lottery over a long deadline and termination to a short deadline because he is unlikely to finish the project (i.e., make two breakthroughs) with a short deadline. The high-type agent strictly prefers the short deadline since he has already made one breakthrough, he expects it will take him less time to complete the project and does not want to risk termination.

We refer to the contract in Theorem 1 as the *Optimal Contract with Minimal Communication* (or OMC-contract) because it minimizes the expected number of reports that the agent is asked to make over the life of the project subject to delivering the maximal payoff to the principal (notice that if $\tau_2 < T_s$, then the agent does not make a report). Because of this property, the OMC-contract is uniquely optimal when the agent must incur (or the principal can impose) a small cost in order to report progress (see Section 4.1).[21]

Optimal Dynamics

The dynamics of agent continuation values under the OMC-contract are illustrated in Figure 1. The horizontal axis corresponds to the continuation value of a low-type agent whereas the vertical axis corresponds to the continuation value of a high-type agent. At $t = 0$, agent continuation values starts at a level that maximizes the principal's ex-ante payoff, which correspond to the asterisk in the upper right corner of Figure 1. Continuation values then drift down along $\mathcal{L}_H = \{(u_1, u_2) : u_2 = u_1 + \phi/\lambda\}$ toward the point labeled κ^*. Note that the incentive compatibility condition inducing low-type agent effort is binding on $\mathcal{L}_H$.

If a breakthrough is not reported upon reaching κ^*, then the principal initiates the probationary phase in which she randomizes over terminating the project (continuation values jump to the origin) and maintaining promised utilities at κ^*. Hence, κ^*, serves as a partially absorbing state until the probationary phase ends with either project termination or a reported breakthrough, at which point, the state variable again evolves toward the origin. Conditional on reaching κ^* and not being terminated, the promised utility to the agent once he (truthfully) reports a breakthrough is independent of when it is reported.

Under the OMC-contract, no communication takes place prior to reaching κ^* and hence promised utilities continue to evolve downward along $\mathcal{L}_H$ until T_s regardless of whether a breakthrough has been made. However, there exist other optimal mechanisms in which, if a breakthrough is reported prior to reaching κ^* the optimal dynamics follow a different

[21] The reward schedule in the OMC-contract is not uniquely pinned down for $\tau_2 > T_s$. Any reward function (including those which are non-monotone in τ_2) satisfying promise keeping and incentive compatibility for the high-type agent will suffice. The reward function given in Theorem 1 is the unique one meeting these criteria that induces piecewise linear promised utility for the high-type agent as a function of time prior to the second breakthrough.

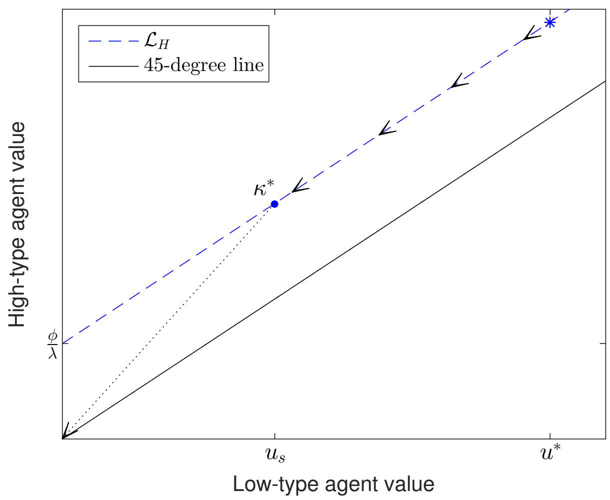

Figure 1: This figure illustrates the optimal dynamics of continuation values under the OMC-contract in Theorem 1. Prior to the soft deadline (i.e., κ^*), continuation values drift down along the dashed line (i.e., $\mathcal{L}_H$). Upon reaching κ^*, the evolution of continuation values depends on whether a breakthrough has been reported. If it has not, they stop drifting and either jump to the origin (if terminated) or remain at κ^*. If a breakthrough is reported at the soft deadline, then the continuation values drift down along the dotted line toward the origin.

trajectory. For example under a direct mechanism, if a breakthrough is reported prior to T_s (i.e., reaching κ^*), continuation values can travel into the region below $\mathcal{L}_H$ and above the 45-degree line. Regardless, prior to a reported breakthrough, continuation values must evolve downward along $\mathcal{L}_H$ under any optimal mechanism that induces truthful reporting—points above $\mathcal{L}_H$ promise the high-type agent excessive rents and points below $\mathcal{L}_H$ induce the low-type agent to shirk.

3.3 Pareto-Efficient Contracts

In the previous subsection, we focused on the principal's optimal contract. In fact, any pair of payoffs on the Pareto frontier can be achieved using a contract that is identical to the one in Theorem 1 simply by varying the length of the soft deadline. In this subsection, we characterize the payoff frontier and show how to implement an arbitrary Pareto-efficient, individually-rational payoff pair. We use recursive techniques to derive the principal's maximal payoff in the first stage of the project given any utility level for the low-type agent (see Appendix A for details). The resultant value function, denoted by F_1, satisfies a first-order, ordinary differential equation prior to the probationary phase (i.e., for agent continuation values above u_s),

$$\lambda F_1(u) = \lambda \underbrace{\left(\left(1 - e^{-\lambda u/\phi}\right)\left(\Pi - \frac{c}{\lambda}\right) - (u + \phi/\lambda)\right)}_{\text{Principal's payoff after a (reported) breakthrough}} - c \;-\; \phi F_1'(u), \tag{6}$$

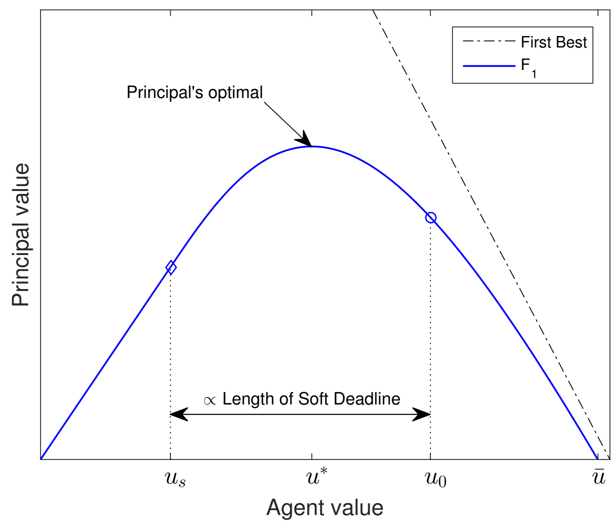

Figure 2: Illustrates the payoff frontier and the relation between the length of the soft deadline and the corresponding (ex-ante) payoffs.

and is linear for $u \in (0, u_s)$. Two boundary conditions pin down the constant in the value function and u_s, which is the level of agent utility at which the probationary phase. Namely,

$$F_1'(u_s) = \frac{F_1(u_s)}{u_s} \tag{7}$$

$$F_1''(u_s) = 0. \tag{8}$$

The first can be interpreted as a local-optimality (or *smooth-pasting*) condition and the second can be interpreted as a global-optimality (or *super-contact*) condition. The free-boundary problem implied by (6)-(8) has a unique solution (Lemma A.4 provides a closed-form expression). Naturally, the solution is useful for characterizing the set of Pareto-efficient payoffs.

Theorem 2. *The set of Pareto-efficient payoffs (satisfying individual rationality for both parties) is given by*

$$X^P \equiv \left\{ (u, v) \in \mathbb{R}^2_+ : u \in [u^*, \bar{u}] \text{ and } v = F_1(u) \right\},$$

where $u^* = \arg\max_u F_1(u)$ *and* $\bar{u}$ *is such that* $F_1(\bar{u}) = 0$. *Any* $(u, v) \in X^P$ *can be achieved using the contract from Theorem 1 with a soft deadline* $T_s = \frac{u - u_s}{\phi}$.

Figure 2 provides a graphical illustration of Theorem 2. The Pareto frontier extends from the principal's optimal contract (agent value of u^*) to the agent's preferred contract (agent value of $\bar{u}$). The larger is the agent's value, the longer is the soft deadline. Notice that, among

Pareto-efficient contracts, the principal's optimal contract involves the shortest soft deadline while the agent's preferred contract involves the longest soft deadline. Thus, the theorem has implications for how the relative bargaining power affects the terms of the contract. When agents with the skills required to manage such projects are relatively scarce (plentiful), we should expect the terms of the contract to involve longer (shorter) horizons. Note also that the probability of project completion is increasing in the length of the soft deadline, so that settings in which agents have more bargaining power generate higher welfare (expected social surplus).

Given the closed-form solution for the principal's value function, we can characterize precisely when a two-stage project is feasible.

Corollary 2. *A two-stage project is feasible if and only if*

$$\lambda\Pi - 2c \geq 2\phi + \lambda\phi T_1^*. \tag{C.2}$$

It is worth noting that (C.2) is considerably stronger than the feasibility condition for a (comparable) one-stage project where the breakthrough arrives at half the rate (which maintains the same first-best value). Specifically, if the arrival rate is $\lambda/2$, then (C.1) is equivalent to $\lambda\Pi - 2c > 2\phi$. Because progress is intangible, the principal must give additional rents of u_s to the low-type agent at the soft deadline in order to induce truthful reporting. When (C.2) holds with equality, $u_s = \phi T_1^*$ and probation begins *immediately* under the OMC-contract (i.e., $T_s = 0$). Thus, the first-best value of the project must exceed not only $2\phi/\lambda$ but $2\phi/\lambda + \phi T_1^*$ for the principal to be willing to undertake it. Nevertheless, as we will see in Section 4.3, the principal is still better off paying the information rents to elicit truthful reports about intangible progress than obscuring the agent's ability to observe it.

4 Project Design

In this section we investigate three variations of the model that have implications for various aspects of project design. First, we consider costly reporting and ask whether there is scope for imposing a cost on the agent in order to submit a progress report. Next, we explore the case of asymmetric stages and ask whether it can be beneficial for the principal to design one of the stages to be more difficult than the other. Finally, we consider the case in which progress is unobservable to both players in order to see whether the principal can benefit from suppressing the agent's access to information about the status of the project. Proofs of the formal results in this section are relegated to an online appendix.

4.1 Costly Reporting

Suppose there are two channels through which the agent may report progress: an informal one (e.g. verbal) and a formal one (e.g., written). Both types of reports can be contracted upon and both are falsifiable. The only operational difference between the two channels is that the formal channel requires the agent to incur a cost $\rho > 0$ (e.g., the time and effort of filling out documentation), whereas the informal channel remains costless for the agent to use (as in Section 3).[22] We will assume that using the formal channel is equally costly for the agent whether he submits a true or false report. Therefore, making a report through the formal channel is more akin to "money burning" than to "signaling," though our results hold *a fortiori* if the cost of reporting is higher for a false report.[23]

Let $F_1(u;\rho)$ denote the principal's first-stage value function when the reporting cost is ρ. Notice that $F_1(u;\rho) \geq F_1(u;0) = F_1(u)$ since the principal can always elect not to use the formal channel. The question is whether it is ever advantageous for her to require the agent to use the costly channel for reporting progress. Proposition 3 shows that indeed the principal benefits by imposing a small reporting cost.

Proposition 3. *If $\rho > 0$ is sufficiently small:*

- *The principal benefits by requiring the agent to use the formal (costly) channel: $F_1(u;\rho) > F_1(u)$ for all $u > 0$.*
- *If the project is feasible, there exists a soft deadline $T_s(\rho) > 0$ such that the optimal contract requires the agent to communicate progress through the costly formal channel (only) after the soft deadline has been reached.*

Using costly reports in this fashion impacts the principal's payoff in two ways, one positive and one negative. On the up side, requiring formal reports at the soft deadline relaxes the truth-telling constraint (by raising the cost of lying) and allows the principal to add extra time to the long clock after a reported breakthrough. On the down side, with positive probability the high-type agent will have to make a costly formal report and the principal must compensate him for this event. In the proof of the proposition, we show that for relatively small values of ρ, the positive effect dominates the negative effect. In other words, the principal is better off if she requires the agent to use the costly channel if the project is not completed before the soft deadline is reached. On the other hand, if ρ is sufficiently

[22] The presence of an informal channel allows us to invoke the revelation principle. To respect limited liability, ρ should be interpreted as a direct loss of utility to the agent rather than a monetary cost.

[23] It is known that money burning can be useful in settings where the principal has private information about the agent's performance (MacLeod, 2003; Fuchs, 2007).

large, then the second effect dominates and it is optimal for the principal to use only the informal channel for reporting progress.

Imposing a small cost of reporting progress on the agent after the soft deadline helps the principal to screen types more effectively. Note that this holds even though we have assumed that the reporting cost is the same whether or not the agent submits an honest report. The intuition is that reporting costs relax the truth-telling constraint with probability one, while the cost is incurred with probability less than one (i.e., the agent avoids paying the reporting cost if he completes the project prior to the soft deadline). Reporting costs, therefore, have greater impact on an agent who has not yet made progress, and it is this differential impact that allows the principal to benefit from making communication costly. To highlight this point further, it is worth noting that the principal is strictly worse off under a direct mechanism that *always* requires costly reports than under a direct mechanism that *never* requires them.

Finally, recall that the optimal contract can be implemented via the OMC-contract (see Section 3.2), where no communication takes place prior to the soft deadline. With costly reporting, communication prior to the soft deadline is also not required. That is, the optimal contract can be implemented with a termination policy and reward scheme that do not depend at all on reports made using the informal channel.Therefore, provided that ρ is not too large, the informal channel is unnecessary; the OMC-contract using only the formal channel can be used to implement the optimum. Further, if the informal channel is unavailable (i.e., any communication requires the agent to incur a cost ρ), then the OMC-contract is uniquely optimal since it minimizes the probability of incurring the reporting cost.[24]

4.2 Asymmetric Stages

In many (probably most) relevant applications, each stage of the project is different. For example, one stage may be expected to take more time (have a smaller λ), require more working capital (higher c), and/or yield greater private benefits to the agent from shirking (higher ϕ). In this subsection, we extend our analysis to a setting with stages that are not identical.

In general, a stage $s \in \{1, 2\}$ can be described by the pair $(\phi_s/\lambda_s, c_s/\lambda_s)$. To fix ideas, we set $\phi_1 = \phi_2 = \phi$ and $c_1 = c_2 = c$ and parameterize the asymmetry of stages by $\alpha \in [-1, 1]$, where

$$\lambda_1 \equiv \frac{\lambda}{1+\alpha} \qquad \text{and} \qquad \lambda_2 \equiv \frac{\lambda}{1-\alpha}$$

for some fixed λ. This parameterization maintains a fixed first-best project value of $\Pi - 2c/\lambda$

[24]Uniqueness is with respect to the communication and termination policy. The reward function after the soft deadline is not uniquely determined. See footnote 21.

as α varies, thereby allowing us to isolate the effect of the asymmetry. Also, note that the probability distribution of τ_2 is symmetric in α. For $\alpha = 0$, the two stages are identical. For $\alpha > 0$, the first stage is expected to take more time and require a larger fraction of the total resources than the second stage. We therefore refer to the first stage as being *harder* if $\alpha > 0$, and *easier* if $\alpha < 0$.

In the online appendix, we extend the formal analysis to this setting. Specifically, with asymmetric stages, the principal's value function and the optimal contract have the same structure as when the stages are symmetric, featuring a soft deadline and a subsequent probationary phase absent a reported breakthrough.

Proposition 4. *Provided the project is feasible, under the optimal contract the following claims hold.*

(i) The principal's payoff is strictly increasing in α for α close to 0.

(ii) The principal's payoff is strictly decreasing (increasing) in α for α close to 1 (-1).

(iii) As the project stages become highly asymmetric (i.e., $\alpha \to \pm 1$), the principal's payoff converges to her payoff in a single-stage project with arrival rate $\lambda/2$.

Part (i) shows that the principal is better off (worse off) if the first stage is moderately more difficult (easier) than the second. The intuition is that under the optimal contract with symmetric stages, the agent is granted more time (in expectation) to complete the first stage than to complete the second one.[25] If $\alpha > 0$, then the first stage is expected to take longer to complete than the second, which coincides with the allocation of time and resources under the optimal contract. However, part (ii) says that making the first stage too difficult relative to the second is not advantageous for the principal. Taken together (i) and (ii) suggests that there is an interior optimal level of difficulty, $\alpha^* \in (0, 1)$, which maximizes the principal's ex-ante payoff.[26] Intuitively, as $\alpha \to \pm 1$, the project effectively has only one stage and the principal's payoff converges accordingly as confirmed by part (iii). Note that such a project essentially has no intermediate breakthrough, which hampers the principals ability to monitor the agent via progress reports and allocate resources accordingly.

These results have several novel implications for the optimal design of projects. For example, it is better to design projects in a way that places somewhat more difficult stages first. Yet, one should be careful not to make the earlier stages too difficult as doing so inhibits the ability to effectively manage agency rents and allocate resources across stages. In particular, it is always better to split a single-stage project into a project with two asymmetric stages even though the principal cannot directly observe progress.

[25] This assertion appeared as an independent result in an earlier draft. The proof is available upon request.

[26] Numerical examples are consistent with this conjecture, though we do not provide a formal proof.

4.3 Unobservable Progress

Because the agent privately observes the state of the project, the principal must give the agent information rents in order to induce truthful reporting. However, the principal also uses the agent's own reports to provide better incentives. In considering how to design projects, it is then natural to ask whether the incentive benefits outweigh the costs of inducing honest reporting. That is, given that the principal cannot observe progress herself, is it better to also restrict the agent's ability to do so? By way of terminology, we refer to progress as being unobservable if τ_1 is not observed by either party and hence cannot be contracted upon either directly or indirectly.

Proposition 5. *When progress is unobservable, it is optimal for the principal to provide incentives through a simple contract. Her ex-ante payoff under this scheme is, however, strictly lower than under the optimal contract when the agent privately observes progress.*

As discussed in Section 3.2, when progress is privately observed by the agent, the principal can do strictly better than a simple contract by using a more sophisticated termination policy in which the project is cancelled at a constant rate if the agent has not reported progress by the soft deadline. This allows the principal to give a low-type agent more expected time to complete the project. Why is it that the principal cannot benefit from a similar strategy with unobservable progress? The reason is that the principal has no way of knowing whether the project is still in the first stage or has progressed to the second. Any policy which uses random termination must do so indiscriminately, which means terminating some projects that have already advanced to the second stage. This highlights the cost of unobservable progress relative to progress privately observed by the agent.

The benefit of unobservable progress is that the agent also does not know whether progress has been made, and it is therefore easier to induce effort.[27] For example, it is no longer the case that, absent progress, the agent necessarily begins shirking as the deadline approaches (i.e., Proposition 2 no longer holds), which is why a simple contract remains optimal when progress is unobservable.

Proposition 5 shows that the principal benefits more from effectively using information about progress than the rents she gives up to acquire it. In other words, it is better for the principal to ask a privately-informed agent for progress reports than to contract with an agent who is unable to observe the evolution of the project. Thus, when designing projects and their incentives schemes, the principal should focus on how best to illicit and effectively use the agent's private information (as in Thoerem 1) rather than trying to suppress his

[27] This benefit is similar to the optimal contract in Fuchs (2007), in the sense that by providing the agent with less information, there are fewer incentive compatibility constraints to satisfy.

access to it.

5 Concluding Remarks

In this paper we study the optimal provision of incentives for two-stage projects in which the agent privately observes intermediate progress. We characterize the principal's optimal contract as well as the set of Pareto-efficient contracts. We also explore the implications for optimal project design.

The optimal contract exhibits features commonly observed in the execution of real-world multistage projects. Specifically, progress reports by the individuals responsible for moving the venture forward are often required, although the consequences for delayed (reported) progress are often vague or imprecise. The optimal contract with the minimal amount of communication involves the use of a soft deadline, prior to which the agent is not required to communicate with the principal and after which breakthroughs must be reported immediately in order to avoid suboptimal project cancelation. This feature closely resembles the project management concept of "slippage," wherein later stages of the project are delayed until either progress is reported or the project is canceled. We also show that any pair of Pareto-efficient payoffs can be achieved by appropriately varying the length of the soft deadline, shorter horizons benefitting the principal relative to the agent.

Regarding the design of projects, we demonstrate three results. First, the principal achieves a higher payoff by imposing a small cost to the agent for reporting a breakthrough after some point. Second, the principal benefits from making the first stage somewhat more difficult than the second. Third, the principal does better if the agent privately observes progress and makes reports than if he does not observe progress himself.

There are numerous avenues for future work in this area. For instance, progress need not be completely unobservable and unverifiable but might be imperfectly observed by the principal, perhaps as the result of a costly audit. Also, we have focused on a setting with two discrete stages. Natural extensions would be to consider a setting with more stages (or model progress as a continuous process) and allow for the possibility of setbacks along the path to project completion. Additionally, in order to isolate progress as an instrument for providing incentives, we have suppressed uncertainty about the underlying value of the project. It would be edifying to study the role of progress in an environment where parties learn about the value of the project. Finally, our analysis is couched in a setting where the principal has full commitment power. Relaxing this assumption may well shed light on other important aspects of the environment.

References

Battaglini, M. (2005). Long-term contracting with markovian consumers. *American Economic Review 95*, 637–658.

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance 22*(6-8), 703–735.

Bergemann, D. and U. Hege (2005). The Financing of Innovation: Learning and Stopping. *Rand Journal of Economics 36*(4), 719–752.

Bergemann, D. and J. Valimaki (2010). The Dynamic Pivot Mechanism. *Econometrica 78*(2), 771–789.

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large risk, limited liability and dynamic moral hazard. *Econometrica 78*, 73–118.

Board, S. (2007). Selling options. *Journal of Economic Theory 136*, 324–340.

Bonatti, A. and J. Hörner (2011). Collaborating. *American Economic Review 101*(2), 632–663.

Clementi, G. L. and H. A. Hopenhayn (2006). A Theory of Financing Constraints and Firm Dynamics. *Quarterly Journal of Economics 121*, 229–265.

DeMarzo, P. M. and M. J. Fishman (2007). Optimal long-term financial contracting. *Review of Financial Studies 20*, 2079–2128.

DeMarzo, P. M., D. Livdan, and A. Tchistyi (2014). Risking Other People's Money: Gambling, Limited Liability, and Optimal Incentives. *Working Paper*.

DeMarzo, P. M. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance 61*, 2681–2724.

Dumas, B. (1991). Super contact and related optimality conditions. *Journal of Economic Dynamics and Control 15*, 675–685.

Edmans, A., X. Gabaix, T. Sadzik, and Y. Sannikov (2012). Dynamic CEO Compensation. *Journal of Finance 67*(5), 1603–1647.

Eso, P. and B. Szentes (2007). Optimal information disclosure in auctions and the handicap auction. *Review of Economic Studies 74*, 705–731.

Ewusi-Mensah, K. and Z. H. Przasnyski (1991). On Information Systems Project Abandonment : An Exploratory Study of Organizational Practices. *MIS Quarterly* (March), 67–87.

Fernandes, A. and C. Phelan (2000). A Recursive Formulation for Repeated Agency with History Dependence. *Journal of Economic Theory 91*, 223–247.

Fuchs, W. (2007). Contracting with Repeated Moral Hazard and Private Evaluations.

American Economic Review 97, 1432–1448.

Glass, R. L., J. Rost, and M. S. Matook (2008). Lying on software projects. *IEEE Software 25*(6), 90–95.

Green, B. and C. Taylor (2015). On Breakthroughs, Deadlines, and the Nature of Progress: Optimal Contracts for Multistage Projects.

Green, E. J. (1987). Lending and the Smoothing of Uninsurable Income. In E. C. P. Wallace and Neil (Eds.), *Minnesota Studies in Macroeconomics, Volume I.* University of Minnesota Press.

Guo, Y. and J. Hörner (2015). Dynamic Mechanisms without Money. *Working Paper*.

Halac, M., N. Kartik, and Q. Liu (2015). Optimal Contracts for Experimentation. *Working Paper*.

Hoffmann, F. and S. Pfeil (2010). Reward for luck in a dynamic agency model. *Review of Financial Studies 23*, 3329–3345.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization 7*, 24–52.

Hopenhayn, H. A. and J. P. Nicolini (1997). Optimal Unemployment Insurance. *Journal of Political Economy 105*, 412.

Hörner, J. and L. Samuelson (2014). Incentives for Experimenting Agents. *RAND Journal of Economics*.

Hu, Z. (2014). Financing Innovation with Unobserved Progress. *Working Paper*.

Inderst, R. and H. M. Mueller (2010). CEO replacement under private information. *Review of Financial Studies 23*, 2935–2969.

Kelley, J. E. (1961). Critical-Path Planning and Scheduling: Mathematical Basis. *Operations Research 9*(3), 296–320.

Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics 32*, 514–526.

Levitt, S. D. and C. M. Snyder (1997). Is no news bad news? Information transmission and the role of "early warning" in the principal-agent model. *RAND Journal of Economics 28*, 641–661.

Lewis, T. R. (2012). A theory of delegated search for the best alternative. *RAND Journal of Economics 43*, 391–416.

MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review 93*(1), 216–240.

Malcolm, D., J.H.Roseboom, C.E.Clark, and W.Fazar (1959). Application of a Technique for

Research and Development Program Evaluation. *Operations Research 7*(n.5, pp 646-669), 25.

Manso, G. (2011). Motivating Innovation. *Journal of Finance 66*, 1823–1860.

Mason, R. and J. Välimäki (2015). Getting it done: dynamic incentives to complete a project. *Journal of the European Economic Association 13*(1), 62–97.

Moroni, S. (2015). Experimentation in Organizations. *Working Paper*.

Myerson, R. B. (1986). Multistage Games with Communication. *Econometrica 54*(2), 323–358.

Pavan, A., I. Segal, and J. Toikka (2014). Dynamic Mechanism Design: A Myersonian Approach. *Econometrica 82*(2), 601–653.

Phelan, C. and R. M. Townsend (1991). Computing Multi-Period, Information-Constrained Optima. *Review of Economic Studies 58*, 853.

Piskorski, T. and A. Tchistyi (2011). Stochastic house appreciation and optimal mortgage lending. *Review of Financial Studies 24*, 1407–1446.

Quadrini, V. (2004). Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics 51*, 713–751.

Radner, R. (1985). Repeated Principal-Agent Games with Discounting. *Econometrica 53*(5), 1173–1198.

Ray, D. (2002). The Time Structure of Self-Enforcing Agreements. *Econometrica 70*(2), 547–582.

Royce, W. W. (1970). Managing the Development of Large Software Systems. *Ieee Wescon* (August), 1–9.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*, 957–984.

Shavell, S. and L. Weiss (1979). The Optimal Payment of Unemployment Insurance Benefits over Time. *Journal of Political Economy 87*(6), 1347.

Spear, S. E. and S. Srivastava (1987). On Repeated Moral Hazard with Discounting. *The Review of Economic Studies 54*, 599.

Tchistyi, A. (2013). Security Design with Correlated Hidden Cash Flows: The Optimality of Performance Pricing. *Working Paper*.

Tirole, J. (2006). *The Theory of Corporate Finance*, Volume 8. Princeton University Press.

Toxvaerd, F. (2006). Time of the essence. *Journal of Economic Theory 129*(1), 252–272.

Varas, F. (2015). Contracting Timely Delivery with Hard to Verify Quality. *Working Paper*.

Williams, N. (2011). Persistent Private Information. *Econometrica 79*(4), 1233–1275.

Zhang, Y. (2009). Dynamic contracting with persistent shocks. *Journal of Economic Theory 144*, 635–675.

A Derivation of the Optimal Two-Stage Contract

In this section, we present the key steps used to prove our main results (Theorems 1 and 2). Formal proofs are located in Appendix B.

A contract is denoted by $\Gamma = \{a, Y, T\}$, where a is the recommended action, T is the termination rule, and dY_t denotes the payment to the agent at time t.[28] Each of the elements of the contract can depend on the history, which includes any reports made by the agent, whether the project has succeeded, and a public randomization device. Clearly, the principal cannot condition payments or the termination rule directly on τ_1, rather, she can only condition on information communicated by the agent and τ_2. Given an arbitrary contract, the agent's continuation value can depend on any information communicated and the actual stage of the project.[29]

By the revelation principle, in searching for the optimal mechanism it is without loss to focus on direct mechanisms (see Myerson (1986) or Pavan et al. (2014) for further discussion). We can therefore restrict attention to mechanisms that induce the agent to report truthfully and immediately. We formulate the principal's problem recursively, making use of three state variables: the project stage $s \in \{1, 2\}$ (as reported by the agent) and the pair of promised continuation values to each type, $(u_1, u_2) \in \mathbb{R}^2_+$. The additional state variable is needed due to the persistent nature of the agent's private information (Fernandes and Phelan, 2000).

For states in which $s = 2$, u_1 can be interpreted as the maximal payoff that a low type could obtain by falsely reporting a breakthrough. For states in which $s = 1$, u_2 can be interpreted either as the maximal payoff that a high type could obtain by not reporting progress or as the promised reward to the low type for making a breakthrough in that state. Importantly, these two payoffs must be identical given the agent is behaving optimally.

Implementable utility pairs

In order to solve the principal's problem, it is useful to first characterize the relevant domain of agent payoffs. That is, the set of (u_1, u_2) that are *implementable*, meaning there is a mapping from the agent's type, s, to a contract Γ_s such that (i) each agent prefers to report his type truthfully, (ii) the recommended action is incentive compatible, and (iii) the contract delivers expected payoff of u_s to an agent who truthfully reports $s \in \{1, 2\}$.

Lemma A.1. *For a two-stage project, the set of implementable utility pairs is given by* $\mathcal{U} = \{(u_1, u_2) \in \mathbb{R}^2_{++} : u_2 \geq u_1\} \cup (0, 0)$.

Intuitively, the high type can always "mimic" the low type, and thus $u_1 > u_2$ is not implementable. Any pair in $\mathcal{U}_L \equiv \mathcal{U} \cap \{u_2 < u_1 + \phi/\lambda\}$ can be implemented with a mechanism in which the agent has a choice between a simple contract and a severance payment with immediate termination (the high type chooses the simple contract while the low type opts for the severance payment). Any pair in $\mathcal{U}_H \equiv \mathcal{U} \cap \{u_2 \geq u_1 + \phi/\lambda\}$ can be implemented through a two-phase mechanism where initially both type of agents work and continuation values decrease along the ray connecting the initial point to the origin. If the project has not succeeded by the time continuation values intersect $\mathcal{L}_H$ (see Figure 1), the simple/severance mechanism is again used.

The principal's problem can then be reduced to choosing an initial point in $\mathcal{U}$ as well as how the continuation values should evolve as a function of the history. We derive the solution using backward

[28]Unlike the one-stage project, we cannot immediately rule out payments being made at dates other than that of project success. Though such payments are not part of the optimal contract, allowing for them facilitates the proof of several intermediate results (e.g., Lemma A.3).

[29]Note that the actual time at which the breakthrough was made is irrelevant for the agent's *continuation value*. Only whether a breakthrough was made and if and when it was reported matters.

induction on the project stage. That is, given an arbitrary implementable utility pair, $(u_1, u_2) \in \mathcal{U}$, we first solve for the principal's value function and optimal policy conditional on reaching the second stage and then use these payoffs to find the optimal first-stage policy. The remaining step (which can be found in the appendix) is to verify that the contract stated in Theorem 1 delivers a payoff to the principal equal to the one under the optimal first-stage policy and satisfies incentive compatibility and truth-telling constraints.

Second stage

Let F_2 denote the principal's second-stage value function, which maximizes her payoff subject to incentive compatibility, delivering the required promised utility u_2 to a high type, and delivering no more than u_1 to the low type. The latter constraint differentiates the second stage problem from the single-stage setting in Section 2. That is, in order to make sure that a low-type agent cannot benefit by falsely reporting progress in the first stage, the principal must limit the maximal payoff that the agent could obtain by doing so, which inherently constrains the amount of time she can give the agent to complete the second stage.

Lemma A.2. *For any $(u_1, u_2) \in \mathcal{U}$, the principal's value function in the second stage is given by*

$$F_2(u_1, u_2) = \left(1 - e^{-\lambda u_1/\phi}\right)\left(\Pi - \frac{c}{\lambda}\right) - u_2. \tag{9}$$

Moreover, conditional on reaching the second stage with promised utilities (u_1, u_2), the optimal continuation contract can be implemented using a simple contract with deadline u_1/ϕ.

The expression for the principal's value function is intuitive in light of the single-stage benchmark (see equation (3)). The larger is u_1, the longer is the deadline the principal can give the agent and therefore the higher is the probability of making the last breakthrough. However, in contrast to the single-stage benchmark and for reasons described above, the link between the promised utility to the agent along the equilibrium path (i.e., u_2) and the maximum amount of time the principal can give the agent to complete the project (i.e., u_1/ϕ) is decoupled.

First stage

We can now turn to the principal's problem in the first stage. Prior to a reported breakthrough, the principal chooses a termination rule, reward scheme and recommended action, as well as how much utility to deliver to each type of agent upon reporting a breakthrough, denoted by W_1, W_2. To induce truth telling, two additional constraints are required. We incorporate random termination (a necessary feature of the optimal contract) by letting the principal choose a distribution over termination dates denoted by S, which combined with a induces a distribution over (τ_1, τ_2, T).[30] Denote the corresponding expectation operator by $\mathbb{E}^{(a,S)}$. The principal's problem in the first stage can be written as follows:

$$\sup_{\Gamma} \mathbb{E}^{(a,S)}\left[F_2(W_1(\tau_1), W_2(\tau_1))\mathbb{1}_{\{\tau_1 \le T\}} - \int_0^{T\wedge\tau_1}\{cdt + dY_t\}\,\Big|\, s = 1\right] \tag{OBJ$_1$}$$

[30] Naturally, we require S to be a right-continuous process and S_t to be measurable with respect to the principal's information set, including whether a breakthrough has been reported at (or prior to) time t.

subject to

$$a \in \arg\max_{\tilde{a}} \mathbb{E}^{(\tilde{a},S)} \left[W_2(\tau_1)\mathbb{1}_{\{\tau_1 \le T\}} + \int_0^{T\wedge\tau_1} \{\phi(1-a_t)dt + dY_t\} \,\Big|\, s=1 \right] \tag{10}$$

and for all $t < T$, the truth-telling constraints are given by

$$W_1(t) \le U_1(t) \equiv \mathbb{E}_t^{(a,S)} \left[W_2(\tau_1)\mathbb{1}_{\{\tau_1 \le T\}} + \int_t^{T\wedge\tau_1} \{\phi(1-a_t)dt + dY_t\} \,\Big|\, s=1 \right] \tag{11}$$

$$W_2(t) \ge \max_{\tilde{a},\tilde{\tau}_1 \ge t} \mathbb{E}_t^{(\tilde{a},S)} \left[W_2(\tilde{\tau}_1)\mathbb{1}_{\{\tilde{\tau}_1 \le T\}} + \int_t^{T\wedge\tilde{\tau}_1} \{\phi(1-\tilde{a}_t)dt + dY_t\} \,\Big|\, s=2 \right]. \tag{12}$$

The first truth-telling constraint ensures that the low-type agent does not want to falsely report a breakthrough and the second ensures that the high-type agent cannot benefit from "hiding" a breakthrough from the principal. To solve the principal's first-stage problem, we first show that it is without loss to focus on contracts in which the agent does not shirk along the equilibrium path (Lemma A.3). We then formulate a recursive version of the problem that relaxes (12) and hence only requires keeping track of the low-type agent's continuation value. The solution to the relaxed program is characterized in Lemma A.4.

Lemma A.3. *It is without loss to focus on contracts such that the agent does not shirk prior to termination.*

The intuition for this result is similar to how a simple contract can be improved upon as described in Section 3.1. That is, it is cheaper to compensate the agent directly with a payment than by letting him shirk. To formulate the problem recursively, we incorporate the principal's ability to randomly terminate the project by letting σ denote the hazard rate of termination. The HJB for the relaxed problem is as follows,

$$0 = \max\left\{ \sup_{w_1,w_2,\sigma} \left\{ \lambda F_2(w_1,w_2) - (\lambda+\sigma)F_1(u_1) - c + F_1'(u_1)\frac{du_1}{dt} \right\},\; u_1 F_1'(u_1) - F_1(u_1) \right\}, \tag{HJB}$$

subject to

$$u_1 \ge w_1 \tag{NFP}$$

$$\lambda(w_2 - u_1) \ge \phi \tag{IC}$$

$$\frac{du_1}{dt} = -\lambda(w_2 - u_1) + \sigma u_1 \tag{PK}$$

$$F_1(0) = 0. \tag{BC}$$

The no-false-progress constraint (NFP) requires that the "reward" to a low-type agent for falsely reporting a breakthrough (w_1) is no more than what she gets for being truthful (u_1). Because F_2 is increasing in w_1 (Lemma A.2), this constraint binds; the principal would like to increase w_1 above u_1 following a reported breakthrough in order to give the agent more time to make the second breakthrough, but doing so would induce a false report. From Lemma A.2, we also know that F_2 is decreasing in w_2. Increasing w_2 simply means giving more rents to a high-type agent who reports a breakthrough without increasing the probability of project success. Clearly, there is no reason for the principal to do this except to incentivize the low-type agent. Hence, (IC) also binds.

Having established these two intuitive features, we state the solution to the relaxed problem and then explain several of its additional characteristics.

Lemma A.4. *The solution to* (HJB) *is given by*

$$F_1(u_1) = \begin{cases} \left(\Pi - \frac{2c}{\lambda}\right)\left[1 - \left(\frac{u_1 - u_s + \frac{2\phi}{\lambda}}{u_s + \frac{2\phi}{\lambda}}\right) e^{-\frac{\lambda}{\phi}(u_1 - u_s)}\right] - u_1 & \text{if } u_1 \geq u_s, \quad (13) \\ \left(\Pi - \frac{2c}{\lambda}\right)\left[\frac{u_1}{u_s + \frac{2\phi}{\lambda}}\right] - u_1 & \text{if } u_1 \in [0, u_s). \quad (14) \end{cases}$$

where u_s is implicitly defined by

$$\left(\frac{\lambda u_s}{\phi} + 2\right) e^{-\lambda u_s/\phi} \equiv \frac{\lambda\Pi - 2c}{\lambda\Pi - c}. \tag{15}$$

For all $u_1 \geq u_s$, the optimal policy involves $w_2 = u_1 + \phi/\lambda$, $w_1 = u_1$, $\sigma = \frac{\phi}{u_s}\mathbb{1}_{\{u_1 = u_s\}}$. For $u_1 < u_s$, the optimal policy involves giving the low-type agent continuation value of u_s with probability u_1/u_s, and terminating the project with probability $1 - u_1/u_s$.

To interpret the principal's value function, recall from Lemma A.3 that the agent never shirks under an optimal contract. The only source of inefficiency, therefore, is the potential for project termination. The social value of the contract, $F_1(u_1) + u_1$, thus equals the first-best value of the project, $\Pi - 2c/\lambda$, times the probability the project is successfully completed, which is the expression in square brackets of (13) and (14).

To see how u_s is determined, note first that for $u_1 > u_s$, F_1 satisfies

$$\lambda F_1(u_1) = \lambda F_2(u_1, u_1 + \tfrac{\phi}{\lambda}) - c - \phi F_1'(u_1), \tag{16}$$

which has a solution of the form

$$F_1(u_1) = \Pi - \frac{2c}{\lambda} - u_1 - u_1 e^{-\frac{\lambda u_1}{\phi}}\left(\frac{\lambda\Pi - c}{\phi}\right) + H_1 e^{-\frac{\lambda u_1}{\phi}}, \tag{17}$$

where H_1 is an arbitrary constant. Next, observe that the net effect of increasing the rate of termination impacts the second term in (HJB), i.e.,

$$\underbrace{u_1 F_1'(u_1)}_{\text{Benefit of preserving } u_1} \quad - \quad \underbrace{F_1(u_1)}_{\text{Opportunity cost of termination}}$$

The first part of this expression is the benefit to the principal of "pushing" u_1 (and hence w_1) upward, which preserves additional time conditional on reaching the second stage. The second part is the opportunity cost of terminating the project. Since the principal's problem is linear in σ, the optimal point at which she begins using a flow rate of termination (i.e., u_s) must be such that the net effect is exactly zero. That is,

$$u_s F_1'(u_s) - F_1(u_s) = 0, \tag{18}$$

which can be interpreted as a local optimality condition. Thus, the boundary condition in (18) implies a constant H_1 for each candidate u_s. It turns out that maximizing H_1 (and therefore F_1)

over all possible u_s is equivalent to requiring that F_1 be twice differentiable at u_s, i.e.,

$$F_1''(u_s) = 0, \tag{19}$$

which is often referred to as the super-contact condition (Dumas, 1991). The value function stated in Lemma A.4 is the solution to the free-boundary problem implied by (17)-(19), which implies an upper bound on the principal's payoff under any contract. The remaining two steps needed to prove Theorem 1 are (i) to verify that the contract stated in the theorem is incentive compatible (i.e., satisfies (10)-(12)), and (ii) show that principal's payoff under the contract is equal to that in the solution to the relaxed program. These steps are straightforward and the details are relegated to a formal proof in the next section.

B Proofs

Before proving, Proposition 1, we state and prove a lemma characterizing incentive compatible contracts.

Lemma B.1. *Given any contract Γ, the optimal action for the agent at time t is*

$$a_t = 1 \iff R_t \geq U_t(\Gamma) + \frac{\phi}{\lambda}. \tag{20}$$

Proof. The HJB equation for the agent's problem can be derived in the usual way.

$$U_t = \sup_{a_t} (\lambda a_t R_t + \phi(1-a_t))dt + (1 - \lambda a_t dt)U_{t+dt}$$

Using a Taylor expansion $U_{t+dt} = U_t + U_t' dt + o(dt)$, canceling U_t on both sides, dividing by dt and taking the limit as $dt \to 0$, we obtain

$$0 = U_t' + \sup_{a_t} \left\{ \phi(1-a_t) + \lambda a_t (R_t - U_t) \right\}. \tag{21}$$

The lemma follows because (i) the HJB in (21) is a necessary condition for a to solve the agent's problem (i.e., maximize (2)) and (ii) it is satisfied if and only if (20) holds. □

Proof of Proposition 1. Fix an arbitrary $u \in \mathbb{R}_+$, we first show that $V(u) \leq \bar{V}(u)$. Since $\max_a U_0(\Gamma) \geq \phi T$, incentive compatibility and promise keeping requires that $T \leq u/\phi$. Because $\lambda\Pi > c \geq \phi$, given any finite deadline T satisfying $T \leq u/\phi$, total surplus is maximized by $T = u/\phi$ and $a_t = 1$ for all $t \in [0, u/\phi]$. That is,

$$V(u) + u \leq \max_{a, T \leq u/\phi} \mathbb{E}^a \left[\Pi \cdot \mathbb{1}_{\{\tau \leq T\}} - \int_0^{T \wedge \tau} c dt \right] = \left(1 - e^{-\lambda \frac{u}{\phi}}\right) \left(\Pi - \frac{c}{\lambda}\right).$$

Hence,

$$V(u) \leq \left(1 - e^{-\lambda \frac{u}{\phi}}\right) \left(\Pi - \frac{c}{\lambda}\right) - u = \bar{V}(u). \tag{22}$$

Next, let $\Gamma(u)$ denote the contract with deadline $T = u/\phi$, reward scheme $R(t) = \phi\left(\frac{1}{\lambda} + u/\phi - t\right)$ and $a_t = 1$ for all $t \in [0, T(u)]$. Notice that $P_0(\Gamma(u)) = \bar{V}(u)$, $U_0(\Gamma(u)) = u$, and $\Gamma(u)$ is incentive compatible. Hence $V(u) \geq \bar{V}(u)$, which combined with (22) implies $V(u) = \bar{V}(u)$. Finally, notice that $\bar{V}(u)$ is strictly concave, which confirms that randomization over the termination deadline is suboptimal and $\bar{V}'(u) \geq -1$, which confirms that incentive compatibility binds. □

Proof of Lemma A.1. Any implementable $\vec{u} = (u_1, u_2)$ such that $u_1 = 0$ must involve immediate termination with probability one, otherwise the low-type agent can shirk and obtain a strictly positive expected payoff. Hence, if $u_1 = 0$ then it must also be that $u_2 = 0$. Also, clearly the high type can achieve a weakly higher payoff then the low type given any Γ. Hence, any implementable pair must satisfy $u_2 \geq u_1$. Therefore, $\mathcal{U} \subseteq \{(u_1, u_2) \in \mathbb{R}^2_{++} : u_2 \geq u_1\} \cup (0, 0)$. The rest of the proof is by construction.

Any $\vec{u} \in \mathcal{U}_L$ can be implemented by providing the agent a choice between two contracts. The first contract terminates the project immediately with a severance payment, $P = u_1$. The second contract has a deadline $T = u_1/\phi$ and a reward function $R(t) = \phi\left(\frac{1}{\lambda} + T - t\right) + q$ where $q = \frac{u_2 - u_1}{1 - e^{-\lambda u_1/\phi}}$. Clearly, the first contract delivers u_1 to both types. Under the second contract, the low type will strictly prefer to shirk (yielding a payoff of u_1) and the high type will strictly prefer to work (yielding an expected payoff of $u_2 \geq u_1$). Hence, each type weakly prefers to report truthfully and any $\vec{u} \in \mathcal{U}_L$ is implementable.

To prove that any $\vec{u} \in \mathcal{U}_H$ can be implemented, we use a contract that is independent of the agent's report for a period of length Δ at which point the continuation utilities are $(\hat{u}_1, \hat{u}_2) \in \mathcal{U}_L$ and the contract in the above paragraph is implemented for $t > \Delta$. To do so, define $S = \frac{u_2}{u_1} > 1$ (since $(u_1, u_2) \in \mathcal{U}_H$). Let $(\hat{u}_1, \hat{u}_2)$ be the point in $\mathcal{L}_H$ that intersects the ray with slope S that goes through the origin and (u_1, u_2):

$$S\hat{u}_1 = \hat{u}_1 + \frac{\phi}{\lambda} \iff \hat{u}_1 = \frac{\phi}{\lambda}\left(\frac{u_1}{u_2 - u_1}\right), \text{ and } \hat{u}_2 = \frac{\phi}{\lambda}\left(\frac{u_2}{u_2 - u_1}\right).$$

Let $U_2(t), U_1(t)$ be the functions jointly satisfying

$$\frac{dU_2}{dt} = \lambda S(U_1(t) - U_2(t)), \quad U_2(0) = u_2$$

$$\frac{dU_1}{dt} = \lambda(U_1(t) - U_2(t)), \quad U_1(0) = u_1,$$

which have unique solutions given by

$$U_2(t) = \frac{1}{S-1}\left(S\left(u_1 + (u_2 - u_1)e^{-\lambda t(S-1)}\right) - u_2\right) \tag{23}$$

$$U_1(t) = \frac{1}{S-1}\left(Su_1 - u_2 + (u_2 - u_1)e^{-\lambda t(S-1)}\right). \tag{24}$$

Note that $U_2(t) - U_1(t) = (u_2 - u_1)e^{-\lambda t(S-1)}$, so define Δ such that $U_2(\Delta) - U_1(\Delta) = \frac{\phi}{\lambda}$. That is

$$\Delta = \frac{1}{\lambda(S-1)} \ln\left(\frac{u_2 - u_1}{\phi/\lambda}\right).$$

Finally, let $\hat{R}(t) = U_2(t) + S(U_2(t) - U_1(t))$ for all $t < \Delta$. Note that for $t < \Delta$, neither type shirks since $\hat{R}(t) \geq U_2(t) + \frac{\phi}{\lambda}$ and $U_2(t) \geq U_1(t) + \frac{\phi}{\lambda}$. By construction, the contract that (i) pays $\hat{R}(t)$ for ultimate success prior to Δ and (ii) implements the contract from the first paragraph at $t = \Delta$ (at which point $(U_1(\Delta), U_2(\Delta)) = (\hat{u}_1, \hat{u}_2) \in \mathcal{U}_L$) delivers the desired expected utilities and induces truth telling, which completes the proof. $\square$

Proof of Lemma A.2. Formally, F_2 solves

$$F_2(u_1,u_2) = \sup_{\Gamma} \mathbb{E}^a \left[\Pi \cdot \mathbb{1}_{\{\tau_2 \le T\}} - \int_0^{T\wedge\tau_2} \{cdt + dY_t\} \,\big|s=2 \right] \tag{25}$$

$$\text{s.t.} \quad a \in \arg\max_{\tilde{a}} \mathbb{E}^{\tilde{a}} \left[\int_0^{T\wedge\tau_2} \{\phi(1-\tilde{a}_t)dt + dY_t\} \,|s=2 \right] \tag{26}$$

$$u_2 = \mathbb{E}^a \left[\int_0^{T\wedge\tau_2} \{\phi(1-a_t)dt + dY_t\} \,\big|s=2 \right] \tag{27}$$

$$u_1 \ge \max_{\tilde{a}} \mathbb{E}^{\tilde{a}} \left[\int_0^{T\wedge\tau_2} \{\phi(1-\tilde{a}_t)dt + dY_t\} \,\big|s=1 \right]. \tag{28}$$

We first show that the expression on the right side of (9) is an upper bound. For $\vec{u} \in \mathcal{U}_L \cup \mathcal{L}_H$, we then show there exists a simple contract that achieves this bound. For $\vec{u} \in \mathcal{U}_H \setminus \mathcal{L}_H$, we construct a sequence of contracts under which the principal's value converges to the right-hand side of (9). To derive the bound, note that

$$\begin{aligned}\max_{\tilde{a}} \mathbb{E}^{\tilde{a}} \left[\int_0^{T\wedge\tau_2} (\phi(1-\tilde{a}_t)dt + dY_t) \,\big|s=1 \right] &\ge \mathbb{E}^{\tilde{a}=0} \left[\int_0^{T\wedge\tau_2} (\phi(1-\tilde{a}_t)dt + dY_t) \,\big|s=1 \right] \\ &\ge \mathbb{E}^{\tilde{a}=0} \left[\int_0^{T\wedge\tau_2} \phi dt \big|s=1 \right] = \phi \mathbb{E}[T].\end{aligned}$$

Therefore, to satisfy (28), the termination policy must be such that $\mathbb{E}[T] \le u_1/\phi$. The promise keeping constraint requires

$$u_2 - \mathbb{E}^a \left[\int_0^{T\wedge\tau_2} \phi(1-a_t)dt \big|s=1 \right] = \mathbb{E}^a \left[\int_0^{T\wedge\tau_2} dY_t \big|s=1 \right],$$

and substituting into (25), we get that

$$F_2(u_1,u_2) \le \sup_{a,T} \mathbb{E}^a \left[\Pi \cdot \mathbb{1}_{\{\tau_2 \le T\}} - \int_0^{T\wedge\tau_2} (c - \phi(1-a_t))dt \big|s=1 \right] - u_2, \qquad \text{s.t. } \mathbb{E}[T] \le u_1/\phi.$$

The right-hand side is increasing in a_t, hence

$$F_2(u_1,u_2) \le \sup_T \left(\Pi - \frac{c}{\lambda} \right) \mathbb{E}\left[1 - e^{-\lambda T} \right] - u_2, \qquad \text{s.t. } \mathbb{E}[T] \le u_1/\phi.$$

By Jensen's inequality we have that $\mathbb{E}\left[1-e^{-\lambda T}\right] \le 1 - e^{-\lambda \mathbb{E}[T]}$, which is strictly increasing in $\mathbb{E}[T]$, and therefore the constraint binds. Inserting $\mathbb{E}[T] = u_1/\phi$ completes the proof of the bound.

For $\vec{u} \in \mathcal{U}_L \cup \mathcal{L}_H$, the bound can be achieved by a simple contract with a deadline of $T = u_1/\phi$ and $R(t) = \phi\left(\frac{1}{\lambda} + T - t\right) + q$ where $q = \frac{u_2 - u_1}{1 - e^{-\lambda u_1/\phi}}$. For $\vec{u} \in \mathcal{U}_H \setminus \mathcal{L}_H$, consider the sequence of simple contracts indexed by S and defined as follows:

$$R_S(t) = \begin{cases} U_2(t) + S(U_2(t) - U_1(t)), & t \in [0, \Delta_S] \\ U_2(t) + \frac{\phi}{\lambda}, & t \in (\Delta_S, T_S] \end{cases}$$
$$\Delta_S = \frac{1}{\lambda(S-1)} \ln\left(\frac{u_2 - u_1}{\frac{\phi}{\lambda}}\right)$$
$$T_S = \frac{U_1(\Delta_S)}{\phi},$$

where $U_2(t)$ and $U_1(t)$ are given by (23) and (24) respectively. By construction, we have that

$$u_2 = \int_0^{T_S} \lambda e^{-\lambda t} R_S(t) dt$$
$$u_1 = \int_0^{\Delta_S} \lambda^2 t e^{-\lambda t} R_S(t) dt + e^{-\lambda \Delta_S}(1 + \lambda\Delta_S)\phi(T_S - \Delta_S)$$
$$R_S(t) \geq U_2(t) + \frac{\phi}{\lambda}.$$

Therefore, it is incentive compatible for the high-type agent to work for all $t \in [0, T_S]$ and his expected payoff from doing so is exactly u_2. Furthermore, given that $U_2(t) = U_1(t) + \frac{\phi}{\lambda}$ for all $t \geq \Delta_S$, the maximal payoff to the low type is exactly u_1. Finally, the expected payoff to the principal under this contract is given by $P_S = \left(1 - e^{-\lambda T_S}\right)\left(\Pi - \frac{c}{\lambda}\right) - u_2$ and $\lim_{S\to\infty} T_S = u_1/\phi$ implies that $\lim_{S\to\infty} P_S = F_2(u_1, u_2)$. □

Proof of Lemma A.3. First, note that Lemma A.2 implies it is optimal to induce effort conditional on reaching the second stage. Therefore, it suffices to prove the same result holds in the first stage. To obtain a contradiction, let (t_1, t_2) denote an arbitrary interval of time over which the agent shirks in the first stage under Γ, where $0 \leq t_1 < t_2 \leq T$. To conserve notation, we assume that Γ involves a deterministic termination rule (the arguments in the proof can easily be extended). Define $\Delta = t_2 - t_1$. Let $\hat{\Gamma}$ be such that prior to a reported breakthrough:

(i) $\hat{T} = T - \Delta$,

(ii) For all $t < t_1$, $(\hat{a}_t, \hat{Y}_t, \hat{W}_1(t), \hat{W}_2(t))$ is identical to $(a_t, Y_t, W_1(t), W_2(t))$,

(iii) For $t \in (t_1, \hat{T})$, let $(\hat{a}_t, \hat{Y}_t, \hat{W}_1(t), \hat{W}_2(t)) = (a_{t+\Delta}, Y_{t+\Delta}, W_1(t+\Delta), W_2(t+\Delta))$,

(iv) At time t_1, the principal makes an unconditional payment to the agent in the amount of $d\hat{Y}_{t_1} = \phi\Delta + \mathbb{E}^{a=0}\left[\int_{t_1}^{t_2} dY_t | s = 1\right]$.

It is straightforward to check that if Γ satisfies (10)-(12) for all $t < T$ then $\hat{\Gamma}$ satisfies (10)-(12) for all $t < \hat{T}$. Prior to t_1, both the agent's action and the principal's payoff conditional on a breakthrough is the same under both contracts. If a breakthrough does not happen prior to t_1, then $P_{t_1}(\hat{\Gamma}) = P_{t_1}(\Gamma) + (c - \phi)\Delta \geq P_{t_1}(\Gamma)$. Hence, $P_0(\hat{\Gamma}) \geq P_0(\Gamma)$. □

Proof of Lemma A.4. First, we construct the value function under the stated policy. Using the boundary conditions (18) and (19), we pin down u_s and show that the value function under the stated policy indeed has the form given by (13)-(14). We then verify that, given the u_s implied by the boundary conditions, the value function in (14)-(13) solves (HJB). That $\max_u F_1(u)$ is an upper bound on the solution to (OBJ$_1$) is immediate (since it relaxes (12)). In the Proof of Theorem 1, we show there exists a contract satisfying this additional constraint that achieves the bound.

For $u > u_s$, F_1 evolves according to (16) and therefore has a solution of the form (17). For $u < u_s$, the principal's value under the stated policy is given by

$$F_1(u_1) = \frac{u_1}{u_s} F_1(u_s). \tag{29}$$

There are two unknowns to pin down: (u_s, H_1). Solving (18) for H_1 gives

$$H_1 = \frac{\lambda u_s \left(-c\lambda \left(\Pi\phi \left(e^{\frac{\lambda u_s}{c\phi}} - 2\right) + u_s\right) + \Pi\lambda^2 u_s + 2c^2\phi \left(e^{\frac{\lambda u_s}{c\phi}} - 1\right)\right)}{c\phi(c\phi + \lambda u_s)^2}.$$

Plugging this expression into (19), (or, equivalently, maximizing H_1 over all possible u_s), we get that u_s is defined implicitly by (15) and hence

$$H_1 = \left(\frac{\lambda u_s}{\phi} - 2\right)\left(\Pi - \frac{c}{\lambda}\right). \tag{30}$$

Substituting (30) into (17), we get (13). Then (14) follows from (29), which verifies that the value function under the stated policy for u_s given by (15) indeed has the stated form.

Finally, we verify that F_1 solves (HJB) subject to the four constraints. Given u_s as defined implicity by (15), one can easily check that $F_1(0) = 0$ and thus (BC) is satisfied. Using arguments already given, subject to (IC), (PK), and (NFP), we have that

$$\sup_{w_1,w_2,\sigma} \left\{\lambda F_2(w_1, w_2) - (\lambda+\sigma)F_1(u_1) - c + F_1'(u_1)\frac{du_1}{dt}\right\}$$
$$= \underbrace{\lambda(F_2(u_1, u_1+\phi/\lambda) - F_1(u_1)) - c - \phi F_1'(u_1)}_{L_s(u_1)} + \sup_{\sigma}\left\{\sigma\left(u_1F_1'(u_1) - F_1(u_1)\right)\right\},$$

and by construction, $L_s(u_1) = 0$ for $u_1 \geq u_s$ and $u_1F_1'(u_1) - F_1(u_1) = 0$ for $u_1 \leq u_s$. Hence, it is sufficient to show that (i) $L_s(u_1) \leq 0$ for $u_1 < u_s$, and (ii) $u_1F_1'(u_1) - F_1(u_1) \leq 0$ for $u_1 > u_s$. For (i), notice from (13) that L_s is concave and hence L_s' is decreasing for all $u_1 \in [0, u_s)$. That $L_s'(u_s) = 0$ implies L_s is increasing below u_s, and that $L_s(u_s) = 0$ then gives the result. For (ii), notice from (13) that $F_1''(u_1) < 0$ and hence $u_1F_1'(u_1) - F_1(u_1)$ is decreasing for all $u_1 > u_s$. The result then follows since $u_sF_1'(u_s) = F_1(u_s)$. □

The following lemma will be used in the proof of Theorems 1 and 2.

Lemma B.2. *Consider F_1 and u_s as defined in Lemma A.4. The following statements hold:*

(i) If (C.2) holds strictly, then $u^* \equiv \arg\max_{u\geq 0} F_1(u) > \phi T_1^* > u_s$ and $F_1(u^*) > 0$.

(ii) If (C.2) holds with equality, then $u_s = \phi T_1^*$, $F_1(u) = 0$ for all $u \in [0, u_s]$ and $F_1(u) < 0$ for all $u > u_s$.

(iii) If (C.2) is violated, then $F_1(u) < 0$ for all $u > 0$.

Proof. Recall that F_1 is concave with $F_1(0) = 0$. Hence, $F_1(u^*) > 0 \iff F_1'(0) > 0$. Using (14), $F_1'(0) = \frac{\Pi - 2c/\lambda}{u_s + 2\phi/\lambda} - 1$, which frome (15) is strictly positive iff $\phi T_1^* > u_s$. To determine when $\phi T_1^* > u_s$, consider the function

$$J(u) \equiv \Pi - \frac{2c}{\lambda} - \left(\Pi - \frac{c}{\lambda}\right)\left(\frac{\lambda u}{\phi} + 2\right) e^{-\lambda u/\phi}.$$

Observe that $J(0) = -\Pi$, $\lim_{u\to\infty} J(u) = \Pi - 2c/\lambda$, and $J'(u) > 0$ for all $u < \infty$. Moreover, from (15), $J(u_s) = 0$. Therefore, $\phi T_1^* > u_s \iff J(\phi T_1^*) > 0$. The latter is equivalent to (C.2) holding strictly. That $u^* > \phi T_1^*$ follows from concavity of F_1 and that $F_1'(\phi T_1^*) = F_1'(u_s) > F_1'(u^*) = 0$, which completes the proof of (i). If (C.2) holds with equality, then by a similar argument, $F_1'(0) = F_1'(u_s) = 0$, which combined with concavity above u_s implies (ii). If (C.2) is violated, then $F_1'(0) < 0$ and concavity implies (iii). □

Proof of Theorem 1. From Lemmas A.1-A.4, $F_1(u_0)$ is an upper bound on the principal's payoff under any incentive compatible contract that delivers an ex-ante expected payoff of u_0 to the agent. From Lemma B.2, if (C.2) is violated, then $F_1(u) < 0$ for all $u \geq 0$ and therefore $T_s = u_s = 0$ (i.e., immediate termination) is uniquely optimal.

For the remainder of the proof, assume that (C.2) holds. Let Γ_s denote the contract stated in the theorem with

$$T_s \equiv \max\left\{\frac{u^* - u_s}{\phi}, 0\right\}.$$

We first show that Γ_s induces the prescribed behavior by the agent (i.e., truthful reporting and no shirking). We then show that $P_0(\Gamma_s) = F_1(u^*) = \max_u F_1(u)$. Start from any $t \geq \tau_1$ (i.e., after a breakthrough has been made), and let $U_2(\tau_1, t)$ denote the high-type agent's equilibrium continuation value at time t from following the prescribed behavior in Γ_s.

- If $\tau_1 \geq T_s$, then

$$U_2(\tau_1, t) = \int_t^{\tau_1 + 1/\sigma} \lambda e^{-\lambda(\tau_2 - t)} R_s(\tau_1, \tau_2) d\tau_2 = \phi\left(1 + \frac{\sigma}{\lambda}\right)\left(\frac{1}{\sigma} + \tau_1 - t\right). \tag{31}$$

  To verify it is optimal for the agent to work (conditional on reporting truthfully) until making the second breakthrough or until running out of time, notice that

$$\lambda(R_s(\tau_1, t) - U_2(\tau_1, t)) = \phi\left(1 + \frac{\sigma}{\lambda}\right) > \phi.$$

  Hence, working is strictly optimal. To verify that it is optimal for the agent to report progress immediately (i.e., that(12) holds), note that for any $t \geq T_s$,

$$W_2(t) = U_2(t, t) = \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma}\right).$$

  Due to the stationarity of the continuation contract, if the agent prefers to delay reporting progress at $t = \tau_1$, then he prefers to delay reporting progress indefinitely (i.e., to never report progress). If he delays a report then it is strictly optimal to shirk (since the reward is zero if $\tau_2 \in (T_s, \hat{\tau}_1)$) and his expected payoff is

$$\int_t^{\infty} \sigma e^{-\sigma(s-t)} \phi(s - t) ds = \frac{\phi}{\sigma} < U_2(t, t).$$

  Hence, the agent strictly prefers to report progress as soon as it arrives for all $\tau_1 \geq T_s$.

- If $\tau_1 < T_s$, then

$$U_2(\tau_1, t) = \int_t^{T_s} \lambda e^{-\lambda(\tau_2 - t)} R_s(\tau_1, \tau_2) d\tau_2 + \int_{T_s}^{T_s + 1/\sigma} \lambda e^{-\lambda(\tau_2 - t)} R_s(\tau_1, \tau_2) d\tau_2 = \phi \left(\frac{1}{\lambda} + \frac{1}{\sigma} + T_s - t \right).$$

Hence, $\lambda(R_s(\tau_1, t) - U_2(\tau_1, t)) = \phi$ and $a_t = 1$ is weakly optimal, regardless of whether progress is reported (since R_s is independent of $\hat{\tau}_1$ for $\tau_2 \leq T_s$). Further, the same argument as in the bullet above shows that the agent prefers to report progress at $t = T_s$ and work from that point forward.

Now consider any $t \leq \tau_1$ (i.e., prior to a breakthrough being made). Let $U_1(t)$ denote the low-type agent's equilibrium continuation value at time t under Γ_s.

- If $t \geq T_s$, then $W_2(t) = \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma}\right)$ and thus

$$U_1(t) = \int_t^{\infty} \lambda e^{-(\lambda+\sigma)(\tau_1 - t)} W_2(\tau_1) d\tau_1 = \frac{\phi}{\sigma}. \tag{32}$$

Since $\lambda(W_2(t) - U_1(t)) = \phi$, working is (weakly) optimal for the agent. Next, we verify that the agent does not want to falsely report progress at any $t \geq T_s$. Due to the stationarity of the continuation contract, if the agent prefers to falsely report progress at $t > T_s$, then he prefers to do so at $t = T_s$. Thus, suppose he falsely reports progress at the soft deadline. Let $\tilde{U}(t)$ be his expected payoff at $t \geq T_s$ from acting optimally henceforth. Consider the following chain

$$\begin{aligned} T_s < t &\iff \phi\left(\frac{\sigma}{\lambda}\right)\left(\frac{1}{\sigma} + T_s - t\right) < \frac{\phi}{\lambda} \\ &\iff \phi\left(1 + \frac{\sigma}{\lambda}\right)\left(\frac{1}{\sigma} + T_s - t\right) - \phi\left(\frac{1}{\sigma} + T_s - t\right) < \frac{\phi}{\lambda} \\ &\iff U_2(\tau_1, t) - \phi\left(\frac{1}{\sigma} + T_s - t\right) < \frac{\phi}{\lambda}, \end{aligned}$$

where the last line follows from (31). Next note that

$$\phi\left(\frac{1}{\sigma} + T_s - t\right) \leq \tilde{U}(t),$$

because the left side is the payoff to the agent from falsely reporting progress at T_s, having no breakthrough, and shirking from date t to $T_s + \frac{1}{\sigma}$ and the right side is his expected payoff from falsely reporting progress at T_s, having no breakthrough, and acting optimally from date t to $T_s + \frac{1}{\sigma}$. Thus, we have

$$U_2(t) - \tilde{U}(t) < \frac{\phi}{\lambda},$$

which says that it is suboptimal for the agent to work at any $t > T_s$ after falsely reporting progress at T_s. Thus, if the agent falsely reports progress at T_s, then it is optimal for him to shirk until time runs out; i.e., $\tilde{U}(t) = \phi\left(\frac{1}{\sigma} + T_s - t\right)$. Finally, recall that ϕ/σ is the agent's expected payoff from honestly reporting no progress at T_s and then working while on probation. Thus, at T_s the agent is indifferent between honestly reporting no progress and then (optimally) working while on probation and falsely reporting progress and then (optimally) shirking until time runs out. Therefore, honestly reporting no progress at T_s is weakly optimal.

- Next, consider $t < T_s$. Noting that

$$W_2(t) = \begin{cases} \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma} + T_s - t\right), & \text{if } t < T_s \\ \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma}\right), & \text{if } t \geq T_s. \end{cases}$$

  and

$$U_1(t) = \int_t^{T_s} \lambda e^{-\lambda(\tau_1 - t)} W_2(\tau_1) d\tau_1 + \int_{T_s}^{\infty} \lambda e^{-(\lambda+\sigma)(\tau_1 - t)} W_2(\tau_1) d\tau_1 = \phi\left(\frac{1}{\sigma} + T_s - t\right), \quad (33)$$

  we get that $\lambda(W_2(t) - U_1(t)) = \phi$ for all $t < T_s$, which shows that working is weakly optimal. Finally note that working until T_s is also optimal for the agent if he plans to falsely report progress at T_s, that is $\tilde{U}(t) = U_1(t)$ for $t \leq T_s$.

We have shown that Γ_s induces truth-telling and no shirking. Next we verify that the principal's expected payoff under Γ_s corresponds to $F_2(\tilde{U}(t), U_2(t))$ in the second stage and $F_1(U_1(t))$ in the first stage.

Second Stage: Suppose the agent has made the first breakthrough. By the above analysis, he immediately and truthfully reports τ_1 and continues to work until he makes the second breakthrough or runs out of time.

For any $t \in [\max\{T_s, \tau_1\}, \max\{T_s, \tau_1\} + 1/\sigma]$ the principal's expected payoff is thus

$$\int_t^{\max\{T_s,\tau_1\}+1/\sigma} \lambda e^{-\lambda(\tau_2 - t)} \left(\Pi - \frac{c}{\lambda} - R_s(\tau_1, \tau_2)\right) d\tau_2$$
$$= \left(\Pi - \frac{c}{\lambda}\right)\left(1 - e^{-\lambda(1/\sigma + \max\{T_s,\tau_1\} - t)}\right) - \phi\left(1 + \frac{\sigma}{\lambda}\right)\left(\frac{1}{\sigma} + \max\{T_s, \tau_1\} - t\right) = F_2(\tilde{U}(t), U_2(t)).$$

Similarly for $t \in [\tau_1, T_s]$, the principal's expected payoff is

$$\int_t^{T_s} \lambda e^{-\lambda(\tau_2 - t)} \left(\Pi - \frac{c}{\lambda} - R_s(\tau_1, \tau_2)\right) d\tau_2 + \int_{T_s}^{T_s + 1/\sigma} \lambda e^{-\lambda(\tau_2 - t)} \left(\Pi - \frac{c}{\lambda} - R_s(\tau_1, \tau_2)\right) d\tau_2$$
$$= \left(\Pi - \frac{c}{\lambda}\right)\left(1 - e^{-\lambda(1/\sigma + T_s - t)}\right) - \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma} + T_s - t\right) = F_2(\tilde{U}(t), U_2(t)).$$

First Stage: Now suppose the agent has not yet made the first breakthrough. For $t \geq T_s$, Γ_s calls for random termination at rate σ, so the principal's expected payoff is

$$\int_t^{\infty} \lambda e^{-(\lambda+\sigma)(\tau_1 - t)} \left(\left(\Pi - \frac{c}{\lambda}\right)\left(1 - e^{-\lambda/\sigma}\right) - \frac{c}{\lambda} - \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma}\right)\right) d\tau_1$$
$$= \left(\frac{\lambda}{\lambda + \sigma}\right)\left(\Pi - \frac{2c}{\lambda} - \left(\Pi - \frac{c}{\lambda}\right)e^{-\lambda/\sigma}\right) - \frac{\phi}{\sigma}$$
$$= F_1(U_1(T_s)),$$

where the last equality follows from $u_s = \frac{\phi}{\sigma}$ and substituting from (15). For $t < T_s$ the principal's expected payoff is

$$\int_t^{T_s} \lambda e^{-\lambda(\tau_1 - t)} \left(\left(\Pi - \frac{c}{\lambda}\right)\left(1 - e^{-(1/\sigma + T_s - \tau_1)}\right) - \phi\left(\frac{1}{\lambda} + \frac{1}{\sigma} + T_s - \tau_1\right) - \frac{c}{\lambda}\right) d\tau_1 + F_1(U_1(T_s))e^{-\lambda(T_s - t)}$$

$$= \left(\Pi - \frac{2c}{\lambda}\right)\left(1 - e^{-\lambda(T_s - t)}\right) + \lambda\left(\frac{1}{\sigma} + T_s - t\right)\left(\Pi - \frac{c}{\lambda}\right) e^{-\lambda(1/\sigma + T_s - t)} - \phi\left(\frac{1}{\sigma} + T_s - t\right)$$

$$= F_1(U_1(t)),$$

where the last equality follows from $U_1(t) = \phi\left(\frac{1}{\sigma} + T_s - t\right)$, $u_s = \frac{\phi}{\sigma}$, and using (15) to substitute. Further, $U_1(0) = \phi(T_s + 1/\sigma) = u^*$, so that $F_1(U_1(0)) = F_1(u^*)$, which we have already argued is an upper bound on the solution to the principal's problem (see Proof of Lemma A.4). Thus, we can conclude that Γ_s is an optimal contract. □

Proof of Theorem 2. We break the proof into the three cases corresponding to those in Lemma B.2.

(i) First suppose (C.2) holds strictly. Then by part (i) of Lemma B.2, u^* is the unique critical point of $F_1(\cdot)$ and it corresponds to a maximum. Thus for $u > u^*$, $F_1'(u) < 0$. Moreover $F_1(u^*) > 0$ and $\lim_{u\to\infty} F_1(u) = -\infty$ so $\bar{u} \in (u^*, \infty)$ exists and is well-defined. Payoffs to the northeast of X^P are infeasible since F_1 is an upper bound and payoffs to the southwest are Pareto dominated.

To prove the statement about implementation, take any $(u, v) \in X^P$ and suppose that the agent is given the contract in Theorem 1 with $T_s = \frac{u - u_s}{\phi}$. By precisely the same arguments as in the proof of Theorem 1, the agent prefers not to shirk and prefers to truthfully report, yielding *ex ante* expected payoffs of u and $v = F_1(u)$.

(ii) Suppose (C.2) holds with equality. Then by part (ii) of Lemma B.2, $F_1(u) = 0$ if $u \in [0, \phi T_1^*]$ and $F_1(u) < 0$ if $u > \phi T_1^*$. Hence, the unique Pareto-efficient, individually-rational pair of payoffs is $u = \phi T_1^*$ and $v = 0$. Because $u_s = \phi T_1^*$, the agent is placed on probation immediately with termination rate $\sigma = 1/T_1^*$. Again the same steps as in the proof of Theorem 1 shows that the agent prefers not to shirk and prefers to truthfully report. It also shows that the *ex ante* expected payoff to the principal is

$$\left(\frac{\lambda}{\lambda + \sigma}\right)\left[\Pi - \frac{2c}{\lambda} - \left(\Pi - \frac{c}{\lambda}\right) e^{-\lambda/\sigma}\right) - \frac{\phi}{\lambda}\left(1 + \frac{\lambda}{\sigma}\right). \tag{34}$$

Substituting T_1^* for $1/\sigma$ and noting that $\left(\Pi - \frac{c}{\lambda}\right) e^{-\lambda T_1^*} = \frac{\phi}{\lambda}$, equation (34) becomes

$$\left(\frac{\lambda T_1^*}{1 + \lambda T_1^*}\right)\left[\Pi - \frac{2c}{\lambda} - \frac{2\phi}{\lambda} - \phi T_1^*\right].$$

The expression in square brackets is 0 when (C.2) holds with equality.

(iii) Suppose (C.2) is violated. Then it follows directly from part (iii) of Lemma B.2 that $u^* = \bar{u} = 0$ and $X^P = \{(0, 0)\}$. □

C Proofs for Section 4

Proof of Proposition 3. The proof takes several steps. In Steps 1 through 3 we assume that the principal *must* use the formal channel, but that she does so optimally. In Step 4 we prove that, for ρ sufficiently small, the principal indeed benefits by using the formal channel.

Step 1. First observe that the only possible reason for requiring a costly report is to relax the no-false-progress constraint so as to add more time to the clock following the first reported breakthrough. This implies that the formal communication channel should only be used to report progress (not lack of progress). Next, observe that it cannot be optimal for the principal to require the costly formal report at date t and take no action until date $t' > t$, because this is dominated by waiting until t' to require the report. (The project might be completed between t and t'.) Thus, it is optimal to putoff requiring a formal report as long as possible. Let y be the highest value of the low type's continuation utility at which the principal requires a formal report, and let $F_2(u_1, u_2; \rho)$ denote the principal's value function in the second stage. This value function is given by

$$F_2(u_1, u_2; \rho) = \left(\Pi - \frac{c}{\lambda}\right)\left(1 - e^{-\lambda(u_1+\rho)/\phi}\right) - u_2 - \mathbb{1}_{\{u_1 \le y\}}\rho e^{-\lambda(u_1-y)/\phi}. \tag{35}$$

This is established using the same method of proof as for Proposition A.2 with two straightforward alterations. First, to satisfy (28), the termination policy must be such that $\mathbb{E}[T] \le (u_1 + \rho)/\phi$. Second, for $u_1 \le y$, there is a chance that the high type will have to pay the reporting cost. Therefore, the promise keeping constraint necessitates

$$u_2 + \rho e^{-\lambda(u_1-y)/\phi} = \mathbb{E}^{a=1}\left[\int_0^{T\wedge\tau_2} dY_t \middle| s = 1\right],$$

Because $F_2(u_1, u_2; \rho)$ is concave in u_1 and linear in u_2, $\sigma = 0$ is optimal.

Step 2. We solve for the first-stage value function assuming $\sigma = 0$ and check for concavity. The HJB is

$$\lambda F_1(u_1; \rho) = \max_{y\ge 0} \lambda F_2(u_1, u_1 + \phi/\lambda; \rho) - c - \phi\frac{dF_1}{du_1}.$$

Because the principal prefers to delay formal reports as long as possible, $y = 0$ is optimal. Therefore we have

$$\frac{\lambda}{\phi}F_1(u_1; \rho) + F_1'(u_1; \rho) = \frac{\lambda}{\phi}\left[\Pi - \frac{2c}{\lambda} - \frac{\phi}{\lambda} - u_1 - \left(\rho + \left(\Pi - \frac{c}{\lambda}\right)e^{-\lambda\rho/\phi}\right)e^{-\lambda u_1/\phi}\right].$$

This has a solution of the form

$$F_1^c(u_1; \rho) = \Pi - \frac{2c}{\lambda} - u_1 - \left(\rho + \left(\Pi - \frac{c}{\lambda}\right)e^{-\lambda\rho/\phi}\right)\frac{\lambda u_1}{\phi}e^{-\lambda u_1/\phi} + Ke^{-\lambda u_1/\phi}. \tag{36}$$

Using the boundary condition $F_1(0; \rho) = 0$ gives

$$F_1^c(u_1; \rho) = \left(\Pi - \frac{2c}{\lambda}\right)\left(1 - e^{-\lambda u_1/\phi}\right) - u_1 - \left(\rho + \left(\Pi - \frac{c}{\lambda}\right)e^{-\lambda\rho/\phi}\right)\frac{\lambda u_1}{\phi}e^{-\lambda u_1/\phi}.$$

For small ρ, this is convex for u_1 sufficiently close to zero, necessitating random termination at some point $u_s(\rho)$.

Step 3. Using the same analysis as in the proof of Lemma A.4 yields

$$F_1(u_1;\rho) = \begin{cases} \left(\Pi - \frac{2c}{\lambda}\right)\left[1 - \left(\frac{u_1 - u_s(\rho) + \frac{2\phi}{\lambda}}{u_s(\rho) + \frac{2\phi}{\lambda}}\right) e^{-\frac{\lambda}{\phi}(u_1 - u_s(\rho))}\right] - u_1 & \text{if } u_1 \geq u_s(\rho), \quad (37) \\ \left(\Pi - \frac{2c}{\lambda}\right)\left[\frac{u_1}{u_s(\rho) + \frac{2\phi}{\lambda}}\right] - u_1 & \text{if } u_1 \in [0, u_s(\rho)). \quad (38) \end{cases}$$

where $u_s(\rho)$ is implicitly defined by

$$\Pi - \frac{2c}{\lambda} - \left(\rho + \left(\Pi - \frac{c}{\lambda}\right)^{-\lambda\rho/\phi}\right)\left(2 + \frac{\lambda u_s(\rho)}{\phi}\right) e^{-\lambda u_s(\rho)/\phi} \equiv 0.$$

Step 4. We wish to show that for ρ sufficiently small and any $u_1 > 0$, $F_1(u_1;\rho) > F_1(u_1;0) = F_1(u_1)$. The claim will follow if we show $\frac{\partial F_1(u;0)}{\partial \rho} > 0$ for all $u > 0$. Note first that

$$u_s'(0) = -\frac{\left(\Pi - \frac{c}{\lambda} - \frac{\phi}{\lambda}\right)\left(2 + \frac{\lambda u_s}{\phi}\right)}{\left(\Pi - \frac{c}{\lambda}\right)\left(1 + \frac{\lambda u_s}{\phi}\right)} < 0,$$

and $F_1(u_1;\rho)$ depends on ρ only through $u_s(\rho)$. Therefore, for all $u > 0$,

$$\text{sign}\left(\frac{\partial F_1(u;0)}{\partial \rho}\right) = -\text{sign}\left(\frac{\partial F_1(u;0)}{\partial u_s}\right) > 0,$$

where the inequality follows from observing that both expressions in (37)-(38) are strictly decreasing in $u_s(\rho)$ for all $u_1 > 0$. □

Proof of Proposition 4. First, we extend the analysis from Appendix A to characterize F_1^α for an arbitrary α. Using the same arguments as in Lemma A.2, we have that

$$F_2^\alpha(u_1, u_2) = \left(1 - e^{-\frac{\lambda u_1}{(1-\alpha)\phi}}\right)\left(\Pi - \frac{c(1-\alpha)}{\lambda}\right) - u_2$$

Replacing F_2 with F_2^α in (HJB) and the appropriately modified (binding) constraints, we find that F_1^α has the form

$$\hat{F}_1^\alpha(u_1) = \left(\Pi - \frac{2c}{\lambda} - u_1\right) + \left(\frac{1-\alpha}{2\alpha}\right)\left(\Pi - \frac{(1-\alpha)c}{\lambda}\right) e^{-\frac{\lambda u_1}{(1-\alpha)\phi}} + C_1^\alpha e^{-\frac{\lambda u_1}{(1+\alpha)\phi}}. \quad (39)$$

If the terminal boundary condition is imposed ($\hat{F}_1^\alpha(0) = 0$) then $C_1^\alpha = \frac{-(1+\alpha)(\lambda\Pi - c(1+\alpha))}{2\alpha\lambda}$ and $\hat{F}_1^{\alpha\prime\prime}(0) = \frac{\lambda^2\Pi}{(1-\alpha^2)\phi} > 0$. Hence, there exists some $u_s(\alpha)$ such that random termination is optimal for $u \in (0, u_s(\alpha)]$. Let $c^\alpha(u)$ denote the constant in the principal's value function that satisfies the smooth-pasting condition (i.e., (18)) at an arbitrary $u > 0$. That is,

$$c^\alpha(u) \equiv \frac{(\alpha+1)e^{\frac{2\alpha\lambda u}{(\alpha^2-1)\phi}}\left(c\left(\phi\left(\alpha^2 - 2\alpha + 4\alpha e^{\frac{\lambda u}{\phi - \alpha\phi}} + 1\right) + u(\lambda - \alpha\lambda)\right) - \lambda\Pi\left(-\alpha\phi + 2\alpha\phi e^{\frac{\lambda u}{\phi - \alpha\phi}} + \lambda u + \phi\right)\right)}{2\alpha\lambda(\alpha\phi + \lambda u + \phi)}. \quad (40)$$

Twice differentiability at $u_s(\alpha)$ (i.e., (19)) is equivalent to $u_s(\alpha) = \max_u c^\alpha(u)$, which requires the first-order condition

$$\frac{e^{\lambda u_s(\alpha)/\phi(1-\alpha)}}{\lambda u_s(\alpha)+2\phi}=\frac{\lambda\Pi-c(1-\alpha)}{\phi(1-\alpha)(\lambda\Pi-2c)}. \tag{41}$$

The right-hand side of the above expression is strictly greater than $1/2\phi$ for all $\alpha\in(-1,1)$. The left-hand side is equal to $1/2\phi$ at $u_s(\alpha)=0$, strictly increasing in $u_s(\alpha)$ and unbounded. This guarantees the existence of a unique $u_s(\alpha)$ satisfying (41), which completes the characterization of F_1^α. To summarize,

- For $u\geq u_s(\alpha)$, F_1^α is of the form given in (39) with $C_1^\alpha=c^\alpha(u_s(\alpha))$, where u_s^α is the unique solution to (41).
- For $u\in[0,u_s(\alpha))$, $F_1^\alpha(u)=\frac{u}{u_s(\alpha)}F_1^\alpha(u_s(\alpha))$.

To prove (i), first note that by the envelope theorem $\frac{d}{d\alpha}c^\alpha(u_s(\alpha))=\frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha))$. Using this fact, evaluating the derivative and taking the limit as $\alpha\to 0$, we get that for $u\geq u_s(0)=u_s$,

$$\lim_{\alpha\to 0}\left(\frac{d}{d\alpha}F_1^\alpha(u)\right)=\left(\frac{e^{-\frac{\lambda u}{\phi}}\left(u(\lambda u_s+\phi)-\lambda u_s^2\right)}{\phi^2(\lambda u_s+\phi)^2}\right)\times$$
$$\left(\Pi\lambda\left(\lambda^2u_s^2-\phi^2\left(e^{\frac{\lambda u_s}{\phi}}-1\right)+\lambda u_s\phi\right)-c\left(\lambda^2u_s^2-2\phi^2\left(e^{\frac{\lambda u_s}{\phi}}-1\right)+2\lambda u_s\phi\right)\right).$$

The first-term on the right hand side is clearly positive for $u\geq u_s$. Using (41), the second term reduces to $\lambda(\lambda u_s+\phi)((\lambda\Pi-c)u_s-\Pi\phi)$, which is also clearly positive if $\lambda u_s/\phi>\frac{\lambda\Pi/c}{\lambda\Pi/c-1}$. We now claim that if u_s solves (41) for $\alpha=0$, then this latter inequality must hold. Let $x\equiv\lambda u_s/\phi\geq 0$, $y\equiv\lambda\Pi/c-2>0$, and $\alpha=0$. The claim is that

$$\frac{e^x}{x+2}=\frac{y+1}{y}\implies x>\frac{y+2}{y+1}.$$

To see that this is true, suppose that $\frac{e^x}{x+2}=\frac{y+1}{y}$ and $x\leq\frac{y+2}{y+1}$. Note that $\frac{e^x}{x+2}$ is strictly increasing. Therefore,

$$\frac{e^x}{x+2}\leq\frac{e^{\frac{y+2}{y+1}}}{\frac{y+2}{y+1}+2}<\frac{y+1}{y},$$

which gives the contradiction. We have thus shown that at $\alpha=0$, the derivative of $F_1^\alpha(u)$ w.r.t. α is strictly positive for all $u\geq u_s(\alpha)$. That the same statement is true for $u\in(0,u_s)$ is immediate by the linearity of the value function below u_s. Since F_1^α is also continuously differentiable in both of its arguments, it must be strictly increasing in a neighborhood around $\alpha=0$ for all $u>0$, which completes the proof of (i).

We prove (ii) and (iii) for the case of $\alpha\to 1$, the proof for $\alpha\to -1$ follows a similar argument. We first show that $\lim_{\alpha\to 1}u_s(\alpha)=0$. To do so, rewrite (41) as

$$\phi(1-\alpha)e^{\lambda u_s(\alpha)/\phi(1-\alpha)}=\frac{(\lambda\Pi-c(1-\alpha))(\lambda u_s(\alpha)+2\phi)}{\lambda\Pi-2c}.$$

Suppose $u_s(1)\equiv\lim_{\alpha\to 1}u_s(\alpha)\in(0,\infty)$. Then $\lim_{\alpha\to 1}\phi(1-\alpha)e^{\lambda u_s(\alpha)/\phi(1-\alpha)}=\infty>\frac{(\lambda\Pi-c(1-\alpha))(\lambda u_s(1)+2\phi)}{\lambda\Pi-2c}$, a contradiction. Also, clearly $u_s(1)<\infty$ otherwise the principal's value function would be arbitrarily negative. The only remaining possibility is $u_s(1)=0$.

From (39), we have that for $u\geq u_s^\alpha$,

$$\lim_{\alpha\to 1}\left(\frac{d}{d\alpha}F_1^\alpha(u)\right) = e^{-\frac{\lambda u}{2\phi}}\lim_{\alpha\to 1}\left(\frac{\lambda u}{4\phi}c^\alpha(u_s(\alpha)) + \frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha))\right).$$

Notice from (40) that $\lim_{\alpha\to 1}(\lim_{u\to 0} c^\alpha(u)) = \lim_{u\to 0}(\lim_{\alpha\to 1} c^\alpha(u)) = -(\Pi - 2c/\lambda)$. Therefore, we can conclude that $\lim_{\alpha\to 1} c^\alpha(u_s(\alpha)) = -(\Pi - 2c/\lambda)$. Hence, to prove (ii), it is sufficient to show that $\lim_{\alpha\to 1}\frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha)) = 0$, for this implies $\frac{d}{d\alpha}F_1^\alpha(u) \approx -e^{-\frac{\lambda u}{2\phi}}(\Pi - 2c/\lambda)\frac{\lambda u}{4\phi} < 0$ for $u \geq u_s(\alpha)$ and α sufficiently close to 1. To see that $\lim_{\alpha\to 1}\frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha)) = 0$, first notice from (41) that

$$\lim_{\alpha\to 1}\frac{e^{-\frac{\lambda u_s(\alpha)}{\phi(1-\alpha)}}}{(1-\alpha)} \in (0,\infty), \tag{42}$$

implying that $u_s(\alpha)$ is $O((1-\alpha)\ln(1-\alpha))$ as $\alpha \to 1$. Differentiating (40) with respect to α and omitting the argument of $u_s(\alpha)$, we get that

$$\begin{aligned}
&\frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha)) = \\
&\quad\frac{e^{\frac{2\alpha\lambda u_s}{(\alpha^2-1)\phi}}}{2(1-\alpha)^2\alpha^2(\alpha+1)\lambda\phi(\alpha\phi+\lambda u_s+\phi)^2} \times \Big[\lambda\Pi\Big((1-\alpha)^2(\alpha+1)^3\phi^3 + 2\lambda^3u_s^3\alpha\left(\alpha^2+1\right) \\
&\quad + \lambda^2u_s^2\phi\left[3\alpha + 2\alpha^4e^{\frac{\lambda u_s}{\phi(1-\alpha)}} + \alpha^3\left(5 - 4e^{\frac{\lambda u_s}{\phi(1-\alpha)}}\right) + \alpha^2\left(2e^{\frac{\lambda u_s}{\phi(1-\alpha)}} - 1\right) + 1\right] + 2(1-\alpha)(\alpha+1)^2\lambda u_s\phi^2\Big) \\
&\quad - (1-\alpha)c\Big[(\alpha-1)^2(\alpha+1)^4\phi^3 + 2\alpha\left(\alpha^2+1\right)\lambda^3u_s^3 - \lambda^2u_s^2\phi\Big(\alpha^4 - 4\alpha + 4\alpha^3\left(e^{\frac{\lambda u_s}{\phi(1-\alpha)}} - 1\right) \\
&\quad - 4\alpha^2e^{\frac{\lambda u_s}{\phi(1-\alpha)}} - 1\Big) + 2(1-\alpha)(\alpha+1)^3\lambda u_s\phi^2\Big]\Big].
\end{aligned}$$

Using (42), we know that $e^{\frac{2\alpha\lambda u_s(\alpha)}{(\alpha^2-1)\phi}}$ is $O(1-\alpha)$ as $\alpha \to 1$. Therefore, any term inside the outermost brackets that goes to zero faster than $O(1-\alpha)$ will converge to zero when scaled by the fraction outside the brackets. By inspection, the only terms that do not clearly go to zero faster than $O(1-\alpha)$ are

$$\lambda^2u_s(\alpha)^2\phi\left[2\alpha^4e^{\frac{\lambda u_s(\alpha)}{\phi(1-\alpha)}} - 4\alpha^3e^{\frac{\lambda u_s(\alpha)}{\phi(1-\alpha)}} + 2\alpha^2e^{\frac{\lambda u_s(\alpha)}{\phi(1-\alpha)}}\right].$$

Thus, we get have that

$$\begin{aligned}
\lim_{\alpha\to 1}\left(\frac{\partial}{\partial\alpha}c^\alpha(u_s(\alpha))\right) &= \lim_{\alpha\to 1}\left(\frac{e^{\frac{2\alpha\lambda u_s(\alpha)}{(\alpha^2-1)\phi}}e^{\frac{\lambda u_s(\alpha)}{\phi(1-\alpha)}}}{(1-\alpha)^2(\alpha+1)(\alpha\phi+\lambda u_s(\alpha)+\phi)^2}\lambda u_s(\alpha)^2\left[\alpha^2 - 2\alpha + 1\right]\right) \\
&= \lim_{\alpha\to 1}\left(\frac{e^{\frac{\lambda u_s(\alpha)}{(1+\alpha)\phi}}}{(\alpha+1)(\alpha\phi+\lambda u_s(\alpha)+\phi)^2}\lambda u_s(\alpha)^2\right) \\
&= \frac{\lambda}{4\phi^2}\left(\lim_{\alpha\to 1}u_s(\alpha)\right)^2 \\
&= 0,
\end{aligned}$$

which completes the proof of (ii). For (iii), we have

$$\hat{F}_1^\alpha(u) - \bar{V}^{\frac{\lambda}{2}}(u) = \left(\frac{1-\alpha}{2\alpha}\right)\left(\Pi - \frac{(1-\alpha)c}{\lambda}\right)e^{-\frac{\lambda u}{(1-\alpha)\phi}} + C_1^\alpha e^{-\frac{\lambda u}{(1+\alpha)\phi}} - \left(\Pi - \frac{2c}{\lambda}\right)e^{-\frac{\lambda u}{2\phi}}$$
$$\leq \left(\frac{1-\alpha}{2\alpha}\right)\left(\Pi - \frac{(1-\alpha)c}{\lambda}\right) + \left|C_1^\alpha - \left(\Pi - \frac{2c}{\lambda}\right)\right| + \left|C_1^\alpha\left(e^{-\frac{\lambda u}{\phi(1+\alpha)}} - e^{-\frac{\lambda u}{2\phi}}\right)\right|$$
$$\leq \left(\frac{1-\alpha}{2\alpha}\right)\Pi + |C_1^\alpha - (\Pi - 2c/\lambda)| + \left|C_1^\alpha\left(e^{-\frac{x^*(\alpha)}{2}} - e^{-\frac{x^*(\alpha)}{(1+\alpha)}}\right)\right|,$$

where the first inequality uses the triangle inequality and $e^{-|x|} \leq 1$ and the second uses the fact that $e^{-\frac{x}{2}} - e^{-\frac{x}{(1+\alpha)}}$ is hump-shaped in x and achieves its maximum at $x^*(\alpha) \equiv 2(1+\alpha)\ln(\frac{1+\alpha}{2})/(\alpha-1)$. Clearly all three terms converge to 0 as $\alpha \to 1$. □

Proof of Proposition 5. Consider first any simple contract with deadline T. If the agent does not shirk,[31] then the probability that the project succeeds at $t \in [0,T]$ is given by $\lambda^2 t e^{-\lambda t}$ and the probability that the project does not succeed prior to T is given by $e^{-\lambda T}(1+\lambda T)$. Therefore, the total surplus is given by

$$S(T) \equiv \int_0^T \lambda^2 t e^{-\lambda t}(\Pi - ct)dt + e^{-\lambda T}(1+\lambda T)(-cT).$$

In order to induce the agent to work, he must be given rents in the amount of at least $u = \phi T$, otherwise he can do better by shirking. Making the change of variables from T to u, we have that the principal's ex-ante expected payoff under a simple contract with deadline $T = u/\phi$ is bounded above by

$$G(u) \equiv S(u/\phi) - u = \left(1 - e^{-\lambda u/\phi}(1+\lambda u/\phi)\right)\Pi - \frac{2c}{\lambda}\left(1 - e^{-\lambda u/\phi}(1+\lambda u/2\phi)\right) - u. \tag{43}$$

Note that $G(u)$ is not the principal's value function under the optimal contract, since her belief about the project stage changes over time. We will construct the value function shortly. To prove the proposition, it suffices to show that

(i) The principal's ex-ante payoff for a project with unobservable progress under *any contract* is bounded above by $\max_u G(u)$.

(ii) There exists an incentive-compatible simple contract under which the principal's ex-ante expected payoff is $\max_u G(u)$.

(iii) For all $u > 0$, $G(u) < F_1(u)$. Therefore, the principal does strictly better with intangible progress than she does with unobservable progress.

For (i), let $w^* = \arg\max_u G(u)$, which is generically unique. We have already argued that the principal's maximal payoff under a simple contract is bounded above by $G(w^*)$. Given that neither player has any information about the status of the project, the only possibility is that the principal randomizes over the termination date. It therefore suffices to show that the principal cannot benefit from such randomization, or equivalently, that the principal's value function (under this simple contract with the optimally chosen deadline) is globally concave in the agent's continuation value.

[31] Arguments similar to those made for a single-stage project can be used to confirm shirking is suboptimal.

Suppose that the principal can implement a simple contract in which the incentive compatibility condition holds with equality for all t (this is the best possible case for the principal). It is most intuitive to construct this value function from the pair of value functions that are conditional on s (i.e., whether a breakthrough has been made) and weight them appropriately by the probability that the principal assigns to each. Given u, the principal's payoff conditional on being in the first stage ($s = 1$) is $G(u)$; i.e.,

$$\begin{aligned} F^{unobs}(u|s=1) &= \int_0^{u/\phi} \lambda^2 t e^{-\lambda t}(\Pi - ct)dt - e^{-\lambda u/\phi}(1 + \lambda u/\phi)cu/\phi - u \\ &= (1 - e^{-\lambda u/\phi})(\Pi - 2c/\lambda) - \frac{\lambda u}{\phi} e^{-\lambda u/\phi}(\Pi - c/\lambda) - u. \end{aligned}$$

Conditional on being in the second stage, the principal's value function is the benchmark payoff $\bar{V}(u)$:

$$F^{unobs}(u|s=2) = \int_0^{u/\phi} \lambda e^{-\lambda t}(\Pi - ct) + e^{-\lambda u/\phi}(-cu/\phi) - u = (1 - e^{-\lambda u/\phi})(\Pi - c/\lambda) - u.$$

Over time, the principal's beliefs will evolve about the state of the project. Conditional on reaching state $u < w^*$ prior to project success, a period of time of length $t(u; w^*) = \frac{w^* - u}{\phi}$ has elapsed. Therefore, the principal's beliefs are given by

$$\mu(u; w^*) = \Pr(\tau_1 \le t(u; w^*)|\tau_2 > t(u; w^*)) = \frac{\lambda\left(\frac{w^*-u}{\phi}\right)}{1 + \lambda\left(\frac{w^*-u}{\phi}\right)}.$$

The principal's value function for $u \le w^*$ is therefore given by

$$F^{unobs}(u; w^*) = \mu(u; w^*)F^{unobs}(u|s=2) + (1 - \mu(u; w^*))F^{unobs}(u|s=1).$$

We will now verify this value function is concave for all $u \le w^*$. Using the functional forms given above and twice differentiating $F^{unobs}(u; w^*)$, we get that

$$\begin{aligned} \frac{d^2}{du^2} F^{unobs}(u; w^*) = \frac{-\lambda e^{-\lambda u/\phi}}{\phi^2 \left(\lambda(w^* - u) + \phi\right)^3} \Big[&(\lambda^3 w^*(w^* - u)^2 + \lambda w^* \phi^2)(\lambda\Pi - c) \\ &+ \lambda^2 \phi (w^* - u)^2 (\lambda \Pi - 2c) + \phi^3 \left(\lambda\Pi + 2(e^{\lambda u} - 1)\right) \Big]. \end{aligned}$$

All three terms inside the brackets are clearly positive, implying the value function is concave in u for all $u \le w^*$, which completes the proof of (i).

We prove (ii) by showing that for any T, there exists a $w : [0, T] \to \mathbb{R}_+$ such that (1) it is incentive compatible for the agent to work for all $t \in [0, T]$, and (2) the agent's continuation utility at date t is $u(t) = \phi(T - t)$. Let $u_2(t)$ be the promised continuation value conditional on being in the second stage and $u(t)$ be the unconditional continuation value at t. Promise keeping requires that

$$\lambda u(t) = \lambda\big(\mu(t)w(t) + (1 - \mu(t))u_2(t)\big) + u'(t).$$

Conditional on progress, the evolution of u_2 is given by

$$\lambda u_2(t) = \lambda w(t) + u_2'(t). \tag{44}$$

We want to find $w(t)$ such that $u(t) = \phi(T - t)$ for all $t \in [0, T]$. Note that this implies that $u'(t) = -\phi$. Using the promise keeping condition,

$$\phi(T - t + 1/\lambda) = (1 - \mu(t))u_2(t) + \mu(t)w(t).$$

Substituting for $w(t)$ from (44), we get that

$$\begin{aligned}\phi\left(T - t + \frac{1}{\lambda}\right) &= (1 - \mu(t))u_2(t) + \mu(t)\left(u_2(t) - \frac{u_2'(t)}{\lambda}\right) \\ &= u_2(t) - \frac{t}{1 + \lambda t}u_2'(t).\end{aligned}$$

Imposing the boundary condition $u_2(T) = 0$, we arrive at a unique solution for $u_2(t)$, which we can then substitute back into (44), to arrive at

$$w(t) = \phi\left(T - t + \frac{1}{\lambda} + \frac{e^{-\lambda(T-t)}}{\lambda^2 T} + \frac{e^{\lambda t}}{\lambda}\left(q(-\lambda T) - q(-\lambda t)\right)\right),$$

where $q(z) = -\int_{-z}^{\infty} e^{-x}/x dx$. It is straightfoward to check that $w(t) > 0$ for all $t \in [0, T]$, which completes the proof of (ii). For (iii), note that $G(u)$ has the same form as $F_1(u)$ (see (17)) with $H_1 = \frac{2c}{\lambda} - \Pi$. The result then follows from the fact that F_1 has a constant strictly larger than $\frac{2c}{\lambda} - \Pi$. □