Informative Milestones in Experimentation*

Christoph Wolf †

April 29, 2018

Abstract

I study a continuous-time moral hazard problem with learning about a two-stage project of unknown quality. The first-stage arrival time is informative but not conclusive about the project's quality. Due to the informativeness, the optimal contract features a combination of continuation value and intermediate bonus payments as a reward. There is a negative correlation between the first success time and the share of bonus payments in the reward. Second-stage deadlines adjust to the first-stage success time: early successes are rewarded with longer deadlines in the second stage. When agent replacement between stages is possible, the principal will replace the agent if the first success arrives late.

JEL Classifications: D82, D83, D86

Keywords: Dynamic Contracts, Two-Stage Experimentation, Milestones

* I am grateful to Volker Nocke and Konrad Stahl for continuous support and guidance. I am particularly thankful to Renato Gomes and Thomas Tröger. I would like to thank Dirk Bergemann, Florian Ederer, Dietmar Harhoff, Bard Harstad, Ulrich Hege, Johannes Hörner, Raphaël Levy, Marco Ottaviani, Larry Samuelson, Johannes Schneider, André Stenzel, Emanuele Tarantino as well as the audiences at Bergen, Berlin, Bocconi, Bologna, Bonn, Helsinki, Mannheim, Oslo, Toulouse, UCL, Warwick, Yale, EEA/ESEM 2016, EARIE Annual Meeting 2016, GAMES 2016, Econometric Society Winter Meeting 2016, SAEe 2016 and many participants of ESSET 2016 for helpful comments and discussions.

† Bocconi University, Department of Economics, and IGIER; E-mail: christoph.wolf@unibocconi.it.

1 Introduction

Intermediate milestones are frequently observed in principal-agent relationships when the feasibility of the project is unknown to both parties. A fundamental reason for this is that the performance in early stages conveys valuable information about subsequent stages to the principal. Given this informational spillover and the agent's ability to affect the observable information, several questions arise: How are the agent's incentives to exert effort affected by the informativeness of the milestone? How does the optimal contract adapt to these incentives? How does the role of bonus payments, deadlines and continuation contracts change with an *informative milestone*?

I show that the informational spillover across stages affects the optimal contract substantially because it introduces an endogenous ratchet effect: By privately shirking and thereby strategically delaying a breakthrough in the first stage, the agent increases the principal's pessimism in the second stage. Conditional on the agent's private, more optimistic, belief, this yields higher rents for the agent through a higher payment tailored to a more pessimistic belief in the second-stage contract. To prevent such a delay in effort, the optimal contract rewards early successes with higher rewards. These additional first-stage rents increase in the level of the continuation value from second-stage experimentation because a higher continuation value amplifies the ratchet effect. Therefore, in contrast to a setting with independent stages, rewarding the agent with a more valuable continuation contract is costly. To reduce the resulting information rents, the optimal contract features bonus payments for first-stage successes. Hence, due to the informativeness of the milestone, full backloading of payments is suboptimal.[1] If the principal has access to a new agent for the second stage, the informative milestone gives rise to replacement of the agent if he succeeds too late in the first stage. Replacing the agent after the first success reduces the incentive to delay effort provision and saves information rents.

As an application of the setup, consider the venture capital industry.[2] Innovative projects involve uncertainty about their quality. They require effort and knowledge of experts as well as substantial amounts of capital. As entrepreneurs rarely have the necessary funds themselves, they contract with financial investors. One of the main funding sources for high-risk startups is venture capital.[3] The considerable risk about the project's future profitability together with the substantial size of the investment may make financiers reluctant to invest. To overcome this problem, intermediate stages, so-called *milestones*, are introduced to gather information about the project's quality at a reduced cost with the option of terminating the project. A typical example of such a milestone is the development of a prototype.

Staging of venture capital contracts is a well-documented feature: Kaplan and Stroemberg

[1] Other papers on staged dynamic moral hazard problems find that all payments should be backloaded when there is only one agent, e.g. Green and Taylor (2016), Moroni (2017).

[2] Other settings also fit the setup: e.g., labor contracts with unknown worker type and probationary periods, within-firm contracts between management and a research division developing several patents.

[3] Gornall and Strebulaev (2015) show that 42% of independent U.S. public firms founded after 1974 are venture-capital backed. Notably, 85% of the R&D expenditures of those independent U.S. public firms founded after 1974 stem from venture-capital backed firms.

(2004) show that 72.8% of contracts in their sample involve staging. It has been argued that staging helps in mitigating agency costs (see, for example, Gompers (1995), Neher (1999) or Cumming (2012)). However, the literature is surprisingly silent about learning and the informational value of milestones.[4] I show that introducing an informational content of early stages has important consequences for the entrepreneur's incentives as well as for the design of the optimal contract. Indeed, Moorman, Wies, Mizik, and Spencer (2012) document that innovative firms understand and exploit that they can affect beliefs by delaying the revelation of innovations and make use of a ratchet-effect motif.

I develop and study a continuous-time principal-agent bandit experimentation model with an informative milestone. A project of uncertain quality has to complete each of two sequential stages to realize its benefits. Any success is immediately and publicly observed. The intensity rate of obtaining a breakthrough depends on the project's quality and on the agent's effort. For a given level of effort, in the first stage stage is higher for a good than for a bad project, while only the good project can succeed in the second stage. Hence, the total effort exerted until the first success is informative but not conclusive about the project's quality.[5] As effort is costly and unobservable to the principal, this is a dynamic moral hazard problem with private learning. I solve for the principal's full-commitment profit-maximizing contract that conditions on the publicly observable success times. I allow for arbitrary payment rules subject to limited liability.

The optimal contract features a deadline for each of the stages because, if the principal becomes too pessimistic, she will terminate the project. Moreover, it is without loss of generality to focus on bonus contracts that have payments to the agent only at success times. Hence, there are at most two payments to the agent. First-stage bonus payments represent short-term incentives that do not condition on the long-run success of the project, and their value is independent of the current belief about the project's quality. However, the second-stage bonus payment and deadline together induce an expected value for the agent that depends on the belief about the project's quality at the beginning of the second stage, which is determined by the first-stage performance. This continuation value can be interpreted as the value of equity given to the agent after the first success.

The informational spillover across the stages implies that effort choices in the first stage affect not only the belief about the project in the first stage but also the initial belief of the second stage. If more effort is exerted before the first success is obtained, players are more pessimistic in the second stage. Unobservable deviations from the expected effort path persistently divert the agent's private belief from the principal's belief. Therefore, the principal holds a different belief *after* the first success than the agent following a deviation. This is the key novelty of this paper and the underlying reason for the main results. The effect of a deviation from the expected effort path on the belief in the current stage gives rise to *procrastination rents*, while

[4]Moroni (2017) studies a related two-stage experimentation model. However, in her setting the first stage does not carry any marginal information about the second stage. Therefore, the belief in the second stage is unaffected by actions in the first stage.

[5]The stochastic process of the breakthrough in the first stage is the same as, for example, in Keller and Rady (2010), while the stochastic process for a breakthrough in the second stage is the same as, for example, in Keller, Rady, and Cripps (2005).

the effect of a deviation on the belief in the following stage gives rise to a novel rent that I call *informativeness rents*.

First, consider the interaction of moral hazard and private learning in the current stage. The agent's incentives are driven by his private belief about the success probability: the reward has to be chosen such that the agent is at least compensated with an *expected* utility that outweighs the cost of effort. However, the agent has the ability to privately shirk and divert the beliefs in a dynamic setting because the principal cannot distinguish whether the absence of a success was due to bad luck or due to a deviation. Therefore, the principal becomes overly pessimistic after a deviation and the reward may be misspecified: if the principal is overly pessimistic, she believes that she has to pay a higher reward to the agent. If the contract does not account for private learning about the current stage, the agent has an incentive to delay effort. This effect is present in both stages, and the agent has to be granted *procrastination rents* to prevent belief manipulation about the current stage.

Second, consider the interaction of moral hazard and private learning about the second stage from first-stage experimentation. By privately shirking in the first stage, the agent induces the principal to be overly pessimistic in the second stage. A low second-stage belief of the principal implies that the continuation contract has to promise the agent a high bonus for a second-stage success: the less likely the principal thinks it is that the agent obtains a success, the higher the bonus payment must be to incentivize the agent to exert effort. Hence, conditional on reaching the second stage and his private belief, the agent wants the principal to be pessimistic about the second stage to enjoy higher payments upon second-stage success.

The intuition for the agent's incentive to manipulate the principal's belief can be related to the *ratchet effect*:[6] the agent wants the principal to be sufficiently optimistic to continue the project; however, conditional on continuation, he wants the principal to be pessimistic to be granted a high bonus after the second success.

This effect is present neither with independent stages nor in a one-stage setting. With independent stages, the belief at the beginning of the second stage is exogenously given and the agent cannot affect this by off-path effort choices in the first stage. In a one-stage setting, the interaction ends after the first success. To prevent the delay in effort, the principal has to reward early successes with higher rents than later successes. I call these rents *informativeness rents* as they only arise due to the informativeness of the milestone. The rate at which the total reward for the first success decreases is exactly the rate at which the agent gains from private information at the beginning of the second stage; that is, the value of holding a marginally more optimistic belief than the principal in the second stage. While the procrastination rents prevent deviations that directly affect the *level* of the reward for a success in the current stage, the informativeness rents prevent deviations that alter the *assessment* of the second-stage contract through the persistence of the agent's private information into the second stage.

Procrastination rents are unaffected by the way in which the reward is delivered. However, the informativeness rent has to be provided because the agent can gain from the persistence in his private information in the continuation contract. I show that the informativeness rent

[6] See, for example, Laffont and Tirole (1988) or, more recently, Bhaskar (2014).

increases in the on-path value of the continuation contract. The higher the value promised from the second stage, the higher is the value of being more optimistic in the second stage. Therefore, the principal faces a tradeoff when choosing how to deliver the first-stage reward. The principal would like to incentivize the agent to work until an extended deadline in the second stage: due to the procrastination rents, experimentation stops inefficiently early in the second stage. By extending the deadline, additional overall surplus is generated, which makes continuation contracts an attractive reward mechanism. However, the downside of delivering utility through more valuable continuation contracts is that these increase the informativeness rents in the first stage. In particular, at all times prior to a particular success time, the agent's incentive to delay effort increases if additional utility is delivered through a continuation contract at that success time: it increases the value of the private information which can be obtained by prior deviations. Hence, the cost of using the long-term reward continuation contract increases in the success time.

In the optimal contract, early successes are rewarded with continuation contracts only, implemented through long second-stage deadlines, because the gain of extending the deadline is large while the cost is very low. As time elapses without a breakthrough, the composition changes such that the reward consists of an increasing share of bonus payments and a decreasing share of continuation value to reduce the informativeness rents for all earlier successes. Hence, compared to other settings with multiple stages and dynamic moral hazard (e.g., Green and Taylor (2016) and Moroni (2017)) full backloading of payments is not optimal if the first stage has a marginal informational value.

In an extension, I consider the possibility of replacing the agent after the first stage. In contrast to other models of experimentation or staged financing, my model can rationalize managerial turnover in young startup firms.[7] Hannan, Burton, and Baron (1996) show that 40% of CEOs are replaced within the first 40 months of a startup. In my model, the presence of the informativeness rent gives rise to turnover which does not occur with independent stages: The principal always wants to introduce two deadlines in the first stage: (i) if the agent succeeds before the first deadline, he is rewarded according to the previously derived contract, (ii) if the agent succeeds after the first and before the second deadline, he receives a payment and is replaced by a new agent in the second stage, (iii) if the agent has not obtained a success before the second deadline, the project is terminated. To see why replacement is optimal, recall that the agent receives the informativeness rent only because his private information is valuable in the second stage. Hence, if the agent is replaced when he succeeds after the first deadline he can be incentivized at a lower cost in the first stage. This implies that the agent receives lower rents in the continuation region because delaying effort becomes less attractive as the replacement deadline approaches. Thus, the principal faces a tradeoff between the cost of more expensively rewarding agents in the replacement region with a bonus payment instead of a continuation contract and the benefit of reduced informativeness rents. I show that the principal always prefers to have both, a replacement and a continuation region, in the optimal contract if there is

[7]Garrett and Pavan (2012) provide a dynamic model of managerial turnover. However, they assume ex ante asymmetric information and that the productivity of the agent is changing.

an agency conflict in the first stage.

I extend the analysis to allow for the principal's choice of informativeness and endogenous staging. Assuming that the principal can choose the intensity rate of the first stage and that the second stage vanishes as the first stage becomes fully informative, I show that the optimal two-stage contract converges to the optimal one-stage contract. Intuitively, a fully informative prototype coincides with the final product and therefore there is no need for a second stage anymore. The principal faces a non-trivial tradeoff in the staging decision: when introducing an informative milestone, the principal gains from the additional information provided in the first stage and can condition the second-stage funding and contract on the first stage outcome. However, introducing an informative milestone generates the informativeness rents. Numerically, I show that the staging decision depends on parameter values, and it cannot be argued that one mode dominates the other generically, but I find that staging is more likely to take place if the initial success probability is low; that is, when the value of additional information is relatively high. This implies that introducing an informative milestone can facilitate funding for projects with low initial success probabilities that would not receive funding as a one-stage project.

The insights I derive can be applied in several other contexts. For example, they could describe the interaction of a CEO with the leader of a research department about work on a risky and expensive project. Alternatively, there could be uncertainty about the worker's type instead of the project's quality. In this case, an employer could offer a contract with a probationary period in which a first signal about the employee's competence can be obtained. The result on the composition of an agent's reward conditional on performance provides a perspective on regulating CEO compensation.

Empirical Implications. My analysis generates several empirical predictions that can be of interest in the different applications mentioned above: (i) The total worth of the reward is higher for the well-performing CEO. (ii) The composition of the agent's compensation changes with performance. In particular, if performance gets worse, the total reward is lower and consists of relatively more short-term than long-term rewards. For example, a well-performing CEO is rewarded with stock options that are tied to future performance. A CEO that performs worse is rewarded with bonus payments and less with stock options. (iii) Deadlines are relatively more responsive to early performance while final-stage bonus payments are less responsive to early performance compared to a setting without informativeness rents. (iv) Early-stage deadlines are relatively shorter if there is a learning spillover to future stages. (v) Successful agents may be replaced if they do not perform sufficiently well although the project is continued and the agent is known to be able to complete future tasks. (vi) Staging occurs more frequently if the initial risk is high.

In addition to the finding in Moorman, Wies, Mizik, and Spencer (2012) that innovative firms understand that the timing of their revealing their innovations can be used to affect beliefs in a ratchet-effect motif, some other of the empirical implications can be found in empirical papers. Lerner and Wulf (2007) find that R&D heads that develop well-performing patents are rewarded more and relatively more with long-term rewards than with cash payments supporting predictions (i) and (ii). Short early-stage deadlines may explain the observation of high failure rates of startups. Gosh is cited in The Wall Street Journal (2012) that only 35% of startups survive to the age of 10 years. High failure rates are not necessarily due to high risk only. If the investors use deadlines and agency conflicts induce deadlines to be inefficiently short, too many startups fail and too few innovations become available to society. Concerning the replacement of successful agents, Wasserman (2008) notes in a Harvard Business Review article that: *"[o]thers invest in a start-up only when they're confident the founder has the skills to lead it in the long term. Even these firms, though, have to replace as many as a quarter of the founder-CEOs in the companies they fund."* Hence, replacement also occurs even though there is no doubt about the agent's qualification to succeed with the project. Although empirical implication (vi) about staging and initial risk is only a numerical outcome, it is fairly intuitive and empirical evidence has been found in Bienz and Hirsch (2011).

Related Literature. My paper contributes to the growing literature on principal-agent models with ex ante symmetric uncertainty about a project's feasibility. Most of the early work focuses on the case where one success suffices to complete the project and in the absence of a success players become pessimistic about the project's quality; see for example, Bergemann and Hege (1998; 2006), Halac, Kartik, and Liu (2016), Hörner and Samuelson (2013). These models apply the exponential bandit model by Keller, Rady, and Cripps (2005), in which one success is fully informative about the project's quality. By contrast, I assume that there may be a first stage that is informative but not conclusive about the quality of the project and a second breakthrough is required to complete the project. First, this modelling approach allows an assessment of the impact of staging that is widely used in contracting relationships when there is uncertainty about the project's value. I show that introducing an additional and informative stage can facilitate funding of projects that would not be undertaken if they were forced to involve only a single stage. Seond, it allows for more flexibility in the learning process compared to the one-stage experimentation literature; during the course of the project, players may become more optimistic instead of increasingly pessimistic.

To the best of my knowledge, three other papers consider staged projects, which are closely related. First, Moroni (2017) considers a principal contracting with several agents on a project that requires multiple breakthroughs to yield the final payoff. She shows that agents have an additional free-riding incentive because another agent may start a subsequent stage. Because she assumes that early stages carry no information about later stages, staging has no informational value, and therefore there is no informativeness rent present. If there is only one agent, her analysis may serve as a benchmark to the present paper without learning across stages and with a fixed second-stage belief. In that case, the agent would be incentivized with a constant continuation value in the first and a constant bonus payment in the second stage. Hence, the continuation contract would be independent of the first-stage performance and there would be no replacement of the agent.

Second, Green and Taylor (2016) and Hu (2014) study dynamic moral hazard problems in which the agent also has to obtain two success. However, the quality of the project is known to be good but the agent has the ability to divert the principal's flow funding. In their case, deadlines arise to prevent the agent from diverting cash. Early successes also have to be rewarded with higher continuation values. However, the reason is fundamentally different: In both papers the agent has a direct benefit from delaying effort, which is the flow benefit from diverting the cash. In my paper, delaying effort creates an informational advantage because it persistently drives a wedge between the principal's and the agent's belief about the project's value. If the project was known to be good, the principal could achieve the first best in my model. While Green and Taylor (2016) focus on the role of communication and private observability of progress after a first success, I study the consequences of learning from a first success.

This paper also relates to the ongoing discussion on staged contracts. Examples of this literature are Bolton and Scharfstein (1990), Neher (1999), Cuny and Talmor (2005) and Booth, Dalgic, and Young (2004). These papers discuss the value of staging contracts to mitigate agency conflicts through the threat of termination and to reduce the hold-up problem. However, these papers do not consider the possibly uncertain feasibility of the project about which the agent can privately learn by exercising effort. While this may be realistic in some cases, learning plays an important role in the financing of innovation. Therefore, I take a different perspective and study the informational value of staging if the project's feasibility is unknown. Pindyck (1993) also discusses the informational value of early investments that can reduce uncertainty over later costs. I show that if learning is private, then there is a tradeoff to introducing informative milestones. On the one hand, an informative milestone can be beneficial because a signal can be generated at a lower cost. On the other hand, the agent's private learning can give rise to an additional agency rent due to the possibility to manipulate the principal's belief. To my knowledge, this is the first paper to address this potential drawback of informative milestones.

Bhaskar (2014) considers a related two-period model with learning about a project's difficulty but without commitment. In the first period, a signal is generated that depends on the agent's effort and the project's type. Similar to the present paper, he shows that the agent has an incentive to manipulate the principal's belief such that he obtains higher payoffs in the second period. Different to his paper, I study the interplay between learning and dynamic moral hazard. Thereby, I can shed light on the use of different reward instruments, deadlines and bonus payments, to incentivize the agent.

On a more abstract level, this paper is related to DeMarzo and Sannikov (2016), who also study a model where agent's private deviations affect the assessment of the promised continuation value. In their model, the principal also has to pay an additional information rent to prevent the agent's deviation to get an informational advantage over the principal. The model differs in the underlying learning process: DeMarzo and Sannikov (2016) study a Brownian model in which informative outputs are produced continuously, while I assume that news arrive at exponentially distributed times. While their model is suitable for analyzing an unknown profitability of a continuously producing firm in which news arrive continuously, my model highlights the aspects of an innovative project in which drastic news arrive at random times. This focus allows me to study staged contracts and the role of deadlines and bonus payments. Prat and Jovanovic (2014) consider a model similar to DeMarzo and Sannikov (2016) with a risk averse agent and a constant quality. As information arrives continuously in their model, early deviations have a stronger impact on the belief diversion than later deviations.

The paper is organized as follows. In Section 2, I present the general two-stage model. In Section 3, I describe the first-best benchmark. The optimal contract is studied in Section 4. Agent replacement is analyzed in Section 5. Extensions are discussed in Section 6. I conclude in Section 7. All proofs are relegated to the Appendix.

2 Model

There is an agent (entrepreneur, he) with access to a project of unknown quality that has to complete two sequential stages, $i \in \{1, 2\}$. The agent has no wealth and contracts with a principal (e.g. a venture capitalist, she) to receive the necessary funds, f_i, that are required to work on stage i. After the completion of the final stage the project immediately generates a value π to the principal. The project can be either good or bad, $\omega \in \{g, b\}$. Only a good project can complete both stages; however, a bad project may complete the first stage. The agent has to undergo experimentation to learn about the quality of the project and to advance it towards completion. Experimentation is modeled as a two-armed bandit in continuous time, $t \in [0, \infty)$. The agent chooses in time interval $[t, t + dt)$ how much effort to exert, i.e., chooses $a_{i,t} \in [0, 1]$, which comes at cost $a_{i,t}c$.

A project of quality ω generates a success in stage i with probability $\lambda_i^\omega a_{i,t}dt$ if effort $a_{i,t}$ has been exerted in time interval $[t, t + dt)$. I assume that the intensity rate of a good project is higher than the intensity of a bad project, $\lambda_i^g > \lambda_i^b$. Moreover, only a good project can succeed in the second stage, i.e., $\lambda_2^g > \lambda_2^b = 0$.[8] The principal and the agent hold a common initial belief $p_0 \in (0, 1)$ that the project is of good quality. Let $p_{i,t}(\{a_{i,s}\}_{0\leq s<t})$ denote the belief that the project is of good quality at time t in stage i given effort path $\{a_{i,s}\}_{0\leq s<t}$.

Breakthroughs are immediately publicly observed. The public history at time t, $h^t \in \mathcal{H}^t$, consists of the success times, i.e., $\{\tau_i\}_{i\in\{1,2\}}$ with $\tau_i \in \{\emptyset \cup \mathbb{R}_+\}$ where the emptyset refers to the case that no success in stage i has been obtained yet. Note that if I did assume that the success is verifiable but not publicly observable, the most profitable deviation of an agent could be to exert effort, but hide a potential success. However, I show in an extension that under the optimal contract with public observability of successes, the agent would have no incentive to hide a success even if he could do so. Denote by $h_\alpha^t \in \mathcal{H}_\alpha^t$ the private history of the agent at time t that consists of the public history as well as the agent's effort choices in each of the stages, $\{a_{i,t}\}_{0\leq s<t}$. The history of past effort choices matters only through its aggregation in each stage $A_{i,t} = \int_0^t a_{i,t}dt$ as this determines the agent's belief. Hence, I can restrict attention

[8]The results would not change qualitatively if a bad project could also succeed in the second stage when a bad project's success probability in the second stage is sufficiently low that it would not be funded if it was known to be of bad quality.

to private histories of the form $\mathcal{H}_\alpha^t \in \{\mathcal{H}^t \times \{\emptyset \cup [0,t]\} \times \{\emptyset \cup [0,t]\}\}$. The agent's strategy is therefore a measurable map from calendar time and his private history into the unit interval, $a_{i,t}(h_\alpha^t) : \mathcal{R}_+ \times \mathcal{H}_\alpha^t \to [0,1]$. To simplify notation, I drop the explicit dependence on the history and keep only the time index t.

The principal offers the agent a profit-maximizing payment process conditioning on the public history to which she is fully committed from time zero on. I restrict attention to deterministic contracts. A payment process consists of a flow payment, $w_f(h^t)$, and a lump-sum payment, $w_l(h^t)$ at every history $h^t \in \mathcal{H}^t$. In the Appendix, I show that it is without loss of generality to restrict attention to bonus contracts; that is, to contracts that have payments only at time zero, and the success times. Denote by $\tilde{\mathcal{H}}^t$ the subset of public histories at time t with a breakthrough at time t. Hence, a bonus contract maps for every history $h^t \in \tilde{\mathcal{H}}^t$ a bonus payment $b(h^t) \in \mathbb{R}$ to the agent and chooses a payment b_0 at time zero. To simplify notation, I drop the dependence on the history of the bonus payment and denote a bonus payment in stage i given the history h^t as $b_{i,t}$. No payments take place at histories $h^t \notin \{\tilde{\mathcal{H}}^t \cup \mathcal{H}^0\}$. I assume that the agent is subject to limited liability. Hence, at every history h^t the agent's bonus payment is nonnegative. Note that without limited liability the principal could obtain the first-best by having the agent make a payment equal to the expected value of the project at time zero.

For expositional purposes and to help building intuition, I consider in the main text that discounting is in the limit $r = 0$. All results remain qualitatively unchanged with a common and positive but sufficiently small discount rate $r > 0$. All proofs are carried out with $r > 0$ in the Appendix.[9]

Given a terminal history h with success times $\{\tau_i\}_{i\in\{1,2\}}$ and bonus payments $b_{i,t}$, the principal's payoff is given by

$$e^{-r\tau_2}\left(\pi - b_{2,\tau_2}\right) - e^{-r\tau_1} b_{1,\tau_1}.$$

Similarly for the agent

$$e^{-r\tau_1} b_{1,\tau_1} - \int_0^{\tau_1} e^{-rt} c a_{1,t} dt + e^{-r\tau_2} b_{2,\tau_2} - e^{-r\tau_1} \int_0^{\tau_2} e^{-rt} c a_{2,t} dt.$$

The agent's outside option is normalized to zero.

[9] To make the analysis of the informative milestone interesting, I assume that discounting is sufficiently small. Strategic incentives in the present setting are driven by the possibility to delay effort. However, if the agent discounts the future more (if r becomes large), the agent becomes less strategic. In the limit case of a myopic agent, the efficient outcome is obtained. Assuming $r = 0$ throughout introduces a technical difficulty in the proof that full effort will be implemente by the principal. This can be circumvented by assuming a strictly positive discount rate.

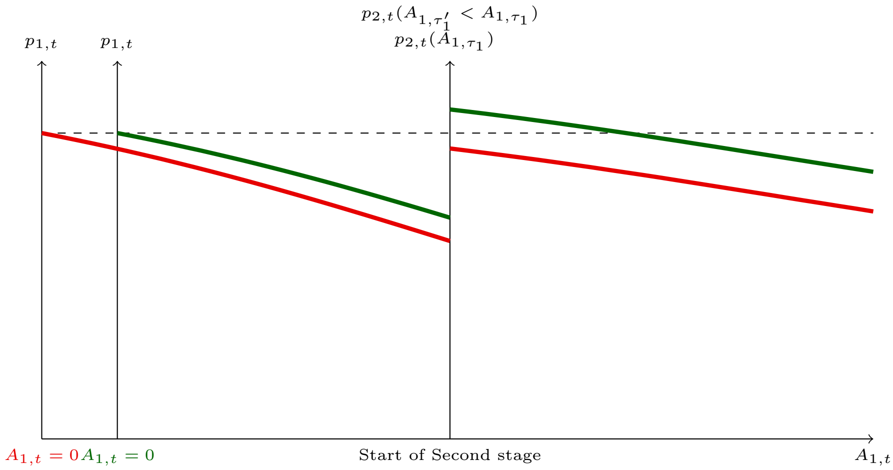

Figure 1: Belief path depending on total amount of effort up to success.
The red line plots the belief path if more effort is required to obtain the first breakthrough, while the green line plots belief path obtains the first breakthrough if less effort is required.

2.1 Learning

A breakthrough is informative but not conclusive. After a success at time τ in stage 1 the belief jumps according to Bayes' rule to[10]

$$p_{2,0}(A_{1,\tau}) = \frac{\lambda_i^g p_{1,\tau}(A_{1,\tau})}{\lambda_i^g p_{1,\tau}(A_{1,\tau}) + \lambda_i^b(1 - p_{1,\tau}(A_{1,\tau}))}.$$

Hence, the less effort has been exerted until a success is achieved, the higher is the upward jump of the belief after a success. When effort is exerted but no breakthrough is observed the agent becomes more pessimistic about project quality as $\lambda_i^g > \lambda_i^b$. The belief follows the differential equation[11]

$$dp_{i,t} = -p_{i,t}(1 - p_{i,t})\Delta\lambda_i a_t$$

where $\Delta\lambda_i \equiv \lambda_i^g - \lambda_i^b$ and initial condition $p_{1,0} = p_0$ and $p_{2,0} = p_{2,0}(A_{1,\tau})$ as defined above. Hence, the belief drifts downwards if the agent exerts effort. To simplify notation, denote $\lambda_1^g = \lambda^g$, $\lambda_1^b = \lambda^b$ and $\lambda_2^g = \lambda$ while $\lambda_2^b = 0$ by assumption. $\lambda^g > \lambda^b$ implies that the absence of a breakthrough makes players more pessimistic about the state of the project.

Note that beliefs do depend on the total effort that has been exerted in each of the stages, but not on how it was distributed over time. Hence, the higher is total effort until t, the lower is the belief about the project quality at t. This is illustrated in Figure 1. Learning is private

[10]Note that to be precise, $p_{1,\tau}$ in this equation is $p_{1,\tau-}$, i.e., the left-limit of the belief held at τ. For almost all t, that is whenever no success occurs, $p_{1,t-} = p_{1,t}$. This is to say that the action at t cannot condition on the arrival of a success at t.

[11]This follows from calculating the belief at $t + dt$ via Bayes' rule and taking the limit $dt \to 0$.

because the agent's effort choices are unobserved by the principal. However, the principal holds a belief about the agent's effort. If the agent's choices coincide with the principal's belief about these, their beliefs about the project's quality coincide. Otherwise, if the agent has exerted less (more) effort than expected by the principal, the principal is more pessimistic (optimistic) than the agent.

3 First-Best Benchmark

As a benchmark consider a social planner that maximizes the sum of payoffs. This optimization is solved by backward induction through the stages. Hence, consider the second stage and assume that the first stage was completed at τ_1. The initial belief at the beginning of the second stage is therefore $p_{2,0}(A_{1,\tau_1})$ with $A_{1,\tau_1} = \int_0^{\tau_1} a_{1,t}dt$. Note first that the probability of reaching time t in the second stage is given by[12]

$$e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_{2,s}ds}$$

and the instantaneous success probability by

$$p_{2,t}(A_{1,\tau_1})\lambda a_{2,t}dt$$

which implies that the probability of a success in $[t, t+dt)$ is given by

$$e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_{2,s}ds}p_{2,t}(A_{1,\tau_1})\lambda a_{2,t}dt.$$

Therefore, in the second stage the social planner chooses $\{a_t\}_{t\geq 0}$ to maximize

$$\int_0^\infty e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_{2,s}ds}a_{2,t}\left(p_{2,t}(A_{1,\tau_1})\lambda\pi - c\right)dt.$$

This gives as optimal choice $a_{2,t} = 1$ for all t such that $p_{2,t}(A_{1,\tau_1})\lambda a_{2,t}\pi \geq c$ and $a_{2,t} = 0$ otherwise. The optimal experimentation duration is given by

$$p_{2,T_2^{FB}(A_{1,\tau_1})}(A_{1,\tau_1})\lambda a_{2,T_2^{FB}}\pi = c$$
$$T_2^{FB}(A_{1,\tau_1}) = \frac{1}{\lambda}\ln\left(\frac{p_{2,0}(A_{1,\tau_1})}{1-p_{2,0}(A_{1,\tau_1})}\frac{\pi\lambda - c}{c}\right).$$

This optimal deadline generates value

$$\Pi_2(A_{1,\tau_1}) = \int_{\tau_1}^{\tau_1+T_2^{FB}(A_{1,\tau_1})} e^{-\int_{\tau_1}^t p_{2,s}(A_{1,\tau_1})\lambda ds}\left(p_{2,t}(A_{1,\tau_1})\lambda\pi - c\right)dt - f_2$$

[12]To ease notation, assume that the clock is restarted when the second stage is reached.

and therefore, in the first stage the principal solves

$$\Pi = \max_{a_{1,t}} \int_0^\infty e^{-(p_{1,s}\lambda^g+(1-p_{1,s})\lambda_b)a_{1,s}ds} a_{1,t}\left((p_{1,t}\lambda^g + (1-p_{1,t})\lambda^b)\Pi_2(A_{1,t}) - c\right) dt - f_1.$$

The optimal effort policy has $a_{1,t} = 1$ if $(p_{1,t}\lambda^g+(1-p_{1,t})\lambda^b)\Pi_2(A_{1,t}) \geq c_2$ and $a_{1,t} = 0$ otherwise.

Note that not only does the probability of succeeding in the first stage decrease in the absence of a success but also the continuation value of reaching the next stage. Because increasing pessimism in the first stage induces also higher pessimism in the second stage, the initial belief of the second stage is lower if more effort was required to reach that stage. This follows from the logic that bad quality projects require more effort to successfully complete the first stage.

As a consequence, the better the project performed in the first stage, the higher is the optimal amount of total experimentation in the second stage. Earlier successes are better news about the project and therefore generate more optimisim about its quality. The first-best experimentation policy therefore uses the first-stage performance to adjust the optimal amount of experimentation in the second stage.

4 Derivation of the Optimal Contract

In this section, I derive the optimal contract via backward induction. That is, I first derive the continuation contract for the second stage. Taking this continuation contract as given, I move to the first stage and study the agent's incentives and solve for the optimal contract. To apply backward induction, I need to ensure that the principal cannot improve upon the optimal continuation contract by committing to a suboptimal continuation contract that reduces deviation incentives in the first stage. I show that the optimal continuation contract subject to promise keeping is indeed the contract that gives the least incentives to deviate which allows me to use backward induction.

4.1 Second-Stage Continuation Contract

I first study the optimal continuation contract after a first-stage success. I proceed in several steps. First, I define the principal's optimization problem. Second, I derive the incentive-compatible bonus payment process that implements any desired effort path. Third, I derive the optimal continuation contract. Finally, I consider the agent's value after a deviation in the first stage.

The principal enters this stage with a belief $p_{2,0}(\hat{A}_{1,\tau_1})$ that depends on her belief about the total effort that has been exerted up to the success time τ_1 in the first stage, where $\hat{A}_{1,\tau_1} \equiv \int_0^{\tau_1} \hat{a}_{1,t}dt$. The principal can only condition on this belief as she can condition on the public history which consists of the success times only. In this subsection, I first assume that the agent has not deviated in the first stage implying that the belief held by principal and agent at the beginning of the stage coincide. To induce the desired effort, the contract has to satisfy the agent's incentive-compatibility condition. Hence, the principal solves the following optimization problem[13]

$$(OBJ_2) \qquad \Pi(\hat{A}_{\tau_1}, v(\tau_1)) = \max_{a_{2,t}, b_{2,t}(\tau_1)} \int_0^\infty e^{-\int_0^t p_{2,s}(\hat{A}_{1,\tau_1})\lambda a_{2,s} ds} a_{2,t} p_{2,t}(\hat{A}_{1,\tau_1})\lambda(\pi - b_{2,t}(\tau_1))dt$$

$$(IC_2) \qquad s.t.\ a_{2,t} \in \arg\max_{\tilde{a}} \int_0^\infty e^{-\int_0^t p_{2,s}(\hat{A}_{1,\tau_1})\lambda \tilde{a}_{2,s} ds} \tilde{a}_{2,t}(p_{2,t}(\hat{A}_{1,\tau_1})\lambda b_{2,t}(\tau_1) - c)dt$$

$$(PK) \qquad v(\hat{A}_{1,\tau_1}) \geq (=) \int_0^\infty e^{-\int_0^t p_{2,s}(\hat{A}_{1,\tau_1})\lambda \tilde{a}_{2,s} ds} \tilde{a}_{2,t}(p_{2,t}(\hat{A}_{1,\tau_1})\lambda b_{2,t}(\tau_1) - c)dt$$

where $v(\hat{A}_{1,\tau_1})$ is the utility the agent is promised from the first stage. Hence, condition PK means that the agent's utility from the second-stage contract has to equal to $v(\hat{A}_{1,\tau_1})$. Note that it depends on the first stage whether the promise-keeping constraint has to hold with equality or inequality. It may be optimal to commit to a value less than the desired level if this reduces first-stage information rents. This will be discussed in the analysis of the first stage. The principal maximizes her payoff by choosing an effort path $\{a_{2,t}\}_{t\geq 0}$ she wants to induce. For the agent to follow that recommendation the bonus payment has to be chosen such that the agent finds it indeed optimal to choose that effort path. That is, $b_{2,t}(\hat{A}_{1,\tau_1})$ has to satisfy IC_2 as well.

Incentive Compatibility. I first study the agent's effort choice and derive the incentive-compatible second-stage bonus payment that induces effort of the agent up to a deadline. The agent's effort choices in the second stage have two effects. First, effort is required to obtain a success at the current instant. Second, effort determines the learning; if more effort has been exerted without a success, the more pessimistic is the agent. Whenever the agent has followed the principal's effort recommendations, their beliefs coincide. However, by deviating from the recommended effort path, the agent can divert his private belief from the principal's belief. In this case, the bonus payment is tailored to the belief the principal holds. This induces a dynamic agency rent. To build intuition, consider a dynamic programming heuristic similar to Bonatti and Hörner (2011). Recall that only a good project can succeed and conditional on a success the value of the project, π realizes and the agent receives bonus $b_{2,t}$.

$$V_t = (1 - e^{-a_{2,t}p_{2,t}\lambda dt})b_{2,t} - ca_{2,t}dt + e^{-a_{2,t}p_{2,t}\lambda dt}V_{t+dt}$$

Using the analogous approximation for V_{t+dt}, approximating the exponentials with a second-order Taylor expansion, dividing by dt^2 and taking the limit as $dt \to 0$ yields for the effect of delaying effort

$$\left(-\frac{\partial V_t}{\partial a_{2,t}} + \frac{\partial V_t}{\partial a_{2,t+dt}}\right)/dt^2 = \dot{b}_{2,t}p_{2,t}\lambda.$$

[13]Note that deadlines are always implemented by the bonus dropping to zero at the desired point in time. Also, note that it is without loss to restart the time variable at the beginning of the second stage. Hence, I use for the second stage the interval $[0, T_2]$ instead of $[\tau_1, \tau_1 + T_2]$.

By shifting effort from today to tomorrow, the agent loses the marginal payoff from effort today, $p_{2,t}\lambda b_{2,t}$, but gains in return the marginal benefit of effort, $p_{2,t}\lambda b_{2,t+dt}$. If the principal were to use an increasing bonus process, the agent had an incentive to delay effort; this induces a *procrastination rent.* If the principal wants to make the agent indifferent between all effort levels, incentive compatibility implies $\dot{b}_{2,t} = 0$. If the bonus process were decreasing, the agent preferred to frontload effort.

Lemma 1. *The time-independent profit-sharing rule $b_{2,t}(A_{1,\tau_1}) = b_2(A_{1,\tau_1})$ that makes the agent indifferent between all effort levels at the second-stage deadline $T_2(A_{1,\tau_1})$, i.e.*

$$b_2(A_{1,\tau_1}) = \frac{c}{\lambda}\frac{1}{p_{2,T_2(A_{1,\tau_1})}(A_{2,T_2})}$$

induces the agent to exert full effort for all $t \in [\tau_1, \tau_1 + T_2(A_{\tau_1})]$. For $t > \tau_1 + T_2(A_{\tau_1})$, $b_2(A_{1,\tau_1}) = 0$.

Note that this bonus payment depends on the total effort that has been required in the first stage. This, as a consequence of the informativeness of the first stage, affects the continuation contract because it determines the belief about the project quality in the second. The bonus chosen by the principal depends on her belief about the effort choices of the agent, i.e., on $\hat{A}_{1,\tau_1}$. However, for now, I assume that the agent has not deviated in the first stage and therefore $\hat{A}_{1,\tau_1} = A_{1,\tau_1}$. With positive discounting, a delay in effort were less attractive and the principal could save on some procrastination rents. The bonus payment was slightly increasing; however, the intuition for the incentives to delay effort were unaltered.

Principal's Optimization. Next, I study the principal's preferred contract subject to the incentive-compatibility and promise-keeping constraints. Because the absence of a success is bad news and the principal as well as the agent become increasingly pessimistic, she will terminate experimentation in finite time. It will turn out that the principal frontloads effort in the second stage. That is, she wants to induce $a_{2,t} = 1$ for all times up to a deadline. Hence, the problem boils down to determining a maximum level of total effort that she wants to induce in the second stage. Because she wants to frontload experimentation and the effort level is at its maximum, total experimentation on path coincides with calendar time, $A_{2,t} = t$.

Recall that incentive compatibility induces a weakly decreasing bonus process. Moreover, promise-keeping requires that the agent's expected utility in the second stage is at least as high as the promise from the first stage, $v(A_{1,\tau_1})$. If the promise-keeping constraint is binding, the principal has to choose how to deliver additional utility to the agent. She can either pay higher bonuses for a success or she can extend the deadline and thereby increase the probability of obtaining the bonus. It is optimal for the principal to deliver additional utility by incentivizing agents to work until extended deadlines. To see why this is optimal, note that the experimentation deadline in the second stage will be distorted downwards from the efficient level derived in Section 3 due to the procrastination rents. Having the agent exert more effort before terminating experimentation increases the total surplus as well as the agent's expected utility generated in the second stage. Therefore, the principal chooses the contract such that the agent receives the promised utility at the highest total surplus. The level of the bonuses is pinned down by the agent's incentive compatibility condition at the deadline.

$$b_{2,T_2}(A_{1,\tau_1}) = \frac{c}{\lambda} \frac{1}{p_{2,T_2}(A_{1,\tau_1})}. \tag{1}$$

If a longer deadline is chosen, the agent is more pessimistic at the deadline and therefore a higher bonus is required to incentivize him to exert effort. The agent's value of a contract with constant bonus process and deadline T_2 is given by

$$v(T_2; A_{1,\tau_1}) = c\,(1 - p_{2,0}(A_{1,\tau_1})) \left(\frac{e^{\lambda T_2} - 1}{\lambda} - T_2 \right). \tag{2}$$

I denote the outcome of a maximization of the principal without promise-keeping constraint by *second-best* contract. The corresponding deadline is denoted by $T_2^{SB}(A_{1,\tau_1})$ and the corresponding utility by $v(T_2^{SB}(A_{1,\tau_1}))$. Note that it may be optimal to commit to a continuation utility that is lower than the second-best utility in the first stage to reduce deviation incentives. When this will occur, will be discussed in the following subsection.

Proposition 1. *The principal-optimal second-stage contract given first-stage success time τ_1, corresponding total effort in the first stage A_{1,τ_1} and agent's promised utility $v(A_{1,\tau_1})$ is given by*

$$T_2(A_{1,\tau_1}) = \begin{cases} T_2^{SB}(A_{1,\tau_1}) & , \text{ if } v(T_2^{SB}(A_{1,\tau_1})) \geq v(A_{1,\tau_1}) \\ T_2(A_{1,\tau_1}, v(A_{1,\tau_1})) & , \text{ if } v(T_2^{SB}(A_{1,\tau_1})) < v(A_{1,\tau_1}) \end{cases}$$

or if the principal commits to providing value less than the second-best by

$$T_2(A_{1,\tau_1}) = T_2(A_{1,\tau_1}, v(A_{1,\tau_1})), \qquad \text{for all } v(A_{1,\tau_1})$$

where $T_2(A_{1,\tau_1}, v(A_{1,\tau_1}))$ is defined as the solution, T, to

$$v(A_{1,\tau_1}) = c(1 - p_{2,0}(A_{1,\tau_1})) \left(\frac{e^{\lambda T} - 1}{\lambda} - T \right)$$

which is given by[14]

$$T_2(A_{1,\tau_1}, v(A_{1,\tau_1})) = -\frac{v(A_{1,\tau_1})}{c(1 - p_{2,0}(A_{1,\tau_1}))} - \frac{1}{\lambda} \left(1 + W_{-1} \left(-e^{-1-\lambda \frac{v(A_{1,\tau_1})}{c(1-p_{2,0}(A_{1,\tau_1}))}} \right) \right).$$

[14] $W_{-1}(x)$ denotes the negative branch of the Lambert-W-function.

$T_2^{SB}(A_{1,\tau_1})$ is given by

$$T_2^{SB}(A_{1,\tau_1}) = \frac{1}{2\lambda} \ln \left(\frac{p_{2,0}(A_{1,\tau_1})}{1 - p_{2,0}(A_{1,\tau_1})} \frac{\pi\lambda - c}{c} \right).$$

The corresponding bonus payment is given by

$$b_{2,t}(A_{1,\tau_1}) = \frac{c}{\pi\lambda} \frac{1}{p_{T_2(A_{1,\tau_1})}}.$$

One important feature of the optimal second-stage continuation contract is that it uses deadlines as the main instrument to deliver utility to the agent: given a deadline, the principal always uses the lowest possible bonus payment that incentivizes the agent to exert effort until that deadline. The underlying reason is that extending deadlines reduces inefficiencies in the total amount of experimentation in the second stage and therefore increases the overall surplus. In addition, by rewarding with extended deadlines, the bonus payment is as low as possible given the promise-keeping condition. Hence, generating the maximum surplus and keeping the bonus payment as low as possible while keeping the promise from the first stage are both obtained through extended deadlines. This observation will lateron lead to the conclusion that this contract is not only profit-maximizing in the second stage but also the contract that yields the lowest incentives to deviate in the first stage.

The main comparative statics that are relevant for the analysis of the first stage are summarized in the following corollary.

Corollary 1 (Comparative Statics of the Continuation Contract.)**.**
The bonus and the deadline in the second stage are (weakly) increasing in the promised utility for a given initial second-stage belief.
The bonus in the second stage is increasing and the deadline decreasing in the initial second-stage belief for a given level of promised utility.

That bonus and deadline are weakly increasing in the promised utility follows from the way the principal provides the agent with additional utility: she extends the deadline and to incentivize the agent to exert effort until the new deadline she has to promise a higher bonus payment conditional on success. The deadline is decreasing in the initial belief because for every deadline, the principal has to provide the agent with higher bonuses to incentivize him. However, she does not want to reduce the deadline too much as this also reduces the probability of obtaining the final breakthrough.

Agent's Continuation Value after a First-Stage Deviation. To study the agent's incentives in the first stage I need to evaluate his continuation payoff after a deviation in the first stage. An agent could deviate by making effort choices that are different from the principal's recommendation. Off-path effort choices have no direct benefit but divert the agent's from the principal's belief. Due to the informativeness of the milestone and the resulting persistence of the private information the deviation has two consequences. First, it affects the agent's belief in the first stage. Second, it affects the initial belief of the second stage because the required effort in the first stage is informative about project quality. If the agent has exerted less effort in the first stage than the principal believes he has, he is more optimistic about the project's quality than the principal. Therefore, the agent will value the continuation contract differently than the principal believes he does. The value of a continuation contract given that the principal beliefs the exerted effort is $\hat{A}_{1,\tau_1}$ while the true exerted effort is A_{1,τ_1} is given by

(3)

$$v(t, A_{1,\tau_1}, \hat{A}_{\tau_1}) = c\left(p_{2,0}(A_{1,\tau_1})\frac{1-p_{2,0}(\hat{A}_{1,\tau_1})}{p_{2,0}(\hat{A}_{1,\tau_1})}\frac{e^{\lambda T_2(A_{1,\tau_1})}-1}{\lambda} - (1-p_{2,0}(\hat{A}_{1,\tau_1}))T_2(A_{1,\tau_1})\right).$$

It is straightforward to show that the agent's value is decreasing in $A_{1,t}$; that is, for every continuation contract he prefers to hold a higher belief than the principal. This already foreshadows that the agent has an incentive in the first stage to shirk in order to become more optimistic than the principal and thereby increase his continuation payoff. I will lateron show that the contract derived in this section is also the contract that yields the smallest incentive to deviate to the agent given the promised utility from the first stage.

4.2 First-Stage Analysis

Given the analysis of the second stage, I now move to the first stage. I analyze the agent's incentives to exert effort first and then study the principal's optimal contract. To incentivize the agent to work in the first stage the principal has to promise a reward in case of a success. As the first stage is followed by the second stage, the principal can use the experimentation assignment in the second stage as a reward instrument. However, the principal can also use a bonus payment to reward the agent for a first-stage success that is independent of future performance. The total reward of the agent consists of both, the bonus payment and the value of the continuation contract

$$w(\tau_1) = b_{1,\tau_1} + v(\tau_1).$$

To understand the agent's incentives in the first stage it is important to note that, in contrast to settings with independent stages, an off-path effort choice has two consequences. First, it diverts the agent's from the principal's belief in the first stage as it is the case in the second stage. Second, it also diverts the initial belief at the beginning of the second stage because the effort required to complete the first stage is informative about the project quality.

To see the impact of the latter on the agent's incentives, consider the implementation of the second-stage contract: the principal chooses the continuation contract such that it delivers in expectation the promised utility from the first stage to the agent. This expectation is calculated based on the principal's belief about the project quality. By first-stage deviations the agent can divert his private belief from the principal's in the second stage. This will affect the agent's true expected value of the continuation contract: recall the off-path value of the agent from equation (3). The gain of holding a marginally more optimistic belief is

$$\frac{\partial v(t, \hat{A}_{1,\tau_1}, A_{1,\tau_1})}{\partial p_{2,0}(A_{1,\tau_1})} = \frac{c}{\lambda} \left(\frac{1 - p_{2,0}(A_{1,\tau_1})}{p_{2,0}(A_{1,\tau_1})} \left(e^{\lambda T_2(A_{1,\tau_1})} - 1 \right) + \lambda T_2(A_{1,\tau_1}) \right) > 0. \tag{4}$$

How this affects the agent's incentives can again be seen in a dynamic programming heuristic:

$$V_t = (1 - e^{-a_{1,t}(p_{1,t}\lambda^g + (1-p_{1,t})\lambda^b)dt}) w_t - c a_{1,t} dt + e^{-a_{1,t}(p_{1,t}\lambda^g + (1-p_{1,t})\lambda^b)dt} V_{t+dt}$$

Using the analogous approximation for V_{t+dt}, approximating the exponentials with a second-order Taylor expansion, dividing by dt^2 and taking the limit as $dt \to 0$ yields for the effect of delaying effort

$$\left(-\frac{\partial V_t}{\partial a_{1,t}} + \frac{\partial V_t}{\partial a_{1,t+dt}} \right) / dt^2 = \left(\dot{w}_t - \frac{\partial v(t, \hat{A}_{1,\tau_1}, A_{1,\tau_1})}{\partial p_{2,0}(A_{1,\tau_1})} \frac{\partial p_{2,0}(A_{1,\tau_1})}{\partial A_{1,\tau_1}} a_{1,t} \right) \left(p_{1,t}\lambda^g + (1 - p_{1,t})\lambda^b \right).$$

This heuristic mirrors the two effects of a deviation: first, a delay in effort affects the level of the total reward, w_t, because the agent can divert the belief in the current stage. Second, the delay in effort also affects the agent's belief in the second stage and thereby the assessment of the continuation contract. Becoming more optimistic than the principal increases the value of a continuation contract that is tailored to a more pessimistic agent. Therefore, the persistence in the private information across stages creates an endogenously arising *ratchet effect*: the agent wants the principal to think that the success probability in the second stage is low because in that case the principal believes that she has to promise high payments conditional on second-stage success to incentivize the agent.

So far, I have assumed that the second-stage contract is implemented as derived in the previous section. By full commitment this is not necessarily the case because it could be better for the principal to commit to a suboptimal second-stage contract that reduces the deviation incentives in the first stage. The following lemma shows that the optimal second-stage contract is the implementation of the promised utility from the first stage that induces the lowest incentive to deviate allowing me to use backward induction.

Lemma 2. *The continuation contract derived in Proposition 1 induces the lowest incentives to deviate in the first stage while satisfying the promise-keeping condition.*

Intuitively, this result holds because the contract in Proposition 1 is the implementation with the lowest bonus payment after a second-stage success. The ratchet effect in the first stage arises because by making the principal more pessimistic the agent is promised a higher bonus conditional on success in the second stage. This effect is increasing in the second-stage bonus and therefore the incentive to deviate is increasing in the bonus payment. Hence, the principal wants to implement the continuation contract such that the bonus payment is as low as possible. As a consequence, she rather extends the deadline further and thereby increases the success probability rather than increasing only the bonus payment keeping the deadline at the second-best level.

Taking this continuation contract as given, I next characterize the minimal total reward process that induces incentive compatibility of an effort path $\{a_{1,t}\}_{t\geq 0}$ in the following proposition.

Proposition 2. *The minimal required continuaton utility to induce effort $\{a_{1,t}\}_{t\geq 0}$ in the first stage solves the following differential equation*

$$\dot{w}(t) = \frac{\partial v(t, \hat{A}_{1,t}, A_{1,t})}{\partial A_{1,t}} a_{1,t}$$

with boundary condition

$$w(T_1) = \frac{c}{p_{1,T_1}\lambda^g + (1 - p_{1,T_1})\lambda^b}$$

and $w_t = 0$ for all $t > T_1$, if there is a T_1 such that $a_{1,t} = 0$ for all $t > T_1$.

This proposition shows that due to the informativeness of the first stage the agent has to receive an additional rent to exert effort if he is assigned experimentation in the second stage. Note that in the absence of the informativeness of the milestone this rent would not be required as then $\frac{\partial v(t,\hat{A}_{1,t},A_{1,t})}{\partial A_{1,t}} = 0$ which is the case, for example, in Moroni (2017). As $\frac{\partial v(t,\hat{A}_{1,t},A_{1,t})}{\partial A_{1,t}} < 0$, the agent's continuation value is decreasing over time and can be disentangled into three components

$$\dot{w}_t = \underbrace{-\frac{c}{\dot{p}_{1,t}}}_{\text{static MH}} + \underbrace{\frac{c}{\dot{p}_{1,t}}}_{\substack{\text{procrastination}\\ \text{rent}}} + \underbrace{\frac{\partial p_{2,0}(A_{1,t})}{\partial A_{1,t}} a_{1,t}}_{\substack{\text{effect of effort}\\ \text{on belief}}} \underbrace{\frac{\partial v(t, \hat{A}_{1,t}, A_{1,t})}{\partial p_{2,0}(A_{1,t})}}_{\substack{\text{effect of belief}\\ \text{on value}}} \tag{5}$$

The first term corresponds to the agent's instantaneous cost of effort that he has to be compensated for. This changes over time as the agent becomes more pessimistic when exerting effort. However, this induces the procrastination incentive to obtain a higher reward and the agent has to be granted a procrastination rent. Moreover, due to the persistence of the private information across the stages, the agent has to obtain the informativeness rent. To incentivize effort the total reward on path has to decrease sufficiently steeply over time. The rate at which it decreases is such that the gain from delaying effort and thereby becoming more optimistic than the principal in the second stage is at most as large as the value the agent loses from not succeeding today. Importantly, it has to decrease more steeply if the value of the continuation contract is higher because it implies that a higher bonus payment is required in the second stage. Hence, the more utility the agent receives through a continuation contract, the higher is the incentive to divert the beliefs. This induces first-stage information rents to increase in the value of the continuation contract at a given success time for all earlier success times. This creates a downside of using continuation contracts and therefore long-term incentives because they induce informativeness rents in the first stage. However, using continuation contracts also

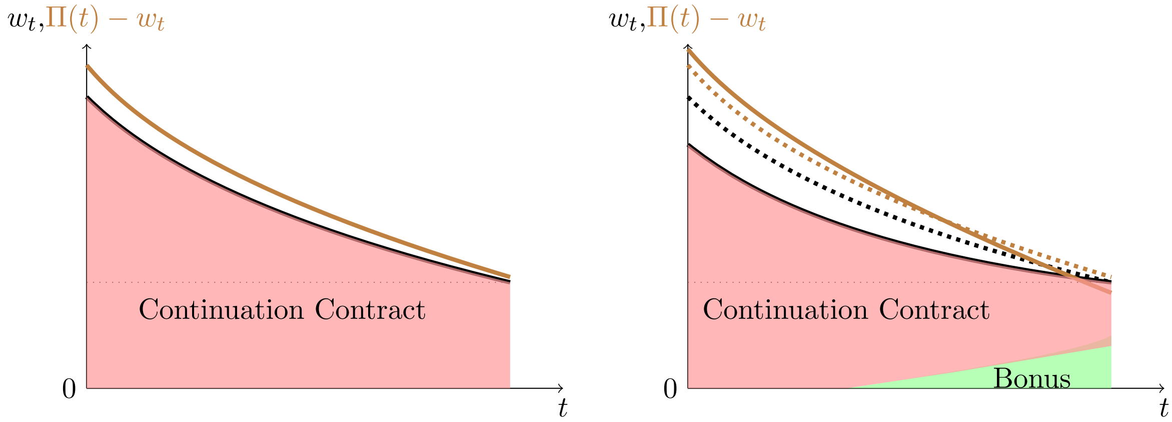

Figure 2: Tradeoff between Continuation Contract and Bonus Payment.
The left panel shows a hypothetical total reward of the agent as a function of the success time if the agent were rewarded with a continuation contract only. The total reward therefore decreases relatively steeply. The brown line illustrates the corresponding profits of the agent.
The right panel shows how the optimal contract improves on the hypothetical contract illustrated in the left panel: by introducing bonus payments at the end, it reduces the informativeness rents for all earlier success. The dotted lines are the total reward and profits from the left panel as benchmark.

has an advantage over bonus payments: by using continuation contracts the principal generates a continuation value to herself as she only receives the benefits of the project if second-stage experimentation is successful. Moreover, recall that second-stage experimentation is inefficiently short due to procrastination rents. Suppose that the principal has to provide the agent with an additional unit of total reward after a first-stage success. Then, she has to choose whether to deliver this through a first-stage bonus payments, which implies lower information rents for earlier successes in the first stage, or through a more valuable continuation contract, which generates additional surplus in the second stage through more efficient second-stage experimentation. This tradeoff is illustrated in Figure 2. Introducing a bonus payment after late first-stage successes reduces the information rents at all previous success times but costs profits at the respective success time.

Therefore, the informativeness of the first stage induces a tradeoff between short-term incentives that only condition on current performance, i.e., bonus payments after the first success, and long-term incentives that condition on future performance as well. This tradeoff does not arise in a setting without persistent information across stages because the agent does not have the ability to divert the beliefs that the second-stage contract terms condition on. Therefore, in a setting with independent stages the principal would never use bonus payments that only condition on first-stage success. In the following I am restricting attention to the case of *costly incentives* defined below.

Definition 1 (Costly Incentives)**.** *First stage incentives are costly if the agent's first-stage* *incentive constraint is binding for all* $t \in [0, T_1]$. *That is*

$$\gamma_t \equiv \dot{w}_t - \frac{\partial v(t, \hat{A}_{1,t}, A_{1,t})}{\partial A_{1,t}} a_{1,t} \tag{6}$$

is such that $\gamma_t = 0$ *for all* $t \in [0, T_1]$. *A sufficient condition for costly incentives to occur is that* $\dot{w}_t \leq \dot{v}_t^{SB}$ *and* $w_{T_1} \geq v_{T_1}^{SB}$.

This implies that the principal does not provide the agent with more utility than necessary to ensure incentive compatibility in the first stage. That is, the first-stage incentive constraint is binding for all $t \in [0, T_1]$. It may be the case that under some parameter values, the principal is willing to give more utility to the agent than necessary. This can occur if the incentives in the first stage are relatively cheap such that a continuation value less than the value of the second stage alone would incentivize the agent to work in the first stage. If incentives are relatively cheap, the incentive constraint still requires $\gamma_t \geq 0$ as in Proposition 2. As the main contribution of my paper lies in the analysis of the first-stage incentives and the corresponding optimal contract, I am assuming that first-stage incentives are costly.

The following theorem shows how the optimal contract solves this tradeoff in the costly incentives case.

Theorem 1. *Suppose that first-stage incentives are costly. The total reward* w_t *the agent receives conditional on completing the first stage at time* t *induces full effort, is strictly decreasing for all* $t \in [0, T_1]$ *and solves*

$$\dot{w}_t = \frac{\partial p_{2,0}(A_t)}{\partial A_t} \frac{\partial v(t, A_t)}{\partial p_{2,0}(A_t)} a_t \quad s.t. \; w(T_1) = \frac{c}{p_{T_1}\lambda^g + (1 - p_{T_1})\lambda^b}$$

and has $w_t = 0$ *for all* $t > T_1$. *There is a* $\hat{t} \in (0, T_1)$ *such that* $w_t = v_t$ *and* $b_{1,t} = 0$ *for all* $t \in [0, \hat{t}]$*; i.e., early successes are rewarded with continuation contracts only.*
For all $t \in (\hat{t}, T_1]$, $w_t > v(t)$ *and* $b_{1,t} > 0$ *with* $\frac{b_{1,t}}{w_t}$ *increasing in* t*; i.e., if the success is obtained after* $\hat{t}$*, the total reward consists of a continuation contract and a bonus payment with the share of the bonus payment in the total reward increasing in the success time.*
The continuation contract, $v(t)$*, is implemented according to the optimal second-stage contract as in Proposition 1.*

Theorem 1 implies that the composition of the reward changes over time. Early successes are rewarded with continuation contracts only and the second-stage contract has deadlines close to the first best. If the success arrives late, the reward consists of bonus payments as well as less valuable experimentation assignment in the second stage. The part of the reward that is provided to the agent with a bonus payment is increasing, the later the breakthrough is obtained.

The reason that a lower share of the reward is provided with continuation contracts over time is that the additional gain in overall surplus from extended deadlines is decreasing in the belief about the project's quality. This belief is decreasing in the first-stage success time. Moreover, the arising ratchet effect implies that higher information rents have to be paid for *all earlier success times* if more utility is delivered through a continuation contract. By choosing the share of the total reward that is delivered through a continuation contract, the principal can control the information rents for all earlier successes. At earlier success times, the agent has to be granted a sufficiently higher reward to prevent him from delaying effort. If a success at a later time is rewarded with a relatively high share of utility through a continuation contract, the gain from holding a more optimistic belief then is high. Therefore, a higher reward for earlier successes is needed to prevent a deviation. Hence, bonus payments become more favorable for later successes for two reasons: the effect on information rents for earlier successes increases and the gain of extending deadlines decreases in the belief.

It is interesting to note that the optimal contract provides a decreasing amount of value through a continuation contract and therefore induces the continuation contract instruments to vary with success times. The second-stage deadline is strongly dependent on the first-stage success time while the second-stage bonus is not as dependent on performance as in the second-best contract. The reason is, as discussed above, that the incentive cost is lower if the agent is rewarded with extended deadlines rather than with higher bonus payments. However, having the bonus payment become less dependent on performance reduces the ratchet effect.

Corollary 2. *Second-stage deadlines are decreasing in the first-stage success time and more responsive to it than in the second-best contract.*
Bonus payments are increasing in the first-stage success time and less responsive to it than in the second-best contract.

How the second-stage contract terms depend on the first stage outcome is illustrated in Figure 3.

5 Endogenous Agent Replacement

In this section, I consider the case in which the principal has access to another agent for the second stage. This gives her the additional choice whether to keep the agent from the first stage to work in the second stage as well or whether she rather has a new agent and get the second-best value in the second stage. If the agent is replaced, he is rewarded with a bonus payment only and no continuation value. As shown in the previous section, this implies that at those success times at which the agent is replaced, he does not need to receive the informativeness rent. Therefore, the principal saves on rents for earlier success times when replacing the agent. However, she foregoes the possibility to gain from extended deadlines in the second stage. Replacement is the most extreme bonus payment and saves the most informativeness rents. Without a new agent this corresponds to terminating the project as no continuation value is granted to the agent and no second-stage experimentation takes place. Because the principal can obtain the second-best value after replacement if she has access to a new agent, I show that she will *always*

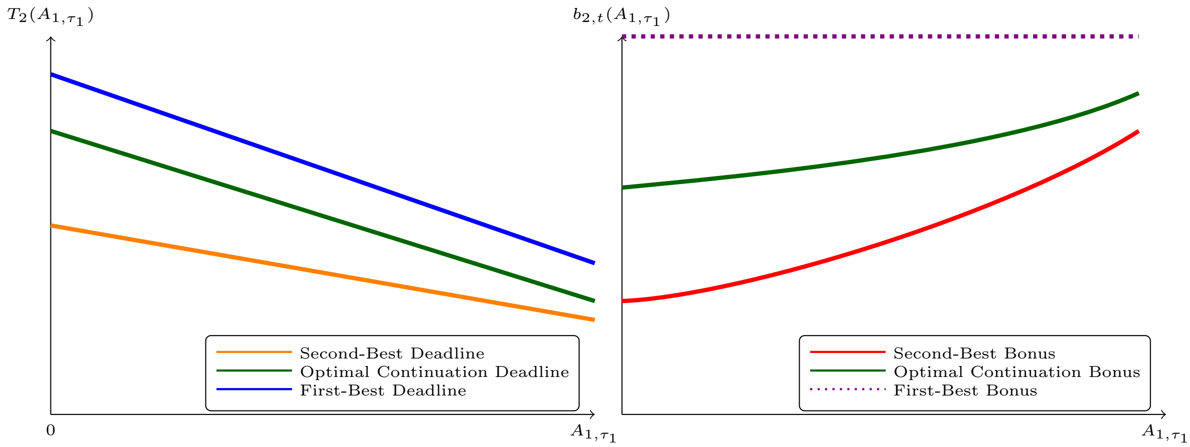

Figure 3: Deadline and Bonus in Optimal Contract.
The left panel shows how the second-stage deadline in the optimal contract varies with the first-stage success time compared to the first-best and second-best deadline.
The right panel show how the second-stage bonus in the optimal contract varies with the first-stage success time compared to the first-best and second-best bonus.

make use of this possibility for success times close to the deadline. The optimal contract with replacement is illustrated in Figure 4. I assume that parameters are such that we are in the costly incentives case. If the agent has not to be granted rents that exceed the second-best value it is straightforward that replacement may not be desirable from the principal's point of view.

Theorem 2. *Suppose that first-stage incentives are costly. If the principal has access to another agent in the second stage, she will choose two deadlines, $\hat{T}$ and T_1 in the first stage. For all $t \in [0, \hat{T}]$, the agent receives w_t upon a breakthrough as in the optimal contract with boundary condition given by $w_{\hat{T}} = \frac{c}{p_{T_1}\lambda^g + (1-p_{T_1})\lambda^b}$ and works in the second stage. For all $t \in (\hat{T}, T_1]$, the agent receives $w_t = b_t = w_{T_1}$ upon a breakthrough and a new agent works on the second-stage contract according to the second-best value and the belief given by $p_{2,0}(A_{1,\tau_1})$. If no success has been obtained by T_1, the project is terminated.*

In light of the optimal contract without replacement this result seems intuitive. However, it is not obvious: there is no learning about the agent's type but still the successful agent gets replaced, although continuation contracts are "cheaper" to provide the required utility than bonus payments. In particular, when the first milestone is not informative about the second stage, replacement between stages never occurs because there are no information rents to be saved by replacement. Moreover, if agents could be continuously replaced even within a stage at no cost, the principal would replace the agent continuously and induce first-best experimentation because no dynamic agency rents have to be paid at all. Theorem 2 shows that informative milestones give rise to replacement of agents that do succeed in their assigned task but took relatively long to do so. The underlying reason is that continuation contracts give rise to the informativeness rents caused by the persistence of learning. Hence, replacement occurs to reduce information rents in the first stage. This may be one explanation for high managerial turnover

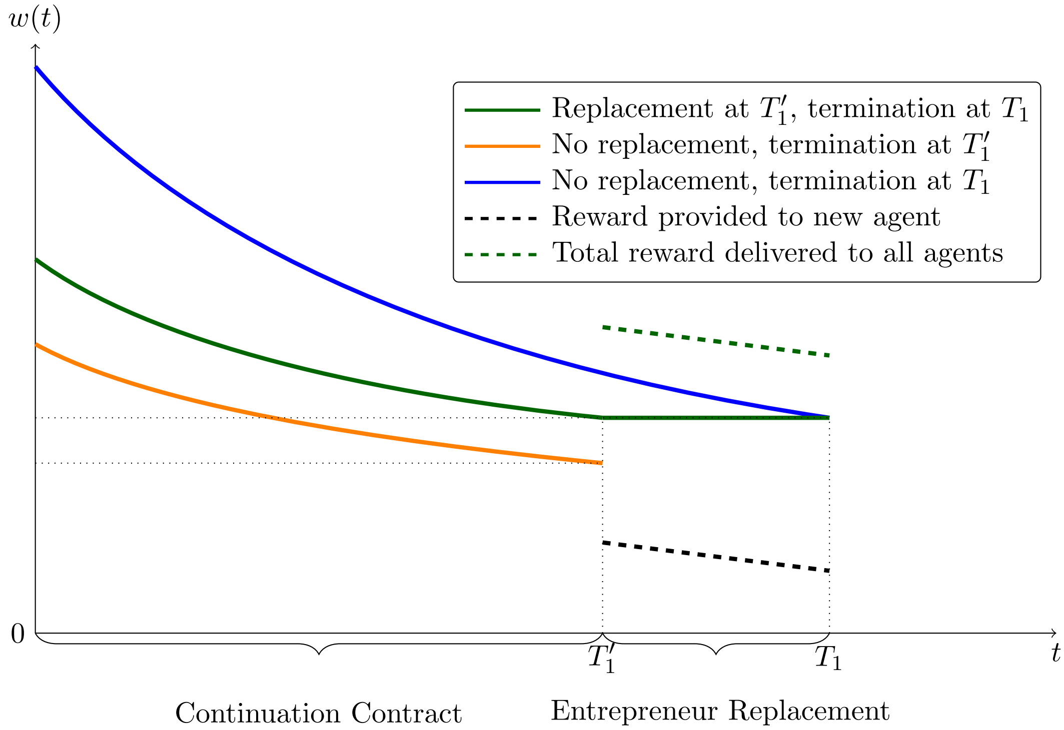

Figure 4: Agent Replacement.

The orange line shows a hypothetical reward process with a short deadline T_1'. The blue line show a hypothetical reward process with longer deadline T_1'. The green line shows the evolution of a contract that rewards the agent at least partially with a continuation contract for a success before T_1' and with a bonus payment only for a success before the deadline T_1. The dashed black line corresponds to the new agent's value in the second stage if the first-stage agent is replaced. The dashed green line is the total reward that both agents receive in the replacement region.

rates in the innovative industries found by Hannan, Burton, and Baron (1996), for example.

6 Extensions

6.1 Project Design

In this section, I study the choice of the informativeness of the first stage and endogenous staging, i.e., the decision of choosing one or two stages for the project. In many instances, a milestone may not be necessary to implement the full project but still the principal requires it: when a prototype is required, it may serve as an informative signal about the project. Typically, the principal is able to decide on the informativeness of the prototype; that is, how many details of the final product should be incorporated. The tradeoff of the principal is that on the one hand she wants the signal to be as informative as possible to prevent funding a bad-quality project in the second stage. On the other hand, if the initial informativeness is low, then if the signal becomes marginally more informative, the agent's ability to divert the beliefs increases and therefore the informativeness rent in the first stage does as well. However, the capital required for the first stage is increasing in the informativeness because the more informative is the first tak, the closer it is to the final product. Hence, the optimal level of informativeness is not obvious and may not be extreme.

The ratio $\frac{\lambda^g}{\lambda^b}$ can be interpreted as the informativeness of the first stage. The higher is the ratio the higher is the upward jump in the belief after a success. If $\frac{\lambda^g}{\lambda^b} \to \infty$, there is certainty after a first-stage success that the project is of good quality. If $\frac{\lambda^g}{\lambda^b} \to 1$, there is no learning at all. I assume that the more has been learned in the first stage, the faster is a success obtained in the first stage. Towards this, I assume that the good project's intensity rate in the second stage is given by $\lambda = \frac{\lambda^g}{\lambda^b}$. This implies that when the first stage is perfectly informative, there is no second stage because $\lambda = \infty$. I can show that the two-stage contract converges to the second-best one-stage contract as $\lambda^b \to 0$. Also, I change λ^g with λ^b such that the expected duration of project completion conditional on the quality being good remains constant, which yields $\lambda^g = 1 + \lambda^b$ when normalizing $\lambda^g(\lambda^b = 0) = 1$. An additional advantage of this formulation is that for all combinations of λ^g and λ^b under this restriction the belief evolution in the first stage is identical because $\lambda^g - \lambda^b = 1$. However, the upwards jump after a success depends on the ratio $\frac{\lambda^g}{\lambda^b}$ as the posterior is given by

$$p_{2,0} = \frac{1}{\frac{1-p_0}{p_0} \frac{\lambda^b}{1+\lambda^b} e^t + 1}$$

which is decreasing in λ^b and goes to one as λ^b goes to zero. To capture the feature of capital infusions that are contingent on milestones, I let the fixed cost per stage, f_i depend on the informativeness of the first stage. The more informative the first stage, the higher is the cost for this stage, $f_i(\frac{\lambda^g}{\lambda^b})$. I assume that if $\frac{\lambda^g}{\lambda^b} \to \infty$ the cost of the second stage converges to zero and the first stage cost converges to the one-stage case. Moreover, I continue to assume that the parameters are such that we are in the costly incentives case. Otherwise, the principal would get the informative first stage signal without having to deliver any additional rents and the staging decision became trivial.

As a first result I show that the optimal two-stage contract converges to the optimal one-stage contract as the first stage becomes perfectly informative.

Lemma 3. *With $\lambda = \frac{\lambda^g}{\lambda^b}$ and $\lambda^b \to 0$, the two-stage optimal contract from the previous sections, converges to the optimal one-stage contract with $\lambda^{onestage} = \lambda^g(\lambda^b = 0)$.*

This result allows me to study endogenous staging as a choice of λ^b numerically quite straightforwardly. When the principal chooses $\lambda^b = 0$ the problem collapses to a one-stage problem. Note that in the present setting, the principal cannot choose an entirely uninformative first stage except for the limit case $\lambda^b \to \infty$ because $\lambda^g = 1 + \lambda^b > \lambda^b$. In this case, again, the problem would collapse to a one-stage problem because the first stage is immediately completed and there is no way to divert beliefs for the agent. It is immediate that the principal would never choose a perfectly uninformative first stage $\lambda^b = \lambda^g$. This would make the first stage a pure moral hazard stage without any signal. Still, the agent has to be incentivized to exert effort. That is, the principal would have to deliver additional rents to the agent without gaining from the first stage.

It follows from the comparative statics of the optimal contract in the informativeness that the informativeness rent is inversely u-shaped in the informativeness. If the first stage is entirely uninformative, then the informativeness rent is zero. If the first stage is fully informative, it is zero as well. In between, it is strictly positive. Hence, moving from a one-stage project to a two-stage project with a somewhat informative first stage has the following effects: (i) a positive informativeness rent has to be delivered to the agent in expectation (ii) the principal can condition the second capital infusion on the first-stage outcome. These two effects work against each other and it depends on the parameters which one dominates. Note that if the initial belief is sufficiently high, investing all capital at once and avoiding the informativeness rent is more attractive. If the initial belief is lower, investing all capital at once is less attractive because it is lost with a high probability. However, the principal may then introduce a first stage at a cost that is lower than investing into the full project immediately to generate an informative signal and condition the second infusion on this signal.

Numerical Results. In a numerical analysis in Mathematica, I study the endogenous choice of staging and the optimal degree of informativeness of the first stage. This reveals that the previously discussed tradeoff between terminating bad projects and agent's information rents is relevant when designing a project. In several specifications, the optimal degree of informativeness is interior and hence the choice of two stages dominates a one-stage project. However, not requiring a milestone may also be optimal under other parameter values. The analysis reveals intuitive comparative statics, as the agency conflict increases, the optimal informativeness decreases. This follows because the cost of the informativeness is increasing in the agency conflict and hence, the optimal informativeness is reduced. Also, the principal chooses an inefficiently low level of informativeness.

The conjecture that projects with lower initial beliefs; i.e., more risky projects are more likely to be staged investments seems to be true in numerical examples. This is in line with the findings in Bienz and Hirsch (2011).

6.2 Privately Observable Successes

One important feature of the optimal contract derived is that even if successes were not publicly but only privately observable, the agent would not make use of the possibility to strategically hide a success. In principle, a profitable deviation of the agent could be to hide a success if it is obtained instead of shirking to divert the beliefs. However, the gain from hiding a success is the same as the gain from shirking for an instant: it alters the principal's belief in the following stage and therefore increases the value of the continuation contract from the agent's perspective. The optimal contract precludes this behavior by rewarding earlier successes with higher rents. Hence, the agent would immediately reveal a private breakthrough.

Corollary 3. *If successes are privately observed by the agent, the agent immediately reveals a success.*

It follows from this corollary that the assumption of publicly observable breakthroughs is without loss of generality. In principle, with private observability the agent has an incentive to strategically delay the arrival of a success due to the arising ratchet effect. The principal needs to impose an additional truthtelling/revelation constraint on her optimization problem. However, by inspecting the incentive to hide a success, it becomes apparent that this condition coincides with the decision to exert effort. Hence, the optimal contract satisfies the revelation constraint as well and the agent immediately reveals a success.

6.3 Learning About Agent's Type

If the learning is about the agent's type instead of the project and the project's quality is known, the optimal contract without replacement is as in Theorem 1. However, replacement is different in this case because a new agent is hired, that is at the beginning of the second stage the initial belief is back at the initial prior p_0. This changes the optimal replacement deadline, but not the incentive to introduce a replacement region. It may even be more attractive to introduce replacement because it allows to increase the belief at the beginning of the second stage if the principal became too pessimistic in the first stage. The continuation value for the principal upon replacement would therefore be unaffected by first stage outcomes and she can always guarantee herself at least this continuation payoff after a first-stage success.

Corollary 4. *If the agent's type is unknown, the optimal contract without replacement is as in Theorem 1. With replacement and independent agents, the continuation value after replacing the agent is independent of the timing of the first breakthrough and given by the second-best value under the initial prior belief* $v^{SB}(p_{2,0} = p_0)$.

This shows that replacement may also occur due to information rent reasons if learning is about the agent's type. That is, replacement does not only occur because the agent is too likely to be of low quality but to reduce informativeness rents that he would have to be provided if he would receive a continuation contract. However, the possibility to obtain a new agent whose quality is drawn according to the initial prior makes replacement attractive as well. It increases the continuation value for the principal if the belief about the current agent's quality is low.

7 Conclusion

In this paper, I study the optimal contract for a two-stage project under full commitment in a dynamic moral hazard setting with ex ante symmetric information and learning within and across stages.

I show that the informativeness of the first stage gives rise to an endogenously arising ratchet effect. As a consequence, the optimal contract has to provide the agent with additional rents for good performance. Moreover, using long-term rewards with continuation contracts that condition on future performance is costly in that they amplify the ratchet effect and therefore increase the agent's information rents. This induces the composition of the total reward to change with performance: bad performance cannot be identified as either bad luck or simple shirking. Thus, the agent receives a higher share of the total reward as a bonus payment rather than a continuation contract if a success is obtained later. Good performance is rewarded more with continuation contracts because these reduce the inefficiencies in the second stage caused by procrastination rents. If the principal has the ability to replace agents after stages, she will make use of this possibility for the latest success times that still induce continuation. By replacing the agent, the principal eliminates the incentive to manipulate the performance within the replacement region and therefore reduces informativeness rents for all success times.

My analysis has several empirical implications: (i) The composition of the agent's compensation changes with performance. In particular, if performance gets worse, the total reward is lower and consists of relatively more short-term than long-term rewards. For example, a well-performing CEO is rewarded with stock options that are tied to future performance. A CEO that performs worse is rewarded with bonus payments and less with stock options. The total worth of the reward is higher for the well-performing CEO. (ii) Deadlines are relatively more responsive to early performance while final-stage bonus payments are less responsive to early performance compared to a setting without informativeness rents. (iii) Early-stage deadlines are relatively short if there is a learning spillover to future stages. (iv) Even successful agents may be replaced if they do not perform sufficiently well although the project is continued and the agent known to be able to complete future tasks. (v) Staging occurs more frequently if the initial risk is high.

Besides studying these empirical implications, there are still open avenues for future research: First, it would be interesting to consider the case of no or only partial commitment of the principal. Second, in a setting with ex ante private information of the agent, one may wonder whether a menu of differently staged contracts can elicit the agent's superior information about the project. Third, analyzing how competition between agents and free-riding interact in the presence of the arising ratchet effect is another potential extension.

8 Appendix

8.1 Preliminaries

The following results will be used frequently throughout the analysis.

Probability that no success has occurred until t. The Poisson distribution implies that no success occurs in an interval $[0,t]$ with probability

$$e^{-\int_0^t (p_{i,s}\lambda^g + (1-p_{i,s})\lambda^b)a_{1,s}ds}.$$

Using the definition of the posterior and its law of motion, $dp_t = -p_t(1-p_t)\Delta\lambda a_{1,t}$, we can rewrite this probability. First, note that the law of the posterior can be written as

$$-p_t\lambda^g a_{1,t} = \frac{dp_{i,t}}{1-p_{i,t}} - p_{i,t}\lambda^b a_{1,t}.$$

Second, I apply this on the probability of no success:

$$\begin{aligned}
&= e^{\int_0^t \frac{dp_{i,s}}{1-p_{i,s}} - p_{i,s}\lambda^b a_{1,s}ds} e^{-\int_0^t [(1-p_{i,s})\lambda^b]a_{1,s}ds} \\
&= e^{\int_0^t \frac{dp_{i,t}}{1-p_{i,t}}} e^{-\int_0^t [(1-p_{i,s})\lambda^b + p_{i,s}\lambda^b]a_{1,s}ds} \\
&= e^{\int_0^t \frac{dp_{i,t}}{1-p_{i,t}}} e^{-\int_0^t \lambda^b a_{1,s}ds} \\
&= e^{-\ln(1-p_{i,t})} e^{-\lambda^b \int_0^t a_{1,s}ds} \\
&= \frac{1-p_{i,0}}{1-p_{i,t}} e^{-\lambda^b \int_0^t a_{1,s}ds}.
\end{aligned}$$

Using the posterior at time t, $1-p_{i,t} = \frac{e^{-\lambda^b \int_0^t a_{1,s}ds}(1-p_{i,0})}{e^{-\lambda^g \int_0^t a_{1,s}ds}p_{i,0} + e^{-\lambda^b \int_0^t a_{1,s}ds}(1-p_{i,0})}$, we get

$$= e^{-\lambda^g \int_0^t a_{1,s}ds}p_{i,0} + e^{-\lambda^b \int_0^t a_{1,s}ds}(1-p_{i,0})$$

This allows me to rewrite the agent's objective functions as follows

$$\begin{aligned}
&\int_0^\infty e^{-rt} e^{-\int_0^t (p_{i,s}\lambda^g + (1-p_{i,s})\lambda^b)a_{1,s}ds} \cdot a_{1,t}\left((p_{i,t}\lambda^g + (1-p_{i,t})\lambda^b)v_i(t) - c\right) dt \\
&= \int_0^\infty e^{-rt} a_{1,t}\left(e^{-\lambda^g \int_0^t a_{1,s}ds}p_{i,0} + e^{-\lambda^b \int_0^t a_{1,s}ds}(1-p_{i,0}) \cdot ((p_t\lambda^g + (1-p_t)\lambda^b)v_i(t) - c)\right) dt \\
&= \int_0^\infty e^{-rt} a_{1,t}\left(\left(p_{i,0}\lambda^g e^{-\int_0^t a_{1,s}\lambda^g ds} + (1-p_{i,0})\lambda^b e^{-\int_0^t \lambda^b a_{1,s}ds}\right) v_i(t)\right. \\
&\left. - \left(e^{-\lambda^g \int_0^t a_{1,s}ds}p_{i,0} + e^{-\lambda^b \int_0^t a_{1,s}ds}(1-p_{i,0})\right) c\right) dt
\end{aligned}$$

where $v_i(t)$ is the value of succeeding in stage i at time t. For the principal, the value is given by

$$\int_0^\infty e^{-rt}\left(p_{i,0}\lambda^g e^{-\int_0^t a_{1,s}\lambda^g ds} + (1-p_{i,0})\lambda^b e^{-\int_0^t \lambda^b a_{1,s} ds}\right)\pi_i(t)dt$$

where $\pi_i(t)$ is the principal's value of a success in stage i at time t.

Posteriors and Odds Ratios. The belief in the second stage depends on the success time in the first stage and parameters of the model. The initial belief in terms of primitives and success time is given by

$$p_{2,0}(\tau_1) = \frac{\lambda^g p_0 e^{-\int_0^{\tau_1}\lambda^g a_{1,s}ds}}{\lambda^g p_0 e^{-\int_0^{\tau_1}\lambda^g a_{1,s}ds} + \lambda^b(1-p_0)e^{-\int_0^{\tau_1}\lambda^b a_{1,s}ds}}$$

and the posterior after an experimentation duration of t in the second stage

$$p_{2,0}(\tau_1, t) = \frac{\lambda^g p_0 e^{-\int_0^{\tau_1}\lambda^g a_{1,s}ds - \int_0^t \lambda a_{1,s}ds}}{\lambda^g p_0 e^{-\int_0^{\tau_1}\lambda^g a_{1,s}ds - \int_0^t \lambda a_{1,s}ds} + \lambda^b(1-p_0)e^{-\int_0^{\tau_1}\lambda^b a_{1,s}ds}}.$$

The odds ratio is then given by

$$\frac{p_{2,0}(\tau_1, t)}{1-p_{2,0}(\tau_1,t)} = \frac{p_0}{1-p_0}\frac{\lambda^g}{\lambda^b}e^{-\int_{\tau_1}^{\tau_1+t}\lambda a_{1,s}ds - (\lambda^g-\lambda^b)\int_0^{\tau_1} a_{1,s}ds}.$$

Bonus contracts are without loss of generality. The same argument as in Moroni (2017) yields the result. Denote the general payment process $\{w_f dt + w_l\}_{t\geq 0}$ by w. w maps histories into payments, $w : \mathcal{H}^t \to \mathbb{R}$. Consider a bonus contract b that only has payments at time zero ($\tau_0 = 0$) and breakthrough times τ_1 and τ_2. Define $w_i(\emptyset, h^{\tau_{i-1}})$ as discounted payoff that payment process w delivers to the agent given the history if the game ended without a breakthrough at h^{τ_i}. Then, let $b_0 = w_1(\emptyset, h^0)$ and $b_{\tau_1}(h^{\tau_1}) = e^{r\tau_1}\left(w_{i+1}(\emptyset, h^{\tau_i}) - w_i(\emptyset, h^{\tau_{i-1}})\right)$. This is a bonus contracts giving the same expected payoff after every history to the agent as the initial contract w. Limited liability is satisfied in the bonus contract as well if no positive payments are made if no breakthrough is obtained in a stage. Such a payment rule is clearly suboptimal for the principal.

8.2 Proofs

If it does not cause confusion, I drop stage indices in the proofs to simplify notation.

Proof of Lemma 1. [15] The proof relies on Pontryagin's maximum principle. The second-stage analysis in my model resembles the one of Moroni (2017). The agent's problem is to choose an effort path $\{a_t\}$ given a contract $b_{2,t}$ to maximize expected payoffs, that is

$$\max_{\{a_t\}} \int_0^\infty e^{-rt} e^{-\int_0^t p_s \lambda a_s ds} a_t \left(p_t \lambda b_{2,t} - c\right) dt.$$

Using the definition of the posterior as well as the differential equation determining its law, this can be rewritten as

$$\max_{\{a_t\}} \int_0^\infty e^{-rt} a_t \left(p_0 e^{-\lambda \int_0^t a_s ds} \lambda b_{2,t} - (p_0 e^{-\lambda \int_0^t a_s ds} + 1 - p_0) c\right) dt.$$

Defining $A_t = \int_0^t a_s ds$, we can rewrite the maximization as optimal control problem

$$\max_{\{a_t\}} \int_0^T e^{-rt} a_t \left(p_0 e^{-\lambda A_t} \lambda b_{2,t} - c p_0 e^{-\lambda A_t} - c(1-p_0)\right) dt$$
$$\text{s. t. } \dot{A}_t = a_t.$$

The Hamiltonian and the costate law are given by

$$\mathcal{H} = e^{-rt} \left(p_0 e^{-\lambda A_t} \lambda b_{2,t} - c p_0 e^{-\lambda A_t} - c(1-p_0)\right) a_t + \eta_t a_t$$
$$\dot{\eta}_t = e^{-rt} p_0 (b_{2,t} \lambda - c) e^{-\lambda A_t} \lambda a_t.$$

Note that the objective is linear in a_t is binary. Let

$$\gamma_t \equiv e^{-rt} \left(p_0 e^{-\lambda A_t} \lambda b_{2,t} - c p_0 e^{-\lambda A_t} - c(1-p_0)\right) + \eta_t,$$

then if $\gamma_t > 0$, the agent will exert effort, $a_t = 1$, and if $\gamma_t < 0$, he will exert no effort $a_t = 0$. If the principal wants to induce effort, she will choose $\gamma_t = 0$ at which the agent is indifferent between working and shirking. This is optimal, because whenever $\gamma_t > 0$, the principal can increase her payoff by slightly reducing $b_{2,t}$ without altering the agent's incentives.
The standard boundary condition gives $\eta_T = 0$ implying $\gamma_T = e^{-rT}\left(p_0 e^{-\lambda A_t} \lambda b_{2,T} - c p_0 e^{-\lambda A_t} - c(1-p_0)\right)$. Choosing $\gamma_t = 0$ implies

$$\eta_t = -e^{-rt} \left(p_0 e^{-\lambda A_t} (\lambda b_{2,t} - c) - c(1-p_0)\right).$$

Differentiating this with respect to time and equating it with (OBJ) delivers

$$e^{-rt} p_0 (b_{2,t} \lambda - c) e^{-\lambda A_t} \lambda a_t = e^{-rt} p_0 (b_{2,t} \lambda - c) e^{-\lambda A_t} \lambda a_t - e^{-rt} p_0 e^{-\lambda A_t} \dot{b}_{2,t} \lambda$$
$$+ r e^{-rt} \left(p_0 e^{-\lambda A_t} (\lambda b_{2,t} - c) - c(1-p_0)\right)$$

[15] The existence and sufficiency results to the optimal control problem analyzed for the second stage in this paper follow directly from Moroni (2017) and the references therein.

and hence, we may conclude that

$$\dot{b}_{2,t} = r\left(b_{2,t} - c\left(1 + \frac{1-p_0}{p_0}\frac{e^{\lambda A_t}}{\lambda}\right)\right)$$

together with the boundary condition

$$b_{2,T} = c\left(1 + \frac{1-p_0}{p_0}\frac{e^{\lambda A_T}}{\lambda}\right)$$

induces effort path $\{a_t\}$ up to time T.
In the limit $r \to 0$, the bonus payment is constant over time $\dot{b}_{2,t} = 0$ and pinned down by the static moral hazard constraint at the deadline.

Proof of Proposition 1. First, consider the second-best second-stage contract without a promise-keeping constraint. Recall that incentive compatibility requires $\dot{b}_{2,t}$ follows from Lemma 1. The principal wants to induce full effort which follows from Moroni (2017) in the second stage. The bonus payment can be integrated to

$$b_{2,t} = \frac{c}{p_{2,0}\lambda(\lambda - r)}\left(\lambda(p_{2,0} + e^{-r(T-t)+\lambda T}(1-p_{2,0})) - r(p_{2,0} + (1-p_{2,0})e^{-r(T-t)+\lambda T})\right).$$

Hence, the principal chooses a total effort A_T equal to calendar time T. The objective of the principal is

$$\max_T \int_0^T e^{-rt}e^{-\lambda A_t}p_0\lambda(\pi - b_{2,t})dt$$
$$\text{s. t. } b_{2,t} = \frac{c}{p_{2,0}\lambda(\lambda - r)}\left(\lambda(p_{2,0} + e^{-r(T-t)+\lambda T}(1-p_{2,0})) - r(p_{2,0} + (1-p_{2,0})e^{-r(T-t)+\lambda T})\right).$$

Applying the constraint and full effort in the objective delivers

$$\max_T \int_0^T e^{-rt}e^{-\lambda t}p_{2,0}\lambda\left(\pi - \frac{c}{p_{2,0}\lambda(\lambda - r)}\left(\lambda(p_{2,0} + e^{-r(T-t)+\lambda T}(1-p_{2,0})) \right.\right.$$
$$\left.\left. - r(p_{2,0} + (1-p_{2,0})e^{-r(T-t)+\lambda T})\right)\right) dt$$

This simplifies to

$$\max_T \frac{1}{r-\lambda}\left(-c(1-p_{2,0})((1+e^{(\lambda-r)T})) + \frac{(\pi\lambda - c)p_{2,0}(r-\lambda)(1-e^{-(r+\lambda)T})}{r+\lambda}\right)$$

and taking the first-order condition delivers

$$T = \frac{1}{2\lambda}\ln\left(\frac{p_{2,0}}{1-p_{2,0}}\frac{\pi - c}{c}\right).$$

Consider the case of the promise-keeping constraint. The promise-keeping constraint is given by

$$v(\tau_1, T_2(\tau_1)) \geq v(\tau_1)$$

where $v(\tau_1)$ is the value promised to the agent. Note that the agent's second-stage value can be written as

$$\frac{ce^{-rT}(1-p_{2,0})}{r(r-\lambda)}\left(r(e^{\lambda T}-1)+\lambda(1-e^{rT})\right)$$

Again Lemma 1 pins down the incentive compatibility condition. With promise-keeping constraint the principal maximizes

$$\text{(OBJ)} \qquad \max_{T_2, b_{2,t}} \int_0^{T_2} e^{-rt} e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_s ds} p_{2,t}(A_{1,\tau_1})\lambda a_t(\pi - b_{2,t})dt.$$

$$\text{(IC)} \qquad s.t. \quad a_t \in \arg\max_{a_t\in\{0,1\}} \int_0^{T} e^{-rt} e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_s ds} a_t(p_{2,t}(A_{1,\tau_1})\lambda b_{2,t} - c)dt$$

$$\text{(PK)} \qquad \max_{a_t\in[0,1]} \int_0^{T_2} e^{-rt} e^{-\int_0^t p_{2,s}(A_{1,\tau_1})\lambda a_s ds} a_t(p_{2,t}(A_{1,\tau_1})\lambda b_{2,t} - c)dt \geq v(\tau_1).$$

Applying Lemma 1 and assuming frontloading of effort, this reduces to

$$\max_{T_2(\tau_1)} \frac{1}{r-\lambda}\left(-c(1-p_{2,0})((1+e^{(\lambda-r)T_2(\tau_1)})) + \frac{(\pi\lambda - c)p_{2,0}(r-\lambda)(1-e^{-(r+\lambda)T_2(\tau_1)})}{r+\lambda}\right)$$

$$s.t. \quad v(\tau_1, T_2(\tau_1)) = \frac{ce^{-rT_2(\tau_1)}(1-p_{2,0})}{r(r-\lambda)}\left(r(e^{\lambda T_2(\tau_1)}-1)+\lambda(1-e^{rT_2(\tau_1)})\right) \geq v(\tau_1).$$

Which yields as Lagrangean

$$\mathcal{L} = \frac{1}{r-\lambda}\left(-c(1-p_{2,0})((1+e^{(\lambda-r)T_2(\tau_1)})) + \frac{(\pi\lambda - c)p_{2,0}(r-\lambda)(1-e^{-(r+\lambda)T_2(\tau_1)})}{r+\lambda}\right)$$
$$- \mu\left(\frac{ce^{-rT_2(\tau_1)}(1-p_{2,0})}{r(r-\lambda)}\left(r(e^{\lambda T_2(\tau_1)}-1)+\lambda(1-e^{rT_2(\tau_1)})\right) - v(\tau_1)\right)$$

that can be solved for $T_2(A_{1,\tau_1})$ by the Kuhn-Tucker Theorem. This is solved by the second-best bonus, whenever $v(\tau_1, T_2(\tau_1)) > v(\tau_1)$ as then $\mu = 0$ and we are in the case of the second best. If the constraint is binding, we require $T_2(\tau_1)$ to be chosen such that

$$\frac{ce^{-rT_2(\tau_1)}(1-p_{2,0})}{r(r-\lambda)}\left(r(e^{\lambda T_2(\tau_1)}-1)+\lambda(1-e^{rT_2(\tau_1)})\right) = v(\tau_1).$$

If $r \to 0$, this is solved by

$$T_2(\tau_1, v(\tau_1)) = -\frac{v(\tau_1)}{c(1-p_{2,0}(\tau_1))} - \frac{1}{\lambda}\left(1 + W_{-1}\left(-e^{-1-\lambda\frac{v(\tau_1)}{c(1-p_{2,0}(\tau_1))}}\right)\right)$$

where W_{-1} denotes the negative branch of the Lambert-W-function.

Proof of Proposition 2. The agent's value, $w_t \equiv b_{1,t} + v_t$, consists of a promised utility from the second stage, v_t, and a bonus payment after the first success, $b_{1,t}$. Note that the continuation value of the agent depends on the agent's private information. The principal promises the continuation value conditional on the expected exerted effort, $\hat{A}_{1,t}$. Conditional on this, she implements a bonus and a deadline in the second stage. However, the true total effort that the agent has exerted is private information, $A_{1,t} = \int_0^t a_s ds$. The promised utility is given by $v(t)$ and implemented through the continuation contract in Proposition 1. The value from the agent's view is given by $v(t, A_{1,t})$ and depends on the true effort because she may hold a different belief than the principal in the second stage. The Hamiltonian of the agent is given by

(7)
$$\mathcal{H} = a_t e^{-rt}\left((p_0\lambda^g e^{-A_t\lambda^g} + (1-p_0)\lambda^b e^{-\lambda^b A_t})(b_{1,t} + v(t, A_t)) - (p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t})c\right) + \eta_t a_t.$$

He will exert effort if

(8)
$$\gamma_t \equiv e^{-rt}\left((p_0\lambda^g e^{-A_t\lambda^g} + (1-p_0)\lambda^b e^{-\lambda^b A_t})(b_{1,t} + v(t, A_t)) - (p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t})c\right) + \eta_t \geq 0$$

and the cheapest way to do so is $\gamma_t = 0$. Hence, if $\gamma_t = 0$

(9)
$$\eta_t = e^{-rt}\left((p_0\lambda^g e^{-A_t\lambda^g} + (1-p_0)\lambda^b e^{-\lambda^b})(b_{1,t} + v(t, A_t)) - (p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t})c\right).$$

The boundary condition is given by $\eta_T = 0$ which yields

$$(b_T + v(T, A_T)) = c\frac{p_0 e^{-A_T\lambda^g} + (1-p_0)e^{-A_T\lambda^b}}{p_0\lambda^g e^{-A_T\lambda^g} + (1-p_0)\lambda^b e^{-A_T\lambda^b}}. \tag{10}$$

The costate evolution is given by $\dot{\eta}_t = -\partial_{A_t}\mathcal{H}$:

(11)
$$\dot{\eta}_t = e^{-rt}\left(a_t(p_0\lambda^{g^2} e^{-A_t\lambda^g} + (1-p_0)\lambda^{b^2} e^{-A_t\lambda^b})(b_{1,t} + v(t, A_t)) - (p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})\frac{\partial v(t, A_t)}{\partial A_t}\right)$$
$$- e^{-rt} a_t(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})c.$$

Differentiating (9) with respect to time delivers

$$\dot{\eta}_t = re^{-rt}\Big((p_0\lambda^g e^{-A_t\lambda^g} + (1-p_0)\lambda^b e^{-\lambda^b})v(t,A_t) - (p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t})c\Big)$$
$$(12) \qquad + a_t((p_0{\lambda^g}^2 e^{-A_t\lambda^g} + (1-p_0){\lambda^b}^2 e^{-A_t\lambda^b})v(t,A_t) - (p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})c$$
$$- (p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})\Big(\dot{b}_t + \dot{v}(t,A_t)\Big))$$

or with γ_t not fixed

$$\dot{\eta}_t = \dot{\gamma}_t + re^{-rt}\Big((p_0\lambda^g e^{-A_t\lambda^g} + (1-p_0)\lambda^b e^{-\lambda^b})v(t,A_t) - (p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t})c\Big)$$
$$(13) \qquad + a_t((p_0{\lambda^g}^2 e^{-A_t\lambda^g} + (1-p_0){\lambda^b}^2 e^{-A_t\lambda^b})v(t,A_t) - (p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})c$$
$$- (p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t})\Big(\dot{b}_t + \dot{v}(t,A_t)\Big))$$

or (8) Equating this with (11) yields

$$(14) \qquad \dot{b}_t + \dot{v}(t,A_t) = r\left(b_{1,t} + v(t,A_t) - \frac{p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t}}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}c\right) + \frac{\partial v(t,A_t)}{\partial A_t}a_t + \frac{\dot{\gamma}_t}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}.$$

Hence, if the promised utility follows (40) together with the boundary equation (10), the principal can induce the effort in the interval $[0,T]$.
If $r \to 0$, this reduces to

$$(15) \qquad \dot{b}_t + \dot{v}(t,A_t) = \frac{\partial v(t,A_t)}{\partial A_t}a_t + \frac{\dot{\gamma}_t}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}.$$

Proof of Lemma 2. Alternatively, the principal could provide the agent with utility by varying the boundary condition, B, for the bonus payment and differing the deadline. The payment rule, however, still has to satisfy incentive-compatibility. The bonus payment for each t under an alternative deadline is given by

$$b(t,B) = Be^{-r(T-t)} - c\left(1 - e^{-r(T-t)} + \frac{r}{\lambda - r}\frac{1-p_{2,0}}{p_{2,0}}\left(e^{\lambda t} - e^{-r(T-t)+\lambda T}\right)\right)$$

This determines the agent's value if, $p_{2,0}(\hat{A}_t)$ is the principal's and $p_{2,0}(A_t)$ the agent's belief, to be given by

(16)

$$v(T,B,p_{2,0}(A_t),p_{2,0}(\hat{A}_t))$$
$$= \frac{1}{p_{2,0}(\hat{A}_t)r(r-\lambda)(r+\lambda)}$$
$$\cdot\Big((\lambda+1)(\lambda-r)rp_{2,0}(\hat{A}_t)p_{2,0}(A_t)+e^{-rT}p_{2,0}(A_t)r(ce^{\lambda T}(p_{2,0}(\hat{A}_t)-1)r+(B+c)p_{2,0}(\hat{A}_t)(r-\lambda))(\lambda+r)$$
$$+(\lambda+r)(c(p_{2,0}(A_t)-1)p_{2,0}(\hat{A}_t)r)+c(p_{2,0}(A_t)r-p_{2,0}(\hat{A}_t)(-1+p_{2,0}(A_t)+rp_{2,0}(A_t)))\lambda$$
$$-(r-\lambda)e^{-(r+\lambda)T}(ce^{\lambda T}((p_{2,0}(A_t)-1)p_{2,0}(\hat{A}_t)+p_{2,0}(A_t)(p_{2,0}(\hat{A}_t)-1)r)(r+\lambda))$$
$$-(r-\lambda)e^{-(r+\lambda)T}(p_{2,0}(A_t)p_{2,0}(\hat{A}_t)r(c(r-1)+B(\lambda+r)))\Big).$$

This delivers as deviation incentive in the first stage

(17)

$$\frac{\partial v(T,B,p_{2,0}(A_t),p_{2,0}(\hat{A}_t))}{\partial A_t}$$
$$= \frac{\partial p_{2,0}(A_t)}{\partial A_t}\frac{1}{p_{2,0}(\hat{A}_t)r(r-\lambda)(r+\lambda)}$$
$$\Big((1+\lambda)cp_{2,0}(\hat{A}_t)r(\lambda-r)+re^{-rT}(ce^{\lambda T}(p_{2,0}(\hat{A}_t)-1)r+(B+c)p_{2,0}(\hat{A}_t)(r-\lambda))(r+\lambda)$$
$$(r+\lambda)(cp_{2,0}(\hat{A}_t)r+c(r-p_{2,0}(\hat{A}_t)(1+r))\lambda)$$
$$(r-\lambda)e^{-(r+\lambda)T}(r-\lambda)(cp_{2,0}(\hat{A}_t)(r-1)r+Bp_{2,0}(\hat{A}_t)r(r+\lambda)+ce^{\lambda T}(p_{2,0}(\hat{A}_t)+(p_{2,0}(\hat{A}_t)-1)r)(r+\lambda))\Big).$$

To see how this varies with the composition of the continuation contract note that it follows from the implicit function theorem applied on the promise-keeping condition that the deadline and the bonus payment vary according to

$$(18)\quad \frac{dT}{dB} = \frac{(1-e^{\lambda T})p_{2,0}(\hat{A}_t)}{(B-c)p_{2,0}(\hat{A}_t)(r+\lambda)+e^{\lambda T}(-c(1-p_{2,0}(\hat{A}_t))(\lambda+re^{\lambda T})+r(p_{2,0}(\hat{A}_t)B-c))} < 0.$$

That is, as intuitive, if the boundary condition increases, the deadline decreases. Considering now the effect of the deadline on the deviation incentive I find that this is affected as follows

(19)

$$\frac{\partial p_{2,0}(A_t)}{\partial A_t}\frac{e^{-(r+\lambda)T}}{p_{2,0}(\hat{A}_t)}\Big(re^{2\lambda T}(1-p_{2,0}(\hat{A}_t))c+e^{\lambda T}(cp_{2,0}(\hat{A}_t)-(c+Bp_{2,0}(\hat{A}_t))r)+p_{2,0}(\hat{A}_t)(c(r-1)+B(\lambda+r))\Big)$$

which is positive. Hence, by extending the deadline, the principal reduces the incentive to deviate in the first stage.

Evolution of Agent's Continuation Utility. The agent's utility from the second stage may evolve different than the $v(t)$ because it depends on his private information about the true effort. The promised utility to the agent is denoted by $v(t)$ which coincides with the agent's continuation utility if he is on path, i.e., if $A_{1,t} = t$. However, if the agent has deviated, $v(t) \neq v(t, A_{1,t})$ where the latter denotes the agent's continuation utility given his private information $A_{1,t}$. We know that the on-path utility evolves according to

$$\dot{w}(t, A_{1,t}) = r\left(b_{1,t} + v(t, A_{1,t}) - \frac{p_0 e^{-\lambda^g A_{1,t}} + (1-p_0)e^{-\lambda^b A_{1,t}}}{p_0\lambda^g e^{-\lambda^g A_{1,t}} + (1-p_0)\lambda^b e^{-\lambda^b A_{1,t}}}c\right) + \frac{\partial v(t, A_{1,t})}{\partial A_{1,t}}a_t. \tag{20}$$

For an agent that has exerted effort $\hat{A}_{1,t}$, the value of succeeding the first stage at t is given by (net the bonus payment, $b_{1,t}$)

$$\begin{aligned} v(\hat{A}_{1,t}, A_{1,t}) = & \frac{ce^{-rT}}{r(r-\lambda)p_{2,0}(A_{1,t})} \\ & \cdot \Big(p_{2,0}(\hat{A}_{1,t})(re^{rT} - \lambda p_{2,0}(A_{1,t})e^{rT} - r(1-p_{2,0}(A_{1,t}))e^{\lambda T}) \\ & + (1-p_{2,0}(\hat{A}_{1,t}))p_{2,0}(A_{1,t})(r-\lambda) + p_{2,0}(A_{1,t})e^{rT}(\lambda - r)\Big) \end{aligned}$$

where $p_{2,0}(A_{1,t})$ is the principal's belief at the beginning of the second stage and $p_{2,0}(\hat{A}_{1,t})$ is the agent's belief. This can be simplified substantially to

$$v(\hat{A}_{1,t}, A_{1,t}) = v(t)\frac{p_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})} - cr(1 - e^{rT_2(A_{1,t})})\frac{p_{2,0}(\hat{A}_{1,t}) - p_{2,0}(A_{1,t})}{p_{2,0}(A_{1,t})}.$$

The total value is evolving according to

$$\dot{w}(\hat{A}_{1,t}, A_{1,t}) = \dot{b}_t + \dot{v}(\hat{A}_{1,t}, A_{1,t}). \tag{21}$$

Hence, if $r \to 0$

$$\dot{v}(\hat{A}_{1,t}, A_{1,t}) = \dot{v}(t)\frac{p_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})} + v(t)\left(\frac{\dot{p}_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})} - \frac{p_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})}\frac{\dot{p}_{2,0}(A_{1,t})}{p_{2,0}(A_{1,t})}\right) \tag{22}$$

Agent's Problem. Given this law of the agent's continuation value, the agent solves the following problem

$$\max_a \int_0^\infty e^{-\int_0^t e^{-rt}(p_s\lambda_1^g + (1-p_s)\lambda_1^b)a_s ds} a_t \left((p_t\lambda_1^g + (1-p_t)\lambda_1^b)v(t, A_{1,t}) - c\right)dt \tag{23}$$

$$s.t. \quad \dot{w}(t, A_{1,t}) = \dot{b}_t + \dot{v}(t, A_{1,t}) \tag{24}$$

$$\dot{w}(t) = r\left(b_{1,t} + v(t, A_{1,t}) - \frac{p_0 e^{-\lambda^g A_{1,t}} + (1-p_0)e^{-\lambda^b A_{1,t}}}{p_0\lambda^g e^{-\lambda^g A_{1,t}} + (1-p_0)\lambda^b e^{-\lambda^b A_{1,t}}}c\right) + \frac{\partial v(t, A_{1,t})}{\partial A_{1,t}}a_t \tag{25}$$

$$\dot{A}_{1,t} = a_t. \tag{26}$$

Existence of Solution to Agent's Problem Existence follows from Clarke (2013), Theorem 23.11. The theorem applies as:

- the laws of motion of the state variables, A_t, $w(t, A_t)$ and $w(t)$ are measurable in t and continuous in A_t
- the control set $a_t \in [0, 1]$ is closed and convex
- the running cost is
  - Lesbesgue measurable in t and (A, a)
  - lower semicontinuous in (A, a)
  - convex in a for any (t, A)
- the effort path $a_t = 0$ for all t and $A_t = 0$ for all t is admissible and delivers a finite value.

Principal's Problem. The Hamiltonian of the principal's problem is given by

(27)
$$\mathcal{H} = e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}\right)(\Pi(T_2(t), t) - c(T_2, t) - w_t) + \gamma_t a_t$$

(28)
$$+\,\eta_t\left(r\left(b_t + v(t, A_t) - \frac{p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t}}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}c\right) + \frac{\partial v(t, A_t)}{\partial A_t}a_t\right) - \zeta_t(w_t - v(T_2(t), t))$$

(29)
$$\dot{A}_t = a_t$$

(30)
$$\dot{w}_t = r\left(b_t + v(t, A_t) - \frac{p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t}}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}c\right) + \frac{\partial v(t, A_t)}{\partial A_t}a_t$$

(31)
$$\dot{\gamma}_t = e^{-rt}\left(p_0{\lambda^g}^2 e^{-\lambda^g A_t} + (1-p_0){\lambda^b}^2 e^{-\lambda^b A_t}\right)(\Pi(T_2(t), t) - c(T_2, t) - w_t) - \eta_t\frac{\partial \dot{w}_t}{\partial A_t}$$

(32)
$$\dot{\eta}_t = \left(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}\right) - \eta_t r + \zeta_t$$

with constraint from agent's problem $w_{T_1} = c\frac{p_0\lambda^g e^{-\lambda^g A_{T_1}} + (1-p_0)\lambda^b e^{-\lambda^b A_{T_1}}}{p_0 e^{-\lambda^g A_{T_1}} + (1-p_0)e^{-\lambda^b A_{T_1}}}$ with associated multiplier μ and moreover boundary conditions $\gamma_T = 0, \eta_0 = 0, \eta_T = \mu, A_0 = 0$. Note that $\Pi(T_2(t), t)$ is the total expected profit from choosing deadline $T_2(t)$ after a success at t, i.e.,

$\int_0^{T_2(t)} p_{2,0}(A_t)e^{-\lambda A_s}\pi ds$ and $c(T_2(t), t)$ is the total expected experimentation cost from choosing deadline $T_2(t)$ after a success at t, i.e., $\int_0^{T_2(t)} p_{2,0}(A_t)e^{-\lambda A_s} c ds$.

Maximization with respect to a_t. To see that the principal wants to implement full effort consider a dynamic programming heuristic

$$
\begin{aligned}
&\left(-\frac{\partial \Pi_t}{\partial a_t} + \frac{\partial \Pi_t}{\partial a_{t+dt}}\right) \\
&= dt\left(\left(p_t\lambda^g + (1-p_t)\lambda^b\right)\left(-\Pi(A_t) - a_t dt \Pi'(A_t) + \Pi(A_{t+dt})\right)\right) \\
&+ dt^2\left(\Pi_{t+2dt}\left(\underbrace{(a_t - a_{t+dt})}_{=0,\text{ if } a_t \text{ continuous}}\ ((1-p_t)^2\lambda^{b^2} + 2(1-p_t)p_t\lambda^b\lambda^g + p_t\lambda^{g^2}) + (1-p_t)\lambda^{b^2} + p_t\lambda^{g^2}\right)\right) \\
&- dt^2\left(c\underbrace{\left(((1-p_t)\lambda^b + p_t\lambda^g) - ((1-p_t)\lambda^b + p_t\lambda^g)\right)}_{=0}\right) \\
&+ dt^2\left(\underbrace{\frac{\partial \Pi(A_{t+dt})}{\partial a_t} - \frac{\partial \Pi(A_{t+dt})}{\partial a_{t+dt}}}_{=0}\right) a_{t+dt}\left(a_t((1-p_t)\lambda^b + p_t\lambda^g) + \frac{1}{2}(1-p_t)^2\lambda^{b^2} p_t(1-p_t)\lambda^g\lambda^b + \frac{1}{2}p_t^2\lambda^{g^2}\right) \\
&- dt^2\ \underbrace{(a_t - a_{t+dt})}_{=0 \text{ if } a_t \text{ continuous}}\ \left((1-p_t)\lambda^{b^2} + p_t\lambda^{g^2}\right)\Pi(A_{t+dt}) \\
&+ dt^2\left((1-p_t)^2\lambda^{b^2} + 2(1-p_t)p_t\lambda^b\lambda^g + p_t\lambda^{g^2}) + (1-p_t)\lambda^{b^2} + p_t\lambda^{g^2}\right) a_t \\
&\left(\Pi(A_t) + \frac{1}{2}a_t dt \Pi'(a_t) - \underbrace{\frac{a_{t+dt}}{a_t}}_{=1 \text{ if } a_t \text{ continuous}}\ \Pi(A_{t+dt})\right) \\
&- dt^2\left(r\left((p_t\lambda^g + (1-p_t)\lambda^b)\Pi(A_{t+dt}) - c + a_{t+dt}(p_t\lambda^g + (1-p_t)\lambda^b)\underbrace{\left(\frac{\partial \Pi(A_{t+dt})}{\partial a_{t+dt}} - \frac{\partial \Pi(A_{t+dt})}{\partial a_{t+}}\right)}_{=0}\right)\right) \\
&+ o(dt^3)
\end{aligned}
$$

Assuming a_t to be continuous this simplifies after division by dt^2 to

$$\begin{aligned}
&\left(-\frac{\partial \Pi_t}{\partial a_t}+\frac{\partial \Pi_t}{\partial a_{t+dt}}\right)/dt^2 \\
&=\left((p_t\lambda^g+(1-p_t)\lambda^b)\left(\frac{\Pi(A_{t+dt})-\Pi(A_t)}{dt}-a_t\Pi'(A_t)\right)\right) \\
&-r\left((p_t\lambda^g+(1-p_t)\lambda^b)\Pi(A_{t+dt})-c\right) \\
&+\left((1-p_t)^2\lambda^{b^2}+2(1-p_t)p_t\lambda^b\lambda^g+p_t\lambda^{g^2})+(1-p_t)\lambda^{b^2}+p_t\lambda^{g^2}\right)a_t \\
&\left(\Pi(A_t)+\frac{1}{2}a_t dt\Pi'(a_t)-\Pi(A_{t+dt})\right).
\end{aligned}$$

Take the limit as $dt \to 0$ and get

$$\begin{aligned}
&\left(-\frac{\partial \Pi_t}{\partial a_t}+\frac{\partial \Pi_t}{\partial a_{t+dt}}\right)/dt^2 \\
&=\left((p_t\lambda^g+(1-p_t)\lambda^b)\left(\dot{\Pi}(A_t)-a_t\Pi'(A_t)\right)\right) \\
&-r\left((p_t\lambda^g+(1-p_t)\lambda^b)\Pi(A_{t+dt})-c\right)
\end{aligned}$$

Note that $\dot{\Pi}(A_t)-a_t\Pi'(A_t)$ is zero up to the second order and therefore the expression is negative as long as t is less than the first-best. Thus, welfare increases if effort is frontloaded. Moreover, note that the expected payment to the agent is decreasing if effort is frontloaded. If discounting dominates learning in a way that the expected payment to the agent is increasing for some measure of time an argument similar to the one in Moroni (2017) delivers optimality of frontloading. Hence, the principal prefers frontloading of effort, that ist $a_t = 1$ for $t \in [0, T_1]$.

Maximization with respect to $T_2(t)$. Note that whenever $w_t - v(T_2(t), t) > 0$ a bonus is paid to the agent and $\zeta_t = 0$. ζ_t follows from the first-order condition with respect to time as

$$\zeta_t = -\frac{\partial \mathcal{H}}{\partial T_2(t)} \Big/ \frac{\partial v(T_2(t), t)}{\partial T_2(t)} \tag{33}$$

$$= -\frac{e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t}+(1-p_0)\lambda^b e^{-\lambda^b A_t}\right)\left(\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)}-\frac{\partial c(T_2(t),t)}{\partial T_2(t)}\right)+\eta_t\frac{\partial \dot{w}_t(A_t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}}. \tag{34}$$

Using this in (32) delivers

(35)

$$\dot{\eta}_t = e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t}+(1-p_0)\lambda^b e^{-\lambda^b A_t}\right)\left(1-\frac{\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)}-\frac{\partial c(T_2(t),t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}}\right)+\eta_t\left(\frac{\frac{\partial \dot{w}_t(A_t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}}-r\right)$$

which is a differential equation of η_t that can be integrated to

(36)

$$\eta_t = -e^{-\int_0^t r+\frac{\partial \dot{w}_s}{\partial T_2(s)}ds} \int_0^t e^{\int_0^s \frac{\partial \dot{w}_\tau}{\partial T_2(\tau)}d\tau} \left(p_0\lambda^g e^{-\lambda^g A_s} + (1-p_0)\lambda^b e^{-\lambda^b A_s}\right) \frac{\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)} - \frac{\partial v_t(A_t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}} ds$$

using the boundary condition $\eta_0 = 0$. Note that η_t is increasing over time as $T_2(t) > T_2^{SB}(t)$ by optimality and costly incentives. Continuity of η_t implies that there is a $\hat{t}$ such that $\zeta_t > 0$ for all $t \in [0, \hat{t}]$ because if $\zeta_t = 0$, we get from (33)

$$e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}\right)\left(\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)}\right) = \eta_t \frac{\partial \dot{w}_t(A_t)}{\partial T_2(t)} \tag{37}$$

which cannot be satisfied in a neighborhood of $t = 0$ if the principal is optimizing. To see why, note that the left-hand side is the first-order condition for the social planner. To induce effort in that solution, the bonus has to be equal to 1 and the principal makes zero profits. It is easy to construct a contract that induces positive profit for the principal. Thus, we have a contradiction and $\zeta_t > 0$ for $t \in [0, \hat{t}]$. Moreover, the left-hand side is decreasing

Moreover, it can be seen that as soon as a positive bonus payment is used, that is, when $\zeta_t = 0$, the bonus of the reward that is given to the agent with a bonus payment is increasing over time because η_t is increasing over time.

$$\eta_t = \frac{e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}\right)\left(\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)}\right)}{\frac{\partial \dot{w}_t(A_t)}{\partial T_2(t)}}. \tag{38}$$

Because η_t measures the marginal cost of changing the state v_t (note that only changing v_t affects the evolution of w_t) it follows that this is increasing over time and hence, for any level w_t of current promised value of succeeding, the principal provides more of this utility through a bonus payment $w_t - v_t(T_2)$.

To see that $\hat{t} < T_1$, equate the condition for $\zeta_{\hat{t}} = 0$ with the equation for η_t which yields

(39)

$$- e^{-\int_0^t r+\frac{\partial \dot{w}_s}{\partial T_2(s)}ds} \int_0^t e^{\int_0^s \frac{\partial \dot{w}_\tau}{\partial T_2(\tau)}d\tau} \left(p_0\lambda^g e^{-\lambda^g A_s} + (1-p_0)\lambda^b e^{-\lambda^b A_s}\right) \frac{\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)} - \frac{\partial v_t(A_t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}} ds$$
$$= \frac{e^{-rt}\left(p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}\right)\left(\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)}\right)}{\frac{\partial \dot{w}_t(A_t)}{\partial T_2(t)}}$$

Note that this is equivalent to a first-order condition for the first-stage deadline if $\frac{\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)} - \frac{\partial v_t(A_t)}{\partial T_2(t)}}{\frac{\partial v_t(A_t)}{\partial T_2(t)}}$ is replaced by $-\left(\frac{\partial \Pi(T_2(t),t)}{\partial T_2(t)} - \frac{\partial c(T_2(t),t)}{\partial T_2(t)}\right)$ and bonus transfers could not be used. However, the latter can be shown to be greater than the former and as the other term is increasing in t, $\hat{t}$ less than the deadline if no bonus payment is used. By giving the principal the additional possibility of a bonus payment she is at least weakly better off and hence, $\hat{t} < T_1$.

Sufficiency of the Necessary Conditions Because we have established the existence of a solution previously, we can conclude that if $\{a_t\}$ is the only effort path that satisfies the necessary conditions, these are also sufficient. Recall that necessity requires an effort path $\{\hat{a}_t\}$ together with a costate $\hat{\gamma}_t$ such that they satisfy (8). Recall that we have from (15)

$$\begin{aligned}\frac{\dot{\gamma}_t}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}} &= -r\left(b_t + v(t, A_t) - \frac{p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t}}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}}c\right) \\ &\quad + \dot{b}_t + \dot{v}(t, A_t) - \frac{\partial v(t, A_t)}{\partial A_t}a_t\end{aligned} \tag{40}$$

which we can, using the on-path and off-path values, rewrite as

$$\begin{aligned}\frac{\dot{\hat{\gamma}}_t}{p_0\lambda^g e^{-\lambda^g \hat{A}_t} + (1-p_0)\lambda^b e^{-\lambda^b \hat{A}_t}} &= r\left(\frac{p_0 e^{-\lambda^g A_t} + (1-p_0)e^{-\lambda^b A_t}}{p_0\lambda^g e^{-\lambda^g A_t} + (1-p_0)\lambda^b e^{-\lambda^b A_t}} - \frac{p_0 e^{-\lambda^g \hat{A}_t} + (1-p_0)e^{-\lambda^b \hat{A}_t}}{p_0\lambda^g e^{-\lambda^g \hat{A}_t} + (1-p_0)\lambda^b e^{-\lambda^b \hat{A}_t}}\right)c \\ &\quad + \left(\dot{v}_t(\hat{A}_t) - \dot{v}_t(A_t)\right) + \frac{\partial v(t, A_t)}{\partial A_t}a_t - \frac{\partial v(t, \hat{A}_t)}{\partial \hat{A}_t}\hat{a}_t\end{aligned} \tag{41}$$

Define $\tau_0 \equiv \inf\{t | \hat{\gamma}_t \neq 0\}$. Suppose $\tau_0 = 0$ and $\hat{\gamma}_{\tau_0} > 0$. By continuity of $\hat{\gamma}_t$, there is an ε such that for $\hat{\gamma}_t < 0$ for $t \in (0, \varepsilon)$. By optimality, we know that $\hat{a}_t = 0$ for $t \in (0, \varepsilon)$. This implies that $\hat{A}_t \leq A_t$ where $\hat{A}_t$ corresponds to the effort of the hypothetical effort path $\{\hat{a}_t\}$ and A_t to the on-path effort path $\{a_t\}$. I want to show that $\dot{\hat{\gamma}}_t < 0$ if $\hat{A}_t \leq A_t$ for r close to zero. Recall (22), then, (41) further reduces to

$$\begin{aligned}\frac{\dot{\hat{\gamma}}_t}{p_0\lambda^g e^{-\lambda^g \hat{A}_t} + (1-p_0)\lambda^b e^{-\lambda^b \hat{A}_t}} &= \dot{v}(t)\left(\frac{p_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})} - 1\right) + v(t)\left(\frac{\dot{p}_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})} - \frac{p_{2,0}(\hat{A}_{1,t})}{p_{2,0}(A_{1,t})}\frac{\dot{p}_{2,0}(A_{1,t})}{p_{2,0}(A_{1,t})}\right) \\ &\quad + \frac{\partial v_t(A_t)}{\partial A_t}a_t - \frac{\partial v_t(\hat{A}_t)}{\partial \hat{A}_t}\hat{a}_t\end{aligned} \tag{42}$$

The right-hand side is now less than zero as $\hat{A}_{1,t} \leq A_{1,t}$ implies $p_{2,0}(\hat{A}_{1,t}) \geq p_{2,0}(A_{1,t})$. Hence, we know that $\dot{\hat{\gamma}}_t < 0$ on $t \in (0, \varepsilon)$ implying that $\hat{\gamma}_t < 0$. Together this implies that $\hat{\gamma}_t < 0$ for all $t \in [0, T_1]$. Recall that the transversality condition implies that

$$\begin{aligned}\hat{\gamma}_T &= \left(p_0\lambda^g e^{-\lambda^g \hat{A}_T} + (1-p_0)\lambda^b e^{-\lambda^b \hat{A}_T}\right)v(t, \hat{A}_t) - \left(p_0 e^{-\lambda^g \hat{A}_T} + (1-p_0)e^{-\lambda^b \hat{A}_T}\right)c \\ &\geq \\ \gamma_T &= \left(p_0\lambda^g e^{-\lambda^g A_T} + (1-p_0)\lambda^b e^{-\lambda^b A_T}\right)v(t, A_t) - \left(p_0 e^{-\lambda^g A_T} + (1-p_0)e^{-\lambda^b A_T}\right)c \\ &= 0\end{aligned} \tag{43}$$

where the inequality follows from $\hat{A}_T \leq A_T$. $\hat{\gamma}_T \geq 0$ contradicts $\hat{\gamma}_t < 0$ for all $t \leq T_1$. An analogous argument applies for $\tau_0 > 0$. For all $t < \tau_0$, $A_t = \hat{A}_t$ and $\hat{\gamma}_t = 0$. Following τ_0 with $\hat{A}_t < A_t$, the reasoning from above yields a contradiction with the transversality condition. Note that $a_t = 1$ will be optimal and this direction suffices to guarantee sufficiency of the necessary conditions in the optimal contract.

Costly Incentives. Note that to complete the solution of the optimal control problem, we need to prove that the principal always sets $\gamma_t = 0$ in the agent's problem. This implies, that the agent's incentive constraint is never slack in the optimal contract. I restrict attention to this case, as the main contribution of the paper lies in the case when first-stage incentives are relevant. In the remaining cas es, it occurs that theprincipal wants to increase the agent's value to get closer to the second-stage second-best value. To do that, she increases γ_t above zero. These cases can occur only if the agent's promised value lies below the second-best value in some regions. The problem can analogously solved for cases with $\gamma_t > 0$. Due to a lack of closed-form solutions there is no sharp characterization of the parametric assumptions for the costly incentives case. However, a sufficient condition on contract terms is that: $\dot{w}_t < \dot{v}^{SB}(t)$ for all $t \in [0, T_1]$ and $w_{T_1} \geq v^{SB}(T_1)$. This implies that at the first-stage deadline the value in the contract is higher than the second-best value of the contract. Moreover, the total reward is decreasing steeper than the value of the second-best contract. Hence, the total reward is always higher than the second-best second-stage contract.

Existence of Solution to Principal's Problem. Recall the requirements from Clarke (2013), Theorem 23.11. Denote the control variables by a and the state variables by A. The theorem applies as

- the laws of motion of the state variables, $g(A)$ are measurable in t and continuous in A
- the control set $a \in \mathcal{A}$ is closed and convex
- the running cost is
  - Lesbesgue measurable in t and (A, a)
  - lower semicontinuous in (A, a)
  - convex in a for any (t, A)
- the effort path $a_t = 0$ for all t and $A_t = 0$ for all t is admissible and delivers a finite value.

The running cost is given by

$$\Lambda(t, A, a) = a_t e^{-rt}(p_0 e^{-\lambda^g A_t} + (1 - p_0)e^{-\lambda^b A_t})\left(\Pi(t, v(t)) - v(t)\right). \tag{44}$$

Convexity in a of the set $\{\Lambda(t, A, \cdot)\}$ has to be established. It suffices to show that Λ is concave in a. Suppose we are in the case with $\gamma_t = 0$. Then, it remains to show that $\Pi(t, v(t))$ is concave in $v(t)$. Recall that

$$\Pi(t, v(t)) = \int_0^{T(v(t))} e^{-rt} p_{2,0}(t) e^{-\lambda s}(p_s \lambda \pi - c) ds \tag{45}$$

$$= \frac{p_{2,0}(t)(\pi\lambda - c)}{r+\lambda}\left(1 - e^{-(r+\lambda)T(v(t))}\right) - \frac{1 - p_{2,0}(t)c}{r}\left(1 - e^{-rT(w(t))}\right). \tag{46}$$

Hence, we get

$$\begin{aligned}\frac{d^2\Pi(t, v(t))}{dv(t)^2} &= \frac{d^2T(v(t))}{dv(t)^2}\left(p_{2,0}(t)(\pi\lambda - c)e^{-(r+\lambda)T(w(t))} - (1 - p_{2,0}(t))ce^{-rT(w(t))}\right)\\ &- \left(\frac{dT(w(t))}{dw(t)}\right)^2\left(p_{2,0}(t)(\pi\lambda - c)(r+\lambda)e^{-(r+\lambda)T(w(t))} - r(1 - p_{2,0}(t))ce^{-rT(w(t))}\right)\end{aligned}$$

which is less than zero. Hence, the running cost is concave in the promised utility and we can conclude that the set $\Lambda(t, A, \cdot)$ is convex.

Proof of Theorem 2. Suppose the principal considers introducing an additional deadline, T_1' before the initial one, T_1 to replace the agent if he succeeds before the second but after the first. To simplify notation denote by Π the surplus in the second stage. Introducing T_1' alters the principal's profits by

$$\begin{aligned}&\int_0^{T_1'} e^{-rt}(1-p_0)\frac{p_t}{1-p_t}\left(\left(\Pi(t, w(t, T_1', T_1)) - \Pi(t, w(t, T_1, T_1))\right) - \left(w(t, T_1', T_1) - w(t, T_1, T_1)\right)\right) dt\\ &+ \int_{T_1'}^{T_1} e^{-rt}(1-p_0)\frac{p_t}{1-p_t}\left(\left(\Pi^{SB}(t) - \Pi(t, w(t, T_1, T_1))\right) - \left(w^{SB}(t) - w(t, T_1, T_1, T_1)\right)\right)\\ &- \int_{T_1'}^{T_1} e^{-rt}(1-p_0)\frac{p_t}{1-p_t}\frac{c}{p_{T_1}\lambda^g + (1-p_{T_1})\lambda^b}dt.\end{aligned}$$

Multiplying by $\frac{1}{T_1 - T_1'}$ and taking the limit $T_1' - T_1 \uparrow 0$ we have

$$\int_0^{T_1} e^{-rt}(1-p_0)\frac{p_t}{1-p_t}\frac{\partial w(t, T_1, T_1', T_1)}{\partial T_1'}\left(1 - \frac{\partial \Pi(t)}{\partial w(t, T_1)}\right) dt > 0 \tag{47}$$

where the inequality follows from $\frac{\partial \Pi(t)}{\partial w(t, T_1, T_1')} < 1$ if the agent's value is above the second-best value and $\frac{\partial w(t, T_1, T_1')}{\partial T_1'} > 0$. So there is an incentive to introduce entrepreneur replacement. Note that if stages are independent $\frac{\partial w(t, T_1, T_1'}{\partial T_1'}|_{T_1'=T_1} = 0$.
On the other hand, suppose $T_1' = 0$ and consider the choice of introducing a reward with continuation contracts instead of replacement, that is marginally increasing T_1'. This yields as

gain

(48)
$$-e^{-rT_1'}(1-p_0)\frac{p_{T_1'}}{1-p_{T_1'}}\left(\Pi^{SB}(T_1')-w^{SB}(T_1')-\left(\Pi(T_1',w(T_1',T_1))-w(T_1',T_1)\right)-\frac{c}{p_{T_1}\lambda^g+(1-p_{T_1})\lambda^b}\right).$$

Note that if $T_1'=0$, $w(0,T_1)=\frac{c}{p_{T_1}\lambda^g+(1-p_{T_1})\lambda^b}$ and hence, we get

$$-(1-p_0)\frac{p_{T_1'}}{1-p_{T_1'}}\left(\Pi^{SB}(T_1')-w^{SB}(T_1')-\Pi(T_1',w(T_1',T_1))\right). \tag{49}$$

If $w^{SB}(T_1')<\frac{c}{p_{T_1}\lambda^g+(1-p_{T_1})\lambda^b}$, then $\Pi(T_1',w(T_1',T_1))>\Pi^{SB}(T_1')$ and increasing T_1' is profitable. However, if $w^{SB}(T_1')>\frac{c}{p_{T_1}\lambda^g+(1-p_{T_1})\lambda^b}$, then $\Pi(T_1',w(T_1',T_1))<\Pi^{SB}(T_1')$, however, as the principal's payoff falls in the parameters that would increase the agent's second-best value, $\Pi^{SB}(T_1')-w^{SB}(T_1')<\Pi(T_1',w(T_1',T_1))$ and hence, a period without replacement would be introduced.

Proof of Lemma 3. To show that the optimal two-stage contract converges to the optimal one-stage contract I consider first the limits of the second-stage contract's instruments, $T(v(\tau_1))$ and $b^2(T(v(\tau_1))$. Recall that

$$T(v(\tau_1))=-\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))}-\frac{1}{\lambda}W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right) \tag{50}$$

We are now interested in the limit as $\lambda\to\infty$.

$$\lim_{\lambda\to\infty}T(v(\tau_1))=\lim_{\lambda\to\infty}\left(-\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))}-\frac{1}{\lambda}W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right)\right) \tag{51}$$

$$=-\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))}-\lim_{\lambda\to\infty}\frac{1}{\lambda}W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right) \tag{52}$$

Applying L'Hôspital's Rule gives

$$= -\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))} - \lim_{\lambda\to\infty}\left(W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right)\right)' \tag{53}$$

$$= -\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))} - \lim_{\lambda\to\infty}\left(-\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))c}\left(1-\frac{1}{1+W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right)}\right)\right) \tag{54}$$

$$= -\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))} - \left(-\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))c}\left(1-\frac{1}{1+\lim_{\lambda\to\infty}W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right)}\right)\right) \tag{55}$$

$$\lim_{\lambda\to\infty} T(v(\tau_1)) = 0 \tag{56}$$

as $W_{-1}(x\uparrow 0)=\infty$.
Recall that the bonus is given by $b^2(v(\tau_1)) = \frac{c}{\pi\lambda}\left(1+\frac{1-p_{2,0}(\tau_1)}{p_{2,0}(\tau_1)}e^{\lambda T(v(\tau_1))}\right)$. We get for the limit

$$\lim_{\lambda\to\infty} b^2(v(\tau_1)) = \lim_{\lambda\to\infty}\frac{c}{\pi\lambda}\left(1+\frac{1-p_{2,0}(\tau_1)}{p_{2,0}(\tau_1)}e^{\lambda T(v(\tau_1))}\right) \tag{57}$$

$$= \frac{c}{\pi}\frac{1-p_{2,0}(\tau_1)}{p_{2,0}(\tau_1)}\lim_{\lambda\to\infty}\frac{e^{\lambda T(v(\tau_1))}}{\lambda} \tag{58}$$

and again by L'Hôspital's rule

$$= \frac{c}{\pi}\frac{1-p_{2,0}(\tau_1)}{p_{2,0}(\tau_1)}\frac{v(\tau_1)}{(1-p_{2,0}(\tau_1))c}\lim_{\lambda\to\infty}\left(1-\frac{1}{1+\lim_{\lambda\to\infty}W_{-1}\left(-e^{-\frac{v(\tau_1)\lambda}{(1-p_{2,0}(\tau_1))}}\right)}\right) \tag{59}$$

$$\lim_{\lambda\to\infty} b^2(v(\tau_1)) = \frac{v(\tau_1)}{\pi p_{2,0}(\tau_1)}. \tag{60}$$

Next, note that we have $\lambda^b \to 0$ and hence

$$\lim_{\lambda^b\to 0} p_{2,0}(\tau_1) = \frac{\lambda^g p_{\tau_1}}{\lambda^g p_{\tau_1} + (1-p_{\tau_1})\lambda^b} = 1. \tag{61}$$

So, that the second stage is immediately and successfully completed in the limit and the agent receives $v(\tau_1)$ as a payment while the principal keeps $\pi - v(\tau_1)$.
As a consequence of the limit of the belief after a success, we have that

$$\lim_{\lambda^b\to 0}\dot{v}(t) = 0 \tag{62}$$

and no informativeness rent is required in the first stage. Hence, we may conclude that the optimal two-stage contract converges to the optimal one-stage contract if $\lambda^b \to 0$.

References

Aghion, P., and J. Tirole (1994): “The Management of Innovation,” *The Quarterly Journal of Economics*, pp. 1185–1209.

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

——— (2006): “The Financing of Innovation: Learning and Stopping,” *RAND Journal of Economics*, 36, 719–752.

Bhaskar, V. (2014): “The Ratchet Effect Re-examined: A Learning Perspective,” (9956).

Bienz, C., and J. Hirsch (2011): “The Dynamics of Venture Capital Contracts,” *Review of Finance*.

Bolton, P., and D. S. Scharfstein (1990): “A Theory of Predation Based on Agency Problems in Financial Contracting,” *The American Economic Review*, 80(1), 93–106.

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

——— (2016): “Career Concerns with Exponential Learning,” *Working Paper*.

Booth, G. G., O. M. Dalgic, and A. Young (2004): “The Staging of Venture Equity Capital and Venture Capitalist Bargaining Power,” *The Journal of Entrepreneurial Finance*, 9(3), 29–40.

Che, Y.-K., E. Iossa, and P. Rey (2016): “Prizes versus Contracts as Incentives for Innovation,” *TSE Working Paper, n. 16-695*.

Clarke, F. (2013): *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264. Springer Science & Business Media.

Cumming, D. (2012): *The Oxford Handbook of Venture Capital*. Oxford University Press.

Cuny, C. J., and E. Talmor (2005): “The Staging of Venture Capital Financing : Milestone vs Rounds,” *EFA 2005 Moscow Meetings Paper*, (314).

DeMarzo, P. M., and Y. Sannikov (2016): “Learning, Termination, and Payout Policy in Dynamic Incentive Contracts,” *Working Paper*.

Gage, D. (2012): “The Venture Capital Secret: 3 Out of 4 Start-Ups Fail,” *The Wall Street Journal*.

Garrett, D. F., and A. Pavan (2012): “Managerial Turnover in a Changing World,” *Journal of Political Economy*, 120(5), 879–925.

Gerardi, D., and L. Maestri (2012): “A principal–agent model of sequential testing,” *Theoretical Economics*, 7(3), 425–463.

Gomes, R., D. Gottlieb, and L. Maestri (2016): "Experimentation and project selection: Screening and learning," *Games and Economic Behavior*, 96, 145 – 169.

Gompers, P. A. (1995): "Optimal Investment, Monitoring, and the Staging of Venture Capital," *The Journal of Finance*, 50(5), 1461–1489.

Gompers, P. A., and J. Lerner (2004): *The Venture Capital Cycle*. MIT press.

Gornall, W., and I. A. Strebulaev (2015): "The economic impact of venture capital: Evidence from public companies," *Available at SSRN 2681841*.

Green, B., and C. R. Taylor (2016): "Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects," *Working Paper*.

Halac, M., N. Kartik, and Q. Liu (2016): "Optimal Contracts for Experimentation," *The Review of Economic Studies*.

Hannan, M. T., M. D. Burton, and J. N. Baron (1996): "Inertia and Change in the Early Years: Employment Relations in Young, High Technology Firms," *Industrial and Corporate Change*, 5(2), 503–536.

Hart, O., and J. Moore (1994): "A Theory of Debt Based on the Inalienability of Human Capital," *The Quarterly Journal of Economics*, 109(4), 841–879.

Hege, U., D. Bergemann, and L. Peng (2012): "Venture Capital and Sequential Investments," *Working Paper*.

Hörner, J., and L. Samuelson (2013): "Incentives for experimenting agents," *RAND Journal of Economics*, 44(4), 632–663.

Hu, Z. (2014): "Financing Innovation with Unobserved Progress," .

Kaplan, S. N., and P. Stroemberg (2004): "Characteristics , Contracts , and Actions: Evidence from Venture Capitalist Analyses," *The Journal of Finance*.

Keller, G., and S. Rady (2010): "Strategic experimentation with Poisson bandits," *Theoretical Economics*, 5(2), 275–311.

Keller, G., S. Rady, and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39–68.

Laffont, J.-J., and J. Tirole (1988): "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), 1153–1175.

Lerner, J., and J. Wulf (2007): "Innovation and Incentives: Evidence from Corporate R&D," *The Review of Economics and Statistics*, 89(4), 634–644.

Malmendier, U., and J. Lerner (2010): "Contractibility and the Design of Research Agreements," *American Economic Review*, 100(1), 214–246.

Moorman, C., S. Wies, N. Mizik, and F. J. Spencer (2012): "Firm Innovation and the Ratchet Effect Among Consumer Packaged Goods Firms," *Marketing Science*, 31(6), 934–951.

Moroni, S. (2017): "Experimentation in Organizations," *Working Paper*.

Neher, D. V. (1999): "Staged Financing: An Agency Perspective," *The Review of Economic Studies*, 66(2), 255–274.

Pindyck, R. S. (1993): "Investments of uncertain cost," *Journal of Financial Economics*, 34(1), 53 – 76.

Prat, J., and B. Jovanovic (2014): "Dynamic contracts when the agent's quality is unknown," *Theoretical Economics*, 9(3), 865–914.

Tirole, J. (2010): *The Theory of Corporate Finance*. Princeton University Press.

Wasserman, N. (2008): "The Founder's Dilemma," *Harvard Business Review*.