Experimentation, Private Observability of Success, and the Timing of Monitoring*

Alexander Rodivilov†

November 2, 2016

For the latest version, please click here.

Abstract

This paper examines the role of monitoring in experimentation when agents may observe success privately and therefore delay announcing success. In the benchmark model without monitoring, private observability of success is inconsequential as we show that the agent never wants to delay announcing success. However, with monitoring of the agent's effort, private observability of success plays a role in choosing the optimal time for monitoring. When success is observed publicly, the optimal time for a principal to hire a monitor is at the start of the relationship. On the contrary, if the agent observes success privately, and the discount factor is high enough, monitoring is performed during the final period.

Keywords: Experimentation; Monitoring; Private observability of success.

JEL classification: D83; D86.

*I am grateful to Fahad Khalil and Jacques Lawarrée for their invaluable suggestions and insightful conversations. I also am thankful for the helpful comments of Nageeb Ali, Philip Bond, Renato Gomes, Quan Wen, participants at Midwest Economic Theory Meetings 2016 and Stony Brook International Conference on Game Theory 2016.

†Department of Economics, University of Washington, Seattle, WA, USA. *Email*: *arodivil@uw.edu*.

1 Introduction

Consider a venture capitalist (principal) who provides funds to an entrepreneur (agent) hoping to succeed in a risky but lucrative project. Both parties are initially unsure whether the project is "good", i.e., if it is possible to implement it successfully at all. The entrepreneur is asked to experiment with the project a certain period of time. The principal needs to determine how long the experimentation will last. She faces two problems: the effort of the agent may not be observable (moral hazard) and the result of each experiment may not be publicly observed.

The principal may therefore not have the correct information about the viability of the project as experimentation proceeds. Consider first the case when success is observed publicly. Suppose the agent secretly shirks and success is not achieved. The principal uses the observed failure to update his beliefs on project quality and becomes more pessimistic from that period on. The agent, in contrast, knowing that the experiment was not successful because he was shirking, will not update his beliefs regarding the project quality. Furthermore, if success is privately observed then the agent might postpone announcing success when the project is successfully implemented. This report makes the principal more pessimistic while the agent knows the project is certainly good.

Another example is a pharmaceutical company that employs a research organization to carry out clinical trials on a group of volunteers to demonstrate the effectiveness of a new drug. If the company does not to observe the agent testing the drug directly, it may doubt whether the agent is exerting effort. If the agent chooses to shirk, he simply can report that the drug was tested unsuccessfully. If the company remains unaware of this falsehood, it will adjust its beliefs about the drug's quality accordingly, becoming more pessimistic. Even if the agent discovers the drug is performing well, he may delay announcing success in favor of personal gain.

In this paper, we ask whether the principal can benefit from hiring a monitor in this environment. In a dynamic relationship, collecting information during every period is prohibitively costly, raising the question of *when* monitoring is most effective. First, we derive the optimal contract without monitoring that determines the length of the principal-agent relationship and solves the agent's moral hazard, ensuring that he works properly during every period of the relationship and announces achieved success promptly. Second, we study the benefits of monitoring of the agent's effort. The optimal contract then becomes contingent on the monitor's reports. Finally, we examine how private observability of success influences the structure of the optimal contract and the optimal timing of monitoring.

To answer these questions, we use a simple two-armed bandit model[1]: The agent can "pull the risky arm" by exerting effort toward implementing the project (achieve success), or he can "pull the safe arm" and shirk. While pulling the risky arm is costly, it allows the project to be implemented successfully if it is good. Pulling the safe arm, however, yields zero return, regardless of project quality.

Our results show that, in the benchmark case without monitoring, the principal commits to terminating the relationship if the agent does not succeed by a certain period. The agent is rewarded only if the project is implemented successfully and receives nothing otherwise. The nominal value of the reward increases to account for rising pessimism as the project does not succeed over time. In particular, the agent is rewarded for earlier success, as the discounted payments are decreasing over time. So even if the agent observes success privately, he will never delay announcing success and the optimal contract is unchanged. Surprisingly, with monitoring of the agent's effort, private observability of success plays a role in choosing the optimal time for monitoring. We show that in the case in which success is observed publicly, the principal should monitor the agent at the beginning of their relationship. If the agent observes success privately, the optimal time for monitoring is affected by patience. For high enough discount factor, monitoring should be performed during the final period.

Consider the first-best scenario, in which effort and all the information the agent learned are observed publicly. Since each attempt to implement the project is costly, and the principal becomes more pessimistic with every period in which success is not announced, the first-best solution is characterized by a stopping rule. The agent is allowed to attempt to implement the project for several periods only.

In the second-best case, the principal faces two problems: the agent chooses effort level privately (moral hazard), and, in addition, success may not be publicly observed. Suppose success is observed publicly. In the case the agent secretly shirks and success is not achieved, the principal becomes relatively more pessimistic from that period on. She then would adjust the agent's reward to induce him to exert effort accordingly: After every period the agent does not succeed, the principal supplies him with larger payments to encourage him to continue working in the next period. The agent, in contrast, understands that the observed failures are uninformative, and at the beginning of the next period would hold the same beliefs as in the previous one.

Besides, private observability of success may exacerbate the problem. In some settings, the principal can observe success easily. In the example of the pharmaceutical company, the clinical research organization may have difficulty hiding a revolutionary drug's success. However, success might be much more difficult to ascertain when information gathering does

[1]See Keller et al. (2005) for more details.

not involve extreme outcomes.

Consider the agent's incentive to announce privately observed success at a certain period of the relationship. This decision is affected not only by the payment tied to success or failure in this particular period, as determined by the optimal contract, but also by payments in all subsequent periods of the relationship. For example, if the discounted value of the promised future reward exceeds the current value, then the agent will postpone an announcement. We show that under the optimal contract, if the agent postpones an announcement of success, the discounted value of his reward will only decrease. Thus, private observability does not worsen the problem, as the agent cannot benefit from hiding success; however, it becomes a crucial factor in defining the optimal time for monitoring.

Given the optimal contract, the agent receives a strictly positive rent, and, as a result, the project is terminated inefficiently early. One way the principal can alleviate this inefficiency is by hiring a monitor who can observe the effort level chosen by the agent. The principal can avoid paying the promised reward since the moral hazard problem vanishes when the monitor is hired. The principal values monitoring as it mitigates the rent paid to the agent and, consequently, allows the relationship to be extended.

The principal benefits from monitoring as she pays the agent a smaller reward in case he succeeds during the monitoring period - we refer to this as the *static effect*. Moreover, recall that the agent can shirk and attempt to implement the project during later periods. As a result, his incentives to work during each period, except for the final one, depend not only on the payment determined by the contract and success of that particular period, but also on the payments for all subsequent periods.[2] The effects of future monitoring reflect in the earlier periods, as monitoring acts as a threat that causes the agent to perceive shirking as less attractive. Thus the *dynamic effect* emerges: The principal can diminish all of the agent's rewards in all periods before he is monitored.

We demonstrate that the dominating effect depends on whether the agent observes success privately or not. The optimal timing of monitoring is governed by the sum of the two effects. Since without monitoring, the agent is rewarded for earlier success, the expected reward is strictly decreasing. This implies that the benefit from the static effect is strictly decreasing in monitoring timing as well. With the dynamic effect, the principal benefits even if the agent does not succeed at the period of monitoring: For earlier periods of monitoring, the benefit from the dynamic effect is strictly increasing, as it allows paying less in several periods. However, as time passes without success, both parties become sufficiently pessimistic, and the benefit from the dynamic effect decreases eventually.

[2] Halac et al. (2016) discuss this dynamic agency effect in their model with moral hazard and adverse selection.

Consider first the case in which success is observed publicly. Because monitoring eradicates the moral hazard problem during the monitoring period, the principal benefits from it only if the agent succeeds in this specific period, which in turn is possible only if the project is good. In earlier periods, the benefit from the static effect increases while the benefit from the dynamic effect goes up, whereas benefits from both effects decline toward the end of the relationship. The benefit from the static effect decreases faster than the benefit from the dynamic effect increases - this result is at the heart of our analysis. Recall that the benefit from the dynamic effect increases only in earlier periods: Because the principal is promising to pay less with each period, the chance that the agent will shirk grows smaller. The principal saves more in these early periods by opting to monitor later due to the dynamic effect. The optimal contract, in contrast, mitigates the moral hazard problem for every period, not only for those in which the benefit from the dynamic effect is increasing, as it induces the agent to exert effort as long as the relationship lasts. Thus, when success is observed publicly, the static effect dominates, and monitoring is optimal during the first period.

When the agent observes success privately, the benefit from the static effect becomes smaller because the principal now still pays some rent to the agent during the monitoring period. Recall that when the agent announces success, he takes into account not only the payment tied to success in this precise period but also payments for success in all subsequent periods of the relationship: If the discounted value of the promised reward in the next period exceeds the current reward, then the agent will postpone an announcement. Given that the optimal contract without monitoring includes a decreasing discounted reward value, the principal can decrease the reward in one period up to the discounted value of the reward in the following period at most. As the discount factor increases, the benefit from the static effect decreases for all periods except the final one.

1.1 Related Literature

Our paper contributes to the literature on incentives for experimentation. Most studies model experimentation based on Keller et al.'s (2005) study, which used a two-armed bandit model with a risky arm that might yield exponentially distributed payoffs and a safe arm that offers a safe payoff. The literature on incentives for experimentation could be divided into two parts, depending on who is initially the owner of the project.

A group of papers considered an entrepreneur who owns a project and is raising the funds necessary to implement a project in a competitive market. In settings with private learning and moral hazard, Bergemann and Hege (1998) considered the provision of venture capital in a dynamic agency model. The optimal share contract operates on the provision that if the entrepreneur succeeds, he conveys a part of the project to the investor. In Bergemann and Hege's study, the share in the earlier periods can rise or fall, but the agent receives the expected value of the project. In our paper, on the other hand, a nominal reward for success is always (weakly) increasing, and the agent receives a positive rent. Bergemann and Hege's (2005) study built on their 1998 study, with one crucially distinct feature - the time horizon is infinite, and the funding decision is renegotiated each period.

Another group of papers considered a principal who owns a research idea but lacks the decisive skills necessary to implement it and must hire an agent in order to do so. Halac et al. (2016) considered the challenges of creating a contract for a project of uncertain feasibility with adverse selection and moral hazard. The optimal contract involves paying the agent initially and penalizing him progressively if success has not been observed. Our research differs from Halac et al.'s (2016) in that we assume the probability of success is known, but the agent is protected by limited liability. The agent thus cannot be penalized for failure; instead, we show how bonus contracts and optimal monitoring discourage the agent from hiding success. Bonatti and Horner (2016) studied a model in which the quality of the project depends on a worker's skill that is revealed through output, and wage is based on the expected output, or assessed ability. However, the authors do not assume limited liability and allow punishment for achieving certain deadlines. Manso (2011) applied a similar model to study a manager who must be incentivized to perform an innovative task. The optimal incentive scheme exhibits significant tolerance and even reward for both early failure and long-term success.

Gerardi and Maestri (2012) analyzed how an agent can be incentivized to obtain and announce information over time. The authors assumed that the principal observes the state of nature with some time lag, and, with the optimal contract, he can reward or punish the agent after comparing the agent's report with the revealed state. In this study, the agent is rewarded only if his report matches the true state, whereas in our study, the true state is learned only if the project is successfully implemented. Mason and Valimaki (2015) considered an infinitely lasting relationship in a model without learning and with moral hazard and continuous effort. They demonstrated that the agent's wage declines over time; however, they did not study monitoring.

When studying monitoring, Bonatti and Horner (2011) considered a team of agents who work together on a project of uncertain feasibility. In their model, agents observe only their own effort levels and form beliefs regarding the effort of other teammates only. An intriguing result is that monitoring (which allows the effort level of all other agents to be known) does not necessarily improve the outcome. A trade-off arises: Though observing other teammates' effort choices prevents unreasonable pessimism, this also may lead to early high-level effort and faster learning, and later low-level effort. However, in a setting with one agent only, we show that the principal benefits from observing the agent's effort choices.

The only researchers who have studied optimal monitoring time sets in optimal contracts for experimentation are Bergemann and Hege (1998). In this study, when success is publicly observed, monitoring is optimal toward the end of the project, whereas in our model, the monitor is hired optimally at the beginning of the relationship. Our paper complements Bergemann and Hege's (1998) result in that it highlights the pivotal role of market structure on the optimal timing of monitoring. In addition, our paper extends Bergemann and Hege's (1998) result, as we show that when the agent observes success privately, patience influences the optimal timing of monitoring, which was not explored in their paper. As the discount factor increases, the principal must promise identical rewards for success at every period except the final one, making monitoring more valuable at the end of the relationship and supporting Bergemann and Hege's (1998) result.

Most of these papers assumed project success would be publicly observed.[3] We, in contrast, assume that the agent could observe success privately. We also assume that information is hard - that is, the agent can either postpone announcing a successful implementation or hide it completely by destroying evidence.[4]

We argue that our findings are not only theoretical, but also empirically significant. Following our example of the pharmaceutical company, it is widely acknowledged that on-site clinical trial monitoring is a source of significant inefficiency in the conduct of clinical trials, and that current monitoring activities do not always lead to increased quality in clinical trials.[5]

The business literature[6] on venture capital for innovation emphasizes the importance of relationship financing and monitoring. For example, Gorman and Sahlman (1989) found in their survey that venture capitalists visit their companies frequently and devote significant amount of time to participating in decision-making. With respect to optimal monitoring time, Lerner (1995) demonstrated that venture capitalists' representation on the board of directors increases over certain periods, especially during chief executive officer turnover. While a venture capitalist may have reasons to participate in the firm's recruitment of a management team and strategic planning, this paper, does not examine such motives and considers monitoring that serves the purpose of eliminating the moral hazard problem only.

[3] Halac et al. (2016) discuss the robustness of an optimal contract to project success being privately observed by the agent.

[4] Henry (2009) makes a similar assumption. Gerardi and Maestri (2012) consider a case of soft information, i.e., when the agent can make up the evidence of success.

[5] Eisenstein et al. (2005) estimated the cost of on-site monitoring to constitute between 25 and 30 percent of clinical trial costs.

[6] See Gorman and Sahlman (1989), Sahlman (1990) and references therein.

The rest of the paper is organized as follows: Section 2 explains the model and the contract space with payoffs, and provides a solution for the first-best benchmark; Section 3 provides a description of the optimal contract with moral hazard and private observability of success; Section 4 extends results for the case in which the principal can hire a monitor with private as well as public observability of success; and Section 5 concludes the paper.

2 Model

2.1 The project

A principal owns a valuable idea that could result in a lucrative project, but he lacks the decisive skills needed to implement it. He hires an agent protected by limited liability to perform the project. Both parties initially are uncertain about the project's quality; that is, the common prior on the project being "good" is $\beta_0 \in (0,1)$.[7] If the project is good, then it can be implemented successfully with a known positive probability, in which case it will yield a fixed return of $V > 0$, which is commonly known at the beginning of the relationship. To implement the good project, the agent must exert effort that is assumed to be subject to a binary choice: $e \in \{0,1\}$. If the project is bad, then it will yield zero, regardless of effort.[8] Exerting effort costs $c > 0$ per period.

The agent's ability, λ, which is the probability of achieving success given that the project is good, conditional on exerting effort, is common knowledge during contracting. Finally, we assume that the effort choice is not observable and that the agent can postpone announcement of or hide a successful implementation.

2.2 Learning the quality of the project

An important feature of this model is learning project quality. When the agent does not succeed, despite exerting effort, he updates[9] his beliefs regarding the quality of the project using Bayes' rule and becomes more pessimistic. Denoting by $\tilde{\beta}_t$, the updated belief of the agent that the project is good at the beginning of period t after $t-1$ failures, we present:

$$\tilde{\beta}_t = \frac{\tilde{\beta}_{t-1}(1-\lambda)}{\tilde{\beta}_{t-1}(1-\lambda) + 1 - \tilde{\beta}_{t-1}}, \text{ which simplifies to } \tilde{\beta}_t = \frac{\beta_0(1-\lambda)^{t-1}}{\beta_0(1-\lambda)^{t-1} + 1 - \beta_0}.$$

[7] It is important that β_0 is strictly positive and strictly less than one. Otherwise, no additional information arrives as the relationship proceeds; in this case, there is no learning regarding the quality of the project, and the problem simplifies to standard dynamic moral hazard.

[8] We refer to an implementation of the project as "success" and to lack of success as "failure."

[9] A failure is more informative if beliefs are close to $\frac{1}{2}$, while beliefs change slowly when parties are relatively certain about the quality of the project. See Bergemann and Hege (1998) for more details.

Since the agent chooses effort level privately, both parties do not share the same beliefs necessarily as their relationship evolves. The principal becomes more pessimistic every period the agent does not announce success. However, if the agent secretly shirks he becomes relatively more optimistic from that period on. Consider a hypothetical scenario in Figure 1 below:

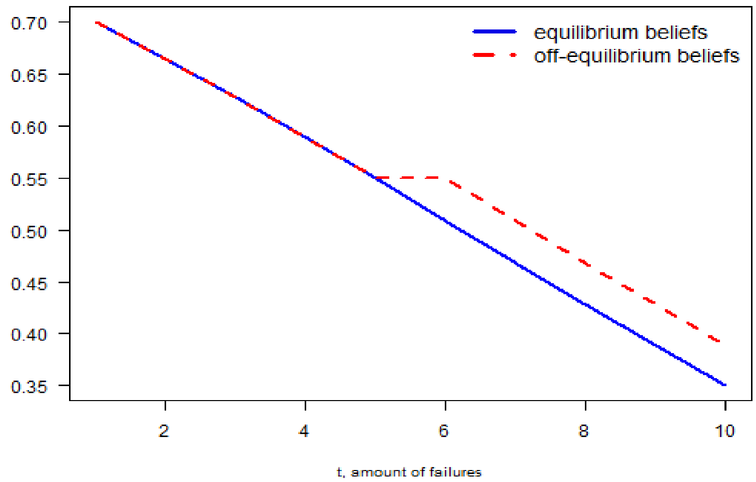

Figure 1. Learning the quality of the project with $\lambda = 0.15$ and $\beta_0 = 0.7$.

Given the parameters, the bold line reflects the evolution of beliefs if the agent continues exerting effort for 10 periods. Suppose the agent secretly shirks at $t = 5$, but reports that success has not been achieved, despite exerting effort. The principal would use this report to update his beliefs and become relatively more pessimistic from period $t = 6$ on. The agent, in contrast, would understand that the reported failure was uninformative, and at the beginning of period $t = 6$ would have the same beliefs as in the previous period. Importantly, this difference in beliefs following one deviation at $t = 6$ would carry into all future periods until the relationship ends.

2.3 Contracts and payoffs

The optimal contract has to take into account four crucial features of the relationship between the principal and the agent: First, the results of each period affects the relationship; with each failure the agent reports, the principal becomes more pessimistic. Second, during each period, the agent chooses privately whether to exert the effort necessary for the project to succeed. Third, the agent is protected by limited liability, so the principal cannot sell the project to the agent. Finally, the agent observes successful project implementation privately. As a result, the payment structure must ensure that the agent neither postpones nor hides the announcement of a successful implementation.

Both parties are risk neutral and share a common discount factor $\delta \in (0, 1]$. An optimal contract must specify how many failures the principal will tolerate and a sequence of transfers as a function of the agent's reports,[10] which in this case is whether or not the agent succeeded. All transfers are from the principal to the agent.

The contract is given by $\varpi = (T, \{b_t\}_{t=1}^{T}, \{w_t\}_{t=1}^{T})$, where $T \in N$ is the duration of the relationship, b_t is the payment to the agent in case he reports success at period $1 \leqslant t \leqslant T$ and w_t is the payment to the agent conditional on reporting failures from the beginning of the relationships up to period $1 \leqslant t \leqslant T$.

Under certain circumstances, the agent can postpone or even hide success. To understand how this fact matters and if it affects the structure of the optimal contract, consider the agent's incentive to announce that he successfully completed the project at $t < T$. This decision is affected not only by the payment tied to success or failure in this particular period, as determined by the optimal contract, but, in addition, by payments in all subsequent periods of the planning stage. For example, if the discounted value of the promised reward for success in the future exceeds the current value, then the agent will postpone an announcement; if the agent is rewarded for consecutive failures,[11] he would benefit from hiding success completely.

To prevent the agent from hiding success, the optimal contract will have to satisfy the following incentive compatibility constraint for every period $1 \leqslant t \leqslant T$:

(IC)
$$b_t \geqslant w_t + \delta b_{t+1} \text{ for } t = 1, ..., T-1,$$
$$b_t \geqslant w_t \text{ for } t = 1, ..., T.$$

Given the optimal contract and effort levels the agent chooses, we can specify the agent's expected utility and the principal's expected profit. The agent's expected utility from accepting contract ϖ at time zero while exerting an effort profile $\vec{e}$ and reporting each project truthfully as a failure or success is:

$$U(\varpi, \vec{e}) = (1 - \beta_0) \sum_{t=1}^{T} \delta^t (w_t - e_t c)$$
$$+\beta_0 \sum_{t=1}^{T} \delta^t (\prod_{s=1}^{t-1} (1 - \lambda e_s))(e_t(\lambda b_t - c) + (1 - \lambda e_t) w_t),$$

where $\vec{e} = (e_1, ..., e_T)$ is an effort profile with $e_t \in \{0, 1\}$[12] for $1 \leqslant t \leqslant T$.

First, the agent has a chance to succeed during the relationship; this occurs only if both the project is good, which is true with probability β_0, and if the agent is exerting effort.

[10]Because success is observed privately, both parties do not necessarily share the same history as the relationship evolves.

[11]For example, Manso (2011) and Chade and Kovrijnykh (2016) explored models where the agent is rewarded for delivering bad news in a different setting.

[12]We refer to $e_t = 1$ as "work" and to $e_t = 0$ as "shirk".

Conditional on the project being good, the relationship lasts for an arbitrary period, $t \leqslant T$, with probability $\prod_{s=1}^{t-1}(1-\lambda e_s)$. If the agent exerts effort at period t, his expected payoff at this period is:

$$\lambda b_t + (1-\lambda)w_t - c,$$

whereas in case he shirks, the agent receives only w_t, as defined by the contract. Second, if the project is bad, which happens with probability $1-\beta_0$, the agent never succeeds, regardless of effort profile.

The principal's expected profit from offering contract ϖ at time zero if the agent exerts an effort profile $\vec{e}$ and reports failures and project success truthfully is:

$$\pi(\varpi, \vec{e}) = -(1-\beta_0)\sum_{t=1}^{T}\delta^t w_t$$
$$+\beta_0 \sum_{t=1}^{T}\delta^t(\prod_{s=1}^{t-1}(1-\lambda e_s))(e_t\lambda(V-b_t) - (1-\lambda e_t)w_t).$$

The optimal contract will have to satisfy the following moral hazard constraint at each period for all possible histories and all possible effort paths in the future:

(MH) $$\vec{1} \in \arg\max_{\vec{e}} U(\varpi, \vec{e}).$$

Given that the *MH* constraint is satisfied, the principal's expected profit from offering contract ϖ at time zero becomes:

$$\pi(\varpi, \vec{1}) = -(1-\beta_0)\sum_{t=1}^{T}\delta^t w_t + \beta_0\sum_{t=1}^{T}\delta^t(1-\lambda)^{t-1}(\lambda(V-b_t) - (1-\lambda)w_t).$$

2.4 The first-best benchmark

Consider the first-best case: The principal observes the effort choice and project outcome. As the relationship proceeds, if success is not being achieved, then every period the marginal benefit, $\lambda\tilde{\beta}_t V$, is the expected value of the project and takes into account both probability of success and current beliefs. Since beliefs are declining as time goes on without success, the marginal benefit decreases strictly. The marginal cost-of-effort, c, is constant. As a result, the first-best solution is characterized by stopping time $T \in N$, such that the agent is allowed to exert effort up until that date only, as follows:

$$T^{FB} = \arg\max_t\{\lambda\tilde{\beta}_t V \geqslant c\}^{13}.$$

Consider the example in Figure 2 below, where $\lambda = 0.15$, $\beta_0 = 0.7$, $V = 20$ and $c = 1$, and where the agent starts with $MB_1 = 2.1$ and continues experimenting with the project for ten periods at most.

[13] Recall that $\tilde{\beta}_t$ are beliefs evolved as a result of the agent exerting effort in all periods until t.

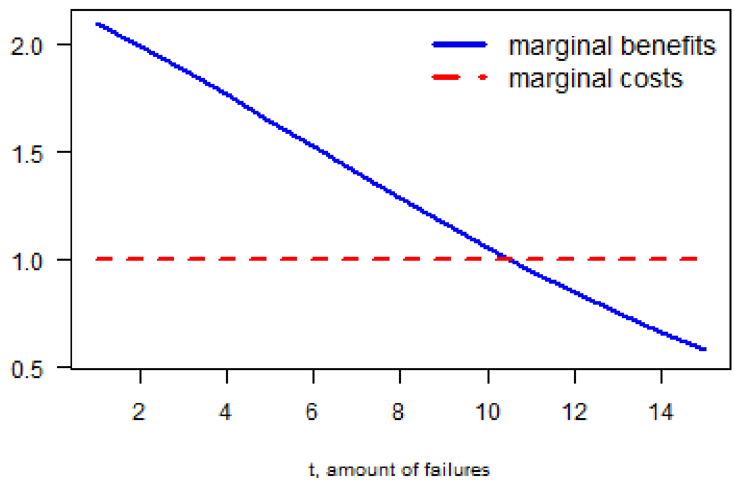

Figure 2. The first-best benchmark with $\lambda = 0.15$, $\beta_0 = 0.7$, $V = 20$ and $c = 1$.

3 The Second-Best Contract

When the agent chooses effort level privately, the optimal contract must ensure the agent works in every period, which is guaranteed by the *MH* constraint. Since the agent is protected by limited liability, he cannot pay the principal, as the *LL* constraint reflects. In addition, the *IC* constraint ensures the agent neither postpones success nor hides it. The principal's optimization problem in this case becomes the following[14]:

$$[P^{SB}]\ \max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH) $$\vec{1} \in \arg\max_{\vec{e}} U(\varpi, \vec{e}),$$

(IC) $$b_t \geqslant w_t + \delta b_{t+1} \text{ for } t = 1, ..., T-1,$$

$$b_t \geqslant w_t \text{ for } t = 1, ..., T,$$

(LL) $$b_t,\, w_t \geqslant 0 \text{ for } t = 1, ..., T.$$

In this model, the moral hazard problem in each period translates into asymmetric information regarding beliefs about the project's quality in all consecutive periods. Before we present a detailed solution to the principal's optimization problem, consider the agent's incentives to deviate at period $t \leqslant T$, assuming that the agent was behaving $1 \leqslant s < t$ without success in all prior periods and will work $t < s \leqslant T$ in all subsequent periods. In case the agent decides to shirk at the beginning of period $t \leqslant T$, his continuation value from the relationship is:

[14]We assume that V is high enough, and it is optimal for the principal when the agent exerts effort in every period.

$$U_t(\varpi,(0,1,...,1)) = w_t + (1-\tilde{\beta}_t)\sum_{s=t+1}^{T}\delta^{s-t}(w_s-c)$$
$$+\tilde{\beta}_t\sum_{s=t+1}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s+(1-\lambda)w_s-c).$$

Note that if the agent follows this one-period deviation, he gets only w_t at period t, since he fails for sure. If the project is good, which, based on history, is true with current beliefs $\tilde{\beta}_t$, then the agent has a chance to succeed in all future periods $s > t$ until the relationship is terminated. If the project is bad, which is true with probability $1-\tilde{\beta}_t$, the agent will receive w_s in all future periods $s > t$ despite exerting effort.

In contrast, if the agent decides to work at period t, his continuation value from the relationship is:

$$U_t(\varpi,(1,1,...,1)) = -c + \lambda\tilde{\beta}_t b_t + (1-\lambda\tilde{\beta}_t)w_t + (1-\tilde{\beta}_t)\sum_{s=t+1}^{T}\delta^{s-t}(w_s-c)$$
$$+\tilde{\beta}_t\sum_{s=t+1}^{T}\delta^{s-t}(1-\lambda)^{s-t}(\lambda b_s+(1-\lambda)w_s-c).$$

Notice that at period t, the agent has a chance to succeed and receive b_t. This occurs either if the project is good or with probability $\lambda\tilde{\beta}_t$. In case the agent is unlucky, with probability $1-\lambda\tilde{\beta}_t$, he gets w_t, despite exerting high effort. As in the case where the agent deviates, if the project is bad, the agent will receive w_s for all future periods.

When the agent deviates at period t, he knows that failure at this period should not change beliefs and make parties more pessimistic regarding the project's quality. However, if this deviation is not observed by the principal, she will consider a failure reported at period t as a signal that the project is more likely to be bad. Importantly, this difference in beliefs reverberates into all future periods. Thus, in this model, the moral hazard problem in each period translates into asymmetrical information regarding beliefs about the project's quality in all consecutive periods.

Combining the two continuation values, the moral hazard constraint at period t (assuming that the agent was behaving in all prior periods $s < t$ and will work in all subsequent periods $s > t$) becomes the following:

$$(MH_t)\ b_t - w_t \geq \frac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s+(1-\lambda)w_s-c).$$

When the agent chooses his effort level privately, he receives a strictly positive rent. The agent could shirk and report that the project failed. The principal can motivate the agent to exert effort by paying a higher reward for success and a lower one for failure. The gap between these payments must be wide enough for the agent to believe it is in his best interest to exert all his efforts after taking into account current beliefs of the project's quality and probability of success. If the agent and principal share the same beliefs about the project's quality, a standard moral hazard problem takes place within each period[15]. Instead, the

[15] To minimize risk, the principal ideally would sell the project to the agent; however, this is not feasible because the agent is protected by limited liability.

agent receives a positive moral hazard rent. Since the principal benefits only if the project is released to market, it gains advantage by awarding little to the agent if failure, and, given the limited liability constraint, the agent is paid nothing if he fails overall.

Moreover, if the agent and the principal do not hold common beliefs, the former receives additional reward (the learning rent). If the agent deviates from project goals at one period, his chance to succeed in all future periods remains. Although the agent will not receive anything if he deviates from his duties during a particular period, he becomes relatively more optimistic than the principal for all future periods. That means a deviation at one period carries into all future periods by creating asymmetric beliefs among parties. In some sense, the agent is relatively more patient[16] than the principal in all periods except the last. During this final period, the agent cannot benefit from shirking, since he will not gain from this, and his rent is contingent on the combination of moral hazard and limited liability only. Because of the positive rent the agent receives, the project could be terminated inefficiently early.

Proposition 1. The agent receives a positive reward if the project is implemented successfully and nothing otherwise. In particular,

$w_t = 0$ and $b_t = \dfrac{c}{\lambda \tilde{\beta}_t} + c \sum_{s=1}^{T-t} \delta^s \frac{1-\beta_0}{\beta_0 (1-\lambda)^{t+s-1}}$ for $1 \leq t \leq T^{SB}$ with the following properties:

- if $\delta = 1$, b_t is *constant*[17];
- if $0 < \delta < 1$, b_t is strictly *increasing* whereas δb_t is strictly *decreasing*.

Moreover, the project is terminated inefficiently early; that is, $T^{SB} \leqslant T^{FB}$.

Proof: See Appendix A.

Our results show that the agent's nominal reward is weakly increasing while the discounted value of the reward is weakly decreasing in time. For $\delta = 1$, similar reward structure holds in Halac et al. (2016) who argue that in the case of no discounting, the principal can be restricted to use constant bonus contracts[18]. When the discount factor is less than one, the agent's reward is strictly increasing while the expected reward is strictly decreasing. This resembles the payment scheme in Gerardi and Maestri (2012) when the agent's report matches the true state observed by the principal. In contrast, in Bergemann and Hege (1998), the agent's reward for earlier success can rise or fall and even becomes strictly decreasing for high enough discount factor.

[16] In Bonatti and Horner (2016) the agent fails to take into account the value of success, which is increasing in the effort; this makes the agent more patient then the principal.

[17] Note that when $\delta = 1$ the optimal contract is unique up to payoff-irrelevant alteration.

[18] Bonatti and Horner (2011) have a similar result in their model with one agent only.

An immediate and perhaps fascinating conclusion from Proposition 1 is that the agent's private observability of success does not exacerbate the problem; that is, the agent will never postpone an announcement of success even if the *IC* constraint was not taken into account directly when solving the principal's optimization problem. The reason is that the optimal contract makes the value of the discounted reward strictly decreasing. This result plays a key role in our analysis, as we will demonstrate that private observability of success becomes critical when it comes to optimal monitoring timing.

To demonstrate the intuition behind Proposition 1, we clarify the dynamics of the moral hazard and learning rents. The first component is always increasing, since the agent becomes more pessimistic as time proceeds without success, and his motivation to exert himself becomes costlier. The second component, however, is non-monotonic and depends on the discount factor. Consider a case in which both the agent and the principal are patient (the discount factor equals one). Under these circumstances, the principal can wait for success indefinitely. Without loss of generality, the principal can offer a contract with constant nominal reward and a deterministic deadline that will ensure the agent will exert effort in every period. Since the moral hazard rent is increasing strictly and the nominal reward is constant, the learning rent decreases strictly. If the principal is patient, the agent benefits less from deviating from project goals since the fixed deadline gives him a smaller horizon to benefit from asymmetric beliefs.

However, if parties to a contract are impatient (the discount factor is less than one), the learning rent becomes non-monotonic. Since $c\sum_{s=1}^{T-t}\delta^s\frac{1-\beta_0}{\beta_0(1-\lambda)^{t+s-1}} = \frac{c\delta(1-\beta_0)}{\beta_0(1-\lambda-\delta)}\frac{(1-\lambda)^{T-t}-\delta^{T-t}}{(1-\lambda)^{T-1}}$, the learning rent is increasing at period $1 \leqslant t \leqslant T$ if and only if

$$\text{either } \delta < 1-\lambda \text{ and } (\tfrac{1-\lambda}{\delta})^{T-t}\ln\delta > \ln(1-\lambda) \text{ or}$$
$$\delta > 1-\lambda \text{ and } (\tfrac{1-\lambda}{\delta})^{T-t}\ln\delta < \ln(1-\lambda).$$

This means that except for the final period, if the agent shirks and does not incur the cost-of-effort, he will have an additional attempt to successfully implement the project and receive a reward. This weakens the agent's incentives to work during each period. The principal, however, benefits if the project is successfully implemented only, and it cannot wait indefinitely, as later success are discounted. This allows the agent to be relatively more patient than the principal. By this logic, the learning rent is increasing. However, the later the agent deviates, the fewer periods remain to exploit the difference in beliefs and, as a result, the learning rent eventually decreases before vanishing completely during the final period. Since the agent has more incentives to deviate at the beginning of the relationship, the optimal contract makes the discounted reward strictly decreasing.

The agent's incentive to announce success at a certain period of the relationship is affected by the payment tied to success or failure in this particular period and, in addition, by payments in all subsequent periods of the relationship. For instance, if the current value of the reward is smaller than the discounted value of the promised future reward, the agent will postpone an announcement of success. We show that the optimal contract makes the discounted reward strictly decreasing; this feature prevents private observability of success from exacerbating the contracting environment.

We present a particular example to better demonstrate the decomposition of the two rents in Figure 3 below. Suppose $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 1$ and $c = 1$. Note that when there is no discounting, the nominal value of b_t is constant, whereas the discounted value of the reward is decreasing, as suggested by Proposition 1. In this case, $(\frac{1-\lambda}{\delta})^{T-t}\ln(\delta) > \ln(1-\lambda)$ and $\delta > 1-\lambda$, and the learning rent is decreasing for all periods.

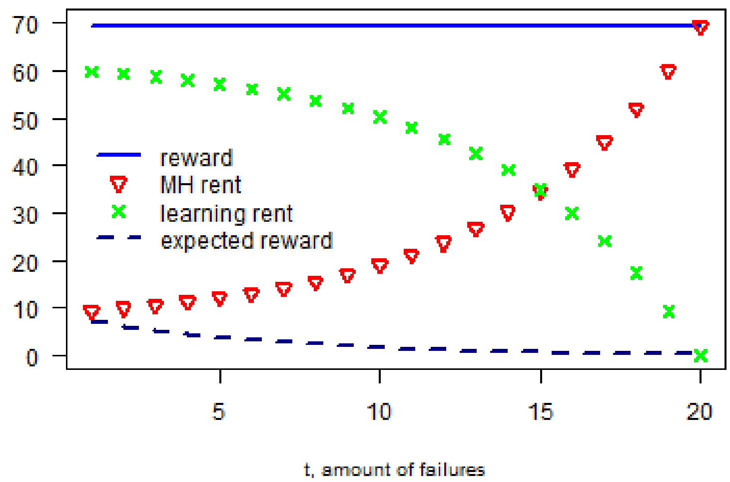

Figure 3. The optimal contract with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 1$ and $c = 1$.

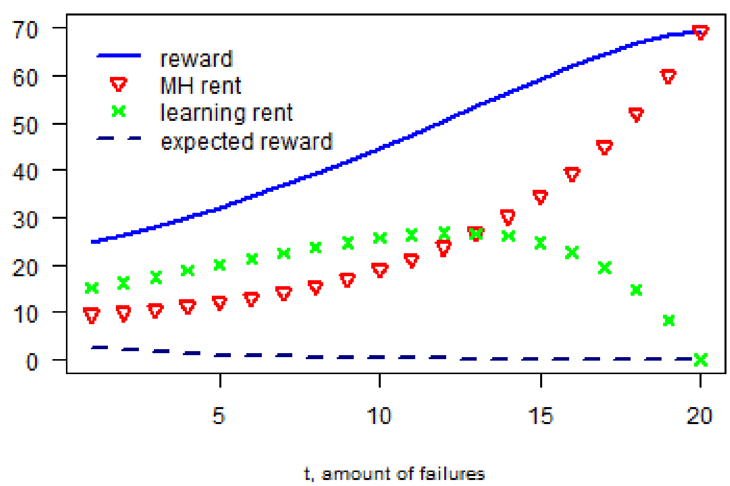

Figure 4. The optimal contract with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$.

Now consider an example with discounting, as depicted in Figure 4 above. Suppose that $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$. In this case, $\delta > 1-\lambda$ and $(\frac{1-\lambda}{\delta})^{T-t}\ln(\delta) < \ln(1-\lambda)$

for $t \leqslant 12$, and the learning rent is increasing for these periods and decreasing thereafter. As in the previous case, the moral hazard rent is increasing strictly to account for the agent's increasing pessimism.

4 Monitoring

Given the optimal contract described in the previous section, the agent receives a strictly positive rent, and the project consequently is terminated inefficiently early. One way the principal can alleviate this inefficiency is by hiring a monitor who assumedly can observe the effort level the agent chooses perfectly. In reality, monitoring is widely used in contracting for experimentation. For example, when running clinical trials, the pharmaceutical company hires an independent data safety monitoring board (DSMB) consisting of experts in the relevant clinical discipline. The DSMB members schedule several meetings as the trials proceed and advise on the conduct of the trial and the integrity of the data. They also evaluate interim analyses and judge efficacy and net clinical effects.

The principal values monitoring because it mitigates the rent paid to the agent and allows the relationship to be extended. For simplicity we assume that monitoring allows a perfect assessment of the agent's effort; however, our results could be extended easily to account for noisy monitoring. In addition, we assume that monitoring costs $\gamma > 0$ per period, the salary of the monitor.

The benefit of hiring a monitor is that for the period the monitor is hired, the principal can promise to pay less since the moral hazard problem is alleviated, inducing the static effect. Recall that without monitoring, the agent is rewarded more for earlier success and, consequently, the expected reward strictly decreases in the optimal contract. This influences the static effect, which strictly decreases in time. In addition, the dynamic effect from monitoring emerges, which reduces the learning rent the agent receives in all periods prior to monitoring. The prospect of future monitoring acts as a threat and makes the agent less likely to shirk in the earlier periods since the benefit of doing so is smaller. As demonstrated previously, the learning rent is non-monotonic, which makes the dynamic effect non-monotonic, as well.

Thus, the optimal time for monitoring is governed by the sum of the two aforementioned effects. We demonstrate that the dominating effect depends on whether the agent observes success privately or not. This is an important result, since without monitoring, private observability does not play any role in the optimal contract. However, it is crucial in determining the optimal monitoring timing.

4.1 Success is publicly observed

We begin the analysis of the optimal timing of monitoring with an example, where the relationship lasts exogenously for two periods ($T = 2$), and the principal can perfectly observe the effort level during one period only when success is publicly observed. The principal's optimization problem (assuming it is optimal when the agent exerts effort in every period) is:

$$\max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH) $$\vec{1} \in \arg\max_{\vec{e}} U(\varpi, \vec{e}),$$

(LL) $$b_1,\, b_2,\, w_1,\, w_2 \geqslant 0.$$

First, suppose the principal chooses to monitor the agent at the beginning of the relationship. In this case, the *MH* constraint could be replaced by:

$$(MH_2)\ \lambda\tilde{\beta}b_2 + (1-\lambda\tilde{\beta})w_2 - c \geqslant w_2,$$

which ensures that the agent behaves at $t = 2$, given that success has not been achieved at period $t = 1$. Then, a solution to the optimization problem with monitoring involves $b_2 = \frac{c}{\lambda\tilde{\beta}}$ and $w_2 = 0$, where $\tilde{\beta} = \dfrac{\beta_0(1-\lambda)}{\beta_0(1-\lambda)+1-\beta_0}$. The principal's expected profit in this case is the following:

$$\begin{aligned}\pi_{m=1} &= \beta_0 \textstyle\sum_{t=1}^{2} \delta^t(1-\lambda)^{t-1}\lambda V - \delta^2\beta_0(1-\lambda)\lambda b_2 - \gamma \\ &= \beta_0 \textstyle\sum_{t=1}^{2} \delta^t(1-\lambda)^{t-1}\lambda V - \delta^2 c(1-\lambda\beta_0) - \gamma.\end{aligned}$$

Second, suppose the principal hires the monitor at $t = 2$. In this case, the MH constraint could be replaced by:

$$(MH_1)\ \lambda\beta_0 b_1 + (1-\lambda\beta_0)w_1 - c \geqslant w_1,$$

which ensures that the agent behaves at $t = 1$, given that he will exert effort at $t = 2$. Then, the solution to the optimization problem involves $w_1 = b_2 = w_2 = 0$ and $b_1 = \frac{c}{\lambda\beta_0}$. The principal's expected profit is the following:

$$\begin{aligned}\pi_{m=2} &= \beta_0 \textstyle\sum_{t=1}^{2} \delta^t(1-\lambda)^{t-1}\lambda V - \delta\beta_0\lambda b_1 - \gamma \\ &= \beta_0 \textstyle\sum_{t=1}^{2} \delta^t(1-\lambda)^{t-1}\lambda V - \delta c - \gamma.\end{aligned}$$

Since $-\delta^2 c(1-\lambda\beta_0) > -\delta c$, it is optimal to monitor at $t = 1$. The intuition is that if monitoring occurs at $t = 1$, the principal expects to pay a reward at $t = 2$, conditional on the agent failing at $t = 1$, despite exerting effort as reflected by $(1-\lambda\beta_0)$ in the principal's expected profit $\pi_{m=1}$. The example above is straightforward but does not capture the main intuition fully, as when the relationship lasts for two periods and the monitor is hired, the agent does not receive any learning rent. We will demonstrate, however, that the result of this example extends to a general setting where the duration of the contract is long enough that the agent is granted a strictly positive learning rent.

Suppose now the principal can monitor the agent perfectly at any period $m \leqslant T^M_{Public}$, where T^M_{Public} is the duration of the contract with monitoring when success is publicly observed. We would like to understand all the benefits from monitoring in this case. First, the principal can avoid paying b_m since the moral hazard problem at period m vanishes. This static effect is at the heart of our analysis, as it will be playing an important role when success is observed privately. Since the static effect alleviates the moral hazard problem at the period of monitoring, the principal benefits from it only if the agent succeeds at period m, which in turn is possible only if the project is good. Thus, the static effect is:

$$SE_m = \delta^m Prob(\text{success at } m) b_m = \delta^m \beta_0 (1-\lambda)^{m-1} \lambda b_m.$$

Second, recall from the MH_m constraint that if the principal decreases a reward for success, b_m, he can scale down all the rewards in all the preceding periods, $1 \leqslant s < m$. This effect, which we call the dynamic effect, will be shown to play an auxiliary role in the environment we consider. Importantly, unlike with the static effect, the principal benefits from the dynamic effect even if the agent does not succeed at some period m. Since the agent always, except for the final period, has a chance to succeed in the later periods, the future rewards make it costlier to ensure the agent behaves at the beginning of the relationship. Intuitively, the promise of future monitoring echoes into the earlier periods, as it acts as a threat and it changes the agent's options if he decides to shirk in the earlier periods.

The dynamic effect is defined as:

$$DE_m = \sum_{t=1}^{m-1} \delta^t Prob(\text{success at } t < m)[\text{nominal decrease in } b_t].$$

What is a nominal decrease in b_t for $t < m$ that is possible because of monitoring that will occur at period m? Under the optimal contract all the MH_t constraints are binding:

$$(MH_t)\ b_t - w_t = \frac{c}{\lambda \tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s + (1-\lambda) w_s - c),$$

and, given that the agent receives nothing if the project does not succeed, we have

$$b_t = \frac{c}{\lambda \tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c).$$

Consequently, a nominal decrease in b_t for $t < m$ due to monitoring at period $1 \leqslant m \leqslant T$ is:

$$\delta^{m-t}(1-\lambda)^{m-t-1}\lambda b_m.$$

Finally, the dynamic effect becomes:

$$DE_m = \delta^m \beta_0 \lambda^2 (1-\lambda)^{m-2}(m-1)b_m.$$

Thus, the *total effect* of monitoring at period m, $TE_m = DE_m + SE_m$, combines the benefit of paying less at period m, which is decreasing in time, and the benefit of scaling down all the rewards in previous periods, which is non-monotonic. It turns out that the former effect is dominant, as stated in Proposition 2.

To understand how monitoring changes rewards consider a numerical example. Suppose $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$. Assume monitoring occurs at period $m = 10$. Since a nominal decrease in b_t for $t < m$ due to monitoring is $\delta^{m-t}(1-\lambda)^{m-t-1}\lambda b_m$, the principal now makes rewards for success smaller for periods $t = 1, ..., 9$, as reflected in Figure 5 below.

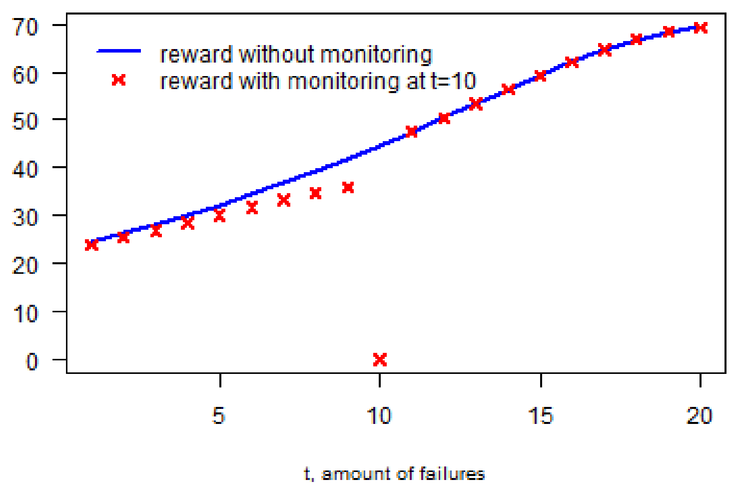

Figure 5. Rewards with monitoring at $t = 10$ when success is publicly observed with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$.

Proposition 2. When success cannot be hidden, monitoring is optimal at the beginning of the relationship. Moreover, the project is terminated inefficiently early:

$$T^{SB} \leq T^M_{Public} \leq T^{FB}.$$

Proof: See Appendix B.

Why does the static effect dominate when success is publicly observed? First, the static effect eradicates the moral hazard problem at the period of monitoring, and this mitigates the moral hazard rent,[19] which was shown to be strictly increasing. What does the dynamic effect accomplish? It mitigates the learning rent, as it makes shirking a less attractive option for

[19] The principal benefits from the static effect only if the agent succeeds at the period when he is monitored, which in turn is possible only if the project is good.

the agent. The dynamic effect increases only during earlier periods: Monitoring influences these periods, but as parties to a contract become increasingly pessimistic, the dynamic effect decreases. Under the optimal contract, however, the payment structure optimally mitigates the learning problem for every period of the relationship, not just those in which the dynamic effect is increasing, as this induces the agent to exert effort throughout the length of the relationship. That is why when success is publicly observed, the static effect dominates.

We illustrate this section with a numerical example. Suppose $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$. The static effect, $SE_t = \delta^t \beta_0 (1-\lambda)^{t-1} \lambda b_t$, is strictly decreasing, as reflected in Figure 6 below. The dynamic effect, $DE_t = \delta^t \beta_0 \lambda^2 (1-\lambda)^{t-2}(t-1) b_t$, is non-monotonic. For early periods $t \leqslant 6$, the agent has to be paid more to behave, since if he deviates once, he can leverage the fact that he is relatively more optimistic until the deadline. However, as time goes by without success, both parties become more pessimistic, and since the expected value of b_t goes down, the dynamic effect diminishes, as well.

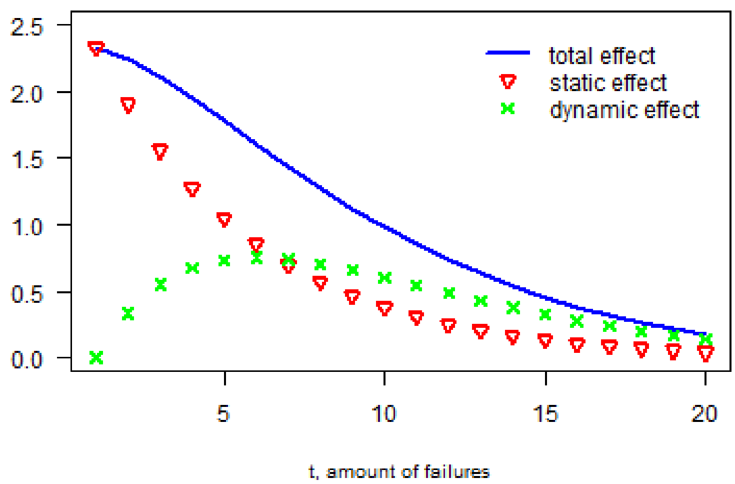

Figure 6. Effect from monitoring when success is publicly observed with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.9$ and $c = 1$.

4.2 Success is privately observed

Suppose the agent can postpone an announcement of a successful implementation. We will denote $T^M_{Private}$ as the duration of the contract with monitoring when success is observed privately. The principal can still benefit from hiring a monitor; however, now the modified reward structure must ensure the agent does not have incentives to postpone or hide success. This is where the additional *IC* constraints:

(IC) $$b_t \geqslant w_t + \delta b_{t+1} \text{ for } t = 1, ..., T-1,$$

$$b_t \geqslant w_t \text{ for } t = 1, ..., T,$$

become relevant and, as we will demonstrate, will be binding for the period when monitoring is implemented.

How can the principal benefit from monitoring when the agent observes success privately? First, as with public observability, the principal can pay less during the monitoring period because the moral hazard problem at period m vanishes. The static effect, however, is different. The reason is that when the agent announces success he takes into account not only the payment tied to success in this precise period but also payments tied to success in all future periods: in case the discounted value of the promised reward in the very next period exceeds the current reward, then the agent will postpone an announcement. For example, if the principal sets $b_m = 0$, then in the case the agent succeeds at this exact period, he will postpone an announcement until the later period as then he gets a positive reward. Given that the optimal contract without monitoring exhibits a decreasing discounted reward value, the principal can decrease the reward in one period at most up to the discounted value of the reward in the following period only. As the discount factor increases, the static effect becomes smaller for all periods except the very last one. Thus, the static effect now has to be modified and becomes: $SE_m = \delta^m \beta_0 (1-\lambda)^{m-1} \lambda (b_m - \delta b_{m+1})$, where $b_{T^M_{Private}+1} = 0$.

In addition, the dynamic effect, has to be modified to take into account the *IC* constraint, as well.

As in the previous case, we will first consider an example where the relationship lasts for two periods ($T = 2$), and the principal can perfectly observe the effort level during one period only when the agent privately observes success. The principal's optimization problem is:

$$\max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH) $$\vec{1} \in \arg\max_{\vec{e}} U(\varpi, \vec{e}),$$

(IC) $$b_1 \geqslant w_1 + \delta b_2,$$

$$b_1 \geqslant w_1,$$

$$b_2 \geqslant w_2,$$

(LL) $$b_1,\, b_2,\, w_1,\, w_2 \geqslant 0.$$

First, suppose the principal monitors the agent at $t = 1$. In this case, the *MH* constraint could be replaced by:

(MH_2) $$\lambda \tilde{\beta} b_2 + (1 - \lambda \tilde{\beta}) w_2 - c \geqslant w_2,$$

(IC) $$b_1 \geqslant w_1 + \delta b_2,$$

that ensures that the agent behaves at $t=2$, given that success has not been achieved at period $t=1$ and, in addition, that he does not postpone an announcement of success from the first period. Then, a solution to the optimization problem involves $w_1 = w_2 = 0$, $b_2 = \frac{c}{\lambda\tilde{\beta}}$ and $b_1 = \delta b_2 = \delta\frac{c}{\lambda\tilde{\beta}}$ where $\tilde{\beta} = \dfrac{\beta_0(1-\lambda)}{\beta_0(1-\lambda)+1-\beta_0}$. The principal's expected profit in this case becomes:

$$\begin{aligned}\pi_{m=1} &= \beta_0\textstyle\sum_{t=1}^{2}\delta^t(1-\lambda)^{t-1}\lambda V - \delta^2 c(1-\lambda\beta_0) - \delta^2\lambda\beta_0\frac{c}{\lambda\tilde{\beta}} - \gamma \\ &= \beta_0\textstyle\sum_{t=1}^{2}\delta^t(1-\lambda)^{t-1}\lambda V - \delta^2 c(1-\lambda\beta_0+\frac{\beta_0}{\tilde{\beta}}) - \gamma.\end{aligned}$$

Second, suppose the principal hires a monitor at the second period. In this case, the MH constraint could be replaced solely by:

$$(MH_1)\ \lambda\beta_0 b_1 + (1-\lambda\beta_0)w_1 - c \geqslant w_1,$$

which ensures that the agent behaves at $t=1$, given that he will exert effort at $t=2$. Note that in this case, the principal does not have to pay anything at the final period, since the agent cannot benefit from hiding his early success.[20] The solution to the optimization problem involves $w_2 = 0$ and $b_1 = \frac{c}{\lambda\beta_0}$. The principal's expected profit is the following:

$$\pi_{m=2} = \beta_0\textstyle\sum_{t=1}^{2}\delta^t(1-\lambda)^{t-1}\lambda V - \delta\beta_0\lambda b_1 - \gamma = \beta_0\sum_{t=1}^{2}\delta^t(1-\lambda)^{t-1}\lambda V - \delta c - \gamma.$$

It is optimal to monitor at $t=2$ if $\delta^2 c(1-\lambda\beta_0+\frac{\beta_0}{\tilde{\beta}}) > \delta c$ or, equivalently, when:

$$\delta > \tfrac{1-\lambda}{(2-\lambda)(1-\lambda\beta_0)},$$

whereas if $\delta < \frac{1-\lambda}{(2-\lambda)(1-\lambda\beta_0)}$, monitoring is performed optimally at the beginning of the relationship. The intuition is straightforward: If monitoring occurs at $t=1$, the principal has to pay a reward to ensure the agent does not postpone announcing success, whereas if monitoring occurs at the final period, no positive reward is needed.

We will demonstrate that when the agent observes success privately, monitoring at the end of the relationship is always optimal when the discount factor is large enough. First, we showed that the dynamic effect is proportional to the change in the reward promised for success during the monitoring period. This is because with private observability, the reduction in the learning rent is smaller and decreases further as the discount increases. Since the agent cannot benefit from hiding success at the final period, the principal can fully eliminate the moral hazard problem in this period. As the discount factor increases enough,

[20] We assume that if the agent is indifferent between announcing success and postponing this announcement, he would choose the former always.

the benefit of a smaller reward at the final period increases due to the static effect. We summarize results in Proposition 3.

To illustrate how monitoring changes rewards when success is private consider a numerical example. Suppose $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.86$ and $c = 1$. Assume monitoring occurs at period $m = 6$: the principal now makes rewards for success smaller for periods $t = 1, ..., 6$, as reflected in Figure 7 below.

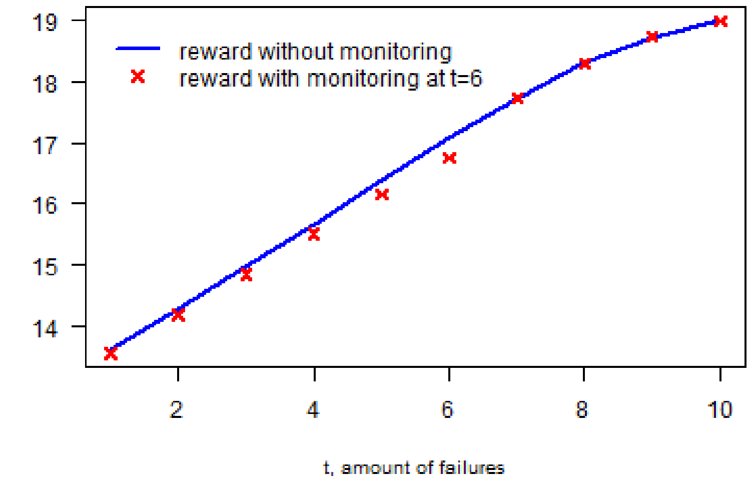

Figure 7. Rewards with monitoring at $t = 6$ when success is private with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.86$ and $c = 1$.

As can be seen, the benefit from monitoring are decreasing in the discount factor. For example, when $\delta = 1$ all the *IC* constraints are binding and monitoring at $t = 6$ becomes completely ineffectual, as reflected in Figure 8 below.

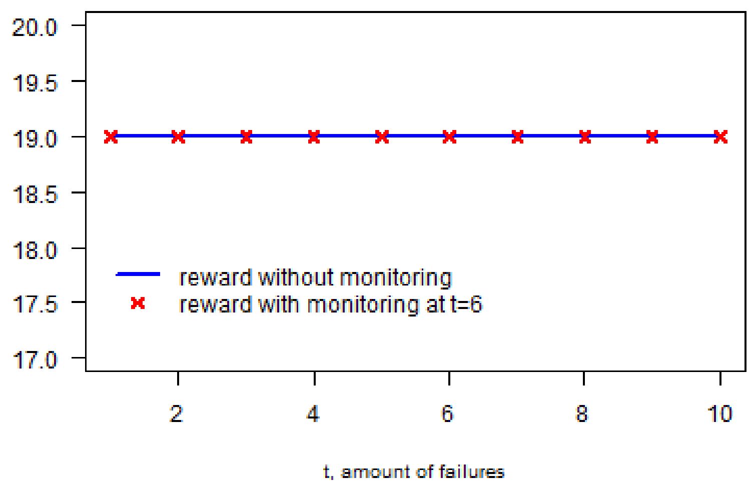

Figure 8. Rewards with monitoring at $t = 6$ when success is private with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 1$ and $c = 1$.

Since now rewards for periods before monitoring are smaller a question arises whether the agent will postpone announcement of success during later periods. Importantly, a nominal decrease in rewards due to the dynamic effect is strictly increasing in time; the agent cannot benefit from postponing early success and announcing it at some period before he is monitored. However, the question remains if the agent will announce success at some period during or after monitoring occurs. We will show that with the modified reward structure the agent will not postpone announcement of success during any later period of the contract.

Proposition 3. When the agent observes success privately the optimal time for monitoring is affected by patience. If the discount factor is high enough, monitoring is used optimally at the end of the relationship. The project is terminated inefficiently early:

$$T^{SB} \leqslant T^{M}_{Private} \leqslant T^{M}_{Public} \leqslant T^{FB}.$$

Proof: See Appendix C.

Consider an example in Figure 9 where the discount factor is not high enough for the monitoring to occur optimally at the end of the relationship.

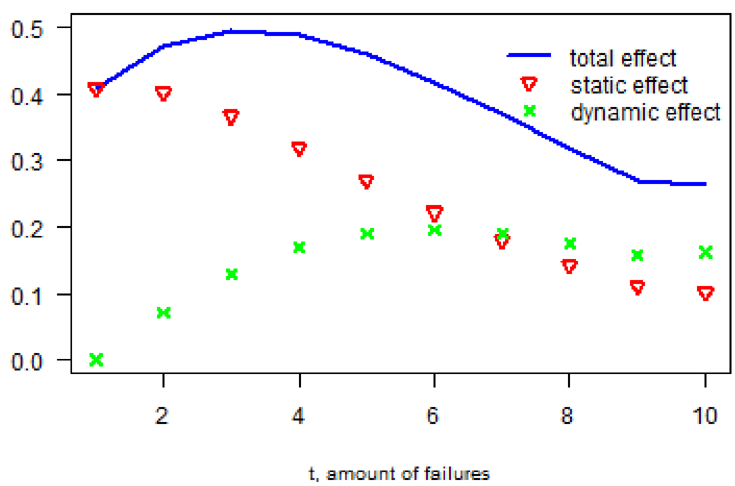

Figure 9. Effect from monitoring when success is private with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.86$ and $c = 1$.

When $\delta = 0.86$, monitoring is employed optimally at the third period. However, if we increase the discount factor up to $\delta = 0.92$, monitoring will be employed optimally toward the end of the relationship. In this case, both the static and dynamic effects are non-monotonic, as Figure 10 demonstrates below.

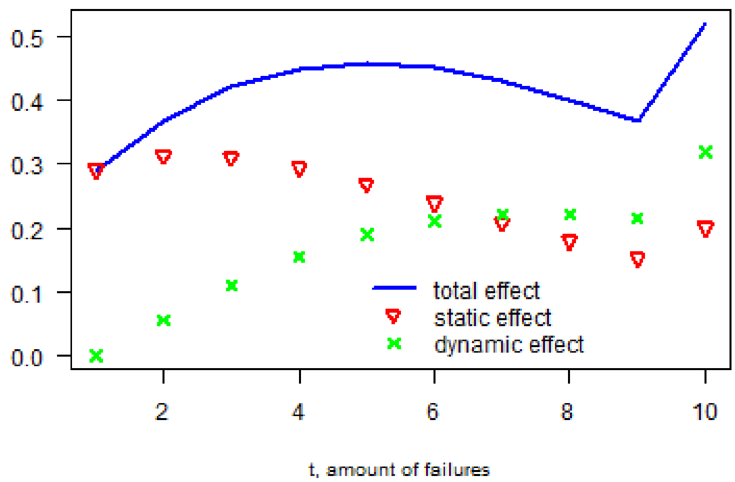

Figure 10. Effect from monitoring when success is private with $\lambda = 0.15$, $\beta_0 = 0.7$, $\delta = 0.92$ and $c = 1$.

In Proposition 3, we proved that when the discount factor is high enough, the monitor is hired optimally at the end of the relationship. Specifically, this implies that when $\delta = 1$, monitoring is optimal at the end of the relationship. However, for a smaller discount factor, optimal timing of monitoring can occur toward the beginning of the relationship. If the discount factor is small enough, monitoring can even be implemented optimally at the start of the relationship.

Throughout the paper we assumed that the monitor is hired for one period only. Nevertheless, our results could be extended easily to examine a case in which it is optimal to hire the monitor for several periods. Consider first the case in which success is publicly observed. Since we proved in Proposition 1 that the total effect of monitoring is strictly decreasing, monitoring is optimal in earlier periods of the relationships. This result is intuitive and supported by real-life observations: The typical financing cycle of a start-up firm in its earlier stages include relationship financing, in which the entrepreneur and the investor share common beliefs regarding the quality of the project because, for example, the entrepreneur spends time on-site, which is a form of monitoring. In later periods of a start-up firm's financing, parties to a contract shift to an arm's-length relationship, and the investor commits to halt funding if the project is not successfully implemented by a specific deadline.

When success is privately observed and the monitor is hired for several periods, the periods of monitoring might be not adjacent; that is, monitoring may occur during some periods, suspended for several periods, then reintroduced during later periods.

The results of our research have empirical implications. In the case that hiding success from the principal is prohibitively costly, if the agent formed the original idea for the project, and investors are competing to fund the project, monitoring should be performed at the end of the relationship. However, if the agent is hired by the owner of the project, monitoring is employed optimally at the beginning of the relationship. If success is enormously costly to observe, monitoring optimally is performed toward the end of the relationship if parties to the contract are patient enough. Thus, we emphasize the pivotal role of private observability and market structure on the optimal monitoring timing.

5 Conclusion

This paper examines optimal contracts and the role of monitoring for experimentation in settings with moral hazard and private observability of success. In the benchmark case without monitoring, we have found that the agent is rewarded only if he succeeds; otherwise, he receives nothing. The nominal value of the payment is strictly increasing to account for increasing pessimism, whereas the discounted values of the optimal rewards decrease over time. Thus, the agent's private observation of success is irrelevant since he will never postpone an announcement of success.

Nevertheless, we demonstrate that private observability factors into the optimal time for monitoring. When success is impossible to hide from the principal, monitoring at the beginning of the relationship improves the efficiency of financial contracting. This contrasts Bergemann and Hege's (1998) results, which demonstrated that monitoring is optimal toward the end of the project. The authors considered a version of our model in which the agent is the owner of the project and raises funds in a competitive market. In their model, the agent's reward for earlier periods can rise or fall with at most one extremum, whereas in our paper, a nominal value of the reward for success is always increasing.

At the same time, when the agent observes success privately, patience influences the optimal time for monitoring. When the discount factor increases, the immediate effect of monitoring decreases, as the principal must pay a high enough reward to prevent the agent from hiding success. As the discount factor increases enough, the principal must promise almost identical rewards for success at every period except the final one, making monitoring more valuable at the end of the relationship and supporting Bergemann and Hege's (1998) result.

Throughout the paper, we assumed that the monitor is honest and does not collude with the agent. An interesting avenue for future research is to study the optimal time for monitoring when the monitor could be colluding with the agent. For example, a pharmaceutical company hires an agent who may considering hiding the results of the clinical trials and reselling them to a rival company. Since the agent's reward is decreasing, this feature could postpone the optimal time for monitoring to later periods.

6 Appendix

6.1 Appendix A

Proof of Proposition 1. The principal's optimization problem is the following:

$$[P^{SB}] \ \max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH) $$\vec{1} \in \arg\max_{\vec{e}} U(\varpi, \vec{e}),$$

(IC) $$b_t \geqslant w_t + \delta b_{t+1} \text{ for } t = 1, ..., T-1,$$

$$b_t \geqslant w_t \text{ for } t = 1, ..., T,$$

(LL) $$b_t,\, w_t \geqslant 0 \text{ for } t = 1, ..., T.$$

We first solve an auxiliary problem $P1$, where the global MH constraint is replaced by a sequence of MH_t constraints for $t = 1, ..., T$ that ensure the agent does not want to deviate at period t, given that he was behaving in all prior periods $s < t$ and will work in all subsequent periods $s > t$. In addition, $P1$ ignores the IC constraint, which will demonstrate automatic satisfaction.

The optimization problem $P1$ is:

$$[P1] \ \max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH_t) $b_t - w_t \geqslant \dfrac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s + (1-\lambda)w_s - c)$ for $t = 1, ..., T$,
(LL) $b_t,\, w_t \geqslant 0$ for $t = 1, ..., T$.

Lemma 1. The following payment sequence solves $P1$:

$$w_t = 0 \text{ and } b_t = \frac{c}{\lambda\tilde{\beta}_t} + c\sum_{s=1}^{T-t} \delta^s \tfrac{1-\beta_0}{\beta_0(1-\lambda)^{t+s-1}} \text{ for } 1 \leqslant t \leqslant T.$$

Proof: Note that increasing w_t makes it more difficult to satisfy the MH_t constraints and lessens the objective function. As a result, the optimal solution must have $w_t = 0$ for $1 \leqslant t \leqslant T$, and the problem can be rewritten as:

$$[P1] \ \max_{\varpi} \pi(\varpi, \vec{1}) \text{ subject to}$$

(MH_t) $b_t \geqslant \dfrac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)$ for $t = 1, ..., T$,
(LL) $b_t \geqslant 0$ for $t = 1, ..., T$.

The auxiliary problem $P1$ has the Lagrangian:

$$L = \beta_0 \sum_{t=1}^{T} \delta^t (1-\lambda)^{t-1}\lambda(V - b_t)$$
$$+ \sum_{t=1}^{T} \mu_t\Big(b_t - \frac{c}{\lambda\tilde{\beta}_t} - \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)\Big) + \sum_{t=1}^{T} \xi_t b_t.$$

The Kuhn-Tucker conditions for the optimization problem are:
$[b_t] : -\beta_0\delta^t(1-\lambda)^{t-1}\lambda + \mu_t - \sum_{j=1}^{t-1}\mu_j\delta^{t-j}(1-\lambda)^{t-j-1}\lambda + \xi_t = 0$ for $1 \leqslant t \leqslant T$, complemented by the constraints of the problem and the corresponding complementary slackness conditions.

If $\xi_t > 0$ for some $1 \leqslant t \leqslant T$, then MH_t would be violated and, as a result, we must have $\xi_t = 0$ for $t = 1, ..., T$.

Consider the first-order conditions with respect to b_t:

$t = 1$: $-\beta_0\delta\lambda + \mu_1 = 0 \Longrightarrow \mu_1 = \beta_0\delta\lambda$;

$t = 2$: $-\beta_0\delta^2\lambda(1-\lambda) + \mu_2 - \mu_1\delta\lambda = 0 \Longrightarrow \mu_2 = \beta_0\delta^2\lambda > 0$;

repeating this procedure until the final period T, we have:

$t = T$: $\mu_T = \beta_0\delta^T\lambda > 0$.

Thus, all MH_t constraints must be binding.

First, we will prove that with $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + c\sum_{s=1}^{T-t}\delta^s\frac{1-\beta_0}{\beta_0(1-\lambda)^{t+s-1}}$, it is the case that:

$$b_t = \frac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c) \text{ for } 1 \leqslant t \leqslant T.$$

For $t = T$, the equality $b_t = \dfrac{c}{\lambda\tilde{\beta}_T}$ trivially follows from the proposed formula itself. For any $t < T$, assume that MH_s holds for $s = t+1$; that is:

$$b_{t+1} = \frac{c}{\lambda\tilde{\beta}_{t+1}} + \sum_{s=t+2}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c).$$

We need to show that $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)$, or, using the line above:

$$b_t = \frac{c}{\lambda\tilde{\beta}_t} + \delta(\lambda b_{t+1} - c) + \sum_{s=t+2}^{T}\delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)$$
$$= \frac{c}{\lambda\tilde{\beta}_t} + \delta(\lambda b_{t+1} - c) + (b_{t+1} - \frac{c}{\lambda\tilde{\beta}_{t+1}}) = \frac{c}{\lambda\tilde{\beta}_t} - \frac{c}{\lambda\tilde{\beta}_{t+1}} - \delta c - (1+\delta\lambda)b_{t+1}.$$

Since $b_{t+1} = \dfrac{c}{\lambda\tilde{\beta}_{t+1}} + \frac{c(1-\beta_0)}{\beta_0(1-\lambda)^t}\frac{\delta}{(1-\delta-\lambda)}(1-(\frac{\delta}{1-\lambda})^{T-t-1})$ and $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + \frac{c(1-\beta_0)}{\beta_0(1-\lambda)^{t-1}}\frac{\delta}{(1-\delta-\lambda)}(1-(\frac{\delta}{1-\lambda})^{T-t})$, it suffices to show that:

$$\frac{c}{\lambda\tilde{\beta}_t} + \frac{c(1-\beta_0)}{\beta_0(1-\lambda)^{t-1}}\frac{\delta}{(1-\delta-\lambda)}(1-(\frac{\delta}{1-\lambda})^{T-t})$$
$$= \frac{c}{\lambda\tilde{\beta}_t} - \frac{c}{\lambda\tilde{\beta}_{t+1}} - \delta c - (1+\delta\lambda)(\frac{c}{\lambda\tilde{\beta}_{t+1}} + \frac{c(1-\beta_0)}{\beta_0(1-\lambda)^t}\frac{\delta}{(1-\delta-\lambda)}(1-(\frac{\delta}{1-\lambda})^{T-t-1})),$$

which is easily verified for any $1 \leqslant t \leqslant T$. *Q.E.D.*

We will demonstrate that with the proposed solution, any period is optimal for the agent to work in, regardless of previous effort history profile. Consider the final period T. Note that if the agent deviates and shirks his duties at some arbitrary period $t < T$, he only can be more optimistic at period T. Thus, for any history of prior effort, the current belief β_T can be higher only than $\tilde{\beta}_T$. Now MH_T $\lambda\tilde{\beta}_T b_T = c$ is satisfied since $\tilde{\beta}_T \geqslant \beta_T$ and $\lambda\beta_T b_T \geqslant c$. Next, assume that working in any period is optimal for the agent, regardless of the previous effort history profile at period $t+1 \leqslant T$. Consider period t as any history of prior effort with current beliefs β_t. Since we already showed that $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)$, for any $\beta_t \geqslant \tilde{\beta}_t$ it is apparent that $b_t \geqslant \dfrac{c}{\lambda\beta_t} + \sum_{s=t+1}^{T} \delta^{s-t}(1-\lambda)^{s-t-1}(\lambda b_s - c)$, and working is optimal for the agent.

Finally, it can be shown by induction that any reward profile that makes every MH_t constraint binding must coincide with $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + c\sum_{s=1}^{T-t} \delta^s \frac{1-\beta_0}{\beta_0(1-\lambda)^{t+s-1}}$ for $1 \leqslant t \leqslant T$. Moreover, since we already proved that working for the agent is optimal in any period, regardless of the previous effort history profile, this also ensures that the agent would find it optimal to work in period t for any possible effort profile before t.

Recall that when solving $P1$, we ignored the IC constraint. Since $w_t = 0$ and $b_t > 0$, the proposed solution obviously satisfies $b_t \geqslant w_t$ for $t = 1, ..., T$. Thus the final condition we must check is $b_t \geqslant w_t + b_{t+1}$ for $t = 1, ..., T-1$.

Given that $\sum_{s=1}^{T-t} \frac{\delta^{s-t}}{(1-\lambda)^{s-1}} = \frac{1-(\frac{\delta}{1-\lambda})^{T-t}}{1-\frac{\delta}{1-\lambda}}$, by performing some algebra, one could verify that:

$$b_t = \frac{c}{\lambda\tilde{\beta}_t} + c\delta\frac{1-\beta_0}{\beta_0}\frac{1-(\frac{\delta}{1-\lambda})^{T-t}}{(1-\lambda)^{t-1}(1-\lambda-\delta)} = \frac{c}{\lambda\tilde{\beta}_t} + c\delta\frac{1-\beta_0}{\beta_0}\frac{(1-\lambda)^{T-t}-(1-\delta)^{T-t}}{(1-\lambda)^{T-1}(1-\lambda-\delta)}.$$

Then, it follows that $b_t \geqslant \delta b_{t+1}$ if and only if:

$$\frac{c}{\lambda}\left(\frac{\beta_0(1-\lambda)^{t-1}+1-\beta_0}{\beta_0(1-\lambda)^{t-1}} - \delta\frac{\beta_0(1-\lambda)^{t}+1-\beta_0}{\beta_0(1-\lambda)^{t}}\right) \geqslant \frac{c\delta(1-\beta_0)(\delta(1-\lambda)^{T-t-1}-\delta^{1+T-t-1}-(1-\lambda)^{T-t}+\delta^{T-t})}{\beta_0(1-\lambda)^{T-1}(1-\lambda-\delta)},$$
$$\frac{c(\beta_0(1-\lambda)^{t}(1-\delta)+(1-\beta_0)(1-\lambda-\delta))}{\lambda\beta_0(1-\lambda)^{t}} \geqslant -\frac{\delta c(1-\beta_0)}{\beta_0(1-\lambda)^{t}},$$

which holds as an equality if $\delta = 1$ (and as a strict inequality as long as $\delta < 1$) for any t.

Thus, $w_t = 0$ and $b_t = \dfrac{c}{\lambda\tilde{\beta}_t} + c\sum_{s=1}^{T-t} \delta^s \frac{1-\beta_0}{\beta_0(1-\lambda)^{t+s-1}}$ for $1 \leqslant t \leqslant T$ is a solution to the principal's optimization problem.

Finally, since $\lambda\tilde{\beta}_t b_t > c$, for all $1 \leqslant t \leqslant T-1$, the project is terminated inefficiently early, $T^{SB} \leqslant T^{FB}$. *Q.E.D.*

6.2 Appendix B

Proof of Proposition 2. We will call T^M_{Public} the duration of the contract when the principal performs monitoring and success is observed publicly. The total benefit, TE_m, from monitoring is a sum of the static and dynamic effects. The static effect from monitoring at period m, SE_m is:

$$SE_m = \delta^m \beta_0 (1-\lambda)^{m-1} \lambda b_m,$$

where $b_m = \dfrac{c}{\lambda \tilde{\beta}_m} + c \sum_{s=1}^{T-m} \delta^s \frac{(1-\beta_0)}{\beta_0 (1-\lambda)^{m+s-1}}$.

The dynamic effect is:

$$DE_m = \sum_{t=1}^{m-1} \delta^t Prob(\text{success at } t < m)[\text{nominal decrease in } b_t].$$

First, we need to define a nominal decrease in b_t for $t < m$ that is possible because of monitoring that will occur at period m. Recall that the optimal payment structure makes all MH_t constraint binding or, equivalently,

$$b_t = \frac{c}{\lambda \tilde{\beta}_t} + \sum_{s=t+1}^{T} \delta^{s-t} (1-\lambda)^{s-t-1} (\lambda b_s - c),$$

and by decreasing a reward in a certain period m (in the right-hand side), the principal can decrease a reward in the left-hand side. Thus, a nominal decrease in b_t for $t < m$ is:

$$\delta^{m-t} (1-\lambda)^{m-t-1} \lambda b_m.$$

As a result, the dynamic effect becomes:

$$DE_m = \sum_{t=1}^{m-1} \delta^t (\beta_0 (1-\lambda)^{t-1} \lambda) \delta^{m-t} (1-\lambda)^{m-t-1} \lambda b_m = \delta^m \beta_0 \lambda^2 (1-\lambda)^{m-2} (m-1) b_m.$$

We can then calculate the total effect from monitoring at period m:

$$\begin{aligned} TE_m &= \delta^m \beta_0 (1-\lambda)^{m-1} \lambda b_m + \delta^m \beta_0 \lambda^2 (1-\lambda)^{m-2} (m-1) b_m \\ &= \delta^m \beta_0 (1-\lambda)^{m-2} \lambda (1-\lambda+\lambda(m-1)) b_m. \end{aligned}$$

We will prove that TE_m is strictly decreasing in m. First, recall form Proposition 1 that b_t was chosen optimally such that $b_t \geqslant \delta b_{t+1}$ for $t = 1, ..., T-1$, and, as a result, $\delta^m b_m$ is decreasing in m.

It suffices to show that $\varphi(m) = (1-\lambda)^{m-2} \lambda (1-\lambda+\lambda(m-1))$ is decreasing in m as well. Notice that $\varphi(1) = \varphi(2) = 1$.

Because $\frac{d\varphi(m)}{dm} = (1-\lambda)^{m-2} (\lambda + (1-\lambda+\lambda(m-1)) ln(1-\lambda))$, it is sufficient to show that $f(\lambda, m) = \lambda + (1-\lambda+\lambda(m-1)) ln(1-\lambda)$ is negative for $m > 1$. Note that $\frac{\partial f}{\partial m} = \lambda ln(1-\lambda) < 0$ for any m and $f(\lambda, 2) = \lambda + ln(1-\lambda)$. Consider $g(\lambda) = f(\lambda, 2) = \lambda + ln(1-\lambda)$, with $\frac{\partial g}{\partial \lambda} = -\frac{\lambda}{1-\lambda}$ being negative for all $0 < \lambda < 1$ and $\lim_{\lambda \to +0} g(\lambda) = 0$. Thus, $f(\lambda, m)$ is negative, and, as a result, $\varphi(m)$ is decreasing in m.

Since TE_m is decreasing in m, monitoring is implemented optimally at the very first period. Given that the principal can promise paying less, he can use these funds to extend the duration of the relationship up to T^M_{Public} and, as a result, $T^{SB} \leqslant T^M_{Public}$. Since the agent still receives a positive rent, the project is still terminated inefficiently early: $T^M_{Public} \leqslant T^{FB}$.

Q.E.D.

6.3 Appendix C

Proof of Proposition 3. We will call $T^M_{Private}$ the duration of the contract when the agent observes success privately and the principal performs monitoring. As in the case when success is publicly observed, the total effect, TE_m, of monitoring is a sum of the static and dynamic effects. However, for this case, we redefine both effects to account for the possibility of hiding success. In particular, the additional *IC* constraints become relevant:

$$b_t \geqslant w_t + \delta b_{t+1} \text{ for } t \leqslant T^M_{Private} - 1,$$

$$b_t \geqslant w_t \text{ for } t \leqslant T^M_{Private}.$$

Since we proved that all $w_t = 0$, the second constraint will not affect either of the two effects, whereas the first constraint will limit the amount of money the principal will save by monitoring, except during the final period of the relationship. The static effect of monitoring at period m, SE_m is:

$$SE_m = \delta^m \beta_0 (1-\lambda)^{m-1} \lambda (b_m - \delta b_{m+1}).$$[21]

To simplify notation, we will define function η_m as follows:

$$\eta_m = \begin{cases} b_m - \delta b_{m+1} & 1 \leqslant m < T^M_{Private} \\ b_m & m = T^M_{Private} \end{cases}.$$

Thus, $SE_m = \delta^m \beta_0 (1-\lambda)^{m-1} \lambda \eta_m$ for $t = 1, ..., T^M_{Private}$.

The dynamic effect is then:

$$DE_m = \sum_{t=1}^{m-1} \delta^t (\beta_0 (1-\lambda)^{t-1} \lambda) \delta^{m-t} (1-\lambda)^{m-t-1} \lambda \eta_m = \delta^m \beta_0 \lambda^2 (1-\lambda)^{m-2} (m-1) \eta_m.$$

Thus, the total effect becomes

$$TE_m = \delta^m \beta_0 (1-\lambda)^{m-2} \lambda (1 - \lambda + \lambda(m-1)) \eta_m = \delta^m \beta_0 \lambda \varphi(m) \eta_m.$$

Before we define the optimal timing of monitoring, consider agent's incentives to postpone announcement of success. Since now rewards for periods before monitoring are smaller some of the *IC* constraints might be violated. First, a nominal decrease in a reward in period t due to the dynamic effect, $\delta^{m-t}(1-\lambda)^{m-t-1}\lambda\eta_m$, is increasing in time; that is, $\frac{\delta^m(1-\lambda)^m\lambda\eta_m}{\delta^t(1-\lambda)^{t+1}}$ is strictly increasing in t. Recall from Proposition 1 that rewards without monitoring were (weakly) decreasing in time. Given this, rewards before monitoring period will be strictly decreasing in time as well and, as a result, the agent cannot benefit from postponing early success and announcing it at some period before he is monitored.

[21] For convenience, we will say that $b_{T^M_{Private}+1} = 0$.

In addition, the discounted value of rewards for success after the period when the monitor is hired are decreasing as well by Proposition 1 since they are not modified. However, the agent might postpone an announcement till the period of monitoring. Given that monitoring occurs at period m the reward for success at this period is δb_{m+1} so that the agent does not postpone success in case it is achieved exactly at period m. Thus, it suffices to guarantee that

$$b_{m-1} - \delta^{m-t}(1-\lambda)^{m-t-1}\lambda\eta_m \geqslant \delta^2 b_{m+1} \text{ for } t = m-1,$$

which can be rewritten as

$$b_{m-1} - \delta\lambda\eta_m \geqslant \delta(b_m - \eta_m),$$

and will be shown to be satisfied for any m in Lemma 2.

Lemma 2. $b_{m-1} - \delta\lambda\eta_m \geqslant \delta(b_m - \eta_m)$ for $1 \leqslant m \leqslant T^M_{Private}$.

Proof: Recall that from Proposition 1, it follows that b_m is increasing strictly, whereas δb_{m+1} is decreasing strictly. In particular, for $t = 1, ..., T^M_{Private} - 1$:

$$\begin{gathered}\eta_m = b_m - \delta b_{m+1} = \\ \tfrac{c}{\lambda}\left(\tfrac{\beta_0(1-\lambda)^{t-1}+1-\beta_0}{\beta_0(1-\lambda)^{t-1}} - \delta\tfrac{\beta_0(1-\lambda)^{t}+1-\beta_0}{\beta_0(1-\lambda)^{t}}\right) - \tfrac{c\delta(1-\beta_0)(\delta(1-\lambda)^{T-t-1}-\delta^{1+T-t-1}-(1-\lambda)^{T-t}+\delta^{T-t})}{\beta_0(1-\lambda)^{T-1}(1-\lambda-\delta)} \\ = \tfrac{c(\beta_0(1-\lambda)^t(1-\delta)+(1-\beta_0)(1-\lambda-\delta))}{\lambda\beta_0(1-\lambda)^t} + \tfrac{\delta c(1-\beta_0)}{\beta_0(1-\lambda)^t} = \tfrac{c(1-\delta)(\beta_0(1-\lambda)^{t-1}+1-\beta_0)}{\lambda\beta_0(1-\lambda)^{t-1}}.\end{gathered}$$

As a result, η_m evolves as follows:

$$\eta_m = \begin{cases} (1-\delta)\frac{c(\beta_0(1-\lambda)^{m-1}+1-\beta_0)}{\lambda\beta_0(1-\lambda)^{m-1}} & 1 \leqslant m < T^M_{Private} \\ \frac{c(\beta_0(1-\lambda)^{m-1}+1-\beta_0)}{\lambda\beta_0(1-\lambda)^{m-1}} & m = T^M_{Private} \end{cases}.$$

Note that $b_{m-1} - \delta\lambda\eta_m \geqslant \delta(b_m - \eta_m)$ can be rewritten using the definition of η_m as follows:

$$\eta_{m-1} \geqslant \delta(1-\lambda)\eta_m,$$

and simplifying we have

$$\begin{gathered}(1-\delta)\tfrac{c(\beta_0(1-\lambda)^{m-2}+1-\beta_0)}{\lambda\beta_0(1-\lambda)^{m-2}} \geqslant \delta(1-\lambda)(1-\delta)\tfrac{c(\beta_0(1-\lambda)^{m-1}+1-\beta_0)}{\lambda\beta_0(1-\lambda)^{m-1}}, \\ \beta_0(1-\lambda)^{m-2} + 1 - \beta_0 \geqslant \delta(\beta_0(1-\lambda)^{m-1} + 1 - \beta_0), \\ \beta_0(1-\lambda)^{m-2}(1-\delta(1-\lambda)) + (1-\delta)(1-\beta_0) \geqslant 0.\end{gathered}$$

which holds as an equality in case $\delta = 1$ (and as a strict inequality as long as $\delta < 1$) for any period of monitoring $1 \leqslant m \leqslant T^M_{Private}$. *Q.E.D.*

First, consider the case when $\delta = 1$. From Proposition 1, it follows that optimal b_m is constant over time, and, as a result, $\eta_m = 0$ for $t \leqslant T^M_{Private} - 1$. For the final period, however, $\eta_{T^M_{Private}} = b_{T^M_{Private}} > 0$. Clearly, monitoring at the final period is optimal.

Consider now the case for $\delta < 1$.

$$TE_m = \delta^m \beta_0 (1-\lambda)^{m-2} \lambda (1-\lambda+\lambda(m-1))\eta_m$$

$$= \begin{cases} (1-\delta)c\frac{\delta^m(\beta_0(1-\lambda)^{m-1}+1-\beta_0)(1-\lambda+\lambda(m-1))}{1-\lambda} & 1 \leqslant m < T^M_{Private} \\ c\frac{\delta^m(\beta_0(1-\lambda)^{m-1}+1-\beta_0)(1-\lambda+\lambda(m-1))}{1-\lambda} & m = T^M_{Private} \end{cases}.$$

We define time period $1 \leqslant j \leqslant T^M_{Private}$ as follows:

$$j \in \arg\max_{1 \leqslant s < T^M_{Private}} c\frac{\delta^s(\beta_0(1-\lambda)^{s-1}+1-\beta_0)(1-\lambda+\lambda(s-1))}{1-\lambda}.$$ [22]

We will show that with a high enough discount factor, monitoring at the final period is always optimal. First, there is a (unique) $\tilde{\delta}$, such that:

$$(1-\tilde{\delta})c\frac{\delta^j(\beta_0(1-\lambda)^{j-1}+1-\beta_0)(1-\lambda+\lambda(j-1))}{1-\lambda} = c\frac{\delta^{T^M_{Private}}(\beta_0(1-\lambda)^{T^M_{Private}-1}+1-\beta_0)(1-\lambda+\lambda(T^M_{Private}-1))}{1-\lambda}$$

$$\text{or, equivalently, } \tilde{\delta} = 1 - \frac{c\frac{\delta^{T^M_{Private}}(\beta_0(1-\lambda)^{T^M_{Private}-1}+1-\beta_0)(1-\lambda+\lambda(T^M_{Private}-1))}{1-\lambda}}{c\frac{\delta^j(\beta_0(1-\lambda)^{j-1}+1-\beta_0)(1-\lambda+\lambda(j-1))}{1-\lambda}} \leqslant 1.$$

For any $\delta > \tilde{\delta}$ the total effect of monitoring at period m, TE_m achieves its highest value at the final period $T^M_{Private}$, whereas when $\delta < \tilde{\delta}$ monitoring is implemented at period j, which, in general, is not the final period. As the discount factor increases, both the static and dynamic effects diminish, and monitoring is implemented optimally at the final period. Since the principal promises paying less, the money that he saves could be used to extend the duration of the relationship beyond $T^M_{Private}$. In comparison with the case in which success is observed publicly, the principal saves less money; however, the second-best outcome is improved marginally. As a result, $T^{SB} \leqslant T^M_{Private} \leqslant T^M_{Public} \leqslant T^{FB}$. *Q.E.D.*

[22] Since $T^M_{Private} < \infty$, a time period j is well defined, although it might be not unique due to the non-monotonicity of $\delta^m(\beta_0(1-\lambda)^{m-1}+1-\beta_0)(1-\lambda+\lambda(m-1))$ in m.

7 References

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

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

Bergemann D. and Hege U. (2005). The Financing of innovation: learning and stopping. *Rand Journal of Economics* 36 (4), 719-752.

Bergemann D. and Valimaki J. (2010). The dynamic pivot mechanism. *Econometrica* 78 (2), 771-789.

Bonatti A. and Horner J. (2011). Collaborating. *American Economic Review*, 101, 632-663.

Bonatti A. and Horner J. (2016). Career concerns with exponential learning. *Theoretical Economics*, forthcoming.

Chade H. and Kovrijnykh N. (2016). Delegated information acquisition with moral hazard. *Journal of Economic Theory* 162, 55-92.

DiMasi J., Hansen R., and Grabowski H. (2003). The price of innovation: new estimates of drug development costs. *Journal of Health Economics* 22: 151-85.

Eisenstein E., Lemons P., Tardiff B., Schulman K., Jolly M. and Califf R. (2005). Reducing the costs of phase III cardiovascular clinical trials. *American Heart Journal* 149, 482-88.

Gerardi D. and Maestri L. (2012). A Principal-Agent Model of Sequential Testing. *Theoretical Economics* 7, 425-463.

Gomes R., Gottlieb D. and Maestri L. (2015). Experimentation and project selection: screening and learning. *Games and Economic Behavior, forthcoming.*

Gorman M. and Sahlman W. (1989). What do venture capitalists do? *Journal of Business Venturing* 4(4), 231-248.

Halac M., Kartik N. and Liu Q. (2016). Optimal contracts for experimentation. *Review of Economic Studies* 83(3), 1040-1091.

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

Henry E. (2009). Strategic disclosure of research results: the cost of proving your honesty. *The Economic Journal* 119, 1036-1064.

Holmstrom B. (1979). Moral hazard and observability. *Bell Journal of Economics* 10, 74-91.

Horner J. and Samuelson L. (2014). Incentives for experimenting agents. *RAND*

Journal of Economics 44(4), 632-663.

Inderst R. and Ottaviani M. (2012). Competition through commissions and kickbacks. *American Economic Review* 102, 780-809.

Keller G., Rady S. and Cripps M. (2005). Strategic experimentation with exponential bandits. *Econometrica* 73, 39-68.

Kwon S. (2013). Dynamic moral hazard with persistent states. *Working Paper.*

Lerner J. (1995). Venture Capitalists and the Oversight of Private Firms. *The Journal of Finance* 50(1), 301-318.

Lerner J. and Malmendier U. (2010). Contractibility and the design of research agreements. *American Economic Review* 100, 214-246.

Manso G. (2011). Motivating innovation. *Journal of Finance* 66 (2011), 1823-1860.

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

Myerson R. (1986). Multistage games with communication. *Econometrica* 54 (2), 323-358.

Pakes A. and Nitzan S. (1983). Optimum contracts for research personnel, research employment, and the establishment of "rival" enterprises. *Journal of Labor Economics* 1(4), 345-365.

Pavan A., Segal I. and Toikka J. (2014). Dynamic mechanism design: a Myersonian approach. *Econometrica* 82(2), 601-653.

Sahlman W. (1990). The structure and governance of venture-capital organizations. *Journal of Financial Economics* 27, 473-521.