Contracting for Experimentation and the Value of Bad News*

Sinem Hidir†

May 2017

This paper studies the optimal provision of incentives in order to acquire information about the unknown quality of a project. News arrives in form of a conclusive good signal or inconclusive bad signals while the agent exerts effort and both effort and signals are private information of the agent. The optimal contract contains history dependant payments upon terminal signals and a deadline which gets extended upon the disclosure of nonterminal bad signals, as long as experimentation is efficient. The extensions of the deadline serve as a way of back-loading the payments and mitigate the inefficiency caused by early stopping while keeping the agent's expected rent constant. If there is stopping before a deadline is reached, this happens at the same stopping belief as in the first best benchmark.

Keywords: Dynamic moral hazard, continuous time principal-agent model, experimentation, Poisson arrivals, private signals, Bayesian learning.

JEL Codes: D82, D83, D86, O32.

*This paper forms the first chapter of my PhD thesis defended in Toulouse School of Economics in June 2015. I am grateful to my advisor Jacques Crémer and also Renato Gomes, Christian Hellwig, Thomas Mariotti and Harry Di Pei for many helpful discussions and also benefited from the comments of George Lukyanov, Lucas Maestri, Teck Yong Tan and seminar participants at Toulouse School of Economics, University of Bocconi, University of St Andrews, University of Warwick and the Royal Economic Society Conference (2016).

†University of Warwick. email: s.hidir@warwick.ac.uk.

1 Introduction

Firms often consult experts in order to acquire information about new and untested technologies. Consider a pharmaceutical company who provides funds to a scientist for performing tests on a new drug. The scientist has the duty of testing different compounds to find out about their efficiency or side effects. In such a setting, both positive and negative information arise only as a result of costly research. Similarly, during the development of a new product, market research may reveal a low potential demand making it less profitable to continue which is interpreted as bad news.

In these examples, information acquisition is dynamic and it is ex-ante unknown how long it may take in order to gather sufficient information before taking a decision, which makes it difficult to monitor effort. The owner of the project would like to acquire information as soon as possible and avoid spending resources inefficiently, potentially diverting them into other projects. However, the agent who is getting paid in order to acquire information does not share the same incentives as he wants to continue getting paid. This paper solves for the optimal way in which a principal can provide incentives to an agent in order to gather information and report it in the presence of these informational constraints.

The agency problem considered features both moral hazard and private information. The agent has the choice between investing in information acquisition or shirking and keeping for himself the resources provided by the principal, which does not result in learning. While the agent exerts effort, a credible good signal which immediately reveals that the project is a good one or a bad signal which leads to a more pessimistic belief about the type of the project arrives through a Poisson process. The agent also chooses whether and when to disclose the signals that he receives in the form of hard information. As the agent values remaining in the project, he might be reluctant to disclose bad signals. Also, as a good signal makes it unnecessary for the principal to continue investing in learning, the agent should be provided with sufficient incentives to disclose it. The principal would like be informed as soon as possible while the agent values remaining in the relationship for as long as possible. The problem of the principal is then to find the optimal length and the features of the contract which incentivizes the agent to work and disclose the signals. This consists of allocating the agency rent between payments and continuation values in order to satisfy the agent's incentive constraints while implementing an optimal termination rule.

The main features of the optimal contract are history dependent payments upon termination, an initial deadline and a rule for extending this deadline after bad signal disclosures. The extensions upon bad signals continue as long as the belief remains above the first best. Stopping before a deadline happens upon the *terminal* bad signal[1] or upon a good signal disclosure at any moment. Hence, even though many times experimentation may stop inefficiently early only due to reaching a deadline without learning, in case there is stopping before a deadline is reached, it happens efficiently.

There are two types of incentive constraints of the agent: the *ex-ante* constraint to work and the *ex-post* constraint which make sure he is willing to disclose the signals. An important feature is that the agent's shirking does not lead to a belief divergence between the principal and the agent, as no signals arrive while the agent shirks. This means that the agent does not get an informational rent from postponing effort, hence the lack of *procrastination rents*.[2] In addition, after having acquired a signal, the best deviation available to the agent is to hide it and shirk until the deadline. This is a nontrivial result given that there are so many different histories and deviations available to the agent. Then, for the *ex-post* constraints, it is sufficient to take into account only one type of deviation which is to shirk until the end of the contract. Hence, the optimal contract is characterized fully by considering only local deviation constraints. The constraint to work defines the total agency rent at any moment and the ex-post constraint puts restrictions on how this rent should be divided between continuation values and payments. At any point in time the agent's incentive constraint to work binds and the total agency rent is given simply by the initial time provided in the contract.

The agent receives payments only upon a terminal signal disclosure which are decreasing over time for a given belief with an upward jump after the disclosure of each additional bad signal. There are two cases in which the contract ends before a deadline: upon the disclosure of a good signal at any point in time, or upon the disclosure of the bad signal which makes the belief drop below the first best threshold.[3]. Although the agent is willing to reveal non terminal bad signals without receiving a reward as long as his continuation value does not decrease, it is optimal to raise his continuation value through extra experimentation time as this relaxes his

[1] a terminal signal is used to denote either a good signal or the first bad signal upon which the belief falls below the first best experimentation threshold.

[2] The procrastination rent is present in other papers on dynamic agency where the belief goes down as long as no success is observed. For example, in Horner and Samuelson (2013). Hence, the principal finds it optimal to downsize the project or take his outside option for some periods in order to decrease the informational rent of the agent. In our setting, the principal does not find it optimal to take the outside option and continue with the project later on. Similarly, in Halac et al. (2016) and Moroni (2016) an informational rent is born because the agent's deviation leads him to hold a different prior than the principal.

[3] This threshold is the belief at which experimentation ends upon a bad signal disclosure and will be solved for in the paper. It will correspond to the same stopping belief as in first best

incentive constraint to work and results in a longer experimentation time at a belief at which experimentation is efficient.

The intuition for the above features of the contract is as follows: there is a direct relation between the deadlines and the continuation value of the agent. A longer experimentation horizon provides more time to experiment on a potentially good project but at the same time the principal has to promise higher payments to the agent in order to make him work and disclose the signals which are terminal. Hence, committing to a deadline allows the principal to control the total agency rent. As time passes without a signal disclosure, it becomes more likely that the agent has been shirking and termination serves as a punishment. At the same time, the deadline causes experimentation to end inefficiently early at times when the belief doesn't go down. The disclosure of a bad signal reveals that the agent has been exerting effort and the extension of the deadline lowers the distortion caused by the deadline, while relaxing the agent's *ex-ante* constraint to work. The payments upon terminal signals are increasing in the time remaining in the contract: in order to induce the agent to reveal a good signal or a terminal bad signal, the principal has to compensate him for the possibility of hiding it and remaining in the contract until the current deadline.

The principal finds it optimal to provide incentives to the agent through experimentation time after the disclosure of bad signals as long as experimentation is efficient[4]. The amount of extension is just sufficient to make the agent's *ex-ante* incentive constraint bind hence does not increase the agency rent compared to a contract in which the deadline remains constant. On the other hand, the agent's *ex-post* incentive constraint to disclose the bad signal is slack. This means there are other incentive compatible contracts with the same agency rent in which the experimentation time doesn't extend upon intermediary bad signals (or extends by a lower amount) that would instead promise higher payments upon the disclosure of terminal signals and would result in less experimentation in expectation. The optimal contract, by minimizing the payments upon good signal disclosures and providing extra experimentation time after bad signals, keeps the agent experimenting longer at beliefs at which experimentation has a strictly positive value.

Finally, an important result is that the extensions of the deadline upon bad signal disclosures continue as long as experimentation is efficient, and at the time when the belief falls below the first-best threshold before a deadline, the principal has to make a payment sufficient to make sure the agent is willing to disclose the signal and end the contract. At this point, it is no longer

[4]Here efficiency is in the first-best sense, when the principal would be experimenting alone

profitable to back-load the payments, as even if the agent's rent doesn't increase, the intrinsic value of experimentation is negative.

This paper relates closely to the literature on learning in dynamic agency in which there is a cash diversion problem, introduced by Bergemann and Hege (1998, 2005) and also studied by Horner and Samuelson (2013) with a main divergence in terms of the learning structure. In the mentioned papers learning happens as in the exponential bandit model by Keller, Rady and Cripps (2005) in which a good project can succeed (a conclusive good signal) with a certain arrival rate and that the lack of success is equivalent to a failure[5]. Hence, even without an agency problem, as time goes by the belief becomes pessimistic enough that the project should be abandoned. In the presence of agency, experimentation ends inefficiently early due to the agency rents and it is not possible to distinguish an agent who is shirking from an agent who is working on a bad project. On the other hand, in my setting of information acquisition, news arrives at certain points in time while the agent experiments in form of a conclusive good signal or nonconclusive bad signals, with a poisson arrival rate. The rate of arrival is independent of the project type, while the type of the signal depends on the underlying type. As experimentation happens in form of information acquisition and leads to learning at certain points in time only while effort is exerted, in the absence of signal arrivals beliefs remain constant. This means that learning is not continuous and belief is not a function of time but of the number of signals already disclosed. Hence, without an agency problem, if learning is initially optimal, there is no stopping until obtaining a good signal or sufficiently many bad signals that the benefit of experimentation falls below its cost. In the presence of agency, although a bad signal leads to a lower belief about the project quality, it happens as a result of the agent's effort. This information structure leads to new contractual features, mainly the termination times as a function of the history of signal disclosures.

The paper most closely related to mine is Maestri and Gerardi (2012) who study contracting for information acquisition when an agent incurs cost to get private signals in each period about a project quality with the difference that information is soft. As the agent is sure to get a signal in each period as long as he incurs cost, there is a fixed date at which learning will be complete at the latest and the agent's rent comes from the possibility of guessing that the state is good without incurring any cost of information acquisition. In my setting, in addition to being hard information, news arrives over time with a certain arrival rate, and there is no predetermined

[5]other experimentation settings where this is the assumption include Keller, Rady and Cripps (2005), and Keller and Rady (2010), Horner and Samuelson (2013), Halac, Kartik and Liu (2016), Bonatti and Horner (2011), Moroni (2016), Guo (2016), Klein (2016)

date at which sufficient information should be acquired.

It is common in the dynamic contracting literature with moral hazard to provide incentives to an agent by modifying his continuation value. Sannikov (2008) is the first who studied the properties of optimal incentive provision in a continuous time principal agent model with moral hazard and no learning. In his setting, the agent's continuation value has an upward drift when wages are back-loaded. In my setting, payments are always back-loaded by increasing the agent's continuation value after a bad signal disclosure through adjusting the termination date and the only payments are made when there is first-best stopping. Demarzo and Sannikov (2016) study a cash diversion model in which both the principal and agent learn about future profitability and agent's information rents are controlled by distorting the termination decision. In their model, shirking by agent lowers the expected future probability hence the agent gets information rents due to this belief manipulation effect. Green and Taylor (2016) study contracting for a two stage project without learning and show how communication by the agent about the completion of one stage affects the time allocated for the completion of the next stage.

Manso (2010) shows in a two period setting that motivating an agent to innovate may require tolerating or even rewarding early failure. In his setting, there is no moral hazard problem and the tradeoff is between exploration and exploitation. I find a similar result in that the agent is rewarded for a bad signal, but instead through experimentation time. In addition, the setting in my setting there is no exogenous deadline and a number of bad signals can be acquired and disclosed before experimentation ends.

Guo (2016) considers a problem of dynamic delegation of experimentation without monetary transfers in which the information structure is as in Keller et al. (2005). She has a result on sliding deadlines when allowing for non-conclusive good news. The reason for the extended deadlines is due to experimentation becoming more profitable after a good signal is realised, which is the opposite case in my setting.

Moroni (2016) studies an experimentation problem with one principal, several agents and two milestones the feasibility of which are both unknown initially and the completion of one stage does not provide information about the feasibility of the next one. She shows that the optimal contract rewards agents for earlier completion of the first stage by giving more time on the second stage rather than a monetary payment. This is also an example of back-loading payments through extended time, but the difference is that there is no learning upon the completion of first stage whereas in my setting time is extended while the profitability of the project goes down.

Akcigit and Liu (2016) consider two firms competing for an innovation on a common research line in which when one firm reaches a dead end on their research line, the other firm which is uninformed may keep experimenting inefficiently on the same arm. Their information structure is similar to mine in that learning happens only upon arrival of news, with the difference that one good or bad signal is enough to conclude about the success or failure of the research line.

Outside the experimentation literature, there are some papers which describe settings with moral hazard where bad news is rewarded. For example, Levitt and Snyder (1997) consider a static setting in which an agent's effort creates a private signal about the project's potential success. Rewarding for bad signal is necessary in order to make sure the agent is truthful and so the principal can avoid investing in the project, in other words the agent would not reveal bad news without a sufficient reward. Lastly, Chade and Kovrijnykh (2016) consider a setting in which an agent is hired to evaluate different options over time where her private effort determines the precision of a signal that is publicly observed. Rewarding for bad news occurs when the prior is low and the agent's effort generates a bad signal, which is due to the signal outcome confirming that effort was put in. Here, the bad signal does not affect the value of the next option, hence there is no learning between different options. The main difference in the rewarding for bad news result in my setting is that the extended time upon bad news is not a necessary condition in order to induce the agent to disclose the non-terminal bad signals, and there are many other incentive compatible contracts having the same expected rent with a fixed deadline in which the agent would be willing to reveal bad signals as they arrive. In those contracts, more payments would be needed upon a good signal disclosure which is terminal.

2 Model

A principal (she) hires an agent (he) in order to learn about an unknown state of the world, which determines the quality of a project. They share a common prior ρ_0 that the project is good and $(1-\rho_0)$ that it is bad. A good project has a net value 1 for the principal and a bad project has a sufficiently negative value that learning is necessary before deciding to implement the project. The state can be learned through costly experimentation requiring resources modeled as a flow investment c by the principal. The agent is hired to use this investment for experimenting to acquire information, but he could also shirk and keep it for his private use. The agent's shirking is not observable by the principal, but it results in no learning about the project.[6] The principal

[6] I have described the setting as one in which the principal provides resources for experimentation and the agent has the possibility to divert benefits to his own use. However, the setting could also be chosen as one in which

and the agent are both risk neutral and share the same discount factor. Outside options are zero and the agent has limited liability, hence any payments made to the agent have to be positive.

Signal structure: Time $t \in [0, \infty)$ is continuous. If the agent puts in effort $a_t \in \{0, 1\}$ over a time interval $[t, t + dt]$ by incurring the cost cdt, a signal privately arrives with probability $a_t \lambda dt$ and is denoted by $z_t \in \{G, B\}$. The arrival rate is independent of the type of the project. While only a bad signal can result from a bad project, the signal from a good project can be a good one with probability θ or a bad one with probability $(1 - \theta)$.[7] Signals are verifiable, they can be hidden but not constructed.

Hence, one good signal is enough to conclude that the project is a good one. Upon the disclosure of a bad signal the belief about project quality goes down as follows:

$$\rho_{k+1} = \frac{(1-\theta)\rho_k}{1-\theta\rho_k}$$

where ρ_k is the public belief at any t up to which exactly k bad signals have been disclosed, which also coincides with the public belief in case the agent reveals the signals as they arrive. It is a crucial feature of our analysis that the belief is not a function of the calendar time but of the number of signals already acquired for the agent and disclosed for the principal. The two will be equivalent in the optimal incentive compatible contract.

Assumption 1. *Experimentation is initially profitable in the absence of an agency problem:*

$$\lambda\theta\rho_0 > c$$

where the left hand side is the probability of a signal arriving which reveals that the project is a good one multiplied by the value of a good project, which is one. This assumption means that without the agency problem the principal would be willing to experiment at least at the initial belief.

First-best Benchmark: First, I consider the principal's problem in case she could carry out experimentation without hiring an agent. Given assumption 1, the first-best consists of a

---

the agent enjoys the benefit c from leisure when putting in low effort and remaining in the contract, without the assumption of investment by the principal. The main results of the analysis would carry on.

[7] The discrete time version of this setting would be one in which each period when the agent incurs the cost, he may receive a signal with probability λ whose type depends on the underlying state, or receive no signal with probability $1 - \lambda$.

stopping belief level where n^* is the minimum n which satisfies:

$$\theta\lambda\rho_{n+1} - c < 0$$

In other words ρ_{n^*} is the lowest belief at which experimentation still has a positive net value. As signals arrive with a Poisson arrival rate only while experimentation is carried out, it is not known ex-ante at which future date the belief will have reached this level. Then, without the agency problem, there is not a stopping time but a stopping level of belief for the principal.

On the other hand, in the presence of agency, the principal needs to choose predetermined deadlines in order to control the agent's rent, hence experimentation may end at a higher belief than in the first best. However, as it will be shown the belief at which experimentation stops before reaching a deadline in the presence of agency will be identical to the first best stopping belief.

Agent's strategy: The agent chooses effort $a_t \in \{0, 1\}$ where 1 is equivalent to putting in effort by incurring the cost of experimentation and 0 the decision to shirk and keep for himself the investment c. The agent also chooses a disclosure plan $x_t \in \{G, B, 0\}$ as a function of his private history.

The agent can disclose any signal or signals among those he has received until t and has not already disclosed. He could possibly keep and disclose a signal later on but cannot construct a fake signal. This implies that the private history of the agent can possibly be very complicated. However, attention is restricted to contracts which induce truthful reporting, hence the agent's private history h^t_A coincides with the public history, h^t. The agent has no possibility of privately learning by shirking, as no signals arrive while he shirks hence the beliefs of the agent and the principal do not diverge.[8]

Assumption 2. *$\lambda\theta\rho_0 > 2c$.*

This assumption ensures that experimentation is profitable at $t = 0$ in the presence of agency given the initial belief ρ_0. The term $\lambda\rho_0\theta$ which is the expected is the benefit of an instant of experimentation, and $2c$ the total cost of an instant of experimentation for the principal: the first c is the flow cost provided by the principal for experimentation and the second c is the *agency cost*, in other words the minimum continuation value per unit time the principal has to promise the agent in order to make sure that he does not shirk.

[8] This is different from most of the literature on experimentation in which the agent's shirking leads to private learning and a more optimistic belief than the principal.

3 The optimal contract

A contract is denoted by $\mathcal{C} = \{w, y\} : H \to \Re_+ \times \{0, 1\}$ where H represents the set of all possible histories. The first component w_t represents the history dependent payments to the agent at time t. The second component $y_t \in \{0, 1\}$ denotes the principal's decision whether to fund experimentation at time t as a function of the public history.

Principal's Problem: As the principal wants to choose the optimal contract which makes the agent work and discloses without delay the signals that arrive, his present expected value at time zero from a contract C is:

$$F_0(\mathcal{C}) = E^a \left[\int_{t=0}^{\infty} e^{-rt} e^{-\lambda\theta \int_0^t a_z p_z dz} y_t (a_t \lambda \theta \rho_t - w_t - c) dt \right] \tag{1}$$

where ρ_t is the belief at time t which depends on how many bad signals have already been disclosed, and $e^{-\lambda\theta \int_0^s a_z p_z dz}$ is the probability that no good signal will be disclosed until t. Once a good signal is disclosed, the profit of the project is realized by the principal and experimentation becomes irrelevant. Given that in our setting the belief only evolves at certain points in time when signals are realized, the belief is a function of the number of bad signals already disclosed, which is k.

The agent's utility is a function of the payments he receives and the resources he keeps in case of shirking:

$$V_0(\mathcal{C}) = E^a \left[\int_{t=0}^{\infty} e^{-rt} y_t [w_t + (1 - a_t) c] dt \right] \tag{2}$$

A contract $\mathcal{C}$ is optimal and incentive compatible if the induced action a^* maximizes (1) and (2) and induces the agent to disclose the signals without delay. The principal's problem is then one of finding the incentive compatible contract which maximizes her payoff subject to satisfying the agent's working and disclosure constraints. First, in an optimal contract $w_t = 0$ should hold as long as no signal is disclosed. This is because given that the signals are the only hard information about the agent's effort there is no need to make a payment as long as no signal has been reported by the agent.[9] Second, as long as the principal keeps investing, it is never optimal to induce the agent to shirk as this implies a waste of resources. While the agent shirks, as no signals arrive the beliefs of the principal and the agent do not diverge. This means in the

[9]The principal has to pay a positive rent in order to make the agent work and reveal the signals, as the agent gets a positive benefit from shirking. Then, given that the signals arrive only while the agent is working, the flow payment should be set to zero in any optimal contract and payments must be only upon the revelation of signals.

optimal contract the agent should be induced to choose $a_t^* = 1$ as long as $y_t = 1$ and disclose $x_t^* = z_t$ whenever there is a signal realization z_t.

Remark 1. *If $y_t(h^t) = 0$, then for all $t' > t$ with public history $h^{t'}$ such that $h^t \prec h^{t'}$, $y_{t'} = 0$.*

If the principal stops investing in experimentation, she will not start investing again. In the absence of new information, the beliefs and hence the opportunity cost of experimentation remain constant. Hence, there is no gain for the principal from starting to invest again after having stopped once. This leads us to conclude that the contract will include a deadline at which, if reached, experimentation will stop once and for all.[10]

Remark 2. *In the optimal contract, the deadline, if ever, should only be updated at times at which bad signals are disclosed by the agent: when $x_t = B$.*

The reason for this is simple: there is no learning in the absence of signal arrivals. Hence, in case the deadline were updated at a certain date before any signal is disclosed, this contract is equivalent to one in which this updated deadline was the initial deadline. Now, we will define the termination rule in the contract.

Definition 1. *A termination rule is denoted by $T = (T^k(h^t))_{k=0}^{n^*}$ where T^k specifies the current deadline at a given time t when the public belief is ρ_k (k bad signals have already been revealed) and the history is h^t. The initial deadline T^0 is such that if reached without any signal disclosure, the contract terminates.*

We will call k, the number of bad signals already disclosed, as the "state". The definition says that the contract ends in 3 different ways: whenever a good signal is disclosed, upon reaching the deadline T^k in any state k, or upon the disclosure of a bad signal in state n^* which is stopping before a deadline due to the terminal bad signal disclosure.

The public history at any moment can be summarized by the times at which bad signals have been disclosed: $h^t = \{t_1, t_2, ..., t_k\}$ as they determine ρ_k as well as the current deadline. These are the only elements of the history that are relevant for the contractual terms. As will be seen, at any moment, the current deadline is indeed a function of the previous deadline and the time of the disclosure of the most recent bad signal.

From now on, we will simplify the notation for the deadline as T^k, the agent's continuation value as $V_{t,k}$ and the principal's continuation value as $F_{t,k}$ at time t and state k. The payments upon disclosure of a good or bad signal at time t and state k are denoted $w_t = (w_{t,k}(G), w_{t,k}(B))$.

[10]it is too costly to provide the agent an infinite amount of time to experiment.

While simplifying the notation, the public history h^t is not omitted, and it will be clear later on how the payments and the deadline depend on the history of disclosures.

Now we rewrite the principal's continuation value in a recursive way at time t. The state k moves to $k+1$ as soon as an additional bad signal is disclosed. At a time t at which the k'th bad signal is disclosed, the present value of the contract to the principal is:

$$(3) \qquad F_{t,k} = \int_{s=t}^{T^k} e^{-(s-t)(\lambda+r)}[\lambda(\theta\rho_k(1-w_{s,k}(G)) + (1-\theta\rho_k)(-w_{s,k}(B) + F_{s,k+1})) - c]ds$$

where $e^{-(s-t)\lambda}$ is the probability of time s being reached with no signals arriving conditional on the agent working, hence that the belief remains ρ_k. The term $e^{-(s-t)r}$ is the discount factor that applies to a signal arriving and being disclosed at time s. During an infinitesimal time period of dt, with probability λdt a signal arrives and is disclosed. With probability $\theta\rho_k$ the signal is a good one and upon its disclosure the contract ends with the principal making the payment $w_{t,k}(G)$, or the signal is a bad one and the state moves on to $k+1$ providing the principal with the continuation value $F_{t,k+1}$, assuming that $k < n^*$. The detailed derivation of equation (3) is provided in the Appendix. Then, the problem of the principal at time zero, given that ρ_n^* is the lowest belief at which experimentation takes place, is:

$$\max_{T^k(h^t)_{k=0}^{n^*}} F_{0,0}$$

subject to $w_{t,k}(G)$ and $w_{t,k}(B)$ satisfying the incentive compatibility conditions of the agent. In other words, the principal's problem is to commit to an optimal termination rule subject to satisfying the agent's constraints. While solving for the optimal contract, attention is initially restricted to local incentive compatibility constraints and it is verified later that these are actually sufficient for global implementability. There are two types of incentive constraints. The first type is the *no shirking constraint* which makes sure that the agent prefers experimenting as induced by the contract to shirking at any moment. The second type of constraints are the *disclosure constraints* which make sure that the agent is willing to reveal the signals he acquires without delay and checks for any possible deviations. Propositions 1 and 2 provide the characteristics of the optimal contract, and the steps of the solution are provided in section 4.

Proposition 1. *In the optimal contract, the only payments to the agent are made in case termination happens before the current deadline either due to the disclosure of a good signal, or* *the terminal ($n^* + 1$'th) bad signal and:*

- $V_{t,k} = \int_t^{T^k} ce^{-r(s-t)}ds$. *The continuation value of the agent at any moment and belief is equal to the payoff he would obtain instead by shirking and keeping the investment c for his own benefit until the current deadline T^k.*

- $w_{t,k}(G) = V_{t,k}$. *The payment upon the disclosure of a good signal in any state is equal to the continuation payoff the agent would obtain by instead remaining in the contract and shirking.*

- $w_{t,k}(B) = 0$ *for* $k < n^*$. *The agent discloses a bad signal without receiving any payment as long as it does not lead to termination.*

- $\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*}$ *and* $w_{t,n^*}(S) \geq V_{t,n^*}$ *for* $S \in (G, B)$. *In state n^*, which is the last state during which experimentation takes place, incentives to work are provided only through payments, and both payments should be set at least equal to what the agent would get by hiding it and remaining in the project until the end. (Proof in section 4)*

There are an infinite number of payment pairs that may satisfy the last item given that in state n^* a good or a bad signal both end the contract. However, the optimal payments in states $k < n^*$ are unique. The principal does not have to make a positive payment upon the disclosure of bad signals which are not terminal: as long as the agent's continuation value, which is a function of the remaining time, does not go down upon disclosing a bad signal, he is willing to disclose it without receiving any payment. On the other hand, as a good signal disclosure ends the contract, a payment is required. The payment upon good signal is equal to what the agent would get by remaining in the project until the current deadline and shirking, and as a result it is an increasing function of the remaining time. Even though there are an infinite number of possible deviations after hiding a signal, the most profitable one is to shirk until the end of the contract.

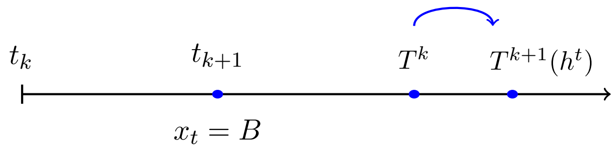
Next proposition describes the increase in the continuation value of the agent which is provided as an extension of the deadline after the disclosure of a nonterminal bad signals.

Proposition 2. *In the optimal contract, the continuation value of the agent increases after the disclosure of a bad signal, as long as this signal is not a terminal one, through an extension of the experimentation time as follows:*

- $V_{t,k+1} - V_{t,k} = \frac{c}{\lambda(1-\theta\rho_k)}$ *for* $k < n^*$. *Upon a bad signal disclosure, the continuation value of the agent increases by an amount which is higher for higher beliefs.*

Figure 1: Updating of T^k

from state k to $k+1$



- $T^{k+1}(t_{k+1}, h^t) > T^k(h^t)$ *for* $t_{k+1} \leq T^k$ *and* $k < n^*$. *When a bad signal is disclosed before the deadline in any state* $k < n^*$, *the current deadline is extended through an increase in the agent's continuation value which satisfies:*

$$\frac{c}{\lambda(1-\theta\rho_k)} = e^{-r(T^k - t)}\frac{c}{r}(1 - e^{-r(T^{k+1}-T^k)})$$

(Proof in section 4*)*

The principal can choose to reward the agent either through bonus payments or experimentation time and prefers the latter as long as experimentation still has a positive value. Given that good signal terminates the project, the extensions happen upon bad signal disclosures. Figure 1 demonstrates the updating rule of the deadlines. The updating of the deadline to $T^{k+1}(t_{k+1})$ depends on the time of disclosure and the current deadline. The extensions continue upon bad signal disclosures until the signal which terminates the contract. The belief ρ_{n^*+1} is the one at which experimentation ends before a deadline is reached and coincides with the stopping level in the benchmark without agency. In a given state, the extension in the time horizon is higher the earlier the bad signal is disclosed: as the extended time is added at the end of the current deadline T^k, due to the discount factor, for the same increase in the agent's continuation value, the extension is higher the more distant T^k is from t. In addition, the extension is lower for more pessimistic beliefs: when it is more likely to receive a bad signal, a lower extension provides the same expected increase in the continuation value.

The principal uses deadlines in order to control the moral hazard rent of the agent. Hence, experimentation may possibly end inefficiently early at a belief level higher than the first best stopping belief. This is why the principal benefits from extended experimentation after the disclosure of bad signals. In other words, the principal minimizes the payments promised upon a good signal and increases the agent's continuation value after bad signal disclosures. The agent's expected payoff from experimentation has three components: the payment upon the disclosure of a good signal, payment upon a bad signal and the continuation value in the more pessimistic state after the disclosure of a bad signal. How much the principal can back load the agency rents into future experimentation time is determined by the disclosure constraints which give the minimum bonus that makes sure the agent discloses the terminal signals. The remaining rent is provided as an extended contract horizon subject to the working constraint binding.

The increase in the continuation value is not a necessary condition for the agent to disclose a bad signal. As long as it does not decrease his continuation value, the agent is willing to disclose a bad signal even without a reward. However, the extension is profitable because it leads to more experimentation time while keeping constant the total agency rent. There are many other incentive compatible contracts, but in any other contract in which agent's continuation value doesn't increase by the same amount after bad signals, the payment upon good signals are higher in order to provide the same agency rent. This implies less experimentation for the same cost. Next section explains in detail these findings.

4 Solving for the optimal contract

This section goes through the steps of finding the contractual features described above. As a first step, the agent's continuation value and the incentive constraints that should be satisfied are provided. In order to find out which constraints bind, we consider the conditions for given deadlines. We find the payments and the updating rule of the deadlines by initially restricting attention to local incentive constraints. It is then verified that the local constraints are indeed sufficient for global incentive compatibility by showing that the only relevant deviation for the agent is the deviation to not work again. Finally, we solve for the only remaining variable which is the initial deadline.

4.1 Agent's Continuation value and incentive constraints

At any moment the agent has the choice between exerting effort or shirking and keeping the resource c for his own benefit. Then, the law of motion of the agent's continuation value at time t and belief ρ_k denoted $V_{t,k}$ is as follows (assuming that the agent discloses the signals as soon as he receives them):

$$V_{t,k} = a_t\lambda dt[\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + (1-rdt)V_{t+dt,k+1})] \\ + (1-a_t)cdt + (1-\lambda a_t dt)(1-rdt)V_{t+dt,k}$$

If the agent chooses a_t, he receives a signal with probability $a_t\lambda dt$ which is a good or a bad signal and he discloses it. Upon disclosing a good signal, he receives the payment $w_{t,k}(G)$ and the contract ends. Upon disclosing a bad signal, he receives the payment $w_{t,k}(B)$ and continues experimenting in state $k+1$ with the continuation value $V_{t+dt,k+1}$. If he does not receive any signals he continues with the continuation value $V_{t+dt,k}$ at $t+dt$, as the state does not change. Letting dt go to 0 leads to the following Hamilton-Jacobi-Bellman equation for the agent:

$$(4)\quad 0 = V'_{t,k} + \max_{a_t}\{-(\lambda a_t + r)V_{t,k} + a_t\lambda[\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1})] + (1-a_t)cdt\}$$

Working Constraint

From equation 4, we get the no shirking incentive constraint for the agent.

Lemma 1. *Given a contract C the agent will choose $a_t = 1$ if and only if:*

$$W_t \geq \frac{c}{\lambda} + V_{t,k} \tag{5}$$

where $W_t = \theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1})$ denotes the incremental value of effort for the agent from a signal arrival, where $V_{t,k+1}$ is the continuation value after an additional bad signal.

Lemma 1 is the local incentive constraint to work. Let us now explain briefly why the local constraint is enough for satisfying global incentive compatibility. When the agent deviates to shirk, he does not get any signals and hence the beliefs of the agent and the principal cannot differ. Then, the future incentives to work are not modified by this deviation. To conclude, the agent's deviation at time t cannot affect his incentives to deviate at a future date. This means if equation (5) is satisfied at any t, it will be satisfied globally as well.

Lemma 2. *The continuation value at any t is then written as:*

$$V_{t,k} = \frac{c}{r}(1 - e^{-r(T^k - t)})$$

Then, for a project with initial deadline T, the total agency rent is given by:

$$V_{0,0} = \frac{c}{r}(1 - e^{-rT})$$

This is the minimum payoff the agent can receive for a contract horizon T, and as will be shown the optimal contract will provide the agent exactly this amount of rent, by making the IC constraint to work bind at any moment.

Disclosure Constraints

Now let us move to the second type of constraints which are the *disclosure* constraints. These constraints make sure that once the agent has acquired a signal he is willing to disclose it without delay. The ability to keep the signals and disclose later on adds the complication of many possible histories after deviations, such as hiding a signal and shirking or experimenting in order to possibly get another one. For now I will restrict attention to local deviations which are limited to delaying the disclosure of a signal and shirking for an instant.

Lemma 3. *The local disclosure constraints which make sure that the agent is not willing delay the revelation of signals are:*

$$w'_{t,k}(G) \leq r w_{t,k}(G) - c \tag{6}$$

$$w'_{t,k}(B) \leq r(w_{t,k}(B) + V_{t,k+1}) - V'_{t,k+1} - c \tag{7}$$

where w' and V' denote derivatives with respect to t. Equation (6) *makes sure that the agent does not want to delay disclosing a good signal, and equation* (7) *makes sure that the agent does not want to delay disclosing a bad signal.*

(Proof in the Appendix.)

If the local constraints are sufficient conditions for global incentive compatibility, then it is sufficient to check that these two conditions are satisfied at each t until the deadline T^k. Now, assuming this is the case, I solve the differential equation (6) using the boundary condition at T^k:

$$w_{t,k}(G) \geq \frac{c}{r}(1 - e^{-r(T^k-t)}) + e^{-r(T^k-t)} w_{T^k,k}(G) \tag{8}$$

which says that the agent should be better off disclosing the signal now rather than shirking until the deadline. Doing the same for the bad signal revelation by integrating equation (7) and using the condition at the deadline T^k:

$$w_{t,k}(B) + V_{t,k+1} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)}(w_{T^k,k}(B) + V_{T^k,k+1}) \tag{9}$$

Next, we find the values at the deadlines which will then enable to find the values at any $t < T^k$.

4.2 The payments and the extension of the deadline

Values at the deadline: We know that $V_{T^k,k} = 0$ as the contract ends as soon as the agent's continuation value hits 0 at T^k. In addition the disclosure constraints are irrelevant at the deadline. The assumption is that the agent chooses to disclose the signals when indifferent.

Lemma 4. *The working constraint is the only condition that should be satisfied at the deadline T^k for any k and hence it binds in the optimal contract:*

$$\theta\rho_k w_{T^k,k}(G) + (1 - \theta\rho_k) w_{T^k,k}(B) = \frac{c}{\lambda} - (1 - \theta\rho_k) V_{T^k,k+1} \tag{10}$$

by making use of the equation (5)*.*

The Appendix shows the non optimality of contracts having any $T^{k+1}(t_{k+1}) < T^k$, in other words contracts whose horizon shortens upon the disclosure of bad signals. This is easy to see: the agent is willing to disclose non terminal bad signals without receiving any payment as long as his continuation value doesn't decrease. If the horizon of the contract shortens due to a disclosure, it must be that some payment was made to the agent upon a bad signal disclosure. However, a payment is a waste as it could have instead been provided as experimentation time to the agent given that experimentation is still efficient, or as a bonus by keeping the horizon constant. Now, I proceed to solve for the optimal contract under the condition $T^{k+1}(t_{k+1}) \geq T^k$ for all k, meaning contracts whose horizon do not become shorter upon the disclosure of a bad signal.

Next lemma provides the main step towards solving for the optimal contract.

Lemma 5. *The payments at the deadline are set to 0: $w_{T^k,k}(G) = w_{T^k,k}(B) = 0$ and the continuation value after a bad signal disclosure is given by:*

$$V_{T^k,k+1} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{11}$$

from the condition at the deadline in lemma 4*. (Proof in the Appendix.)*

This lemma says that in the optimal contract the payments upon disclosures at the deadline are set to zero and the continuation value after the disclosure of a bad signal is strictly positive. The incentive to work near the deadline is provided solely through the increased continuation values after a bad signal is disclosed. As this increase is provided while the incentive constraint is binding, it is easy to see that the agency rent is kept constant.

Values for $t < T^k$:

Now, by the same reasoning as in Lemma 5, at any $t < T^k$, $w_{t,k}(G)$ and $w_{t,k}(B)$ are set to the values which make the disclosure constraints bind, and the the remaining rent will be provided through extra experimentation time upon the disclosure of bad signals.

Lemma 6. *The payment upon the disclosure of bad signals is zero: $w_{t,k}(B) = 0$ for $k < n^*$ and the continuation value after a bad signal disclosure is:*

$$V_{t,k+1} = V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)}$$

We will now find the rule for the extension of the deadline. At the deadline, replacing $V_{T^k,k+1} = \frac{c}{r}(1 - e^{-r(T^{k+1}(T^k))})$ in the condition (11) and taking logs:

$$T^{k+1}(T^k) - T^k = \frac{ln(\frac{\lambda(1-\theta\rho_k)}{\lambda(1-\theta\rho_k)-r})}{r}$$

Doing the same for $t < T^k$:

$$\frac{c}{r}(e^{-r(T^k(t)-t)} - e^{-r(T^{k+1}(t)-t)}) = \frac{c}{\lambda(1-\theta\rho_k)} \tag{12}$$

From the above equation, it is easy to conclude that $T^{k+1}(t) - T^k$ is decreasing in k: as the belief becomes more pessimistic, the extension in the deadline after a bad signal is less. The reason for this is that when the belief is more pessimistic, acquiring a bad signal is more likely hence a lower extension in the time provides the same incentives for effort. Also, in a given state k, the extended time decreases in t. The reason is that the farther t is from T^k, less costly the extension in the time horizon from a time t point of view.

Now it can be concluded that the continuation value of the agent (or the current deadline) and the current public belief are sufficient statistics for history dependency. The agent's effort affects the history only through the realization of signals. Then, at a given time, what matters for the agent's incentives is his current belief (and not the times at which the signals were received) and his continuation value. As the disclosure of signals are already reflected in the agent's continuation value (equivalently the current deadline), the only relevant part of the history at time t is summarized by k.

Discussion: The agent is promised an increase in his continuation value upon disclosing a bad signal in any state $k < n^*$ by an amount to make his incentive constraint to work bind together with the payment upon good signal disclosure. This means the total agency rent is kept constant while extending the deadline. The term $\frac{c}{\lambda(1-\theta\rho_k)}$ which is the flow opportunity cost of effort for the agent divided by the probability of acquiring a bad signal is the increase in the agent's continuation value provided upon disclosing a bad signal.

The disclosure of a bad signal causes the belief about the project quality to go down, but it happens only while the agent exerts effort. The agent's expected payoff from working consists of the payment upon good news, payment upon bad news and the change in his continuation value after a bad signal, and how this rent is decomposed among the three is irrelevant for his incentives as long as his constraints are satisfied. Hence, the principal chooses the payments and continuation values in an optimal way which is through minimizing the payments and increasing the continuation value of the agent in the more pessimistic state which is reached after the disclosure of a bad signal, while keeping the total agency rent constant. The principal is better off providing incentives through extended time as long as the intrinsic value of experimentation is positive.

State n^*:

As already mentioned n^* is the last state in which experimentation can take place, and hence $V_{t,n^*+1} = 0$ for any t as the contract ends once the n^*+1'th bad signal is disclosed. Proposition 3 will show that this is the stopping belief.

Lemma 7. *The payments in state n^* should satisfy the following conditions:*

$$\theta\rho_{n^*} w_{t,n^*}(G) + (1-\theta\rho_n^*) w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*} \tag{13}$$

$$w_{t,n^*}(S) \geq V_{t,n^*} + e^{-r(T^{n^*}-t)} w_{T,n^*}(S) \tag{14}$$

for $S \in \{G, B\}$.

The first equation is the incentive constraint to work, while the second one is the incentive constraint to disclose the signals. As the continuation value is zero after the disclosure of either type of signal in this terminal state, the payments need to be strictly positive upon the disclosure of either. In addition, both should be at least equal to the continuation value of the agent at that instant. There are many different pairs of payments that satisfy this and are feasible in the optimal contract.

4.3 The stopping belief before a deadline

Next proposition provides the belief at which experimentation ends before a deadline is reached.

Proposition 3. *Experimentation ends whenever the belief falls down to ρ_{n^*+1} before a deadline, which is identical to the stopping belief in the benchmark setting without agency:*

$$\lambda\theta\rho_{n^*+1} < c$$

where n^* is the lowest value which satisfies this condition.

Let us give the intuition why this belief coincides with the first best benchmark without agency. The total cost of experimentation to the principal per unit time is $2c$ which can be decomposed as follows: the first c is the investment per unit time promised to the agent over the remaining time horizon and the second c is the reward per unit time required to prevent the agent from shirking. By initially committing to a deadline, the principal promises the agent the investment c over a certain time horizon. Then, the agent's disclosure constraint requires that for a signal that will terminate the relationship, he should be compensated for the value of his opportunity cost of remaining in the project until the deadline and shirking, which is equal to $\frac{c}{r}(1-e^{-r(T^k-t)})$. This is due to the ex-post incentive compatibility. The second c is due to the ex-ante flow opportunity cost of experimentation for the agent which should be promised in order to make him work at any instant, before a signal is already acquired. The optimal way to provide this rent is through experimentation time after bad signal disclosures while keeping constant the total agency rent as long as $\lambda\theta\rho_k \geq c$, which corresponds to the benchmark case without agency. As the probability of getting a bad signal is $\lambda(1-\theta\rho_k)$, this leads to an increase of $\frac{c}{\lambda(1-\theta\rho_k)}$ in the agent's continuation value. Once the belief is such that $\lambda\theta\rho_k < c$, an extension is no longer profitable and the principal will make a payment to the agent even upon a bad signal disclosure in order to end the contract.

4.4 Sufficiency of local disclosure constraints:

This section verifies that the local constraints for disclosure in Lemma 1 are indeed sufficient to account for global incentive compatibility. Let us check that there is no profitable deviation for the agent after receiving a bad signal, such as hiding one or more signals. The agent should be compensated at least for the change in his continuation value upon disclosing a bad signal when the state moves from k to $k+1$:

$$w_{t,k}(B) \geq max[0, V_{t,k}^B - V_{t,k+1}] \tag{15}$$

where $V_{t,k}^B$ is the continuation value of the agent after hiding a bad signal. Now, let us find the continuation value of the agent after hiding this signal. The incentive constraint to work before deviation binds:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) = \frac{c}{\lambda} + V_{t,k} - (1-\theta\rho_k)V_{t,k+1}$$

After the deviation, the gain in the continuation value of the agent upon the disclosure of the next bad signals will be respectively $\frac{c}{\lambda(1-\theta\rho_k)}$ and $\frac{c}{\lambda(1-\theta\rho_{k+1})}$ which are independent of when they are disclosed. Also, the reward upon the disclosure of a good signal is always higher for higher k: $w_{t,k}(G) < w_{t,k+1}(G)$, which means hiding a bad signal can only decrease the payment from a good signal disclosure. Then, hiding and delaying the disclosure of a bad signal is weakly dominated by a strategy in which the agent reveals the $k+1$'th bad signal upon receiving it and obtains the increase in his continuation value $\frac{c}{\lambda(1-\theta\rho_k)}$ earlier. Then, equation (15) holds and the agent cannot gain by hiding or delaying a bad signal disclosure.

Finally, we will verify that the deviation to hide a good signal and continue experimenting in order to get another bad signal is not profitable either. This deviation could only be attractive if $w_{t,k+1}(G)$ were sufficiently high compared to $w_{t,k}(G)$. In other words, only if the reward to a good signal was much higher for pessimistic beliefs. The best deviation of this kind would be to hide the good signal and experiment in order to get a bad signal and in case it arrives, disclose it before the good signal at some time $\hat{t}$ in state $k+1$ in order to receive $w_{\hat{t},k+1}(G)$. It is enough to look at an instantaneous deviation of this kind, meaning deviating to experiment for an additional dt and in case a bad signal arrives, disclosing it at some point in the future. This local deviation is not profitable if the following holds:

$$\begin{aligned} w_{t,k}(G) \geq [\lambda dt(1-\theta)(1-rdt)(\frac{c}{r}(1-e^{-r(\hat{t}-t)}) + w_{\hat{t},k+1}(G))] \\ + (1-\lambda(1-\theta)dt)(1-rdt)w_{t+dt,k}(G) \end{aligned} \tag{16}$$

for any $\hat{t}$. The first term denotes the possibility of getting a bad signal after deviating to experiment for an additional instant dt. Then, the agent will shirk until a time $\hat{t}$ at which he reveals the good signal hence terminates the contract. I consider the *best* deviation of this kind, as in case of hiding a bad signal the agent can reveal the signals at any moment until the deadline, and work or shirk in the meantime. However, as the reward upon good signal revelation is higher the higher the number of bad signals already disclosed, the agent would always disclose all the bad signals he has received before the good one. By the revelation constraint in equation (7), the agent does not want to delay revealing a bad signal. In addition, equation (6) makes sure that the agent does not delay the revelation of a good signal at a given belief. Then, it is sufficient to look at the condition when $\hat{t}$ is replaced by t, as this gives the upper bound on the agent's possible deviation to work. Then, if the condition in (16) is satisfied at t, it should also be satisfied at any $\hat{t} > t$. By making dt go to zero in equation (16), the condition becomes:

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq \lambda(1-\theta)(w_{t,k+1}(G) - w_{t,k}(G)) \tag{17}$$

where the left hand side is greater than c. It is sufficient to check for one shot deviations of this kind: if after finding a good signal it is not profitable to experiment in state k in order to get a signal B and reach state $k+1$, it will not be profitable to get the bad signal two times either. To see this, check that the change in $w_{t,k}(G)$ due to an increase in k multiplied by the probability of getting the bad signal is constant: $\lambda(1-\theta\rho_k)[w_{t,k+1}(G) - w_{t,k}(G)] = c$, which is just equal to the benefit from shirking. This implies that:

$$w_{t,k+1}(G) - w_{t,k}(G) = \frac{c}{\lambda(1-\theta\rho_k)}$$

and as $\lambda(1-\theta) < \lambda(1-\rho_k\theta)$, the constraint (17) is satisfied. After hiding a good signal the agent is better off shirking rather than working. The gain from an additional B is not high enough that an agent who has already acquired a good signal would be willing to hide it and experiment in order to get an additional bad signal. Finally, the local constraints are sufficient to satisfy global incentive compatibility.

4.5 The optimal T^0

Now that the updating scheme of the deadlines is provided, once T^0 is found, the other deadlines can be written as a function of the updating rule.

Proposition 4. T^0 *is the time initially allocated to the agent to experiment such that if reached without any signal disclosure the contract gets terminated and is given by:*

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0+(1-\theta\rho_0)F_{T^0,1})-(1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{18}$$

The value $F_{T^0,1}$ is positive and less than 1 (given that the value of a project discovered as good is equal to 1), in addition it depends only on $T^1(T^0) - T^0$ which can be found from the agent's continuation value $V_{T^0,1} = \frac{c}{\lambda(1-\theta\rho_0)}$. Then, $F_{T^0,1}$ can be found by using the updating rule of the deadlines. It is easy to see that T^0 is decreasing in c. As the cost of experimentation increases the principal allocates a shorter time period for the agent to experiment. The sign of the derivative with respect to λ can be positive or negative. The reason is that a higher arrival rate makes experimentation more profitable due to faster learning but it also means that the lack of any signal is more suggestive of the agent's shirking. The derivatives with respect to θ and ρ_0 could be either positive or negative depending on the value of $F_{T^0,1}$. An increase in these parameters have 2 opposing effects: first, a higher probability of good signal arrival implies that an instant of experimentation is more profitable in case the project is a good one. On the other hand, given that the rate of arrival rate is high, there is less incentive to allocate more time because the agent is likely to receive a signal earlier. Hence, the sign of the derivative with respect to θ and ρ_0 depend on which one of these two effects dominate. The common tradeoff in these comparative statics is between the value of experimentation due to faster learning versus more agency rents that should be provided to the agent.

5 Extensions

In this section, I consider some extensions to the original model.

5.1 Public Signals

I consider the case in which the signals are publicly observed both by the agent and the principal. In this setting, only one of the two types of frictions seen in the previous model is present: the moral hazard due to the agent's possibility to shirk. When the signals are publicly observed, the disclosure constraints are no longer relevant. This means the agent's rent is a pure moral hazard rent. As the agent cannot choose whether to disclose the signals or not, the only constraint that should be satisfied is the one which makes sure that he works.

Proposition 5. *An optimal contract in the presence of publicly observed signals has the following* *features:*

- $w_{t,k}(G) = w_{t,k}(B) = 0$ *for all* $k < n^*$. *The payments promised upon the realizations of either type of signal are zero as long as it doesn't happen in the terminal state* n^*.
- $V_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + \frac{V_{t,k}}{(1-\theta\rho_k)}$ *for* $k < n^*$. *Incentives to the agent to exert effort are provided through increased continuation values upon the disclosure of bad signals.*
- $\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*}$. *In state* n^*, *as experimentation will end with one additional signal, the payment upon disclosure of any signal must be positive.*

Proof. The incentive constraint which makes sure that the agent works is identical to the one with private signals:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) = \frac{c}{\lambda} + V_{t,k} - (1-\theta\rho_k)V_{t,k+1}$$

The difference is that there are no more disclosure constraints, which implies that the above condition will bind in the optimal contract. In addition, it is still optimal for the principal to extend the horizon of experimentation upon the revelation of bad signals, and even more this time by setting $w_{t,k}(G) = w_{t,k}(B) = 0$ for $k \leq n^*$, which leads to:

$$V_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + \frac{V_{t,k}}{(1-\theta\rho_k)}$$

and in state n^*, as $V_{t,n^*+1} = 0$:

$$\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*}$$

In the optimal contract $w_{t,n^*}(G)$ and $w_{t,n^*}(B)$ can take on any values that satisfy this equation.

□

In the presence of public signals, the only positive payments are made when experimentation ends in state n^* due to the arrival of a good or a bad signal. In all the previous states, due to the public observability of signals, the principal is able to set the payment even upon a success equal to zero, and incentivize the agent solely through the disclosure of bad signals and hence an extended experimentation time. Indeed, realization of a good signal at a state $k < n^*$ is not favorable for the agent as it ends the contract without providing any positive payment, but in overall his expected benefit from working is high enough that he is willing to work. The agent gets a positive payment only if a signal is realized while the belief is ρ_n^*. The reason is that the principal no longer finds it optimal to extend the deadline upon receiving the n^*+1'th bad signal, hence the only way to provide incentives to the agent is through the bonus payments upon the realization of signals. The division of this payment between good and bad signals does not matter as the agent cannot choose to hide a signal.

Now let us compare the cases of public and privately observed signals. Call $\tilde{V}_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + \frac{\tilde{V}_{t,k}}{(1-\theta\rho_k)}$ where $\tilde{V}_{t,k+1}$ denotes the public signal case, and $V_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + V_{t,k}$ the private case. It is easy to see that the increase in the continuation value is higher in the public signals case. This is because the principal does not have to pay the informational rent due to the private observation of signals and can set the payments to zero for $k < n^*$, and incentivize the agent only through increased continuation values. Hence, the incentives are completely back loaded in the case of public signals.

In the public signals case, the principal makes payments less often: in states $k < n^*$ he never makes a payment. Instead, the horizon of experimentation is extended by more upon the revelation of each bad signal (the jump in continuation value is higher), and the agent gets a positive payment only if a signal is realized in state n^*. This means the agent gets the bonus payments less often, but in case he does get paid due to success, this may be a higher payment given that the horizon of the contract extends more upon each bad signal revelation.

5.2 Case of no moral hazard

Consider the case in which the agent's decision to experiment or shirk is perfectly observed by the principal, but not the arrival of the signals. Although the agent has the option to hide the signal, there is no gain in doing so as his expected payoff in the contract is always zero when the principal can monitor his effort. In addition, as the agent cannot lie about the realization of the signals, he does not get informational rents either and the principal's problem is identical to the case in which she experiments alone. This means that private observation of the signals alone does not lead to any distortions in the principal's problem compared to the first best.

5.3 Signals are lost if not revealed right away

This is a special case of the original setting considered in the paper and does not modify the results. Indeed, under this assumption, the possible deviations after receiving a signal are much simpler. If the agent hides a good signal, he will not find it optimal to work again in this setting either. To see why this is the case: the possible actions after hiding a good signal are either to stay in the project and shirk, or to work in order to get a bad signal. However, given that now the agent knows the state is good, the probability of getting a bad signal is low enough that he does not find it profitable to work. The most profitable deviation which involves hiding the good signal is to shirk until the deadline, which provides the same minimum reward, $w_{t,k}(G) = V_{t,k}$, as in the original contract. The condition for the revelation of G is:

$$w_{t,k}(G) \geq V_{t,k}$$

which implies that the payment upon a good signal can be chosen to be the same as in the original problem. On the other hand, the agent cannot do better by hiding a bad signal either, as revelation of a bad signal increases his continuation value and does not end the relationship. Then, it is possible to set $w_{t,k}(B) = 0$ as long as $V_{t,k+1} \geq V_{t,k}$. So, the original contract remains optimal under this assumption.

6 Conclusion

This paper studied the optimal provision of incentives in a setting in which an agent gets signals through a poisson arrival process while effort is exerted. This informational structure enables distinguishing an agent who is unlucky from an agent who shirks. The possibility of not receiving any signals over a period of time even while experimenting implies that monitoring the agent's effort is not trivial. As a result, the optimal contract punishes the agent through early stopping whenever enough time passes without any news. In addition, the principal promises payments and continuation values to makes sure the agent has incentives to disclose the signals. As the signals are hard information about the agent's effort, payments and rewards are promised only upon their disclosure.

The optimal contract specifies a termination date until which the agent is allowed to experiment, and provides extra experimentation time after the disclosure of bad signals as long as experimentation remains efficient. The principal extends the contract horizon while keeping constant the expected agency rent. Even though a bad signal leads to a more pessimistic belief, the principal prefers incentivizing the agent through experimentation time as much as possible. The results of this paper suggest that incentives should be optimally provided through endogenous deadlines when learning about a project quality happens over time through costly effort.

There are some questions related to the paper that may be of interest for future research.

One such question is that of introducing career concerns into the model, hence that the agent may be less eager to disclose bad signals because it reveals something about his type which is important in the long run. Another interesting related topic is communication incentives in teams, mainly how rewards should be structured in order to induce the agents in a team to share their information. These questions remain for future research.

Appendix

Derivation of the principal's value function

First, at $t = 0$, if the principal invests and the agent works as induced by the contract, the expected continuation value of the principal is:

$$F_{0,0} = (\lambda(\theta\rho_0(1 - w_{0,0}(G)) + (1 - \theta\rho_0)(-w_{0,0}(B) + F_{dt,1})) - c)dt + (1 - \lambda\theta dt)F_{dt,0}$$

when $dt => 0$ and replacing t instead of $t = 0$:

$$-\dot{F} + (r + \lambda)F = \lambda(\theta\rho_0(1 - w_{t,0}(G)) + (1 - \theta\rho_0)(-w_{t,0}(B) + F_{t,1})) - c$$

As T^0 is the terminating time, $F_{T^0,0} = 0$. Solving the differential equation yields:

$$F_{0,0} = \int_0^{T^0} e^{-t(\lambda+r)}[\lambda(\theta\rho_0(1 - w_{t,0}(G)) + (1 - \theta\rho_0)(-w_{t,0}(B) + F_{t,1})) - c]dt$$

At any moment t in state 0 when a bad signal arrives and is revealed, the continuation value of the principal at that moment becomes $F_{t,1}$. This gives the principal's value function.

Proof of Lemma 1

If the agent shirks for an infinitesimal time period of dt, no signal arrives and the index k does not change. He enjoys cdt and obtains the continuation value $V_{t+dt,k}$. The agent prefers to work rather than shirk iff:

(19)

$$\lambda[\theta\rho_k w_{t,k}(G) + (1 - \theta\rho_k)(w_{t,k}(B) + V_{t,k+1})]dt + (1 - \lambda dt)(1 - rdt)V_{t+dt,k} \geq cdt + (1 - rdt)V_{t+dt,k}$$

which, letting dt go to 0, leads to:

(20) $$\lambda[\theta\rho_k w_{t,k}(G) + (1 - \theta\rho_k)(w_{t,k}(B) + V_{t,k+1})] \geq c + \lambda V_{t,k}$$

dividing everything by λ provides the result.

Proof of lemma 2

The condition the contract should satisfy in order to make sure that the agent prefers to work rather than shirk for an infinitesimal time period dt is:

$$V_{t,k} \geq cdt + (1 - rdt)V_{t+dt,k}$$

Letting dt go to 0:

$$-V'_{t,k} + rV_{t,k} \geq c \tag{21}$$

In state k, as the contract will end as soon as the deadline T^k is reached (hence the $k+1$'th bad signal has not been disclosed), $V_{T^k,k} = 0$. Then, solving the differential equation $-V'_{t,k} + rV_{t,k} \geq c$ by making use of the boundary condition $V_{T^k,k} = 0$ we get:

$$V_{t,k} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) \tag{22}$$

Hence, equation (22) gives the minimum continuation value that should be provided to the agent in order to make him work. As the principal has no incentive to provide more continuation value than needed, this equation will bind.

Proof of Lemma 3

First, I derive equation (6) which makes sure that the agent is willing to reveal a good signal upon receiving it instead of delaying revelation to $t + dt$ and shirking in the meantime, which is:

$$w_{t,k}(G) \geq cdt + (1 - rdt)w_{t+dt,k}(G)$$

which as dt goes to 0, leads to:

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq c \tag{23}$$

Next, I will derive the equation (7). The constraint which makes sure that it is not profitable to wait for an infinitesimal time dt before revealing a bad signal is:

$$w_{t,k}(B) + V_{t,k+1} \geq cdt + (1 - rdt)(w_{t+dt,k}(B) + V_{t+dt,k+1})$$

Letting dt go to 0:

$$- w'_{t,k}(B) - V'_{t,k+1} + r(w_{t,k}(B) + V_{t,k+1}) \geq c \tag{24}$$

Proof of Lemma 4

Consider the incentive constraint to work just before T^k:

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) \geq \frac{c}{\lambda} + V_{T^k,k} - (1-\theta\rho_k)V_{T^k,k+1}$$

By replacing $V_{T^k,k} = 0$, this constraint simplifies to:

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) \geq \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1}$$

At T^k, the disclosure constraints are irrelevant: if a signal arrives just before the deadline, the agent is willing to reveal it without receiving any payment as in case of no disclosure his continuation value is zero. Then, this condition is the only one that should be satisfied, and hence it will bind. In case it were slack, the principal could decrease his expected payment while still satisfying this constraint.

Proof of Lemma 5

I will prove that $V_{T^k,k+1} = \frac{c}{\lambda(1-\theta\rho_k)}$ and hence $w_{T^k,k}(B) = w_{T^k,k}(G) = 0$ in the optimal contract. In order to show this, I will show that $V_{t,k+1} = V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)}$ for any t. This will be done in 2 steps.

Step 1: First, let us show that $V_{t,k+1} \leq V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)}$. Assume the contrary, $V_{t,k+1} = V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)} + \Delta$. The incentive constraint at t is:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1}) = \frac{c}{\lambda} + V_{t,k} \tag{25}$$

Then we have: $\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) < \theta V_{t,k}$. This means, even when $w_{t,k}(B) = 0$, $w_{t,k}(G) < V_{t,k}$ hence the disclosure constraint is violated. Then, it cannot be possible that $V_{t,k+1} > V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)}$.

Step 2:

Now I will show that $V_{t,k+1} \geq V_{t,k} + \frac{c}{\lambda(1-\theta\rho_k)}$. Assume the contrary, that $V_{t,k+1} = V_{t,k} +$

$\frac{c}{\lambda(1-\theta\rho_k)} - \Delta$. Look at the incentive constraint:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1}) = \frac{c}{\lambda} + V_{t,k}$$

For this to hold, we have $\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) = \theta\rho_k V_{t,k} + (1-\theta\rho_k)\Delta$. This means $w_{t,k}(G) > V_{t,k}$ or $w_{t,k}(B) > 0$. Hence, at least one disclosure constraint is now slack. Then, consider increasing $V_{t,k+1}$ by Δ while keeping the constraint binding, hence decreasing the expected payment. The right hand side of the incentive constraint decreases by $\Delta(1-\theta\rho_k)$, implying the expected payments at t decreases by $\lambda\Delta(1-\theta\rho_k)$. As the payments are decreased by the same amount for any t, the revelation constraints are still satisfied. I will show that the decrease in payments at time t is exactly equal to the expected agency rent during the extended horizon, and hence the net effect of this extended time period is positive given that the value of experimentation in state $k+1$ is positive. Then, I will conclude that this modification at any t increases profits.

Before the extension of time at T, the profit during a period $(t, T^{k+1}(t))$ when the state is $k+1$ is:

$$\int_t^{T^{k+1}(t)} e^{-(s-t)(\lambda+r)}[\lambda(\theta\rho_{k+1}(1-\frac{c}{r}(1-e^{-r(T^{k+1}-t)})) + (1-\theta\rho_{k+1})F_{s,k+2}) - c]ds$$

which is equal to:

$$(26)\quad (1-e^{-(T^{k+1}-t)(\lambda+r)})\frac{[\lambda\theta\rho_{k+1}]}{\lambda+r} + \int_t^{T^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})F_{s,k+2}ds - \frac{c}{r}(1-e^{-r(T^{k+1}-t)})$$

where $\frac{c}{r}(1-e^{-r(T^{k+1}-t)}) = V_{t,k+1}$. After increasing $V_{t,k+1}$ by Δ, $\hat{T}^{k+1}$ is such that $\frac{c}{r}(1-e^{-r(\hat{T}^{k+1}-t)}) = V_{t,k+1} + \Delta$.

Then, the principal's profit during the extended horizon is:

$$(27)\quad (1-e^{-(\hat{T}^{k+1}-t)(\lambda+r)})\frac{\lambda\theta\rho_{k+1}}{\lambda+r} + \int_t^{\hat{T}^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})\hat{F}_{s,k+2}ds - (V_{t,k+1}+\Delta)$$

the increase in cost at t, $\Delta(1-\theta\rho_k)$, is equal to the expected payment to the agent during the extended horizon. Finally, we show that the expected profit has increased. $(1-e^{-(\hat{T}^{k+1}-t)(\lambda+r)})\frac{[\lambda\theta\rho_{k+1}+(1-\theta\rho_{k+1})]}{\lambda+r}$ increases in T^{k+1}, hence the first term has increased as $\hat{T}^{k+1} > T^{k+1}$. Then, $\int_t^{\hat{T}^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})\hat{F}_{s,k+2}ds > \int_t^{T^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})F_{s,k+2}ds$, because $\hat{T}^{k+1} > T^{k+1}$, and $\hat{F}_{s,k+2} > F_{s,k+2}$.

Proof of Lemma 6

Proof. I make the revelation constraint for G in equation (8) bind in order to find the minimum $w_{t,k}(G)$:

$$w_{t,k}(G) = \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)} w_{T^k,k}(G) \tag{28}$$

where $w_{T^k,k}(G) = 0$, hence $w_{t,k}(G) = \frac{c}{r}(1 - e^{-r(T^k - t)})$.

Given that the incentive constraint for working binds for any (t, k), the agent's continuation value is given by:

$$V_{t,k} = \frac{c}{r}(1 - e^{-r(T^k - t)})$$

which implies: $w_{t,k}(G) = V_{t,k}$. Then, the working constraint given in lemma 1 becomes:

$$(1 - \theta\rho_k)(w_{t,k}(B) + V_{t,k+1}) \geq \frac{c}{\lambda} + (1 - \theta\rho_k)V_{t,k}$$

It is optimal that this constraint binds. Now I claim that it is optimal to set $w_{t,k}(B) = 0$ and increase the continuation value as experimentation is still profitable when the belief is ρ_{k+1}, which implies:

$$V_{t,k+1} = V_{t,k} + \frac{c}{\lambda(1 - \theta\rho_k)}$$

To see that $w_{t,k}(B) = 0$ satisfies the disclosure constraint, replace $w_{T^k,k}(B) = 0$ in the revelation constraint (9):

$$w_{t,k}(B) + V_{t,k+1} \geq V_{t,k} + e^{-r(T^k - t)} V_{T^k,k+1}$$

after replacing $V_{T^k,k+1} = \frac{c}{\lambda(1-\theta\rho_k)}$:

$$w_{t,k}(B) \geq V_{t,k} - V_{t,k+1} + (1 - e^{-r(T^k - t)}) \frac{c}{\lambda(1 - \theta\rho_k)} \tag{29}$$

As $V_{t,k+1} - V_{t,k} = \frac{c}{\lambda(1-\theta\rho_k)}$, the right hand side is negative, which means this constraint is slack. It then concludes that it is optimal to set $w_{t,k}(B) = 0$ for any t and $k \leq n^*$.

□

Proof of Lemma 7

Consider the incentive constraint right before the deadline T^{n^*} which is the last moment of experimentation. By replacing $V_{T^{n^*},n^*} = V_{T^{n^*},n^*+1} = 0$, the no shirking constraint in (5) simplifies to:

$$\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_{n^*})w_{T^{n^*},n^*}(B) \geq \frac{c}{\lambda} \tag{30}$$

As the revelation constraints are irrelevant at the deadline T^{n^*}, the principal will make the incentive constraint (30) bind. In case this constraint were slack, the payments could be decreased without modifying the agent's incentives and the principal's profits would have increased. Then, for $t < T^{n^*}$:

$$\theta\rho_{n^*} w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{\lambda} + V_{t,n^*} \tag{31}$$

after replacing $V_{t,n^*+1} = 0$ in the incentive constraint given by lemma 1. The first term, $\frac{c}{\lambda}$, represents the compensation for the instantaneous flow cost of working for the agent, and V_{t,n^*} is the future payoff foregone after revealing a signal that leads to project termination. Before concluding that equation (31) binds, it is necessary to check the *disclosure constraints*. Multiplying the constraint for the revelation of G given in equation (8) and the constraint for B given in equation (9) respectively by their probabilities $\theta\rho_n^*$ and $1-\theta\rho_n^*$ gives:

$$\begin{aligned}\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) + \\ e^{-r(T^{n^*}-t)}(w_{T,n^*}(S)\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_n^*)w_{T^{n^*},n^*}(B))\end{aligned} \tag{32}$$

It is easy to conclude that equation (32) is slack when the constraint in equation (31) binds, given that $V_{t,n^*} = \frac{c}{r}(1-e^{-r(T^{n^*}-t)})$.

Finally, the payments can be set to any value which make the no shirking constraint (31) bind and satisfy the *disclosure constraints*. Then, the payments in state n^* in the optimal contract are as given by lemma (7).

Proof of Proposition 3

Initially I assume that there is a belief ρ_n^* at which experimentation ends upon the revelation of an additional signal B, then will show that this condition is indeed independent of the calendar time t and the history. The principal's profit at t when the belief is ρ_n^* is:

$$F_{t,n^*} = \int_0^{T^{n^*}} e^{-t(\lambda+r)}(\lambda(\theta\rho_n^* - w_{t,n^*}) - c)dt \tag{33}$$

in case experimentation ends at the $n^* + 1$'th good signal where $w_{t,n^*} = \theta\rho_n^* w_{t,n^*}(G) + (1 - \theta\rho_n^*) w_{t,n^*}(B) = \frac{c}{\lambda} + \frac{c}{r}(1 - e^{-r(T^{n^*} - t)})$ and n^* is the last state. If the contract does not end at $n^* + 1$'th bad signal, then it will end at the bad signal $n^* + 2$. In that case, F_{t,n^*} becomes:

$$F_{t,n^*} = \int_0^{T^{n^*}} e^{-t(\lambda+r)}[\lambda(\theta\rho_n^*(1 - w_{t,n^*}(G)) + (1 - \theta\rho_{n^*})F_{t,n^*+1}) - c]dt \tag{34}$$

where $w_{t,n^*}(G) = \frac{c}{r}(1 - e^{-r(T^{n^*} - t)}) = V_{t,n^*}$. Now, the condition for 33 > 34 is:

$$\int_0^{T^{n^*}} -c - \lambda V_{t,n^*} dt \geq \int_0^{T^{n^*}} -\lambda\theta\rho_{n^*} V_{t,n^*} + \lambda(1 - \theta\rho_{n^*})F_{t,n^*+1} dt$$

which, after replacing the payments and integrating, simplifies to:

$$-\frac{c}{\lambda(1 - \theta\rho_{n^*})} - V_{t,n^*} \geq F_{t,n^*+1}$$

if this holds, then it is optimal to end experimentation at the $n^* + 1$'st bad signal at any t. Replacing $V_{t,n^*+1} = \frac{c}{\lambda(1-\theta\rho_{n^*})} + V_{t,n^*}$:

$$- V_{t,n^*+1} \geq F_{t,n^*+1} \tag{35}$$

Let us calculate the right hand side:

$$F_{t,n^*+1} = \int_t^{T^{n^*+1}} e^{-(s-t)(\lambda+r)}[\lambda(\theta\rho_{n^*+1} - \frac{c}{\lambda} - w_{s,n^*+1}) - c]ds$$

where $\frac{c}{r}(1 - e^{-r(T^{n^*+1} - t)}) = w_{t,n^*+1}$. Integrating this expression:

$$\frac{(1 - e^{-(T^{n^*+1} - t)(\lambda+r)})}{\lambda + r}(\lambda\theta\rho_{n^*+1} - c) - \frac{c}{r}(1 - e^{-r(T^{n^*+1} - t)})$$

hence the condition (35) holds if and only if:

$$\frac{(1 - e^{-(T^{n^*+1} - t)(\lambda+r)})}{\lambda + r}(\lambda\theta\rho_{n^*+1} - c) \leq 0$$

Finally, the following is the stopping belief:

$$\rho_{n^*+1} \leq \frac{c}{\lambda\theta}$$

Proof of proposition 4

Let us write down the principal's problem:

$$F_{0,0} = \int_0^{T^0} e^{-t(\lambda+r)}[\lambda(\theta\rho_0(1-w_{t,0}(G)) + (1-\theta\rho_0)F_{t,1}) - c]dt$$

where $w_{t,k}(G) = \frac{c}{r}(1-e^{-r(T^0-t)})$.

$$F_{t,1} = \int_t^{T^1(t)} e^{-(s-t)(\lambda+r)}[\lambda(\theta\rho_1(1-w_{s,1}(G)) + (1-\theta\rho_1)F_{s,2}) - c]ds$$

and it continues for all k until $k = n^*$. Then the derivative of $F_{0,0}$ with respect to T_0 is:

$$e^{-rT^0}[e^{-T^0\lambda}(\lambda(\theta\rho_0 - (1-\theta\rho_0)(c + F_{T^0,1})) - \theta\rho_0 c)] + (1 - e^{-T^0(\lambda+r)})\frac{\lambda(1-\theta\rho_0)F'_{T^0,1}}{\lambda+r}$$

Here, $F'_{T^0,1} = 0$, which is the derivative of $F_{T^0,1}$ with respect to T^0 as $F_{T^0,1}$ does not depend on T^0. Then, after rearranging we have:

$$e^{-T^0(\lambda+r)}[\lambda(\theta\rho_0 + (1-\theta\rho_0)F_{T^0,1}) - c] - \theta\rho_0 c e^{-rT^0}(1 - e^{-T^0\lambda})$$

where the first term denotes the marginal benefit from extending experimentation for an instant at T^0, and the second term denotes the cost of increasing experimentation time due to increased payments that should be promised in all the previous periods. After rearranging:

$$e^{-rT^0}[e^{-T^0\lambda}(\lambda(\theta\rho_0 + (1-\theta\rho_0)F_{T^0,1}) - (1-\theta\rho_0)c) - \theta\rho_0 c]$$

second derivative:

$$-(\lambda+r)e^{-T^0(r+\lambda)}[\lambda\theta\rho_0 + \lambda(1-\theta\rho_0)F_{T^0,1} - (1-\theta\rho_0)c] + \theta\rho_0 e^{-rT^0} rc$$

The first derivative has a single root, and it can be verified that at this point, the second derivative is negative which means it is a local maximum. As the first order condition has no other root, this function has only one reflection point, hence T^0 is indeed a global maximum. The optimal T^0 is then found as:

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0 + (1-\theta\rho_0)F_{T^0,1}) - (1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{36}$$

It can be checked that the second derivative is negative at $T = 0$ and at the optimal T^0, which implies that the value function of the principal is not convex in any region until T^0. This also proves that randomizing on the stopping time cannot be optimal for the principal, justifying the initial restriction to deterministic deadlines. In addition, the value function is decreasing for $t \geq T^0$. The second derivative becomes positive as T goes to ∞. However, as there is no other point at which the first derivative is zero and the second derivative is negative for $T > T^0$, I conclude that this value can never go above T^0.

Non optimality of contracts having k such that $T^{k+1}(t_{k+1}) \leq T^k$

In this section I will verify that it is never optimal to have $T^{k+1}(t_{k+1}) \leq T^k$. First, I will solve for the optimal payment schedule under this condition. Then, by replacing the payments in the principal's objective function, I will find that the profits always increase in the deadline T^{k+1} justifying the initial restriction to contracts with $T^{k+1}(t_{k+1}) \geq T^k$.

The optimal payments and continuation values in a contract in which $T^{k+1}(t_{k+1}) \leq T^k$ are:

$$\begin{aligned}
&\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k) w_{t,k}(B) = \frac{c}{\lambda} + V_{t,k} \\
&V_{t,k} = V_{t,k+1} \\
&\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k) w_{T^k,k}(B) = \frac{c}{\lambda} \\
&T^{k+1}(t_{k+1}) = T^k
\end{aligned}$$

where $V_{t,k} = \frac{c}{r}(1 - e^{-r(T^k - t)})$. It is not optimal to shorten the horizon of experimentation, hence when restricted to $T^{k+1}(t_{k+1}) \geq T^k$, it is found that $T^{k+1} = T^k$.

Now let us show this result. When the deadline is T^k and $T^{k+1} \leq T^k$, $V_{T^k,k+1} = V_{T^k,k} = 0$ and the no shirking condition (5) simplifies to:

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k) w_{T^k,k}(B) \geq \frac{c}{\lambda} \tag{37}$$

The revelation constraints are irrelevant at the deadline T^k. It is then optimal that the equation (37) binds, which provides the payments at the deadline. Then, for $t < T^k$, using $V_{T^k,k+1} = V_{T^k,k} = 0$, the revelation constraint for B becomes:

$$w_{t,k}(B) + V_{t,k+1} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)} w_{T^k,k}(B)$$

The revelation constraint for G is:

$$w_{t,k}(G) \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)} w_{T^k,k}(G) \tag{38}$$

multiplying $w_{t,k}(G)$ and $w_{t,k}(B)$ respectively by their weights $\theta\rho_k$ and $1 - \theta\rho_k$, we get:

$$\begin{aligned} \theta\rho_k w_{t,k}(G) + (1 - \theta\rho_k) w_{t,k}(B) \geq V_{t,k} + e^{-r(T^k - t)} (\theta\rho_k w_{T^k,k}(G) \\ + (1 - \theta\rho_k) w_{T^k,k}(B)) - (1 - \theta\rho_k) V_{t,k+1} \end{aligned} \tag{39}$$

where the right hand side is equal to $V_{t,k} + e^{-r(T^k - t)} \frac{c}{\lambda} - (1 - \theta\rho_k) V_{t,k+1}$. The incentive constraint to work is:

$$\theta\rho_k w_{t,k}(G) + (1 - \theta\rho_k) w_{t,k}(B) \geq \frac{c}{\lambda} + V_{t,k} - (1 - \theta\rho_k) V_{t,k+1} \tag{40}$$

Comparing the equations (39) and (40), it is easy to conclude that the incentive constraint is the binding one. Then, the payments $w_{T^k,k}(G)$ and $w_{T^k,k}(B)$ should be chosen such that the incentive constraint (40) binds and the revelation constraints are satisfied. Next I will show that $T_{k+1}(t_{k+1}) < T_k$ cannot be optimal for any k by replacing the payments into the principal's value function. First, I will look at F_{0,n^*-1} (normalizing the starting time of state $n^* - 1$ to 0) and F_{t,n^*} to show that $T^{n^*}(t_{n^*}) \geq T^{n^*-1}$. After this, I will show that it also holds for any $k < n^*$. The optimal payment schedule in state n^* was already provided for any optimal contract in subsection (4.2), given that it is the last possible state. Replacing the values $w_{t,k}(G)$ and $w_{t,k}(B)$ into the principal's problem as well as $V_{t,k} = \frac{c}{r}(1 - e^{-r(T^k - t)})$ and rearranging:

$$\begin{aligned} F_{0,n^*-1} = \int_0^{T^{n^*-1}} e^{-t(\lambda + r)} \Big[\lambda(\theta\rho_{n^*-1} - \frac{c}{\lambda} - \theta\rho_k \frac{c}{r}(1 - e^{-r(T^{n^*} - t)}) \\ - \frac{c}{r}(e^{-r(T^{n^*} - t)} - e^{-r(T^{n^*-1} - t)}) + F_{t,n^*}) - c \Big] dt \end{aligned} \tag{41}$$

In state n^*:

$$F_{t,n^*} = \int_t^{T^{n^*}(t)} e^{-(s-t)(\lambda + r)} \left[\lambda(\theta\rho_{n^*} - \frac{c}{\lambda} - \frac{c}{r}(1 - e^{-r(T^{n^*} - s)})) - c \right]$$

maximizing F_{0,n^*-1} with respect to T^{n^*}:

$$\int_0^{T^{n^*-1}} e^{-t(\lambda+r)}(1-\theta\rho_{n^*-1})\Big[ce^{-r(T^{n^*}-t)}+ e^{-(T^{n^*}-t)(\lambda+r)}[\lambda\theta\rho_{n^*}-c]-ce^{-r(T^{n^*}-t)}\Big]dt>0$$

The first and the third terms cancel out, and we are left with $e^{-(T^{n^*}-t)(\lambda+r)}[\lambda\theta\rho_{n^*}-c]$. Then, as $\lambda\theta\rho_{n^*}>c$ this expression is always positive for $k<n^*$. Hence the profit of the principal is always increasing in T^{n^*} when restricted to $T^{k+1}(t_{k+1})\leq T^k$. This implies that the optimal contract has the feature that $T^{n^*}(t)\geq T^{n^*-1}$. Finally, I need to show that a shortening of the time horizon is not optimal when the next state, $k+1$ is such that $T^{k+2}(t_{k+2})\geq T_{k+1}$ either, in other words given that in the next state the deadline is extended upon revelation of a bad signal. I replace the payment schedule for state $k+2$ which is given by the proposition 1:

$$F_{0,k}=\int_0^{T^k} e^{-t(\lambda+r)}[\lambda(\theta\rho_k(1-\frac{c}{\lambda\theta\rho_k}-\frac{c}{r}(1-e^{-r(T^k-t)}))+ (1-\theta\rho_k)[-\frac{c}{r}e^{rt}(e^{-rT^{k+1}}-e^{-rT^k})+F_{t,k+1}])-c]dt$$

where:

$$F_{t,k+1}=\int_0^{T^{k+1}} e^{-(s-t)(\lambda+r)}[\lambda(\theta\rho_{k+1}(1-\frac{c}{r}(1-e^{-r(T^{k+1}-s)}))+ (1-\theta\rho_{k+1})[F_{s,k+2}])-c]ds$$

The derivative of this whole term with respect to T^{k+1}:

$$\int_0^{T^k} e^{-t(\lambda+r)}(1-\theta\rho_k)\big[ce^{-r(T^{k+1}-t)}+e^{-(T^{k+1}-t)(\lambda+r)} [\lambda\theta\rho_{k+1}+\lambda(1-\theta\rho_{k+1})F_{T^{k+1},k+2}-c]-\theta\rho_{k+1}c(e^{-(T^{k+1}-t)(\lambda+r)}-e^{-r(T^{k+1})+t(\lambda+r)})\big]dt$$

the terms $ce^{-r(T^{k+1}-t)}$ and $\theta\rho_{k+1}c(e^{-(T^{k+1}-t)(\lambda+r)}-e^{-r(T^{k+1})+t(\lambda+r)})$ are positive and hence the whole expression is also positive as long as $\lambda\theta\rho_{k+1}+\lambda(1-\theta\rho_{k+1})F_{T^{k+1},k+2}-c\geq 0$ which is the case as $\lambda\theta\rho_{k+1}-c\geq 0$. The final one is the condition for experimentation to be profitable initially. I then conclude that it is always profit enhancing to increase T^{k+1} in the region when $T^{k+1}\leq T^k$. This means, $T^{k+1}(t_{k+1})=T^k$ and $V_{t,k+1}=V_{t,k}$. The optimal payment schedule follows from the constraints. Hence, there cannot be an optimal contract whose time horizon shortens after the release of a bad signal. Now I can conclude that in the optimal contract there cannot be any $T^{k+1}(t_{k+1}) < T^k$, which justifies the initial restriction to contracts having $T^{k+1}(t_{k+1}) \geq T^k$.

References

[1] Ufuk Akcigit and Qingmin Liu. 2016. "The Role of Information in Innovation and Competition." *Journal of the European Economic Association*, 14(4), 828-870.

[2] BBC News Technology. (2014) ''Secret Google lab 'rewards staff for failure."

[3] Dirk Bergemann and Ulrich Hege. 1998. "Venture Capital Financing, Moral Hazard, and Learning." *Journal of Banking and Finance,* 22, 700-735.

[4] Dirk Bergemann and Ulrich Hege. 2005. "The Financing of Innovation: Learning and Stopping." *RAND Journal of Economics*, 36, 719-752.

[5] Patrick Bolton and Christopher Harris (1999): "Strategic Experimentation" *Econometrica.*

[6] Alessandro Bonatti and Johannes Horner (2011): "Collaborating." *American Economic Review*, 101, 632-663.

[7] Alessandro Bonatti and Johannes Horner (2016): "Career Concerns with Exponential Learning." *Theoretical Economics.*

[8] Hector Chade and Natalia Kovrijnykh (2016): "Information Acquisition, Moral Hazard, and Rewarding for Bad News". *Journal of Economic Theory.*

[9] Peter DeMarzo and Yuliy Sannikov (2016):"Learning, Termination and Payout Policy in Dynamic Incentive Contracts.", Working paper.

[10] Dino Gerardi and Lucas Maestri (2012): "A Principal-Agent Model of Sequential Testing." *Theoretical Economics*, 7, 425-463.

[11] Renato Gomes, Daniel Gottlieb and Lucas Maestri (2016): "Experimentation and Project Selection: Screening and Learning." *Games and Economic Behavior*,96, 145-169.

[12] Brett Green and Curtis R.Taylor (2016): "Breakthroughs, Deadlines and Severance: Contracting for Multistage Projects". *American Economic Review.*

[13] Yingni Guo (2016): "Dynamic Delegation of Experimentation". *American Economic Review*, 106(8).

[14] Godfrey Keller , Sven Rady and Martin Cripps. (2005): "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73, 39-68.

[15] Nicolas Klein. (2016):"The Importance of Being Honest." *Theoretical Economics*, 11,773-811.

[16] Marina Halac, Navin Kartik, and Qingmin Liu (2016): "Optimal Contracts for Experimentation" *Review of Economic Studies*, 83, 1040-1091.

[17] Johannes Horner and Larry Samuelson (2013): "Incentives for Experimenting Agents." *RAND Journal of Economics*, 4,632-663.

[18] Godfrey Keller and Sven Rady (2010): "Strategic Experimentation with Poisson Bandits." *Theoretical Economics*, 5: 275-311.

[19] Godfrey Keller, Sven Rady and Martin Cripps (2005). "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73: 39-68.

[20] Nicolas Klein (2016): "The importance of being honest." *Theoretical Economics.*

[21] Suehyun Kwon (2014): "Dynamic Moral Hazard with Persistent States." Working paper.

[22] Jean-Jacques Laffont, and Jean Tirole (1988): "The Dynamics of Incentive Contracts." *Econometrica*, 56,1153-1175.

[23] Tracy R. Lewis. (2011): "A Theory of Delegated Search for the Best Alternative." *RAND Journal of Economics.*

[24] Tracy R. Lewis. and M. Ottaviani (2008): "Search Agency." Working paper.

[25] Gustavo Manso (2011): "Motivating Innovation." *Journal of Finance*, 66, 1823-1860.

[26] Robin Mason and Juuso Valimaki. (2011): "Dynamic Moral Hazard and Stopping". Working Paper.

[27] Sofia Moroni. (2016): "Experimentation in Organizations". Working paper.

[28] Pauli Murto and Juuso Valimaki. (2011): "Learning and Information Aggregation in an Exit Game". *Review of Economic Studies.*

[29] Julien Prat, and Boran Jovanovic. (2013): "Dynamic Contracts when the Agent's Quality is Unknown". *Theoretical Economics*, 9, 865-914.

[30] Dinah Rosenberg, Eilon Solan and Nicolas Vieille, 2007. "Social Learning in One-Arm Bandit Problems". *Econometrica*, 75(6).

[31] Yuliy Sannikov (2008): "A Continuous-Time Version of the Principal-Agent Problem." *The Review of Economic Studies*, 75, 957-984.

[32] Stephen E. Spear and Sanjay Srivastava. (1987): "On Repeated Moral Hazard with Discounting." *The Review of Economic Studies*, 54, 599.