Contracting for Experimentation and the Value of Bad News*

Sinem Hidir†

Job Market Paper

January 2015

Abstract

I study optimal contracting in a model in which a principal hires an agent in order to experiment on a project of unknown quality. The principal provides the resources needed for experimentation and at each moment the agent has the choice between working or keeping the benefits for himself. While the agent experiments, news arrives in form of good or bad signals about the underlying state. Lack of signals may be either due to the agent's shirking or due to the fact that it is taking time for the project to yield results. The optimal contract incentivizes the agent to work and reveal the signals as they arrive. It consists of history dependent bonus payments and a termination rule in which the current deadline is updated each time a bad signal is revealed. The principal minimizes the bonus payments and rewards the agent through increased continuation values, hence extended experimentation time, upon revelation of bad signals. If experimentation stops before a deadline is reached, it stops at the belief which is the same as in the first best benchmark.

Keywords: Dynamic moral hazard, principal-agent model, innovation, experimentation, private signals, Bayesian learning.

JEL Codes: D82, D83, D86, O32.

*I am grateful to my advisor Jacques Crémer, Renato Gomes, Christian Hellwig, and Thomas Mariotti. I have also benefited from the comments of George Lukyanov, Lucas Maestri, Teck Yong Tan and the participants at the Applied Theory Workshop Toulouse. Mistakes remain mine.

†Toulouse School of Economics. email: sinem.hidir@tse-fr.eu. Please find the latest version at https://sites.google.com/site/sinemhidirecon

1 Introduction

Innovative activities are carried out under a lot of uncertainty. It is not possible to know at the outset whether a promising idea will indeed turn into a successful project or not. To learn about this requires costly resources and experimentation, and abandoning once the belief about its success becomes sufficiently low. In addition, many times experimentation is carried out by third parties, introducing an agency problem. This kind of situation is very relevant in innovation intensive industries, such as a pharmaceutical company hiring a scientist to test the effects of a drug.

This paper studies how a principal should optimally provide incentives to an agent to experiment in presence of dynamic moral hazard and private information that arises over time. Only when the agent chooses to put in costly effort, good or bad signals may arrive about the project quality which is a new feature compared to the current literature.

When the effort choice and progress are private information of the agent, I show that incentives should be provided to experiment and reveal the signals (*outcomes*) through a contract consisting of bonus payments and a termination rule. In this contract, the initial time allocated to the agent for experimenting is extended each time an additional bad signal is revealed by the agent until the final one which concludes that experimentation no more profitable.

The agent is hired to work on an innovative project which requires constant investment by the principal. At any moment, the agent has the choice between working (experimenting) or shirking and keeping for himself the resources meant for experimentation. Working is equivalent to choosing to pull the risky arm and experimenting by incurring the cost and shirking is equivalent to pulling a safe arm which lets the agent save the cost. The agent also has the choice of revealing or not the signals upon arrival, and in case of keeping a signal, revealing in the future or not. The signal structure comes from the "exponential bandit" with Poisson arrival rates[1] which has been used in the literature on experimentation. When the project is good, it can succeed or fail with some probability whereas a bad project always fails, hence one success is conclusive. In the current literature, lack of success is considered equivalent to a failure. When time passes without the

[1]Bolton and Harris (1999), Keller, Rady and Cripps, (2005), and Keller and Rady (2010)

realization of a success, at some point the belief becomes pessimistic enough and the project is abandoned. This means even without the agency problem, there is a deterministic deadline at which experimentation is abandoned. The information structure is the same when the agent works on a bad project as when he shirks: no signal is observed in both cases. I introduce the novel feature that news arrives over time with an arrival rate and only while the agent experiments in form of good (success) and bad (failure) signals. This signal structure implies that there may be a period of time when no signal is realized, which means learning is not continuous. The lack of signals can be either because the agent is not working or because the project did not provide results as yet. The possibility of getting bad signals help distinguish between an agent who works on a bad project and an agent who shirks. The belief about project quality only goes down upon the revelation of bad signals. This implies that without the presence of agency, the principal would continue until collecting enough bad signals, hence there is no deterministic deadline. In addition, the fact that signals arrive only while the agent experiments implies that when the agent shirks, the beliefs of the principal and the agent do not differ.[2]

The assumption of good and bad signals is especially relevant in environments where an innovative project is being tested. In addition, it highlights the fact that it is not known a priori long it will take until enough information will be gathered in order to make a decision.

There are two layers of frictions in this problem: the moral hazard due to the possibility of the agent to shirk and keep the resources provided by the principal for his own use and the information rent due to the private observation of the signals, which result in two types of incentive constraints. The first type makes sure that the agent prefers to work rather than shirk at any moment. The second type of constraints make sure that the agent does not want to hide or delay revealing a signal. The question then is one of finding how the agency rent should be optimally

[2]This is the opposite case in the current literature on experimentation where the belief goes down as long as no success is observed such as Horner and Samuelson (2013) or when the underlying state has a Markov transition, such as Kwon (2014) where an informational rent is born because the agent's deviation leads him to hold a different prior than the principal. 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 this setting, the principal does not find it optimal to take the outside option and continue with the project later on.

allocated between bonus payments and continuation values subject to satisfying these constraints. I show that the principal prefers paying the agent through continuation values, which translates to longer experimentation time, rather than bonus payments whenever possible.

I find that the optimal contract consists of history dependent bonus payments, an initial deadline and a rule for extending this deadline upon revelation of bad signals. The principal initially lets the agent experiment for some time, and every time the agent reveals a bad signal, extra experimentation time is allocated. This is true until the belief falls down to the level at which experimentation would also stop in the first best benchmark which happens which I call upon revelation of *terminal* bad signal. The only positive payments are made when the contractual relationship terminates due to the revelation of a good signal or the terminal bad signal. The bonus payment upon good signal decreases while getting closer to the deadline at a given belief (or *state*), and has an upward jump after the revelation of each bad signal when the belief becomes more pessimistic. The agent is willing to reveal the bad signals (except the last one which terminates the relationship) without receiving any rewards, but the principal finds it optimal allocate more experimentation time by shifting the current agency cost to expected payments in future experimentation time.

The intuition is as follows: the principal commits to a deadline in order to control the moral hazard rent of the agent, in other words how much the agent can shirk. This is a distortion as experimentation may end at a belief which is optimistic due to the presence of the deadline. As time passes and no signal arrives, it becomes more likely that the agent has been shirking and the termination of the relationship serves as a punishment. On the other hand, the revelation of a bad signal shows that the agent has actually been experimenting. Then, the extension in the time horizon is equivalent to a decrease in the initial distortion due to the deadline. While doing this, the principal does not increase the expected payment to the agent, as she is only shifting the cost from immediate bonus payments to expected payments in future experimentation horizon.

While the principal increases the agent's continuation value through longer experimentation horizon after revelation of a bad signal, there are two effects. First, by decreasing the current bonus payments promised to make the agent reveal the signals, she back loads the cost of incentive to the extra experimentation time after the release of a bad signal. [3] While the principal does this she keeps the total agency rent of the agent constant. Second, the time horizon of the contract is extended, which implies higher expected benefit due to longer experimentation time at a belief at which experimentation has still positive value.

The extension in the time horizon is larger the earlier the agent reveals a bad signal, and decreases as getting closer to the deadline. The reason is that extending the horizon is less costly the farther the current deadline is found from today. The increase in the continuation value is lower for more pessimistic beliefs, when it is more likely for the agent to receive a bad signal.

I initially find the optimal contract using local incentive constraints, and verify in the end that these are sufficient for global incentive compatibility. This justifies the use of the first order approach, hence the optimal contract can be fully characterized using the local incentive compatibility constraints. Then, it is sufficient to make sure that the incentive constraint for working binds and the revelation constraints which make sure that the benefit to revealing a signal decreases while getting closer to the deadline are also satisfied.

Future rewards distort today's incentives because when the agent reveals a success, he is giving up the benefits from remaining in the contract. Hence, the rent of the agent is increasing in the amount of time remaining which is the reason that the principal commits to a termination rule. In order to make the agent reveal a good signal or the terminal bad signal, the principal has to compensate him for his outside option of hiding it and remaining in the project. Then, given that the agency cost is already incurred, it is profitable to experiment longer as long as the intrinsic value of experimentation is positive. The principal prefers to provide incentives for the agent to work through continuation values after the revelation of bad signals rather than through bonus payments. Experimentation often stops inefficiently early due to deadlines but in case it stops before a deadline is reached, it stops either due to the release of a good signal or at the same stopping level of belief as in the first best benchmark.

This paper relates to the literature on dynamic incentives for experimentation in presence of agency. Bergemann and Hege (1998, 2005) are the first to examine

[3] This is the opposite case in a model in which there is only one type of signal which is a success that would end experimentation, where a higher continuation value tomorrow makes the agent less eager to work as he risks realizing a success and giving up future benefits.

incentives for experimentation in a principal agent model with no commitment by considering an entrepreneur seeking funding from an investor to carry out a risky project. Another paper with a similar setting is Horner and Samuelson (2013) who find downsizing the project in order to decrease the continuation value of the agent under the assumption that the principal cannot commit to a contract. The possibility of a success makes the agent less eager to work as he risks giving up future benefits of remaining in the project. In both of these models, the agency problem comes from the non observability of effort. The belief goes down as long as there is no success and experimentation ends too soon compared to the first best.

There is a more recent literature on contracting for experimentation. A paper that is related to mine is Maestri and Gerardi (2012) who study contracting for information acquisition when the agent incurs cost to get private signals in each period in form of soft information about the project quality. As the agent is sure to get a signal in any period in which he incurs the cost, there is a fixed deadline and the agent's rent is purely an information rent due to the possibility of guessing a good state without incurring any cost. In Mason and Valimaki (2011), the agent's cost of effort is convex hence has tendency to smooth his effort over time, which implies his continuation value always goes down. This changes in the setting of this paper as working can also result in a bad signal hence a higher continuation value which makes working today more attractive.

Halac, Kartik and Liu (2013) study optimal contracts for experimentation in presence of both moral hazard and adverse selection on agent's type, when both types of the agent can succeed on a good project. Gomes, Gottlieb and Maestri (2013) consider the presence of two dimensional adverse selection. Bonatti and Horner (2013) consider an agent who has career concerns. Klein (2014) studies how a principal should incentivize an agent to choose the honest way of experimenting when he also has access to a cheating option. Kwon (2014) studies a dynamic moral hazard problem in which an informational rent is endogenously born due to the persistent underlying state and private effort choice of the agent.

Guo (2014) studies dynamic delegated experimentation without transfers when the agent has a private prior on the project quality. She finds *sliding deadlines* when considering the case of inconclusive success and hence that the deadline for experimentation is shifted forward upon each success. The reason is that the belief about the project quality increases and it is optimal to let the agent experiment longer, whereas the dynamics are very different in the current paper. I find that there is extended time upon revelation of bad signals even though the belief becomes more pessimistic, and the underlying reason is that the principal can extend the horizon of experimentation without increasing the agency cost. The main assumption driving this result is that the belief is only updated upon revelation of signals, hence as long as the belief is not low enough experimenting longer is always profitable.

A recent working paper by Green and Taylor (2014) studies a multistage project which is of a certain value and and shows that the completion of one stage leads to allocation of extra time for the completion of the next stage. Finally, this paper could also relate to the literature on delegated search, such as Lewis and Ottaviani (2008) and Lewis (2011).

Finally, a contemporaneous working paper is Akcigit and Liu (2014) who consider two firms competing for an innovation. They focus on the fact that when one firm reaches a dead end, the other may keep experimenting on that arm inefficiently and explore what a social planner would do. The similarity to this paper is in the signal structure that they also have good and bad signal, however one signal is conclusive about whether the research line is a good or bad one.

The outline of the paper is as follows. Section 2 explains the model, payoffs and strategies, section 3 provides the main results, section 4 leads through the solution of the optimal contract, section 5 provides some extensions and section 6 concludes.

2 Model

I consider a continuous time model in which a principal (she) hires an agent (he) in order to learn about an uncertain state of the world (project quality) through experimentation. The principal owns a project whose quality depends on the state initially unknown to both. The common prior that the state is good is ρ_0 and bad is $(1 - \rho_0)$. A good project has net value normalized to 1 and a bad project has sufficiently negative value. The state can be learned through costly experimentation which requires a flow investment c by the principal. The agent is hired to experiment, but he could also shirk and divert benefits to his own use, which is unobservable to the principal.[4] The principal and the agent are both risk neutral and share the same discount factor. Outside options are zero and the agent has limited liability.

Signal structure: While experimenting, signals about project quality arrive over time with a Poisson arrival process. When costly effort is exerted over a time interval $[t, t + dt]$ by incurring cdt, a signal denoted by z_t (or an *outcome*) arrives with probability λdt. The arrival rate is independent of the project or signal type. The signal is always B if the state is bad. If the state is good, the signal is G with probability θ or B with probability $(1 - \theta)$. Signals are verifiable, they can be hidden but not constructed or modified.[5]

The signal G reveals that the project is good and upon its revelation, experimentation ends and a net benefit of 1 is realized by the principal. Upon revelation of a signal B the belief is updated as follows:

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

where ρ_k is the belief when k bad signals have been revealed, which will be the public belief in case the agent reveals the signals as they arrive.

Contracts: The principal offers and commits to a long term contract at $t = 0$. In case the agent rejects, both sides receive their outside option equal to zero. The revelation principle holds, hence contracts will be contingent on the public history of signal revelations by the agent.

The public history at t, h^t consists of the signal revelations by the agent:

$$h^t : \{x_s \in \{0, G, B\}, s \leq t\}$$

Where x_s is the signal revelation of the agent at time s. A contract is denoted by $(w, y) : \mathcal{H} \to \Re_+ \times \{0, 1\}$, where $\mathcal{H}$ denotes the set of possible public histories.

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

[5] The discrete time analogue of this setting would be one in which each period that 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 $1 - \lambda$.

The first component, $w_t(h^t)$ represents the history dependent sequence of bonus payments to the agent at time t. The second element $y_t(h^t)$ denotes the decision of the principal whether to fund the project or not at t as a function of the history.

As the revelation of a good signal causes the relationship to end, the public history can be simplified as the revelations of bad signals until time t and the number of signals already revealed is given by k where $\int_0^t \mathbb{1}(x_t = B) = k$.

Agent's strategy: The agent's private history h_A^t consists of his past decisions to put in effort or not and the signals that he realized until time t:

$$h_A^t = \{(a_s, z_s), s \leq t\}$$

Where $z_t \in \{0, G, B\}$ is the realization of a signal at time s. I denote the pure strategy of the agent by $\sigma = (a, x)$:

- $a_t : h_A^t \rightarrow \{0, 1\}$
- $x_t : h_A^t \rightarrow \{0, G, B\}$

The first component is the binary effort strategy where 1 denotes the decision to put in effort and 0 the decision to shirk and keep the benefit c. The second term is the reporting strategy of the agent: he can reveal any signal or signals among the ones he has received until t and that he has not already revealed. The agent has the possibility to keep and reveal a signal later on, he can hide but cannot produce a fake signal. This implies that the private history of the agent can possibly be extremely complicated. However, under the optimal contract the agent will reveal the signals as they arrive. The agent's private history h_A^t coincides with the public history h^t as long as he reveals the signals upon receiving them. Even if the agent shirks, as no signals arrive during that time, the beliefs will not be updated. The beliefs of the principal and the agent can only differ in case the agent receives and hides a signal. I denote by σ^* the strategy under which the agent works as long as the principal keeps investing and reveals the signals as they arrive. This means, $a_t^* = 1$ as long as $y_t = 1$ and $x_t^* = z_t$ for any signal realization.

Agent's payoff: The agent's expected utility is additive in the bonus payments he receives and the benefit he gets from shirking:

$$V_0(\sigma, w, y) = E\left[\int_{t=0}^{\infty} e^{-rt} y_t [w_t + (1 - a_t)c] dt\right]$$

Principal's Problem: The present expected value of the principal from the contractual relationship at time zero is:

$$F_0(\sigma, w, y) = E\left[\int_{t=0}^{\infty} e^{-rt} y_t (a_t \lambda \theta \rho^t \chi_{n_t=0} - w_t - c) dt\right]$$

ρ^t is the belief at t which depends on how many bad signals have already been revealed and n_t is the number of good signals already realized until t. The term $\chi_{n_t=0}$ is a binary variable and takes into account that the principal gets the benefit of 1 from the project only upon the realization of the first good signal. The principal's problem is then:

$$\max_{w,y} F_0$$

subject to satisfying the agent's incentive compatibility constraint to follow the recommended strategy:

$$V(\sigma^*, w, y) \geq V(\hat{\sigma}, w, y)$$

for any other $\hat{\sigma}$.

Benchmark without agency: I consider the principal's problem when she carries out experimentation without the agent. The first best solution is a stopping belief level where n^* is the lowest n such that:

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

where ρ_n^* is the lowest belief at which experimentation is profitable. This condition says that experimentation continues as long as the benefit from an instant of experimentation minus cost c is positive, and stops as soon as the belief falls below this level, which happens upon realization of the $n^* + 1$'th bad signal. As signals arrive with Poisson arrival while the agent experiments, there is predetermined date at which the belief is sure to be low enough. This implies that without the agency problem, there is no stopping time but a belief level. The belief ρ_n^* is the lowest belief at which experimentation is profitable and will be shown to coincide with the belief at which experimentation stops before a deadline is reached in the presence of agency.

Assumption 1. $\lambda\theta\rho_0 - 2c \geq 0$. This assumption ensures that experimentation is profitable at least for an instant dt in the presence of agency. The term $\lambda\rho_0\theta$ is the benefit from an instant of experimentation, and $2c$ is the total cost of an instant of experimentation for the principal: the first c is the cost incurred 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 works. The reason for this last point is that by working the agent gives up the opportunity to keep the benefit c for himself.

3 The Optimal Contract and the Deadline Schedule

This section provides the main results of the paper. Section 4 will lead through the details of the solution. First, I will show that the principal's problem simplifies to choosing a deadline schedule subject to satisfying the incentive constraints of the agent.

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

This says that once the principal stops investing in experimentation, he will not start again in the future.

Lemma 2. *If the deadline is ever updated during the relationship, it should be at times at which bad signals are revealed: at t such that $x_t = B$.*

Proof. Consider a termination rule $T(h^t)$ where T denotes the deadline as a function of the public history h^t. The good signal is equivalent to a success and ends experimentation. Then, the only possible histories h^t such that the deadline is updated at t either have $z_t = B$ or $z_t = 0$ (no signal). Call a history $\hat{h}^t$ such that no bad signal has yet been realized, and the deadline T gets updated to $\hat{T}$ at t. Then, this contract is equivalent to the following one: the initial deadline is $\hat{T}$ and the updating rule $T(h^t)$ for any other history h^t and t is kept constant. Then, it is without loss to restrict the possibility of updating the deadline to the times at which bad signals are revealed. $\square$

The public history can now be simplified as the times at which bad signals are revealed: $h^t = \{t_1, t_2, ..., t_k\}$ when the belief is ρ_k. These are the only elements of history that will be relevant for the contract. The initial deadline is T^0. If the first bad signal is revealed at t_1, the deadline becomes $T^1(t_1, T^0)$ as t_1 and T^0 are the only relevant elements of the public history. Then, when the second bad signal is revealed at t_2, the new deadline can be denoted as $T^2(t_2, T^1)$ as t_1 and T^0 have already been taken into account while determining T^1. Then, at any moment, the current deadline is actually a function of the previous deadline and the time of revelation of the last bad signal.

Definition 1. *A deadline schedule is denoted by $T = (T^k(h^t))_{k=0}^{n^*}$ where T^k specifies the stopping time at a given moment t when the public belief is ρ_k and the history is h^t. The deadline T^0 determines the time initially allocated to the agent for experimenting such that if reached without any signal revelation, the contract terminates.*

The termination rule says that at any belief ρ_k, the relationship will end as soon as a signal G is revealed, at T^k if no other bad signal is revealed, or when the $n^* + 1$'th bad signal is revealed.

I will use the simplified notations T^k for the deadline, $V_{t,k}$ and $F_{t,k}$ respectively for the agent's and the principal's continuation values, and $w = (w_{t,k}(G), w_{t,k}(B))_{t=0}^{\infty}$ respectively for the bonus payments upon revelation of a good and a bad signal at time t and state k (the belief is ρ_k). I am not omitting the public history, h^t, while doing this. Any optimal contract should have $w_{t,k}(0) = 0$: the payment to the agent is equal to zero as long as no signal has been reported.[6]

Now I will rewrite the principal's problem. The index k which denotes ρ_k changes to $k+1$ as soon as another bad signal is revealed. If the k'th bad signal is revealed at time t and the current deadline is T^k, the present value of the contract to the principal at that moment is:

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

[6]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 bonuses payed only upon revelation of signals.

where $e^{-(s-t)\lambda}$ is the probability of time s being reached with no signal arriving conditional on the agent working, hence that the state is still k. The term $e^{-r(s-t)}$ is the discount factor that applies when a signal arrives at time s. In an infinitesimal time period of dt, with probability λdt a signal arrives and it is revealed. With probability $\theta\rho_k$ the signal is G and if the agent reveals it the contract ends while the principal makes the payment $w_{t,k}(G)$, or the signal is B and the state moves to $k+1$ providing the principal the continuation value $F_{t,k+1}$. The detailed derivation of equation (1) is provided in the Appendix. Then, the problem of the principal at time zero is:

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

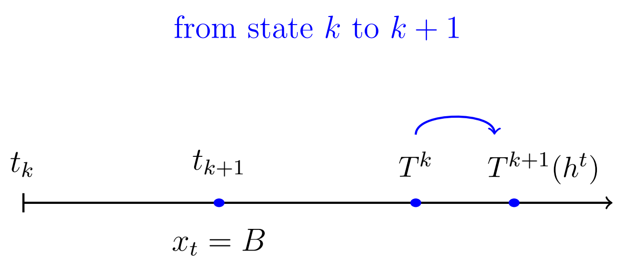
subject to $w_{t,k}(G)$ and $w_{t,k}(B)$ satisfying incentive compatibility. The incentive compatibility constraints make sure that the agent is willing to work and reveal the signals as they arrive. While solving for the optimal contract, initially I restrict attention to the local incentive compatibility constraints and verify in the end 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 in the contract rather than shirking at any moment. The second type of constraints are the *revelation constraints* (or disclosure) which make sure that the agent is actually willing to reveal the signals he acquires without delay and checks for any possible deviations. These will be explored in more detail in the next section. Propositions 1 and 2 provide the optimal contract, and I go over the proof of the solution in section 4.

Proposition 1. *In any optimal contract, the only positive payments are made when termination happens before the current deadline either due to the revelation of a success, or upon the revelation of the terminal bad signal:*

- $V_{t,k} = \int_t^{T^k} ce^{-r(s-t)}ds$. *The continuation value of the agent at any moment and state is equal to his option of shirking and keeping the benefit c until the current deadline* T^k.

- $w_{t,k}(G) = V_{t,k}$. *The payment upon revelation of a success in any state is set to its minimum value which is the continuation value the agent would get by hiding it and remaining in the project until the deadline.*

Figure 1: Updating of T^k



- $w_{t,k}(B) = 0$ *for* $k < n^*$. *The agent is induced to reveal the bad signals without any payment as long as this signal 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 of experimentation, as an additional (good or bad) signal will terminate the contract, both payments are set at least equal to the value that the agent would get by hiding the signal and remaining in the project until the end. (Proof in section* 4*)*

The reason I use the term “any optimal contract” is due to the infinite number of payment pairs that may satisfy the last item. However, the optimal payments for states $k < n^*$ are uniquely determined. The principal does not have to make a positive payment upon revelation of bad signals as long as this signal does not lead to the termination of the contract. The reason is that as long as the agent's continuation value does not decrease after revealing a bad signal, he is willing to reveal it without any payment. The payment upon good signal, $w_{t,k}(G)$, is equal to the agent's outside option which is the benefit he would get from remaining in the project until the current deadline in state k and is independent of the next state. This is because upon receiving and hiding a success, the agent does not have an incentive to work again and hence the state cannot change. Even though initially there are infinite number of possible deviations, I show that the only relevant one is to not work again and keep the investment. The earlier from the deadline the agent reveals a success, higher will be his payment, as by revealing this signal and ending the relationship, he is giving up the opportunity to keep the investment c until the current deadline. Next proposition provides the change in the continuation value and hence the deadline upon revelation of a bad signal.

Proposition 2. *In any optimal contract, the continuation value of the agent is increased upon revelation of a bad signal, as long as this signal is not terminal, through a longer time horizon of experimentation as follows:*

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

- $T^{k+1}(t_{k+1}, h^t) > T_k(h^t)$ *for* $t_{k+1} \leq T^k$ *and* $k < n^*$. *The jump in the continuation value is provided as an extension in the time allocated to the agent for experimentation:*

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

(Proof in section 4)

The revelation of a bad signal leads to an increase in the continuation value of the agent which is translated into extra experimentation time, hence an extension of the current deadline. This is because the principal can provide incentives either through bonus payments or experimentation time and he prefers the latter as long as possible and as long as experimentation is still profitable. As the agent always has the possibility of getting benefit by not working, longer experimentation time is equivalent to higher benefit for the agent. Figure 1 demonstrates the updating of the deadlines. In the termination rule, T^k denotes the stopping time when the belief is ρ_k as a function of the history. When the $k+1$'th bad signal is revealed, the deadline becomes $T^{k+1}(t_{k+1})$. The deadline is updated every time a bad signal is revealed until the last one which ends the relationship. The belief ρ_{n^*+1} is the threshold at which experimentation ends before a deadline is reached and coincides with the stopping level of belief in the benchmark case without agency. The change in the deadline is higher the earlier the bad signal is revealed: the extended time is added at the end of the current deadline T^k, and due to the discount factor, for the same cost the change in the time horizon is higher the more distant T^k is from t. Let us provide some intuition of these results. The principal uses deadlines in order to control the moral hazard rent of the agent, in other words the maximum benefit he could get by shirking, which is given by $V_{t,k} = \int_t^{T^k} ce^{-r(s-t)}ds$. Hence, experimentation may end inefficiently early at a belief level at which it would be profitable to continue. This is the reason for the increase in the time horizon upon revelation of a bad signal: by minimizing the bonus payments, the principal provides incentives by allocating more time for experimentation upon revelation of bad signals while keeping constant the expected payment to the agent. The agent's gain from experimenting has three components: the payment upon revelation of a good signal, upon revelation of a bad signal and the continuation value in the more pessimistic state after the revelation of a bad signal, and what matters for the agent's incentives is his expected payoff from experimenting and not the decomposition of it. By increasing the agent's continuation value in the more pessimistic future state which is reached when he reveals a bad signal, the principal decreases the current bonus payments that should be promised to the agent. However, there is a limit to how much the principal can back load the agency cost in to future experimentation time. This is given by the revelation constraints which determine the minimum bonuses that make sure the agent reveals the signals. Then, the rest of the moral hazard rent is allocated in terms of extended contract horizon to the agent. While extending the horizon the principal is shifting the current bonus payments to future expected payments.

The increase in the continuation value is not a necessary condition for inducing the agent to reveal a bad signal. As long as the revelation of the bad signal does not decrease his continuation value, the agent is willing to reveal it at no cost. However, the extension in the deadline upon revelation of a bad signal is optimal because it provides more experimentation time while keeping constant the total agency cost.

4 Solving for the Optimal Contract

This section leads through the solution of the optimal contract provided by propositions 1 and 2. The principal's problem is to choose an optimal termination schedule while minimizing the expected payment to the agent. As a first step, the agent's continuation value is provided and then, the incentive constraints that should be satisfied for making him work and reveal the signals. In order to find out which constraints bind, I consider the conditions at the deadlines. Then, the rest of the values are obtained by making use of the conditions at the deadlines.

Initially, I restrict attention to local constraints. This provides the result on the bonus payments and the updating rule of the deadlines. Finally, I conclude by showing how the initial deadline T^0 is computed. In the end, I verify that local constraints are indeed sufficient for global incentive compatibility. Even though the types of deviations for the agent are huge, I show that there is only one that is relevant, which is the deviation to not work.

4.1 Agent's continuation value

At any moment while the agent experiments, he has a probability to receive a good or a bad signal or no signal. Then, the evolution of his continuation value if he follows the strategy induced by the contract is as follows:

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

The agent receives a signal with probability λdt which is good or bad and he reveals it. In case of a good signal, he receives the payment $w_{t,k}(G)$ and the contract ends. In case the signal is bad 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 gets the continuation value $V_{t+dt,k}$ at $t+dt$, as the state does not change when the agent shirks. Letting dt go to 0 gives:

$$-V'_{t,k} + (\lambda + r)V_{t,k} = \lambda[\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1})] \tag{2}$$

4.2 Incentive Compatibility

Now I will provide the incentive compatibility constraints. First one is the one for working and second one for revealing the signals.

The no shirking constraint

This is the type of constraint which makes sure that the agent does not want to deviate to shirk at any moment.

Lemma 3. *The local incentive constraint which makes sure that the agent works is:*

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) \geq \frac{c}{\lambda} + V_{t,k} - (1-\theta\rho_k)V_{t,k+1} \tag{3}$$

Proof. If the agent shirks for an infinitesimal time period of dt, no signal arrives and the state k does not change. He gains the benefit cdt and gets the continuation value $V_{t+dt,k}$. Following is the condition for the agent to prefer to work than shirk for an infinitesimal time period:

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

Letting dt go to 0:

$$-V'_{t,k} + rV_{t,k} \geq c \tag{4}$$

After replacing this condition in the agent's continuation value in equation (2), I obtain the no shirking constraint (3). □

Lemma 3 makes sure that the agent does not benefit from deviating to shirk for an instant. I will now explain why the local constraint for working in equation (3) is enough for global incentive compatibility for working. The reason is that when the agent deviates and shirks, he does not get any signals and the beliefs of the agent and the principal will not differ. Then, his future incentives to work are not modified by this deviation. Consider 2 different histories at t, h^t and $\hat{h}^t$ which share the same history until time $t_1 < t$ given by h^{t_1} and hence the belief ρ_k. In the first history h^t, the agent shirks from t_1 until t. In the history $\hat{h}^t$, the agent works until t but does not receive any signals. Hence, given that no signals arrive from t_1 until t, the public belief at t in both histories is ρ_k. Then, these two histories are actually equivalent at t: $h^t = \hat{h}^t$, as the history consists of the times at which signals are revealed. As the agent's belief and continuation value at t are the same for the two histories, this deviation did not lead to any informational difference between the agent and the principal. Hence, the agent's deviation at a time t does not affect his incentives for deviating at a future date. Then, if equation (6) is satisfied for any t, it will also be satisfied globally.

In state k, as the contract ends as soon as T^k is reached (hence the $k+1$'th bad signal has not been revealed), $V_{T^k,k} = 0$. Then, the condition for $V_{t,k}$ can be found by integrating $-V'_{t,k} + rV_{t,k} \geq c$ and using the boundary condition:

$$V_{t,k} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) \tag{5}$$

Hence equation (5) is a necessary condition for the continuation value of the agent. If the constraint in equation (4) binds at any t, then equation (5) will also hold as an equality.

Revelation constraints

Second type of constraints that should be satisfied are the *revelation constraints* which make sure that the agent is willing to reveal the signals upon having received them. The ability to keep the signals and reveal later on adds the complication of infinitely many possible histories after deviations, such as hiding a signal and experimenting in order to possibly get another signal. After hiding a signal, the agent could either shirk and reveal it in the future, or work in order to get another signal or signals. For now I will restrict attention to the simplest deviation to hide a signal and shirk, in other words local constraints, and I will show in subsection 4.5 that these conditions are indeed sufficient to account for all possible deviations.

Lemma 4. *Following are the local revelation constraints which make sure that the agent is willing to reveal the signals that arrive without delay:*

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq c \tag{6}$$

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

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

Proof. First, I derive equation (6) which makes sure that the agent is willing to reveal G upon receiving it instead of hiding to reveal it at $t + dt$ and shirking in the meantime. The constraint for not waiting to reveal G is:

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

which as dt goes to 0, leads to:

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq c \tag{8}$$

Second, I will derive equation (7). The constraint which makes sure that it is not profitable to wait before revealing a signal B 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:

$$-w'_{t,k}(B) - V'_{t,k+1} + r(w_{t,k}(B) + V_{t,k+1}) \geq c \tag{9}$$

□

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. Now, assuming this is the case, I can solve the differential equation (6) using the boundary condition at T^k:

$$w_{t,k}(G) \geq \frac{c}{r}(1 - e^{-r(T^k-t)}) + e^{-r(T^k-t)} w_{T^k,k}(G) \tag{10}$$

I will do the same for B by integrating equation (7) and using the condition at the deadline T^k:

$$w_{t,k}(B) + V_{t,k+1} \geq \frac{c}{r}(1 - e^{-r(T^k-t)}) + e^{-r(T^k-t)}(w_{T^k,k}(B) + V_{T^k,k+1}) \tag{11}$$

where $w_{T^k,k}(B) = 0$. The equations (10) and (11) will be used after having found the values at the deadlines T^k in order to get the conditions on the payments $w_{t,k}(G)$ and $w_{t,k}(B)$.

4.3 The condition at the deadlines T^k

I will now derive the condition at T^k for any k. The reason for focusing on the deadlines is, first it is already known that $V_{T^k,k} = 0$ as the contract will terminate at T^k, and second the revelation constraints are irrelevant as the relationship will terminate in case no more signals are revealed.

Lemma 5. *The no shirking constraint is the only condition that should be satisfied at the deadline T^k for any k and it binds in the optimal contract:*

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k) w_{T^k,k}(B) = \frac{c}{\lambda} - (1-\theta\rho_k) V_{T^k,k+1} \tag{12}$$

Proof. Consider the incentive constraint right 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:

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) \geq \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1} \tag{13}$$

At T^k, the revelation constraints are irrelevant: if a signal arrives at that moment, the agent is willing to reveal it without receiving any payment. Then, the condition in equation (13) is the only constraint that should be satisfied, and it will bind. In case it were slack, then the principal could decrease the expected payment while still satisfying this constraint. □

4.4 Finding $w_{t,k}(G)$, $w_{t,k}(B)$, $V_{t,k}$ and T^k

I will start by finding the values in state n^* which is the last state in which experimentation can be carried out and hence $V_{t,n^*+1} = 0$ for any t as the contract ends once the $n^* + 1$'th bad signal is revealed. The threshold belief for stopping will be provided in proposition (3).

Lemma 6. *The payments in state n^* in any optimal contract should satisfy the following conditions:*

$$\theta\rho_{n^*} w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*} \tag{14}$$

$$w_{t,n^*}(S) \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) + e^{-r(T^{n^*}-t)}w_{t,n^*}(S) \tag{15}$$

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

Proof. Look at the incentive constraint just before 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 (3) simplifies to:

$$\lambda[\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_{n^*})w_{T^{n^*},n^*}(B)] \geq c \tag{16}$$

As the revelation constraints are irrelevant at the deadline T^{n^*}, it is optimal to make the incentive constraint in equation (16) bind. In case this constraint were slack, the payments could be decreased without modifying the incentives and the principal's profit would have increased (decreasing these payments will only relax the earlier constraints). Then, for $t < T^{n^*}$:

$$\theta\rho_{n^*} w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{\lambda} + V_{t,n^*} \tag{17}$$

after replacing $V_{t,n^*+1} = 0$ in the incentive constraint given by lemma 3. The first term, $\frac{c}{\lambda}$, represents the compensation for the instantaneous flow that the agent could obtain by shirking, and V_{t,n^*} is the future payoff foregone after revealing a signal that leads to the termination of the project. Before concluding that equation (17) binds, it is necessary to check the *revelation constraints*. Multiplying the constraint for the revelation of G given by equation (10) and the constraint for B given by equation (11) respectively by their probabilities $\theta\rho_n^*$ and $1-\theta\rho_n^*$ gives:

$$\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) + \\ e^{-r(T^{n^*}-t)}(\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_n^*)w_{T^{n^*},n^*}(B)) \quad (18)$$

It is easy to conclude that equation (18) is slack when the constraint in equation (17) binds, given that $V_{t,n^*} \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)})$. Finally, the payments can be set to any value which make the no shirking constraint (17) bind and satisfy each of the *revelation constraints*. Then, the payments in state n^* in the optimal contract are as given by lemma (6). □

The Appendix shows the non optimality of contracts having any $T^{k+1}(t_{k+1}) < T^k$. Now, taking this as given, I proceed to solve for the optimal contract under the restriction to $T^{k+1}(t_{k+1}) \geq T^k$ for all k, meaning contracts whose horizon may not shorten upon revelation of a bad signal. I will make use of the incentive constraint at the deadline provided in lemma 5. The next lemma provides the main step in solving for the optimal contract.

Lemma 7. *The optimal contract should have $w_{T^k,k}(G)$ and $w_{T^k,k}(B)$ set to 0, and the continuation value $V_{T^k,k+1}$ chosen such that equation (12) binds (for $k \leq n^*$):*

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) = \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1} \quad (19)$$

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

(Proof in the Appendix.)

This lemma says that in the optimal contract the payments upon revelation of signals at the deadline are set to zero and the continuation value upon revelation of a bad signal is strictly positive. The incentive cost necessary to make the agent work is back loaded to the extended time horizon which is provided after the revelation of a bad signal. Now, by the same reasoning as in Lemma 7, at any t, $w_{t,k}(G)$ and $w_{t,k}(B)$ should be set to the values which make the revelation constraints bind, and the rest of the incentives should be provided through extra experimentation time upon the revelation of the bad signal.

Lemma 8. *The payments upon revelation of the first n^* bad signals are zero: $w_{t,k}(B) = 0$ for $k < n^*$.*

Proof. By replacing $w_{T^k,k}(B) = 0$, the revelation constraint for signal B given by (7) becomes:

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

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

$$w_{t,k}(B) \geq \frac{c}{r}(1 - e^{-r(T^k-t)}) - V_{t,k+1} + (1 - e^{-r(T^k-t)})\frac{c}{\lambda(1 - \theta\rho_k)} \tag{20}$$

I know that $V_{t,k+1} \geq \frac{c}{r}(1-e^{-r(T^k-t)})$ due to $V_{t,k} \geq \frac{c}{r}(1-e^{-r(T^k-t)})$ and $V_{t,k+1} \geq V_{t,k}$. Hence, the right hand side of equation (37) is negative, which means this constraint is slack. I then conclude that it is optimal to set $w_{t,k}(B) = 0$ for any t and $k \leq n^*$. □

Finally, I will find $w_{t,k}(G)$.

Lemma 9. *The payment upon revelation of a success is set to its minimum value which leaves the agent indifferent to revealing it or hiding and remaining in the project in order to shirk until the deadline:*

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

Proof. I will make the revelation constraint for G in equation (6) bind in order to get the minimum $w_{t,k}(G)$:

$$w_{t,k}(G) = \frac{c}{r}(1 - e^{-r(T^k-t)}) + e^{-r(T^k-t)}w_{T^k,k}(G) \tag{21}$$

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

Now, I can replace the payments $w_{t,k}(B)$ and $w_{t,k}(G)$ into the no shirking condition:

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

This constraint binds for the same reason as in lemma (7). Given that the incentive constraint for working binds at any (t, k), I conclude that the condition given in equation (5) binds and the agent's continuation value is given by:

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

Then, by replacing $\theta\rho_k V_{t,k}$ on the left hand side of equation (22):

$$V_{t,k+1} - V_{t,k} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{23}$$

At the deadline, the condition was given by lemma 7:

$$V_{T^k,k+1} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{24}$$

Then, by replacing $V_{T^k,k+1} = \frac{c}{r}(1 - e^{-r(T^{k+1}(T^k)})$ 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 $t < T^k$ in equation (23):

$$e^{-r(T^k(t)-t)} - e^{-r(T^{k+1}(t)-t)} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{25}$$

From the above equation, it is easy to conclude that $T^{k+1}(t) - T^k$ is decreasing in k: as the belief gets more pessimistic, the change in the deadline is lower. Also, in a given state k, $T^{k+1}(t)$ decreases in t for fixed k. The reason is that the cost is transferred into the extended horizon starting at the initial deadline T^k, and the farther T^k is from t, 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 variables for summarizing the history dependence. The reason is that 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 the number of bad signals received which determines his current belief (and not the times at which they were received) in addition to his continuation value. The times of revelation of signals are already reflected in the agent's bonus payments and continuation value (equivalently the current deadline). Hence, the relevant history at time t can be summarized by T^k and ρ_k which can be shortened to k. The updating of the deadline only depends on the time of arrival of the last bad signal and the current deadline.

Discussion: In any state $k < n^*$, when the agent reveals a bad signal, the continuation value is increased just enough to make the incentive constraint for working bind. This is equivalent to saying that the total agency rent is kept constant. The term $\frac{c}{\lambda(1-\theta\rho_k)}$, which is a part of the agent's incentive cost, is added into the continuation value $V_{t,k+1}$ that the agent obtains upon revealing B.

The revelation of a bad signal causes the belief to go down, but at the same time signals only arrive while the agent is working. The agent's expected benefit from working consists of the payment upon good news, payment upon bad news and the continuation value in the next state after a bad news is revealed, and how this is decomposed into these three is irrelevant for his incentives as long as the revelation constraints are satisfied. As what matters for the agent's incentives is the continuation value $V_{t,k}$ and that the revelation constraints are satisfied, the principal chooses the components of the continuation value in an optimal way. This is achieved by minimizing the bonus payments and increasing the continuation value in the more pessimistic state which is reached after the release of a bad signal, while keeping the total rent of the agent constant. The principal benefits from experimenting longer as long as $\rho_k \geq \frac{c}{\lambda\theta}$, in other words as long as the intrinsic value of experimentation is positive and the total expected payment to the agent does not increase.

4.5 Sufficiency of local revelation constraints:

Now I will verify that the two local constraints for revelation of signals provided by Lemma 3 are indeed sufficient for global implementability. First, I will show that the agent cannot benefit from hiding a bad signal. After, I will show that upon hiding a good signal, he does not have an incentive to work again. This implies that the local constraints which ensure that the agent does not want to delay revealing the signals are sufficient. The same results also hold in case the signals get lost if not revealed right away, which is discussed in the section 5. First, I will verify that there is no profitable deviation for the agent after receiving a signal B, such as hiding one or more signals. The agent should be compensated at least for the change in his continuation value upon revealing a signal B when the state moves from k to $k+1$:

$$w_{t,k}(B) \geq max[0, V_{t,k}^{B} - V_{t,k+1}] \tag{26}$$

where $V_{t,k}^{B}$ is the maximum continuation value after hiding a signal B. Now, let us show what is the best deviation upon hiding a bad signal. Initially, the incentive constraint to work binds:

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

Now, given that the agent holds the belief ρ_{k+1}, he is strictly better off working. Let us consider the future payoff of the agent from this moment on. The payment upon revelation of a good signal is always higher for higher k: $w_{t,k}(G) < w_{t,k+1}(G)$. The increase in the continuation value of the agent upon revelation of the next 2 bad signals will be respectively $\frac{c}{\lambda(1-\theta\rho_k)}$ and $\frac{c}{\lambda(1-\theta\rho_{k+1})}$ which is independent of when he reveals them. Then, this strategy is dominated by one in which he reveals the $k+1$'th bad signal upon receiving and gets the increase in his continuation value $\frac{c}{\lambda(1-\theta\rho_k)}$ earlier, and follows the strategy induced by the contract afterward. Then, equation (26) holds and the agent cannot gain by hiding or delaying the revelation of a signal B. Lastly, I will verify the deviation to hide a success and continue experimenting in order to get a signal B and reach state $k+1$. This might be profitable in case $w_{t,k+1}(G)$ is high enough compared to $w_{t,k}(G)$, in other words if the revelation of a good signal is much more profitable when the belief is more pessimistic. The best deviation of this kind would be to experiment in order to get a signal B and in case it arrives, reveal B first and then G at some $\hat{t}$ in state $k+1$ in order to receive $w_{\hat{t},k+1}(G)$ where $\hat{t} \leq T^{k+1}$. This constraint can then be written as follows:

$$\begin{aligned} w_{t,k}(G) \geq [\lambda dt(1-\theta)(1-rdt)(\frac{c}{r}(1-e^{-r(\hat{t}-t)}) + w_{\hat{t},k+1}(G))] \\ + (1-\lambda(1-\theta))(1-rdt)w_{t+dt,k}(G) \end{aligned} \tag{27}$$

for any $\hat{t}$. I consider the *best* deviation of this kind, as in case of getting a bad signal the agent can reveal the signals at any moment in the next state until the deadline, and work or shirk in the meantime. 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 G in a given state. Then, it is sufficient to look at the limit as $\hat{t}$ goes to t, as this deviation gives the upper bound on the profit from possible deviations. Then, if the condition in (27) is satisfied at t, it should also be satisfied at any $\hat{t} > t$. By making dt go to zero in equation (27):

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq \lambda(1-\theta)(w_{t,k+1}(G) - w_{t,k}(G)) \tag{28}$$

where the left hand side is greater than c due to the constraint in (8). It is sufficient to check for one shot deviations of this kind: if after finding G 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 signal B twice either. The reason is that the change in $w_{t,k}(G)$ due to an increase in k multiplied by the probability of getting a bad signal is constant: $\lambda(1-\theta\rho_k)[w_{t,k+1}(G) - w_{t,k}(G)] = c$. Hence, it is enough to verify that the agent does not find it profitable to deviate once and get a signal B:

$$w_{t,k+1}(G) - w_{t,k}(G) \leq \frac{c}{\lambda(1-\theta)} \tag{29}$$

where $w_{t,k}(G) = V_{t,k}$. Then, as $V_{t,k+1} - V_{t,k} = \frac{c}{\lambda(1-\theta\rho_k)}$ and $\lambda(1-\theta\rho_k) < \lambda(1-\theta)$, the constraint (29) is indeed slack. The agent is better off shirking after hiding a good signal than working and this deviation is indeed not relevant in the optimal contract. The gain from an additional B in terms of the increase in the payment upon revealing G is not so high that even an agent who has already acquired a success would be willing to hide it and experiment to get a bad signal. Finally, I conclude that the local constraints are sufficient for global incentive compatibility.

4.6 The belief for ending experimentation

Next proposition provides the threshold belief at which experimentation ends.

Proposition 3. *Experimentation ends as soon as the belief falls to ρ_{n^*+1}, which is equal to the stopping level of belief in the benchmark setting without agency:*

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

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

The intuition for why this belief coincides with the first best benchmark case without agency is as follows: the total cost of experimentation to the principal per unit time under agency is $2c$. By initially committing to a deadline, the principal promises the agent the benefit c over the remaining horizon which he can obtain by shirking. This means, any time the agent reveals a signal that will terminate the relationship, he is compensated for the value of remaining in the project until the deadline, which is $\frac{c}{r}(1-e^{-r(T^k-t)})$. Hence, given that the principal already committed to pay the agency cost c, it is optimal to continue experimentation (keeping constant the total agency cost) as long as $\lambda\theta\rho_k \geq c$ which corresponds to the benchmark case without agency.

4.7 The optimal T^0

Now that the updating scheme of the deadlines is provided, once T^0 is found, the rest of the deadlines can be recovered from the updating rule given by equation (23). Next proposition provides the initial deadline T^0.

Proposition 4. T^0 *is the initial deadline which is the date such that if reached without any signal revelation the project gets terminated:*

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0+(1-\theta\rho_0)F_{T^0,1})-(1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{30}$$

The value $F_{T^0,1}$ is positive and less than 1 (as the value of a success is equal to 1), in addition it is independent of T^0 as it only depends on $T^1(T^0)-T^0$ which can be found from $V_{T^0,1} = \frac{c}{\lambda(1-\theta\rho_0)}$. Then, $F_{T^0,1}$ can be found using the updating scheme of the deadlines. It is easy to see that T^0 is decreasing in c. The derivative wrt λ is $\frac{(1-\theta\rho_0)}{\theta\rho_0\lambda^2}$, which is positive. However, the derivative with respect to θ and ρ_0 is equal to $c-\lambda F_{T^0,1}$, which could be either positive or negative depending on the value of $F_{T^0,1}$. The reason for this is that these parameters have 2 opposing effects. First, a higher probability of success implies that extra time of experimentation is more profitable. On the other hand, given that the rate of arrival of success is high, there is less incentive to allocate the agent more time because given that he experiments he is likely to receive a signal early. Hence, the sign of the derivative with respect to θ and ρ_0 depends on which of these two effects dominate.

5 Extensions

In this section, I consider some modifications to the original model. First I look at the case with public signals, second the case without moral hazard and third, I consider what happens when signals cannot be kept and are lost when not revealed right away. Finally, I consider the game with no commitment to a contract.

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 from the previous model exists: the moral hazard due to the agent's private decision to experiment or shirk. When the signals are publicly observed, the revelation constraints are no longer relevant. This implies that 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 upon realizations of either type of signal are zero as long as it is not realized 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 for exerting effort are provided completely in form of increased continuation values upon revelation 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 when only one more signal is realized, the expected payment upon realization of signals must be positive.*

Proof. The incentive constraint which makes sure that the agent works is identical to the case of privately observed 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 now there are no revelation constraints, which implies that the above condition will bind. In addition, it is still optimal for the principal to extend the horizon of experimentation upon revelation of bad signals, and set $w_{t,k}(G) = w_{t,k}(B) = 0$ for $k \leq n^*$, leading 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 $V_{t,n^*+1} = 0$:

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

The optimal contract will consist of any $w_{t,n^*}(G)$ and $w_{t,n^*}(B)$ that satisfy this equation. □

In the presence of public signals, the only positive payments are made when experimentation ends in state n^* 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 only through the possibility of getting a bad signal 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 experimentation 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 tool left for providing incentives to the agent is through the bonus payments upon realization of signals. The division of this payment between good and bad signal 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 due to the fact that the principal does not have to pay the information which was due to the private observation of signals and can set the bonus 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.

When the signals are public, the principal makes payments less often upon realization of a success: in states $k < n^*$ he never makes a positive payment. Instead, the horizon of experimentation increases faster upon 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 payed, 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 benefit 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 rent 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 cause any distortions in the principal's problem compared to the first best.

5.3 Signals get lost after hiding

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 simplified. If the agent hides a good signal, he will not find it optimal to work again also in this setting. To see why this is the case: the possible actions after hiding a good signal is 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 results in the same minimum payment, $w_{t,k}(G) = V_{t,k}$, 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 $V_{t,k+1} > V_{t,k}$. So, the original contract still remains optimal under this assumption.

5.4 The Case without Commitment

Let us discuss the game with no commitment. For this part, I will consider discrete time. In each period the principal can make an offer to the agent, consisting of promised payments which induce him to work and reveal the signals upon receiving them. As the principal cannot commit to stopping, she will continue making offers in each period as long as the belief remains above a threshold. This means the principal will stop making offers only when enough bad signals is revealed. Until this last signal, bad signals can be acquired from the agent at zero cost just as before. However, it will be too costly to make the agent reveal the good signal, as this implies that the game comes to an end as well as the terminal bad signal. The incentive constraint to work at period t is:

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

Then, given that $w_{t,k}(B) = 0$, the above equation simplifies to:

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

where the continuation values satisfy $V_{t+1,k} = V_{t+1,k+1} = \frac{c}{1-\delta}$, due to the fact that the agent can always shirk and pretend that no signal arrived, and the principal would continue making offers. As the principal has no commitment power, the agent anticipates that she will continue making offers as long as the belief is high enough. It is optimal to make the above equation bind, hence $w_{t,k}(G) = \frac{c}{\lambda\theta\rho_k} + \delta\frac{c}{1-\delta}$. Finally, let us check that the revelation constraint for G is also satisfied. If the agent hides the good signal and shirks forever:

$$w_{t,k}(G) \geq \delta\frac{c}{1-\delta}$$

When the belief is ρ_n^*, the incentive constraint for working is:

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

These payments coincide with the payments in a contract with an infinite horizon, as in that case the contract will never end unless the n^*+1'th bad signal is revealed. Hence, without the use of deadlines, the ability to commit to a contract alone does not increase the profits of the seller compared to the case without commitment.

6 Conclusion

This paper studies how incentives should be provided optimally in a principal agent model of experimentation. An agent is hired to get signals about the underlying quality of an innovative project. The uncertainty on the arrival time of signals even while the agent is working implies that the principal is unable to monitor whether the agent is actually working or diverting benefits to his own use. In addition, the private observation of signals by the agent implies that the agent should be provided enough incentives to reveal them. The novelty of the model is the presence of good and bad signals which arrive randomly over time. This means news only arrives with some probability and in the absence of signals the principal cannot know for sure whether the agent is shirking or not. As the signals only arrive while the agent actually experiments, bonuses are payed only upon the revelation of signals.

I find that the optimal contract initially allocates some time for the agent to experiment, and provides extra time upon revelation of bad signals. The main point is that while doing so, the principal does not increase the total expected payment to the agent but he just back loads the payments to extended horizon. When the agent experiments, he has the possibility to receive a good or a bad signal and a continuation value in the more pessimistic state and he cares about the overall benefit from working rather than the composition of it. Increasing the continuation value of the agent in the future state in which the belief is more pessimistic decreases the incentive cost of making him work today. Even though a bad signal leads to a more pessimistic belief, the principal prefers to experiment longer as long as the incentive cost does not increase and the intrinsic value of experimentation is still positive. The principal back loads the agency cost in such a way that the experimentation time is extended as much as possible while the expected payment to the agent is kept constant. The results of this paper suggests that incentives can be optimally provided through endogenous deadlines in innovative activities when news about project quality arrive over time in form of good or bad signals and only while effort is put in.

There are many questions related to the paper that are of interest for future research. One such question is what may happen if the underlying quality of the project is persistent but not constant over time, if there is a learning effect which means that even an initially bad project can succeed in the future depending on the effort put in. One could also consider a modification to the model in which the agent has career concerns, hence he is less eager to reveal bad signals because it reveals information about his type which he cares about in the long run. This would be a model in which the agent is hired for producing projects over time which are of good or bad quality depending on the agent's type which is unknown at the start and revealed over time. Finally, an important and promising topic is to explore communication incentives in teams, mainly how rewards should be structured in order to induce the agents in a team to collaborate by sharing their information. These questions remain for future research.

Appendix

Proof of Lemma 6

Consider there is a history h^t such that at t, the principal stops investing in the project until the future date $t+\Delta$. Call the value of the principal at $t+\Delta$ as $F_{t+\bar{\Delta}}$. Then, in case experimentation stops at t, the current expected value of experimentation from time $t+\Delta$ is discounted by $e^{-r(\Delta)}$. If the principal starts experimenting again at time $t+\Delta$, this means $F_{t+\Delta}>0$. Then, it is inefficient to stop financing the project at a date t and start again at $t+\Delta$ as it leads to discounting without changing the future profit.

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 value of the principal is:

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

when $dt => 0$:

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

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

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

Then, at any moment t in state 0 when a bad signal arrives, the continuation value of the principal at that moment becomes $F_{t,1}$.

Proof of lemma 7

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. This will be shown 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, $V_{t,k+1}=V_{t,k}+\frac{c}{\lambda(1-\theta\rho_k)}+\Delta$. This means that the incentive constraint at t is slack:

$$\theta\rho_k w_{t,k}(G)+(1-\theta\rho_k)w_{t,k}(B)\geq 0>\frac{c}{\lambda}+V_{t,k}-(1-\theta\rho_k)V_{t,k+1} \tag{31}$$

given that $w_{t,k}(B) = 0$ and $w_{t,k}(G) = V_{t,k}$ still hold (δ is small enough), then we have:

$$(1 - \theta\rho_k)V_{t,k+1} \geq \frac{c}{\lambda} + V_{t,k} - \theta\rho_k w_{t,k}(G) \tag{32}$$

If we increase $V_{t,k+1}$ by Δ, and increase $V_{t,k}$ by $(1-\theta\rho_k)\Delta$, calling it now $\hat{V}_{t,k}$, while T^k remains the same. Then, $w_{t,k}(G)$ is not modified, $w_{t,k}(G) = \frac{c}{r}(1 - e^{-r(T^k-t)})$, and $\hat{V}_{t,k} = \frac{c}{r}(1-e^{-r(T^k-t)})+(1-\theta\rho_k)\Delta$. (As Δ is low enough the deviation to work after G is not relevant). Hence, the expected agency cost to the agent in state k has not been modified. However, if state $k+1$ is reached, the cost increases by Δ, which happens with probability $(1 - \theta\rho_k)$. Call the updated deadline $T^{k+1}(t)$ initially when $V_{t,k+1} = V_{t,k} + \frac{c}{\lambda(1-\theta)}$. Then, when $V_{t,k}$ is modified to $\hat{V}_{t,k}$, the updated deadline is modified to $\hat{T}^{k+1}(t) > T^{k+1}(t) > T^k$. This means the cost of experimentation increases conditional on reaching state $k+1$. Then, if it is optimal to incur this cost in state $k+1$, it would be better to incur this cost in state k as well, by increasing T^k to $\hat{T}^k$ such that $\frac{c}{r}(1 - e^{r(\hat{T}^k-T^k)}) = \Delta$, hence $\hat{V}_{t,k} = \hat{V}_{t,k} + \Delta$. This means the cost Δ should also be incurred in state k when it is more profitable to experiment than in state $k+1$.

Step 2:

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

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

It means this constraint is slack. Then, increase $V_{t,k+1}$ by Δ (where Δ is small enough). The right hand side of the incentive constraint decreases by $\Delta(1 - \theta\rho_k)$, implying the expected payments at t decrease 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 cost to the agent during the extended horizon, and hence the net effect of this extended time period is positive. Then, I will conclude that this modification increases profits at any t. Before the extension, the profit during $(t, T^{k+1}(t))$ when the state is $k+1$:

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

which is equal to:

$$(1-e^{-(T^{k+1}-t)(\lambda+r)})\frac{[\lambda\theta\rho_{k+1}]}{\lambda+r}+\int_t^{T^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})F_{s,k+2}ds-\frac{c}{r}(1-e^{-r(T^{k+1}-t)}) \tag{33}$$

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 profit during the extended horizon is:

$$(1-e^{-(\hat{T}^{k+1}-t)(\lambda+r)})\frac{\lambda\theta\rho_{k+1}}{\lambda+r}+\int_t^{\hat{T}^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})\hat{F}_{s,k+2}ds-(V_{t,k+1}+\Delta) \tag{34}$$

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, it is necessary to show that the 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 Proposition 3

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

$$F_{t,n^*} = \int_0^{T^{n^*}} e^{-t(\lambda+r)}(\lambda(\theta\rho_n^* - \frac{c}{\lambda} - w_{t,n^*}) - c)dt \tag{35}$$

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:

$$F_{t,n^*} = \int_0^{T^{n^*}} e^{-t(\lambda+r)}[\lambda(\theta\rho_n^*(1-w_{t,n^*}(G)) + (1-\theta\rho_{n^*})F_{t,n^*+1}) - c]dt \tag{36}$$

where $w_{t,n^*}(G) = \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) = V_{t,n^*}$. Now, the condition for 35 > 36 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 $V_{t,n^*+1} = \frac{c}{\lambda(1-\theta\rho_{n^*})} + V_{t,n^*}$:

$$-V_{t,n^*+1} \geq F_{t,n^*+1} \tag{37}$$

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 (37) 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 condition:

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

Proof of proposition 4

Let us write down the principal's problem:

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

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

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

and it continues for all k until $k = n^*$. Then the derivative of $F_{0,0}$ wrt 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 $F_{T^0,1}$ wrt 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:

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0 + (1-\theta\rho_0)F_{T^0,1}) - (1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{38}$$

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 cannot be optimal for the principal, justifying the initial restriction to deterministic deadlines. In addition, as there is only one reflection point, the value function is decreasing after 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 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:

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

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

Now I will prove 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 (3) simplifies to:

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) \geq \frac{c}{\lambda} \tag{39}$$

The revelation constraints are irrelevant at the deadline T^k. It is then optimal that the equation (39) 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:

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

The revelation constraint for G is:

$$w_{t,k}(G) \geq \frac{c}{r}(1 - e^{-r(T^k-t)}) + e^{-r(T^k-t)}w_{T^k,k}(G) \tag{40}$$

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

$$\begin{aligned}\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) \geq V_{t,k} + e^{-r(T^k-t)}(\theta\rho_k w_{T^k,k}(G) \\ + (1-\theta\rho_k)w_{T^k,k}(B)) - (1-\theta\rho_k)V_{t,k+1}\end{aligned} \tag{41}$$

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:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) \geq \frac{c}{\lambda} + V_{t,k} - (1-\theta\rho_k)V_{t,k+1} \tag{42}$$

Comparing the equations (41) and (42), 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 (42) binds and the revelation constraints are satisfied. Now it will be shown 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 (4.4), 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:

$$F_{0,n^*-1} = \int_0^{T^{n^*-1}} e^{-t(\lambda+r)}\Big[\lambda(\theta\rho_{n^*-1} - \frac{c}{\lambda} - \theta\rho_k\frac{c}{r}(1 - e^{-r(T^{n^*}-t)}) - \frac{c}{r}(e^{-r(T^{n^*}-t)} - e^{-r(T^{n^*-1}-t)}) + F_{t,n^*}) - c\Big]dt \quad (43)$$

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} wrt 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 the next state is a state in which 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:

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

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 wrt 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 $\lambda\theta\rho_{k+1} - c \geq 0$. This final one is the condition for experimentation to be profitable initially. I 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. 2014. "The Role of Information in Innovation and Competition." Working paper.

[2] Marco Battaglini and Rohit Lamba. 2014. "Optimal Dynamic Contracting: the First-Order Approach and Beyond." Working paper.

[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] Bruno Biais, Thomas Mariotti, Jean Charles Rochet and Stephen Villeneuve (2010): "Large Risks, Limited Liability, and Dynamic Moral Hazard," *Econometrica*, 78, 73-118.

[6] Patrick Bolton and Christopher Harris. 1999: "Strategic Experimentation" *Econometrica*, 67, 349â374.

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

[8] Alessandro Bonatti and Johannes Horner (2013): "Career Concerns with Coarse Information." Working paper.

[9] Hector Chade and Natalia Kovrijnykh (2012): "Information Acquisition, Moral Hazard, and Rewarding for Bad News". Working paper.

[10] Peter DeMarzo and Michael J. Fishman (2007): "Optimal Long-Term Financial Contracting." *The Review of Financial Studies*, 20, 2079-2128.

[11] Umberto Garfagnini (2011): "Delegated Experimentation." Working paper.

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

[13] Renato Gomes, Daniel Gottlieb and Lucas Maestri (2013): "Experimentation and Project Selection: Screening and Learning." Working paper.

[14] Brett Green and Curtis R.Taylor (2014): "Breakthroughs, Deadlines and Severance: Contracting for Multistage Projects". Working paper.

[15] Yingni Guo (2014): "Dynamic Delegation of Experimentation". Working paper.

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

[17] Marina Halac, Navin Kartik, and Qingmin Liu (2013): "Optimal Contracts for Experimentation" Working Paper.

[18] Bengt Holmstrom (1979): "Moral Hazard and Observability", *Bell Journal of Economics*,10, 74-91.

[19] Bengt Holmstrom and Paul Milgrom (1987): "Aggregation and Linearity in The Provision of Intertemporal Incentives", *Econometrica*, 55, 303-328.

[20] Bengt Holmstrom and Paul Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *The Journal of Law, Economics Organization*, 7, Special Issue: 24-52.

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

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

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

[24] Nicolas Klein (2014): "The importance of being honest." Working paper.

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

[26] Suehyun Kwon (2014): "Recursive Formulation with Learning." Working paper.

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

[28] Tracy R. Lewis. (2011): "A Theory of Delegated Search for the Best Alternative." *RAND Journal of Economics*, 43, 391â416.

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

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

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

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

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

[34] Dinah Rosenberg, Eilon Solan and Nicolas Vieille, 2007. "Social Learning in One-Arm Bandit Problems". *Econometrica*, 75(6), 1591â1611.

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

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

[37] Felipe Varas (2014): "Contracting Timely Delivery with Hard to Verify Quality." Working Paper.