Contracting for Experimentation and the Value of Bad News*

Sinem Hidir†

April 15, 2016

Abstract

I consider a dynamic problem in which a principal hires an agent in order to learn the underlying quality of a project. While the agent exerts costly effort, news arrives in form of good or bad signals about the underlying state. Lack of signals may be due to the agent's shirking or 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 payments and a termination rule in which the current deadline is updated each time a bad signal is revealed. The principal rewards the agent through increased continuation values, equivalent to extended experimentation time, upon revelation of bad signals. If the contract induces stopping before a deadline is reached, it stops at the belief which is the same as the stopping level of belief 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.

†University of Warwick. email: s.hidir@warwick.ac.uk. Please find the latest version at https://sites.google.com/site/sinemhidirecon

1 Introduction

This paper considers the problem of a principal who contracts with an agent in order to learn about the unknown quality of a project. While the agent puts in effort, good or bad signals arrive over time and are privately observed. The arrival rate of the signals depend on the agent's effort who could instead divert funds for private use. As signals arrive with a poisson rate, there is uncertainty about how much time and costly effort is required before a decision should be taken. The principal would like to obtain the necessary information as soon as possible in the least costly way in order to stop spending resources.

We study this contracting problem with moral hazard and private arrival of information. The fact that signals about the project quality arrive only while the agent puts in effort is a crucial feature of the paper. For instance, during the development a new product, existence of a similar product may become known making it less profitable to continue which is interpreted as a bad signal. For a researcher trying to prove a hypothesis, a bad signal can be thought of as trying a solution concept which does not work.

Most real world settings involving new and untested technologies rely on agency relationships. This is because expertise and funding do not usually belong to the same person. The incentives of the party who provides the funding and benefits from a success (the principal) and that of the one who is hired to carry out the research (agent) are not usually aligned. The principal would like this procedure to be as fast as possible given that she is providing funds while the agent values remaining in the relationship as long as possible. In addition, it is not possible to assess from the beginning how much time it should take the agent to get enough information. This is captured by the poisson arrival rate of the signals while the agent is putting in effort. The principal is unable to guess how long it will take the agent to acquire enough information even while he is working. The agent is not willing to reveal news unless he is given sufficient incentives.

An example of this situation is a new pharmaceutical drug which needs to pass clinical trials before being approved. The drug has to prove to be both safe and effective, after which the company submits a New Drug Application or Biologics License Application to the FDA. After weighing the benefits and risks of the medicine, a decision is made on whether to grant approval. This implies there are two types of information, positive and negative, which are collected as a result of experiments and used in order to make a decision.

In our setting, the agent has the choice between experimenting or shirking and keeping for himself the resources invested by the principal. Working is equivalent to pulling a risky arm and experimenting by incurring the cost and possibly acquiring a signal whereas shirking is equivalent to pulling a safe arm which lets the agent keep the resources but does not lead to learning. Second, the agent also chooses whether and when to reveal the signals that arrive. The signals arrive with a Poisson process which is common in the literature on learning about a project quality through experiments.[1] In the literature, a good arm can yield a success or a failure whereas a bad arm always fails, hence one success is conclusive and when enough time passes with no success, the belief becomes pessimistic enough that the project is abandoned. In contrast, in our setting there can be a period of time during which no signal is realized and there is no learning.

The possibility of getting bad signals allows an agent who works on a bad project to distinguish himself from an agent who does not get any signals due to shirking. In addition, as no signals arrive while the agent shirks, there is no possibility of private learning for the agent which simplifies our analysis. The agent does not get an informational rent from postponing effort, but he has incentives to shirk and divert resources for private benefit.[2] It is possible to distinguish an agent who is shirking from one who obtains negative results. In this paper I capture this by assuming that the signals, whether good or bad, only arrive while the agent experiments and are credibly revealed. This assumption rules out the agent's private learning by shirking, as when the agent shirks his belief will not differ from the principal's. In addition, this information structure gives a stationary feature to the beliefs while no signals are realized.

Two types of informational frictions are present in this contracting problem:

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

[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.

moral hazard due to the possibility of the agent to shirk and keep the resources provided by the principal for his own use and private observation of signals, which lead to two types of incentive constraints. The first one is the working constraint which makes sure that the agent prefers to work rather than shirk at any moment. The second type of constraints are the disclosure constraints which make sure that the agent does not want to hide or delay the revelation of signals. The principal's problem is one of finding how the agency rent should be optimally allocated between bonus payments and continuation values while satisfying the agent's constraints.

I show that the optimal contract uses bonus payments and an extendable deadline in order to incentivize the agent to work and reveal the signals that arrive. The initial time allocated to the agent for experimenting gets extended after each bad signal revelation until the final one which terminates the contract. It is beneficial to postpone the payment to the agent by increasing his continuation value, because by doing so the risk of early termination is reduced. Increasing the continuation value is equivalent to an extension of the initial deadline of the contract.

There is a direct relation between the deadline and the continuation value of the agent. As the agent values remaining in the project, his continuation value is increasing in the contract horizon. In addition, there is a direct relation between the termination date and bonus payments promised to the agent. In order to induce the agent to reveal a good signal or a terminal bad signal, the principal has to compensate him for his possible deviation of hiding it and remaining in the contract. Hence, the bonus promised should be sufficient to make him give up the opportunity of remaining in the contract. A longer experimentation horizon provides more time to experiment on a potentially successful project but at the same time the principal has to promise higher payments to the agent in order to induce him to reveal the signals. Due to this, the optimal contract leads to stopping too early.

The optimal contract features history dependent bonus payments, an initial deadline and a rule for extending this deadline after the revelation of bad signals. The principal initially lets the agent experiment for some time, and extra time is allocated after the revelation of bad signals. This is true until the belief falls down to a level at which experimentation would have stopped even in the absence of an agency problem, which happens upon the revelation of the *terminal* bad signal. The only payments to the agent are made when termination happens due to a signal revelation. The payment to the agent for the revelation a good signal decreases over time, and has an upward jump after the disclosure of each bad signal as the belief becomes more pessimistic. The agent is willing to reveal non terminal bad signals at no cost, but it optimal to extend more experimentation time as this relaxes the agent's incentive constraint while decreasing the inefficiency due to stopping inefficiently early.

The deadline trades off longer experimentation time versus higher agency rents. The principal prefers to provide incentives to the agent through continuation values (experimentation time) after the revelation of bad signals rather than through bonus payments as long as it is still profitable to experiment. Even though experimentation stops inefficiently early in case of reaching the deadline, when stopping happens before the deadline, it is either due to the release of a good signal or due to reaching the same stopping level of belief as in the first best benchmark.

The intuition for the above features of the contract is as follows: the use of a pre-determined deadline allows the principal to control the moral hazard rent of the agent, in other words how much the agent can remain in the contract without working. This implies that experimentation may end at an optimistic belief due to reaching the deadline. As time passes and no signal arrives, it becomes more likely that the agent has been shirking and termination serves as a punishment. On the other hand, the revelation of a bad signal shows that the agent has been experimenting and the extension in the time horizon leads to a decrease in the distortion caused by the presence of the deadline itself. Extending the time horizon is not costly as it relaxes the incentive constraint of the agent. The principal back loads the agency rent into extra experimentation time and keeps the agent experimenting longer at a belief at which it has a positive value.

An interesting feature is that even though one could consider infinitely many ways the agent could deviate by hiding a signal, in the optimal contract there is a single deviation that should be taken into account. I initially characterize the optimal contract by using local incentive constraints, and verify in the end that these are indeed sufficient for global incentive compatibility. Hence, the optimal contract can be fully characterized by taking into account only local constraints. The reason is that after deviating to hide a signal the agent does not find it optimal to work again. It is shown that the incentive constraint for working binds and the revelation constraints are also satisfied.

This paper relates to the literature on experimentation in principal-agency settings. Bergemann and Hege (1998, 2005) study the incentives for experimentation in a principal agent model without 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). In these papers, by working and increasing the possibility of a success he risks giving up the benefits of remaining in the project. The non observability of effort leads to *procrastination rents*. The belief goes down as long as there is no success and experimentation ends too soon compared to the first best benchmark.

Another related paper 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 a project quality. As the agent gets a signal in each period during which he incurs the cost, there is a fixed deadline at which learning will be complete and the agent is allocated a rent due to the possibility of guessing a good state without incurring any cost.

Halac, Kartik and Liu (2015) study optimal contracts for experimentation in presence of both moral hazard and adverse selection on the agent's type, when both types of the agent can succeed on a good project but with different probabilities. 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) asks 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 (2015) studies dynamic delegation of experimentation without transfers to an agent who has a private prior about the project quality. One of the results is on *sliding deadlines* in the case of inconclusive success and hence that the deadline for experimentation is extended 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. 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 is better off paying the agent through contracting horizon rather than bonuses whenever she can.

Green and Taylor (2015) study a multistage project and show that the completion of one stage leads to the allocation of extra time for the completion of the next stage. In contrast to my setting, there is no learning about the project quality.

In DeMarzo and Fishman (2007), the principal optimally chooses among two types of tools in order to provide incentives to the agent. Similarly, this is either through instant payments which corresponds to bonuses in my paper or by affecting the continuation value of the agent which corresponds to allocating longer experimentation time. This cash diversion feature is also present in DeMarzo and Sannikov (2015) in which the question is how to design a contract which provides incentives while controlling the agent's information rent.

Lastly, Akcigit and Liu (2014) consider two firms competing for an innovation on a common research line, and one good or bad signal is conclusive about whether the research line is a good or a bad one. They focus on the fact that when one firm reaches a dead end on their research, the other firm which is uninformed keeps experimenting inefficiently on that arm. The question is how much a social planner could do to improve efficiency in this setting.

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

A principal (she) hires an agent (he) in order to learn about an unknown state of the world, which determines the profitability of the project initially unknown to both. The common prior that the state is good is ρ_0 and that it is bad is $(1 - \rho_0)$ where the index 0 implies that no bad signals have yet been realized. A good project has a net value normalized to 1 for the principal[3]. The project type can be learned through costly experiments requiring resources modeled as a flow

[3]a bad project has sufficiently negative value that learning is profitable

investment c by the principal. The agent is hired to experiment, but he could also shirk and keep the funds for his private use.[4] Here "shirking" indicates diverting the funds invested by the principal for personal use. 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.

Signals:

Time $t \in [0, \infty]$ is continuous. If the agent puts in effort $a \in \{0, 1\}$ over a time interval $[t, t + dt]$ incurring the cost cdt, a signal (or an *outcome*) arrives with probability $(a\lambda)dt$ and is denoted by $z_t \in \{G, B\}$. It is privately observed by the agent. The arrival rate is independent of the type of the project. While only a bad signal can arrive from a bad project, the signal coming from a good project can be a good one with probability θ or a bad one with probability $(1 - \theta)$. Signals are verifiable, they can be hidden but not constructed or modified. [5]

The revelation of one good signal concludes that the project is a good one and a net benefit of 1 is realized by the principal. Upon the revelation of a bad signal the belief about project quality goes down as follows:

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

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

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

$$\lambda\theta\rho_0 \geq c \tag{1}$$

This assumption makes sure that without the agency problem the principal would be willing to experiment at least initially.

[4]I have described the setting as one in which the principal provides resources for experimentation and the agent has the possibility to divert benefits to his own use. However, the setting could also be chosen as one in which 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 version of this setting would be one in which each period when the agent incurs the cost, he may receive a signal with probability λ whose type depends on the underlying state, or receive no signal with probability $1 - \lambda$.

The strategy of the agent is denoted by $\sigma = (a, x)$. The agent chooses effort $a : \tilde{\mathcal{H}} \to \{0, 1\}$. Effort choice 1 denotes the decision to put in effort and 0 the decision to shirk and keep the investment c. The agent also chooses a disclosure plan as a function of his private history $x : \tilde{\mathcal{H}} \to G, B$.

The agent can reveal any signal or signals among those he has received until t and has not already revealed. He could possibly keep and reveal a signal later on but cannot construct a fake one. This implies that the private history of the agent can possibly be extremely complicated. However, it is without loss to focus on contracts in which the agent reveals the signals as they arrive, hence the agent's private history h^t_A coincides with the public history h^t. The agent has no possibility to privately learn by shirking, as no signals arrive when he shirks. The beliefs of the principal and the agent can only differ in case the agent receives and hides a signal. This is different from most of the literature on experimentation where the agent's shirking leads to private learning and a more optimistic belief than the principal.

Assumption 2. $\lambda\theta\rho_0 - 2c \geq 0$.

This assumption ensures that experimentation is profitable at $t = 0$ in the presence of agency given the initial belief ρ_0. The term $\lambda\rho_0\theta$ is the benefit of an instant of experimentation, and $2c$ is the total cost of an instant of experimentation for the principal: c is the flow 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.

Contracts: At $t = 0$ the principal offers and commits to a contract. In case the agent refuses, both take their outside options of zero. The revelation principle implies that it is without loss to restrict attention to contracts in which the agent discloses the signals as they arrive. The principal wants to provide incentives to the agent to put in effort and stop once it is no longer profitable to invest in the project. Hence, the principal's problem is one of finding the incentive compatible contract which minimizes the expected payment to the agent and implements an optimal stopping rule.

A contract is denoted by $\mathcal{C} = \{W, y\} : H \to \Re_+ \times \{0, 1\}$. The first component

W_t represents the history dependent payments to the agent at time t. The second component $y_t = Pr(Invest|h^t)$ denotes the principal's decision whether to fund experimentation at time t as a function of the public history.

Given that there is a flow cost to the principal, and that no learning happens when the agent stops working, it is without loss of generality to restrict attention to contracts in which the agent is induced to work and disclose the arriving signals.

Principal's Problem: Assuming the agent adheres to the recommended action, the present expected value of the principal at time zero from a contract C is:

$$F_0(\mathcal{C}) = E^a\left[\int_{t=0}^{\infty} e^{-rt} y_t (a_t \lambda \theta \rho^t \mathbb{1}_{n_t=0} - dW_t - c) dt\right]$$

where ρ^t is the belief at time t which depends on how many bad signals have already been revealed and n_t the number of good signals revealed up to time t. The term $\mathbb{1}_{n_t=0}$ takes the value 1 as long as no good signal has been revealed, as one good signal indicates the project is a good one.

The agent gets utility from the payments he receives and the resources he keeps from shirking:

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

A contract $\mathcal{C}$ is incentive compatible if the induced action a^* maximizes the agent's expected utility.

The principal's problem is then one of finding the incentive compatible contract which maximizes equation 2 subject to satisfying the agent's disclosure and working constraints. First, in an optimal contract $W_t > 0$ arises only at t at which signals are revealed by the agent, in other words at t such that $x_t \in \{G, B\}$. This is because the signals arrive only while the agent puts in effort. Second, as long as the principal keeps investing, it is never optimal to induce the agent to shirk as this would imply a waste of resources invested for the principal. I denote by $\sigma^* = (a^*, x^*)$ the strategy under which the agent works as long as the principal invests 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 z_t.

First-best Benchmark: Now I consider the principal's problem in case she could carry out experimentation without delegating it to the agent. The first best stopping problem is a belief level where n^* is the lowest n satisfying:

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

In other words ρ_n^* is the lowest belief at which experimentation is still valuable. This condition says that experimentation lasts as long as its instantaneous benefit minus cost is positive, and stops as soon as the $n^* + 1$'th bad signal is revealed after which experimentation has negative value. As signals arrive with a Poisson arrival rate while the agent puts in effort, it is not known ex-ante when the belief will reach this stopping level. Then, without the agency problem, there is not a stopping time but a stopping level of belief. The belief at which experimentation ends before a deadline ρ_{n^*+1} will be shown to coincide with the belief at which experimentation stops before reaching a deadline.

3 The optimal contract

This section summarises the characteristics of the optimal contract. Section 4 goes through the steps of the solution. First, I will show that the principal's problem simplifies to choosing a termination rule 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 $h^t \prec h^{t'}$, $y_{t'} = 0$.*

When the principal stops investing in experimentation, she will not start again at a future date. In the absence of new information, the beliefs are not modified hence the opportunity cost of experimentation remains constant. Then, there is no gain for the principal to start investing again after having stopped once, which implies that there is a date at which experimentation will stop once and for all.

Lemma 2. *The deadline, if ever, should be updated 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. A good signal is equivalent to a success and ends experimentation. Then, the only possible histories h^t such that the deadline gets updated at t should have either $x_t = B$ or $x_t = 0$ (no signal). Let us denote a history $\hat{h}^t$ as one in which no bad signal has yet been realized until t, and suppose the initial deadline T gets updated to $\hat{T}$ at time t. Then, this contract is equivalent to the following one: the deadline at $t = 0$ is set to $\hat{T}$ initially and the updating rule

$T(h^t)$ for any other history h^t and t is kept constant. It is without loss to restrict the updating of the deadline to the times at which bad signals are disclosed by the agent. □

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

The termination rule says that the contract will end as soon as a good signal is revealed, at T^k if no other bad signal is revealed, or upon the revelation of the $n^* + 1$'th bad signal. The public history can be summarized by the times at which bad signals are revealed: $h^t = \{t_1, t_2, ..., t_k\}$ for a current belief ρ_k. These are the only elements of the history that are relevant for the contractual terms. The initial deadline is T^0. If the first bad signal is revealed at t_1, the deadline gets updated to $T^1(t_1, T^0)$ as t_1 and T^0 are the only relevant elements of the public history. When the second bad signal is revealed at t_2, the deadline becomes $T^2(t_2, T^1)$ as t_1 and T^0 were already 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 the revelation of the last bad signal.

I will now simplify the notation as T^k for the deadline, $V_{t,k}$ and $F_{t,k}$ respectively for the agent's and the principal's continuation values at time t and belief ρ_k and $w = (w_{t,k}(G), w_{t,k}(B))_{t=0}^{\infty}$ respectively for the bonus payments upon the revelation of a good and a bad signal. While simplifying the notation, I do not omit the public history h^t. The optimal contract should have $w_{t,k}(0) = 0$: there is no need to make a payment to the agent as long as no signal has been reported.[6]

Now I will rewrite the principal's problem in a recursive way. The index k in the public belief ρ_k changes to $k+1$ as soon as an additional bad signal is revealed. If the k'th bad signal is revealed at time t and the deadline becomes T^k, the present value of the contract to the principal denoted by $F_{t,k}$ can be recursively written

[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 must be paid only upon the revelation of signals.

as:

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

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 belief remains ρ_k. The term $e^{-r(s-t)}$ is the discount factor that applies when a signal arrives at time s. During an infinitesimal time period of dt, with probability λdt a signal arrives and is disclosed. With probability $\theta\rho_k$ the signal is a good one and upon its disclosure the contract ends 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 with the continuation value $F_{t,k+1}$. The detailed derivation of equation (2) is provided in the Appendix. Then, the problem of the principal at time zero, given that ρ_n^* is the lowest belief at which there can be experimentation, is:

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

subject to $w_{t,k}(G)$ and $w_{t,k}(B)$ satisfying the incentive compatibility conditions of the agent. While solving for the optimal contract, I initially restrict attention to local incentive compatibility constraints and verify later that these are actually sufficient for global implementability. There are two types of incentive constraints. The first type is the *no shirking constraint* which makes sure that the agent prefers experimenting as induced in the contract rather than shirking at any moment. The second type of constraints are the *disclosure constraints* which make sure that the agent is willing to reveal the signals he acquires without delay and checks for any possible deviations. Propositions 1 and 2 provide the characteristics of the optimal contract, and the details of the solution are provided in section 4.

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

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

- $w_{t,k}(G) = V_{t,k}$. *The payment upon revelation of a success in any state is just equal to the continuation payoff the agent would obtain by hiding it and remaining in the contract.*

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

- $\theta\rho_n^* w_{t,n^*}(G) + (1 - \theta\rho_n^*) w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*}$ *and* $w_{t,n^*}(S) \geq V_{t,n^*}$ *for* $S \in (G, B)$. *When the belief is* ρ_n^*, *which is the lowest belief for experimentation, both payments are set at least equal to the value that the agent would get by hiding it and remaining in the project until the end. (Proof provided in section 4)*

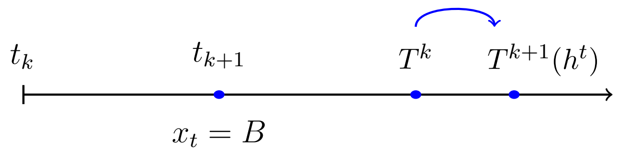
There are an infinite number of payment pairs that may satisfy the last item. This is because the agent cares about the expected payoff from putting in effort. However, the optimal payments in states $k < n^*$ are uniquely determined. The principal does not have to make a positive payment upon the revelation of bad signals in this region. As long as the agent's continuation value, which is a function of the remaining time, does not go down due to revealing a bad signal, he is willing to reveal it without receiving 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 by remaining in the project until the current deadline in state k and is increasing in the remaining time. This is because upon receiving and hiding a success, the agent does not have an incentive to work again. Even though there are an infinite number of possible deviations, the only relevant one is to shirk until the current deadline. The earlier the agent reveals a success, the higher will be his payment, as by revealing a signal and ending the relationship, he gives up the opportunity to keep the investment c until the deadline. Next proposition provides the increase in the continuation value of the agent and hence the change in the deadline after the disclosure of a bad signal.

Proposition 2. *In any optimal contract, the continuation value of the agent increases after the revelation of a bad signal, as long as this signal is not a terminal one, by allocating him a longer experimentation time as follows:*

- $V_{t,k+1} - V_{t,k} = \frac{c}{\lambda(1-\theta\rho_k)}$ *for* $k < n^*$. *When a bad signal is revealed, the*

Figure 1: Updating of T^k

from state k to $k+1$



continuation value of the agent increases by an amount which is higher for higher public beliefs.

- $T^{k+1}(t_{k+1}, h^t) > T_k(h^t)$ *for* $t_{k+1} \leq T^k$ *and* $k < n^*$. *The increase in the continuation value is due to an extension in the time allocated 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 a rise in the continuation value of the agent which translates into extra experimentation time, hence an extension of the current deadline. This is because the principal can reward the agent either through bonus payments or experimentation time and prefers the latter as long as possible and as long as experimentation has positive value. As the agent always has the possibility to enjoy positive surplus by not working, longer experimentation time is equivalent to higher benefits for the agent. Figure 1 demonstrates the updating of the deadlines. The termination rule T^k denotes the stopping time as a function of the history when the belief is ρ_k. When the $k+1$'th bad signal is revealed, the deadline becomes $T^{k+1}(t_{k+1})$. This continues until the last bad signal which terminates the contract. The belief ρ_{n^*+1} is the belief at which experimentation ends before a deadline is reached and coincides with the stopping level of belief in the benchmark case without agency. The extension in the deadline is higher the earlier the bad signal is revealed: as the extended time is added to the end of the current deadline T^k, 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.

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 can possibly end inefficiently early at a belief level at which the principal would have continued without the agency problem. This is why the time horizon gets extended upon the revelation of a bad signal: the principal provides incentives by allocating more time for experimentation after the revelation of bad signals while keeping constant the expected payment to the agent. In other words, the principal minimizes the current bonus payments and raises the agent's continuation value after bad signal revelation. The agent's benefit from experimenting has three components: the reward upon the revelation of a good signal, the reward upon revelation of a bad signal and the continuation value in the more pessimistic state after the revelation of a bad signal. What matters for the agent's incentives is the expected value of experimentation and not the decomposition of it.

While increasing the agent's continuation value in the more pessimistic state which is reached once he reveals a bad signal, the principal can propose lower bonus payments upon revelation of signals. However, how much the principal can back load the agency cost into future experimentation time in this way is determined by the disclosure constraints which determine the minimum bonuses that make sure the agent discloses the signals. The rest of the agency rent is allocated in form of extended contract horizon to the agent. While extending the contract horizon, the principal transfers the bonus payments into future expected payments.

The increase in the continuation value is not a necessary condition to induce the agent to reveal a bad signal. Actually, 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 of the deadline upon the revelation of a bad signal is optimal because it leads to more experimentation time while keeping constant the total agency rent.

4 Solving for the Optimal Contract

The agent's continuation value is written as:

$$V_t = \lambda a_t dt[\theta\rho_k W_{t,g} + (1-\theta\rho_k)W_{t,b}] \\ + (1-a_t)cdt + (1-\lambda a_t dt)(1-rdt)V_{t+dt,k}$$

The agent chooses $a_t = 1$ iff the following equation is maximized at $a_t = 1$.:

$$a_t\lambda\theta dt(\rho_k W_{t,g} + (1-\rho_k)W_{t,b}) + (1-a_t)cdt + (1-\lambda a_t dt)(1-rdt)V_{t+dt,k} \qquad (3)$$

This section leads through the steps of finding the optimal contract. The principal's problem is one of choosing an optimal termination rule while minimizing the expected payments to the agent subject to satisfying his incentive compatibility conditions. As a first step, the agent's continuation value is provided and then, the incentive constraints that should be satisfied to make him work and reveal the signals. In order to find out which constraints bind, I consider the incentive compatibility conditions at the deadlines. Initially, I restrict attention to local incentive constraints. This provides the bonus payments and the updating rule of the deadlines. Then, I conclude by showing how the initial deadline T^0 is computed. Finally, I verify that the local constraints are indeed sufficient for global incentive compatibility. I show that even though there are many types of deviations for the agent, the only relevant one is the deviation to not work again.

4.1 Agent's continuation value

While the agent puts in effort, he may receive a good or a bad signal or no signal at all. Then, the law of motion of the agent's continuation value at time t and belief ρ_k which I call $V_{t,k}$ is as follows (given that he discloses the signals as soon as he receives them):

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

If the agent chooses a_t, he receives a signal with probability $a_t\lambda dt$ which is a good or a bad one 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 a bad one and he reveals it, he receives the payment $w_{t,k}(B)$ and continues experimenting in state $k+1$ getting the continuation value $V_{t+dt,k+1}$, until the n^*+1'th bad signal is revealed. 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 while the agent shirks. Letting dt go to 0 gives the Hamilton-Jacobi-Bellman (HJB) equation for the agent:

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

Lemma 3. *Given a contract C the agent will choose $a_t = 1$ iff:*

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

where $W_t = \theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) + V_{t,k+1})$ which is the expected payoff to the agent due to the probability of receiving and revealing a signal.

Lemma 3 makes sure that the agent does not benefit from deviating to shirk at any moment. Let us now explain why the local constraint for working in equation (5) is enough for satisfying global incentive compatibility. When the agent deviates to shirk, he does not get any signals and the beliefs of the agent and the principal do not differ. Then, 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 up to time $t_1 < t$ given by h^{t_1} and hence the belief ρ_k. The first h^t is such that the agent shirks from t_1 until t. The second history $\hat{h}^t$ is such that the agent works from t_1 until t but does not receive any signals. Hence, given that no additional 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 is summarized by the times at which signals are revealed. As the agent's belief and the continuation value at t are the same in the two histories, this deviation does 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 to deviate at a future date. Then, in case equation (8) is satisfied at any t, it will be satisfied globally as well.

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

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

Letting dt go to 0:

$$-V'_{t,k} + rV_{t,k} \geq c \tag{6}$$

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

$$V_{t,k} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) \tag{7}$$

Hence equation (7) gives the minimum continuation value that should be provided to the agent. If the constraint in equation (6) binds at any t, then equation (7) will also hold as an equality. The continuation value is found assuming that this constraint is sufficient for global incentive compatibility. It will be verified later on that there is actually no other profitable deviations.

The second type of constraints that should be satisfied are the *disclosure* constraints, which ensure that the agent reveals the signals upon receiving them. The ability to keep the signals and reveal later on adds the complication of many possible histories after deviations, such as hiding a signal and experimenting in order to possibly get another one. 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 which consists of hiding a signal and shirking afterwards, in other words to local deviations.

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

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

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

where w' and V' denote derivatives with respect to t. Equation (8) *makes sure that the agent does not want to delay revealing a good signal, and equation* (9) *makes sure that the agent does not want to delay revealing a bad signal.*

Proof. First, I derive equation (8) which makes sure that the agent is willing to reveal a good signal upon receiving it instead of hiding it to reveal at $t + dt$ and shirking in the meantime, which is:

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

which as dt goes to 0, leads to:

$$-w'_{t,k}(G) + rw_{t,k}(G) \geq c \tag{10}$$

Next I will derive equation (9). The constraint which makes sure that it is not profitable to wait for an infinitesimal time dt before revealing a bad signal is:

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

Letting dt go to 0:

$$-w'_{t,k}(B) - V'_{t,k+1} + r(w_{t,k}(B) + V_{t,k+1}) \geq c \tag{11}$$

□

If the local constraints are sufficient conditions for global incentive compatibility, then it is sufficient that these two conditions are satisfied at each t. Now, assuming this is the case, I solve the differential equation (8) using the boundary condition at T^k:

$$w_{t,k}(G) \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)}w_{T^k,k}(G) \tag{12}$$

Doing the same for the bad signal revelation by integrating equation (9) and using the condition at the deadline T^k:

$$w_{t,k}(B) + V_{t,k+1} \geq \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)}(w_{T^k,k}(B) + V_{T^k,k+1}) \tag{13}$$

where $w_{T^k,k}(B) = 0$.

4.2 Conditions at the deadline T^k

I will now derive the incentive conditions that should be satisfied at the deadline T^k for any k. The reason for focusing on the deadlines is first it is known that $V_{T^k,k} = 0$ as the contract ends when T^k is reached, and second the disclosure constraints are irrelevant at the deadline as the relationship will terminate in case no other 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:*

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) = \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1} \tag{14}$$

by using equation (5)*.*

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

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

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

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)(w_{t,k}(B) \geq \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1} \tag{15}$$

At T^k, the revelation constraints are irrelevant: if a signal arrives just before the deadline, the agent is willing to reveal it without receiving any payment as it has no effect on his continuation value which is zero. Then, the condition in equation (15) is the only one that should be satisfied, and hence it will bind. In case it were slack, then the principal could decrease the expected payment while still satisfying this constraint. □

4.3 The values $w_{t,k}(G)$, $w_{t,k}(B)$, $V_{t,k}$ and T^k

I will start by finding the payments in state n^* which is the last state in which experimentation happens 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 by proposition (3).

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

$$\theta\rho_{n^*} w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) = \frac{c}{\lambda} + V_{t,n^*} \tag{16}$$

$$w_{t,n^*}(S) \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) + e^{-r(T^{n^*}-t)}w_{T,n^*}(S) \tag{17}$$

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

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

$$\lambda[\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_{n^*})w_{T^{n^*},n^*}(B)] \geq c \tag{18}$$

As the revelation constraints are irrelevant at the deadline T^{n^*}, the principal will make the incentive constraint in equation (18) bind. In case this constraint were slack, it would mean the payments could be decreased without modifying the incentives and the principal's profits would have increased. Then, for $t < T^{n^*}$:

$$\theta\rho_{n^*}w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{\lambda} + V_{t,n^*} \tag{19}$$

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 benefit that the agent could obtain by shirking, and V_{t,n^*} is the future payoff foregone after revealing a signal that leads to project termination. Before concluding that equation (19) binds, it is necessary to check the *disclosure constraints*. Multiplying the constraint for the revelation of G given in equation (12) and the constraint for B given in equation (13) respectively by their probabilities $\theta\rho_n^*$ and $1-\theta\rho_n^*$ gives:

$$\theta\rho_n^* w_{t,n^*}(G) + (1-\theta\rho_n^*)w_{t,n^*}(B) \geq \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) + e^{-r(T^{n^*}-t)}(\theta\rho_n^* w_{T^{n^*},n^*}(G) + (1-\theta\rho_n^*)w_{T^{n^*},n^*}(B)) \tag{20}$$

It is easy to conclude that equation (20) is slack when the constraint in equation (19) 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 (19) bind and satisfy 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 $T^{k+1}(t_{k+1}) < T^k$, meaning contracts whose horizon decreases upon the revelation of bad signals. Now, 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 will not shorten upon the 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 towards solving for the optimal contract.

Lemma 7. *The condition to work at the deadline should bind:*

$$\theta\rho_k w_{T^k,k}(G) + (1-\theta\rho_k)w_{T^k,k}(B) = \frac{c}{\lambda} - (1-\theta\rho_k)V_{T^k,k+1} \tag{21}$$

The optimal contract should have $w_{T^k,k}(G)$ and $w_{T^k,k}(B)$ set to 0, and the continuation value $V_{T^k,k+1}$ is given by:

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

(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 after the revelation of a bad signal is strictly positive. This means incentive to work near the deadline is provided only through higher continuation values.

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 bad signals.

Lemma 8. *The payments upon the 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 (9) 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 $V_{T^k,k+1} = \frac{c}{\lambda(1-\theta\rho_k)}$:

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

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 (38) 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, $w_{t,k}(G)$ will be found.

Lemma 9. *The payment upon the revelation of a success is set to its minimum value which leaves the agent indifferent between revealing it or hiding it and remaining in the project and shirking 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 (8) bind in order to find the minimum $w_{t,k}(G)$:

$$w_{t,k}(G) = \frac{c}{r}(1 - e^{-r(T^k - t)}) + e^{-r(T^k - t)} w_{T^k,k}(G) \tag{23}$$

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:

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

This constraint binds for the same reason as in lemma (7). Given that the incentive constraint for working binds for any (t, k), I conclude that the condition given in equation (7) 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 (24):

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

At the deadline, the condition given by lemma 7:

$$V_{T^k,k+1} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{26}$$

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 (25):

$$e^{-r(T^k(t)-t)} - e^{-r(T^{k+1}(t)-t)} = \frac{c}{\lambda(1 - \theta\rho_k)} \tag{27}$$

From the above equation, it is easy to conclude that $T^{k+1}(t) - T^k$ is decreasing in k: as the belief becomes more pessimistic, the extension in the deadline after a bad signal becomes less. Also, in a given state k, $T^{k+1}(t)$ decreases in t for a 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 the view point of time t.

Now it can be concluded that the continuation value of the agent (or the current deadline) and the current public belief are sufficient for summarizing the history. 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 which determines his current belief (and not the times at which they were received) and his continuation value. As revelation of signals are already reflected in the agent's bonus payments and continuation value (equivalently the current deadline), the only relevant history at time t can be summarized by k.

Discussion: In any state $k < n^*$, when the agent reveals a bad signal, his continuation value increases by an amount which makes 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 disclosing B.

The revelation of a bad signal causes the belief about the project quality to go down, but at the same time signals only arrive while the agent is working. The agent's expected payoff from working consists of the payment upon good news, payment upon bad news and the change in continuation value after the revelation of a bad signal, and how this rent is decomposed into these three is irrelevant for his incentives as long as the constraints to work and disclose are satisfied. Hence, the principal chooses the payments and continuation values in an optimal way, which is through minimizing the bonus payments and increasing the continuation value of the agent in the more pessimistic state which is reached after the revelation of a bad signal, while keeping constant the total rent. The principal benefits from longer experimentation as long as $\rho_k \geq \frac{c}{\lambda\theta}$, in other words as long as the intrinsic value of experimentation is positive.

4.4 Sufficiency of local revelation constraints:

In this section I verify that the local constraints for the disclosure in Lemma 3 are indeed sufficient to account for global incentive compatibility. First, I will show that the agent cannot benefit from hiding a bad signal. Then, I will show that upon hiding a good signal, he does not have an incentive to work again. The results remain robust in case the signals are lost when not revealed right away, which is discussed in the section 5.

As a first step I check that there is no profitable deviation for the agent after receiving a bad signal, such as hiding one or more signals. The agent should be compensated at least for the change in his continuation value upon revealing a bad signal when the state moves from k to $k+1$:

$$w_{t,k}(B) \geq max[0, V^B_{t,k} - V_{t,k+1}] \tag{28}$$

where $V^B_{t,k}$ is the continuation value of the agent after hiding the bad signal. Now, let us consider what is the best deviation of the agent after hiding a bad signal. The incentive constraint to work initially binds:

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

After deviation, the gain in the continuation value of the agent upon the revelation of the next bad signals will be respectively $\frac{c}{\lambda(1-\theta\rho_k)}$ and $\frac{c}{\lambda(1-\theta\rho_{k+1})}$ which are independent of when he reveals them. Also, the reward upon the revelation of a good signal is always higher for higher k: $w_{t,k}(G) < w_{t,k+1}(G)$, which means the agent would reveal the good signal only after having revealed all the bad signals that he has already received. Then, hiding and delaying the revelation of a bad signal is weakly dominated by a strategy in which he reveals the $k+1$'th bad signal upon receiving it and obtains the same increase in his continuation value of $\frac{c}{\lambda(1-\theta\rho_k)}$ earlier. Then, equation (28) holds and the agent cannot gain by hiding or delaying the revelation of a bad signal.

Finally, I will verify that the deviation to hide a good signal and continue experimenting in order to get another bad signal is not profitable either. This might have been profitable if $w_{t,k+1}(G)$ were sufficiently high compared to $w_{t,k}(G)$. In other words, only if the reward to a good signal is much more profitable while the belief is more pessimistic. The best deviation of this kind would be to deviate to hide the good signal and experiment in order to get a signal B and once it arrives, reveal it and then reveal G at some time $\hat{t}$ in state $k+1$ in order to receive $w_{\hat{t},k+1}(G)$ where $\hat{t} \leq T^{k+1}$. It is enough to look at an instantaneous deviation of this kind, meaning deviating to experiment for a moment dt and in case a signal arrives, revealing it at some point in the future. This local deviation is not profitable if the following holds:

$$w_{t,k}(G) \geq [\lambda dt(1-\theta)(1-rdt)(\frac{c}{r}(1-e^{-r(\hat{t}-t)})+w_{\hat{t},k+1}(G))] \\ +(1-\lambda(1-\theta)dt)(1-rdt)w_{t+dt,k}(G) \quad (29)$$

for any $\hat{t}$. The first term denotes the possibility of getting a bad signal after deviating to experiment for an additional instant of dt. Then, the agent will shirk until a time $\hat{t}$ at which he reveals the good signal hence terminates the contractual relation. I consider the *best* deviation of this kind, as in case of hiding a bad signal the agent can reveal the signals at any moment until the deadline, and work or shirk in the meantime. However, as the reward upon good signal revelation is higher when the number of bad signals already revealed is higher, the agent would always reveal all the bad signals he has received before revealing the good one. By the revelation constraint in equation (9), the agent does not want to delay revealing a bad signal. In addition, equation (8) makes sure that the agent does not delay the revelation of a good signal at a given belief. Then, it is sufficient to look at the condition when $\hat{t}$ is replaced by t, as this gives the upper bound on the agent's possible deviation to work. Then, if the condition in (29) is satisfied at t, it should also be satisfied at any $\hat{t} > t$. By making dt go to zero in equation (29), the condition becomes:

$$-w'_{t,k}(G)+rw_{t,k}(G) \geq \lambda(1-\theta)(w_{t,k+1}(G)-w_{t,k}(G)) \quad (30)$$

where the left hand side is greater than c. It is sufficient to check for one shot deviations of this kind: if after finding a good signal it is not profitable to experiment in state k in order to get a signal B and reach state $k+1$, it will not be profitable to get the bad signal two times either. To see this, check that the change in $w_{t,k}(G)$ due to an increase in k multiplied by the probability of getting the bad signal is constant: $\lambda(1-\theta\rho_k)[w_{t,k+1}(G)-w_{t,k}(G)] = c$, which is just equal to the benefit from shirking. This implies that: $w_{t,k+1}(G)-w_{t,k}(G) \leq \frac{c}{\lambda(1-} - V_{t,k} = \frac{c}{\lambda(1-\theta\rho_k)}$ and $\lambda(1-\theta\rho_k) < \lambda(1-\theta)$, the right hand side is less than c and this constraint is indeed slack. After hiding a good signal the agent is better off shirking instead of working and this type of deviation is indeed not relevant for the optimal contract. The gain from an additional B in terms of the higher reward 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 in order to get a bad signal. Finally, I conclude that the local constraints are sufficient to satisfy global incentive compatibility.

4.5 The stopping belief

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

Proposition 3. *Experimentation ends as soon as the belief falls down 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 without agency is as follows: the total cost of experimentation to the principal per unit time under agency is $2c$: the first c which should be provided to invest in experimentation and the second c as a reward in order to prevent the agent from shirking. By initially committing to a deadline, the principal promises the agent the benefit c over the remaining horizon which he can obtain without working. This means, any time the agent reveals a signal that will terminate the relationship, he should be compensated for the value of his deviation to shirk and remain in the project until the deadline, which is equal to $\frac{c}{r}(1 - e^{-r(T^k - t)})$. Hence, given that the principal has already committed to pay the agency cost c, it is optimal to continue experimentation while keeping constant the total agency cost as long as $\lambda\theta\rho_k \geq c$ which corresponds to the benchmark case without agency.

4.6 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 equation (25). Next proposition provides the initial deadline T^0.

Proposition 4. *T^0 is the time allocated such that if reached without any signal revelation the project gets terminated.*

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0 + (1-\theta\rho_0)F_{T^0,1}) - (1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{31}$$

The value $F_{T^0,1}$ is positive and less than 1 (given that the value of a project known to be good is equal to 1), in addition it depends only on $T^1(T^0) - T^0$ which can be found from the agent's utility $V_{T^0,1} = \frac{c}{\lambda(1-\theta\rho_0)}$. Then, $F_{T^0,1}$ can be recovered by using the updating rule of the deadlines. It is easy to see that T^0 is decreasing in c. The sign of the derivative with respect to λ can be positive or negative. The derivatives with respect to θ and ρ_0 have the same sign as $\theta(c - \lambda F_{T^0,1})$, which could be either positive or negative depending on the value of $F_{T^0,1}$. The reason for this is that an increase in 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 enough. Hence, the sign of the derivative with respect to θ and ρ_0 depend on which one 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 if not revealed right away. Finally, I consider the game without 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 in 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 to the agent to exert effort are provided completely through increased continuation values upon the 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 any signal must be positive.*

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

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

The difference is that now there are no revelation constraints, which implies that the above condition will bind in the optimal contract. In addition, it is still optimal for the principal to extend the horizon of experimentation upon the revelation of bad signals, and this time setting $w_{t,k}(G) = w_{t,k}(B) = 0$ 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 can 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^* due to the arrival of a good or a bad signal. In all the previous states, due to the public observability of signals, the principal is able to set the payment even upon a success equal to zero, and incentivize the agent completely 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 the contract without providing any positive payment, but in overall his expected benefit from working is high enough that he is willing to work. The agent gets a positive payment only if a signal is realized when 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 to provide incentives to the agent is through the bonus payments upon the realization of signals. The division of this payment between good and bad signals does not matter as the agent cannot choose to hide a signal.

Now let us compare the cases of public and privately observed signals. Call $\tilde{V}_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + \frac{\tilde{V}_{t,k}}{(1-\theta\rho_k)}$ where $\tilde{V}_{t,k+1}$ denotes the public signal case, and $V_{t,k+1} = \frac{c}{\lambda(1-\theta\rho_k)} + V_{t,k}$ the private case. It is easy to see that the increase in the continuation value is higher in the public signals case. This is due to the fact that the principal does not have to pay the informational rent 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.

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

5.2 Case of no moral hazard

Consider the case in which the agent's decision to experiment or shirk is perfectly observed by the principal, but not the arrival of the signals. Although the agent has the option to hide the signal, there is no gain in doing so as his expected 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 rents either and the principal's problem is identical to the case in which she experiments alone. This means that private observation of the signals alone does not lead to any distortions in the principal's problem compared to the first best.

5.3 Signals get lost if not revealed right away

This is a special case of the original setting considered in the paper and does not modify the results. Indeed, under this assumption, the possible deviations after receiving a signal are much simplified. If the agent hides a good signal, he will not find it optimal to work again in this setting either. To see why this is the case: the possible actions after hiding a good signal 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 provides the same minimum reward, $w_{t,k}(G) = V_{t,k}$, as in the original contract. The condition for the revelation of G is:

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

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

5.4 The Case without Commitment

Let us discuss the game with no commitment. For this part, I will focus on 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, and similar for 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 remains high enough. In the optimal contract the above equation will 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 contributes to the literature on agency models of experimentation by introducing a different learning process. The novelty is that the agent can obtain good or bad signals only while he exerts costly effort and hence it is possible to distinguish an agent who is unlucky from an agent who shirks. The possibility of not receiving any signals even while the agent experiments implies that monitoring the agent's effort is difficult and requires punishing the agent through stopping too early at times. Punishment through terminating the contract happens when enough time passes without the agent providing any news. In addition, the private observation of signals by the agent implies that he should be provided enough incentives to reveal them. As the signals are a result of the agent's effort, bonuses are promised only upon the revelation of signals.

I find that the optimal contract initially allocates a certain amount of time for the agent to experiment, and provides extra experimentation time after the revelation of bad signals. The principal extends the contract horizon while keeping constant the expected agency rent. While the agent experiments, he has a possibility to receive a good signal and a bonus upon its revelation or a bad signal and a higher continuation value after the revelation of it. Increasing the continuation value is a way of paying the agent. Even though a bad signal leads to a more pessimistic belief, the principal prefers experimenting longer as long as the agency rent does not increase and the value of experimentation is still positive. The principal back loads the agency rent in such a way that the experimentation time is extended as much as possible while the expected payment to the agent remains constant. The results of this paper suggest that incentives can be provided optimally through endogenous deadlines when news about project quality arrive over time in form of good or bad news.

There are some questions related to the paper that may be of interest for future research. One such question is what could happen if the underlying quality of the project is persistent but not constant over time, in other words if there is a learning effect which means that even a project which is bad can succeed in the future depending on the effort put in by the agent. 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 is important in the long run. This would be a model in which the agent is hired for creating 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 interesting future 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

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

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

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

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:

$$\theta\rho_k w_{t,k}(G) + (1 - \theta\rho_k)w_{t,k}(B) \geq 0 > \frac{c}{\lambda} + V_{t,k} - (1 - \theta\rho_k)V_{t,k+1} \qquad (32)$$

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

$$(1 - \theta\rho_k)V_{t,k+1} \geq \frac{c}{\lambda} + V_{t,k} - \theta\rho_k w_{t,k}(G) \qquad (33)$$

If we decrease $V_{t,k+1}$ 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 be payed 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 now becomes $\hat{T}^{k+1}(t) > T^{k+1}(t) > T^k$. 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 as 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, consider increasing $V_{t,k+1}$ 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 of time at T, 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:

$$(1-e^{-(T^{k+1}-t)(\lambda+r)})\frac{[\lambda\theta\rho_{k+1}]}{\lambda+r} + \int_t^{T^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})F_{s,k+2}ds - \frac{c}{r}(1-e^{-r(T^{k+1}-t)}) \tag{34}$$

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

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

$$(1-e^{-(\hat{T}^{k+1}-t)(\lambda+r)})\frac{\lambda\theta\rho_{k+1}}{\lambda+r} + \int_t^{\hat{T}^{k+1}(t)} e^{-(s-t)(\lambda+r)}(1-\theta\rho_{k+1})\hat{F}_{s,k+2}ds - (V_{t,k+1}+\Delta) \tag{35}$$

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

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

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

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:

$$F_{t,n^*} = \int_0^{T^{n^*}} e^{-t(\lambda+r)}[\lambda(\theta\rho_n^*(1-w_{t,n^*}(G)) + (1-\theta\rho_{n^*})F_{t,n^*+1}) - c]dt \tag{37}$$

where $w_{t,n^*}(G) = \frac{c}{r}(1-e^{-r(T^{n^*}-t)}) = V_{t,n^*}$. Now, the condition for $36 > 37$ 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 $V_{t,n^*+1} = \frac{c}{\lambda(1-\theta\rho_{n^*})} + V_{t,n^*}$:

$$-V_{t,n^*+1} \geq F_{t,n^*+1} \tag{38}$$

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 (38) 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:

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

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

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

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

$$T^0 = \frac{ln(\frac{\lambda(\theta\rho_0+(1-\theta\rho_0)F_{T^0,1})-(1-\theta\rho_0)c}{\theta\rho_0 c})}{\lambda} \tag{39}$$

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, the value function is decreasing for $t \geq T^0$. The second derivative becomes positive as T goes to ∞. However, as there is no other point at which the first derivative is zero and the second derivative is negative for $T > T^0$, I conclude that this value can never go above T^0.

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

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

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

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

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

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

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

The revelation constraints are irrelevant at the deadline T^k. It is then optimal that the equation (40) binds, which provides the payments at the deadline. Then, for $t < T^k$, using $V_{T^k,k+1} = V_{T^k,k} = 0$, the revelation constraint for B becomes:

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

The revelation constraint for G is:

$$w_{t,k}(G) \geq \frac{c}{r}(1-e^{-r(T^k-t)}) + e^{-r(T^k-t)}w_{T^k,k}(G) \tag{41}$$

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

$$\begin{aligned}
\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) \geq V_{t,k} + e^{-r(T^k-t)}(\theta\rho_k w_{T^k,k}(G) \\
+ (1-\theta\rho_k)w_{T^k,k}(B)) - (1-\theta\rho_k)V_{t,k+1}
\end{aligned} \tag{42}$$

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:

$$\theta\rho_k w_{t,k}(G) + (1-\theta\rho_k)w_{t,k}(B) \geq \frac{c}{\lambda} + V_{t,k} - (1-\theta\rho_k)V_{t,k+1} \tag{43}$$

Comparing the equations (42) and (43), 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 (43) binds and the revelation constraints are satisfied. Next I will show that $T_{k+1}(t_{k+1}) < T_k$ cannot be optimal for any k by replacing the payments into the principal's value function. First, I will look at F_{0,n^*-1} (normalizing the starting time of state n^*-1 to 0) and F_{t,n^*} to show that $T^{n^*}(t_{n^*}) \geq T^{n^*-1}$. After this, I will show that it also holds for any $k < n^*$. The optimal payment schedule in state n^* was already provided for any optimal contract in subsection (4.3), 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:

$$F_{0,n^*-1} = \int_0^{T^{n^*-1}} e^{-t(\lambda+r)} \Big[\lambda(\theta\rho_{n^*-1} - \frac{c}{\lambda} - \theta\rho_k \frac{c}{r}(1 - e^{-r(T^{n^*}-t)}) - \frac{c}{r}(e^{-r(T^{n^*}-t)} - e^{-r(T^{n^*-1}-t)}) + F_{t,n^*}) - c\Big] dt \quad (44)$$

In state n^*:

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

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 in the next state the deadline is extended upon revelation of a bad signal. I replace the payment schedule for state $k+2$ which is given by the proposition 1:

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

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