ESSAYS ON DYNAMIC GAMES AND CONTRACTS WITH LEARNING

Nishant Ravi

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Co-Supervisor of Dissertation

George J. Mailath
Professor of Economics

Co-Supervisor of Dissertation

Steven A. Matthews
Professor of Economics

Graduate Group Chairperson

Jesus Fernandez-Villaverde, Professor of Economics

Dissertation Committee:

George J. Mailath, Professor of Economics

Steven A. Matthews, Professor of Economics

Mallesh M. Pai, Assistant Professor of Economics

Dedicated to my parents, Veena and Bhola.

ACKNOWLEDGMENTS

I am deeply indebted to my advisers, Prof. George Mailath and Prof. Steven Matthews for their unwavering encouragement, support, and guidance in the tough process of learning and self-discovery that has culminated into this thesis. They believed in me and were available for me during my darkest hours. Without their support, this thesis wouldn't be possible.

I am also deeply grateful to Prof. Mallesh Pai for his constant support, encouragement, and guidance since my early days in the graduate program. He has stood resolutely behind me in my academic endeavors and his advice and support at critical junctures have been immensely helpful. I sincerely thank my dear friend and co-author Aditya Kuvalekar for his support and guidance without which this dissertation would not be possible.

I thank other faculty members of the theory group at Penn, including Aislinn Bohren, David Dillenberger, Annie Liang, Andrew Postlewaite, Rakesh Vohra, and Yuichi Yamamoto for their encouragement and support at various stages of the dissertation and during the job market process. I would also like to thank my advisor from my master's program, Prof. Arunava Sen who encouraged me to pursue a career in academia and has supported me throughout in this endeavor. I would also like to thank my fellow students and friends Himanshu Sharma, Daniel Hauser, Behrang Kamali, Tim Hursey, Raphael Galvao, Kunhee Kim, Tanvi Rai and Manvi Bhatnagar for their support and company during my days at Penn.

I finally thank my family for their endless love and support, and Natalie for accompanying me through these years with unconditional understanding, encouragement and affection.

ABSTRACT

ESSAYS ON DYNAMIC GAMES AND CONTRACTS WITH LEARNING

Nishant Ravi

George J. Mailath

Steven A. Matthews

In the first chapter of the dissertation, “When and How to Reward Bad News”(joint with Aditya Kuvalekar), we explore when and how to reward the bearer of bad news in a dynamic principal-agent relationship with experimentation. The agent receives flow rents from experimentation, and divides his time between searching for conclusive good news and conclusive bad news about project quality. The principal commits in advance to rewards conditional on the type of news. At each instant, the principal makes a firing decision. We explore two environments: when the principal observes the agent's allocation and when she does not. We show that, in both the environments, the principal's optimal equilibrium features a stark reward structure—either the principal does not reward the bearer of bad news at all or rewards the bearer of either news equally.

In the second chapter of the dissertation “Supervising to Motivate”, I study a dynamic principal-agent relationship in which the principal invests costly resources in a project of uncertain quality to induce costly effort from an agent. The principal observes the output from the project privately and can be either informed (has learned that project quality is high) or uninformed. The agent learns about project quality through the investments made by the principal. The principal wants to invest less when pessimistic about project quality; however, the agent demands higher investment when pessimistic to exert effort. The principal faces the trade-off between investing optimally and transmitting information about project quality to the agent. The principal's optimal equilibrium features full information transmission when the uninformed principal has high beliefs (probability that project quality is high) and no information transmission at low beliefs. The informed principal may invest at sub-optimally high levels early in the relationship, but eventually, optimality is restored. That is, the principal's optimal equilibrium may exhibit distortions in the short run but not in the long run.

Contents

List of Figures

Chapter 1

When and How to Reward Bad News

Joint with Aditya Kuvalekar

1.1 Introduction

When faced with projects of uncertain feasibility, individuals or organizations engage in experimentation to acquire information about the prospects of such projects. Typically, there are multiple ways to acquire information. For example, a researcher with a conjecture in hand may attempt to develop a constructive proof or search for a counterexample disproving the conjecture. A scientist in a tech firm may look for information that confirms that a prototype satisfies all requirements to be put to production, or may look for a fatal flaw in the prototype. That is, different strategies of acquiring information about a project may produce different types of news, such as "good news" establishing that the project is successful or "bad news" establishing that the project will fail. Timely bad news helps organizations save future costs and better allocate resources to other activities. However, when the entity performing experimentation is different from the one bearing the costs of experimentation, as is often the case, the incentives of the two parties may differ. For example, an R&D department tasked with the development of a new product may not want to provide bad news to management fearing closure of projects, funding cuts etc. An obvious remedy—one that is advocated in the literature from economics (Levitt and Snyder (1997)) and finance (Manso (2011)) to organizational behavior (Stefflre (1985))—is to reward the bearer of bad news.

Accepting the idea behind rewarding employees for bearing bad news, the goal of this paper is to explore when and how much to reward the bearer of bad news in principal-agent relationships with experimentation.

There are two important issues we need to address in this regard. First, rewarding the bearer of bad news is costly to the firm. Second, such rewards could, in principle, also create perverse incentives leading to employees spending an inefficiently large proportion of their time searching for bad news.[1]

We develop a simple dynamic principal-agent model to pursue these issues. An agent performs experimentation to assess the quality of a project while an investor (the principal) bears the costs of experimentation and can terminate the relationship at will. The agent allocates his resources across two different sources of information—a good news source that can produce a signal only if the project is of high quality and a bad news source that can produce a signal only if the project is of low quality. The principal commits to a news contingent reward but cannot commit to a termination policy. The model delivers three main insights. First, we find that the principal should do one of two things—terminate the relationship with no severance payment upon producing bad news, or reward the agent the same amount for producing bad news or for producing good news. Second, it may be optimal to reward the agent for producing bad news if the initial assessment about the project quality is sufficiently high but not if the initial assessment is low. Moreover, it will never be optimal to reward the agent for producing bad news if the initial assessment is low while rewarding him if the initial assessment is high. The intuition behind this, seemingly counterintuitive, observation is that rewarding for bad news is costly for the principal and, when the initial assessment about the project quality is low, the costs may not offset the benefits. The third insight is that, while the above two observations hold *regardless* of whether the agent's action is observable to the principal or not, the principal may do strictly better if she does not observe the agent's action.

More specifically, we study a continuous time principal-agent relationship where the players seek to learn about a project of unknown quality (state), either high or low. Both

[1]There is also the possibility of agent fabricating bad news or sabotaging the project. While these aspects are important in certain circumstances, they are not the focus of this work.

players are equally informed about the quality of the project. In the beginning of the game, the principal commits to a reward structure, specifying rewards for the agent upon revealing the project quality. After accepting the reward structure, the agent experiments. At each instant, the principal chooses whether to continue or terminate the relationship by firing the agent. Conditional on the principal continuing the relationship, the agent chooses how to allocate his one unit of effort across two arms—a good news arm and a bad news arm. The agent's choice is observable to the principal but is not verifiable in an outside court. Therefore, the parties cannot write contracts contingent on agent's actions. The good (bad) news arm produces a conclusive signal—"good (bad) news"—at an arrival rate proportional to the effort allocated to it if and only if the project quality is high (low). Hence, a signal on either arm fully resolves all the uncertainty. Good news also provides the knowledge needed to implement the project which results in a lump-sum payoff to the principal while bad news is costless in and of itself. Experimentation is costly and the costs are borne by the principal. The agent earns flow rents while experimenting.

The initial reward structure has two restrictions: First, limited liability, i.e., the agent cannot be forced to pay the principal under any circumstance. Second, the agent's reward upon obtaining good news is bounded from below by the flow rent of the agent. The motivation behind this assumption arises from our interpretation of good news as the principal adopting the project and employing the agent to work on it to implement the project. After having produced good news, the agent continues to receive the flow rent he receives during experimentation implying that the reward to the agent on producing good news cannot be less than the flow rent.

Both parties have the same discounting rate and have outside options that are normalized to zero. We study the Markov Perfect Equilibria, henceforth equilibria, of this game using the natural state variable—the posterior probability that the project quality is high.

An important feature of our model is that the agent's allocation is perfectly observable but not verifiable in an outside court. We make two points in this regard. First, in Section 1.5.1, we show that our main results continue to hold even if the agent's allocation is not observed by the principal. Second, this is a realistic assumption in relationships of experimentation such as startups financed by venture capitalists. A wide body of evidence suggests that venture capitalists closely monitor the firms they invest in by having more board seats (Lerner and Tåg (2013)) and this leads to an increase in innovation (Bernstein, Giroud, and Townsend (2016)).

Should the principal offer a reward upon producing bad news? If yes, how much should it be relative to the reward upon producing good news? Notice that the absence of a signal when searching for bad news makes the players more optimistic. Hence, if the principal chooses to keep the agent employed when her belief is above a cutoff belief—as will be the case in our principal-optimal equilibrium—then the agent has an incentive to look for bad news closer to this belief to ensure that the beliefs do not drift below the cutoff. However, for this very reason, it may be the case that the agent searches for bad news only to avoid termination and nothing more.

Lemma 1 shows that this is indeed the case when the reward upon producing bad news is less than the agent's flow rent. When the principal does not fire the agent when the posterior belief is above a cutoff belief, say $\underline{p}$, the agent searches for good news everywhere except at $\underline{p}$, where he combines the search for good and bad news in a way that the beliefs remain at $\underline{p}$ in the absence of a signal. This choice is called "freezing beliefs". The intuition behind this agent behavior is as follows. If project quality is high, bad news can never arrive and delaying the search for bad news increases the probability of producing good news and earning a reward. On the other hand, if the project quality is bad, he remains employed for a longer time as it takes longer for bad news to arrive. When employed, he collects the flow rent which is larger than the reward for producing bad news. In summary, regardless of the project quality, it is optimal to delay the search for bad news when the reward for producing bad news is lower than his flow rent.

When the reward upon producing bad news is higher than the flow rent, the agent may search for bad news when sufficiently pessimistic because the likelihood of obtaining bad news, and thereby its associated reward, may be higher than obtaining good news at low beliefs. As a result, there is a cutoff belief p^f such that, the agent searches for bad news below it and searches for good news above.

With the agent's behavior fully understood, we turn our attention to the principal's value to ask: is it ever optimal to reward bad news? If yes, when, and how? Toward this answer, we make the following simple, but important, observation. Whenever the rewards for producing good and bad news are larger than the flow rent, the resulting agent's best response can be supported by choosing strictly lower rewards so long as they are larger than the flow rent. As consequence, we obtain that the optimal reward structure is stark—either do not reward the agent for producing bad news or reward him equally for producing good or bad news. In fact, both rewards should be equal to the flow rent.

When the principal rewards the agent for producing both types of news by setting the rewards equal to the flow rent, the agent's best response is not unique. In fact, any behavior of the agent that results in a non-negative drift of beliefs at the cutoff belief below which the principal fires, is a best response for the agent. Therefore, the principal's problem reduces to finding an optimal allocation policy within the set of best responses of the agent. The resulting optimal policy for the principal is characterized by a switching belief p^s such that the agent looks for bad news below p^s, and good news above.

Finally, we find the optimal reward structure by comparing the two values to the principal at the initial prior about project quality: one by not rewarding the agent for bad news, and the other where the reward for bad news is equal to the flow rent. Depending on the primitives of the model, the optimal reward structure can fall into one of three cases. In the first case, the agent is rewarded for bad news regardless of the initial prior. Second case is when the agent is not rewarded for bad news regardless of the initial prior. Unlike the first two cases, the optimal reward structure is sensitive to the initial prior in the third case. Here the agent is rewarded for bad news when the initial prior is sufficiently high, and not rewarded otherwise. The intuition behind this seemingly counterintuitive reward structure of rewarding for bad news if starting at high priors is related to the observation we made earlier—rewarding for bad news is costly to the principal. However, when the initial prior is high, the principal is willing to incur the costs of rewarding bad news, for he expects that to be a less likely outcome.

In proposition 3, we provide sufficient conditions for the reward structure to be of the first or the second type. In particular, we show that when the bad news technology is sufficiently informative and the cost of experimentation is high enough, it is optimal to reward bad news for all initial priors. In contrast, when the good news technology is sufficiently informative, it is optimal to not reward bad news for all initial priors.

Lastly, we explore the robustness of our findings. Of particular interest is the case when the agent's allocation choice is private—not observed by the principal. Dynamic games where actions affect learning about the underlying state are often intractable due to the possibility of the deviating player (agent in this case) possessing persistent private information, and thereby private beliefs. However, we find that the main forces that drive the results when allocation choice is observed continue to apply when allocation choice is private. We compute the principal's optimal policy in this case, and show constructively that it can be supported in equilibrium by either having the rewards for both good news and bad news to be equal to the flow rent, or by not rewarding the agent for producing bad news at all. Interestingly, the principal may strictly benefit when the agent's action is not observable to her. The reason is that when there are no rewards for bad news, the agent will exclusively search for good news until getting fired, a policy that the principal prefers over one in which the agent searches for a good news until a cutoff belief, where he freezes. We also prove that the main insights of our model remain unchanged even if the bad news technology was not fully revealing, i.e., one bad news would not mean that the project quality is necessarily low.

The paper is organized as follows. We next comment on our connection with the literature. In Section 2.4 we present the model and then present results in Section 1.3. Lastly we discuss extensions in Section 2.6. All proofs are relegated to the appendix.

<u>*Related Literature:*</u> On the problem of rewarding the agent for bad news, the literature has focused on incentivizing the agent to reveal bad news that he observes privately. For example, Levitt and Snyder (1997) show that rewarding for bad news may be optimal when the agent receives a private signal about the project quality. Hidir (2017) and Chade and Kovrijnykh (2016)) are examples of dynamic contracting problems where the agent has the freedom to disclose bad news. We complement this literature by showing that, even though both actions and signals are public, the inability of the parties to write contracts contingent on actions can deter the agent from searching for bad news. Note that the choice of specifically searching for bad news is absent in the above mentioned papers. Manso (2011)

shows in a two period setting with full commitment, that motivating an agent to innovate may require tolerating or even rewarding early bad news. Like the ones mentioned above, this model also does not allow for a technology to search for bad news.

Our model builds on the exponential bandit models of Keller, Rady, and Cripps (2005) and Keller and Rady (2015), which study good and bad news arms resepectively. Technically, the paper closest to ours is Che and Mierendorff (2016). They study a single agent decision problem (as opposed to a two player game we have) of experimentation where the agent has the choice to look for good news and bad news. In a related single agent decision problem, Damiano, Li, and Suen (2017) introduce an auxiliary learning process that allows for looking for both good and bad news while experimenting on a one arm bandit in lines of Keller et al. (2005).

Garfagnini (2011) and Guo (2016) also study a delegation game between a principal and an agent where the agent carries out experimentation. While the contracting and payoff environment differs, the key distinction is our focus on how the agent's incentives shape the dynamics when the choice of both good and bad news is available. This tradeoff is absent in both Garfagnini (2011) and Guo (2016). As an agency problem of collective experimentation, this paper also relates to Kuvalekar and Lipnowski (2018). However, the efforts there are ranked in the sense of Blackwell (1953) making the agent's choice, when not getting fired, straightforward—choose the least informative action. Since the good news and bad news sources are not ranked in the sense of Blackwell (1953), the dynamics are richer in our environment. Halac, Kartik, and Liu (2016), Bergemann and Hege (2005) and Hörner and Samuelson (2013) are other instances of contracting problems with delegated expeirmentation with moral hazard and (or) adverse selection.

Recently, the question of information acquisition in the presence of multiple information sources has been pursued among others by Che and Mierendorff (2016), Liang, Mu, and Syrgkanis (2017), Liang and Mu (2018), Fudenberg, Strack, and Strzalecki (2017), and Mayskaya (2017). In contrast, in this paper we explore information acquisition from multiple sources of information in a principal-agent setting where the incentives of the two parties differ.

1.2 Model

Players: There are two players, a principal (she) and an agent (he). Time t is continuous with an infinite horizon. The principal hires the agent to work on a project of unknown quality. The quality of the project is high, $\theta = 1$, or low, $\theta = 0$. At time 0 both players have a common prior on the underlying project quality: $\mathbb{E}_0\theta = p_0 \in (0, 1)$.

Actions: At each instant, the principal chooses whether to fire, $s = 0$ or not to fire the agent, $s = 1$. Firing is irreversible and ends the game.[2] Conditional on not firing, the agent divides a unit of effort between a good news technology and a bad news technology. The agent's allocation to the good news technology at time t is $a_t \in [0, 1]$, and $(1 - a_t)$ is the effort devoted to the bad news technology. The agent's allocation is *observable* to the principal but not contractible.

Information: The agent's allocation affects the arrival rate of two exponentially distributed signals (news). A realized good (bad) signal is denoted by G(B). The arrival rate of a G signal is $\lambda_g a_t \theta$, and that of a B signal is $\lambda_b(1-a_t)(1-\theta)$. Both signals are publicly observed. Since actions and outcomes are public, there is no private information: players have the same posterior belief about θ on or off-path. Also, notice that either signal, G or B, resolves all the uncertainty: the realization of G(B) gives both players the belief $p = 1 (p = 0)$. A G signal, apart from confirming that the project is of high quality, also provides the knowledge needed to implement the project. We denote by $y_t \in \{\phi, G, B\}$ the news at time t, where ϕ denotes no news.

Payoffs: At the beginning of the relationship, the principal commits to a reward structure which specifies a payment of R to the agent if a G signal arrives and F if a B signal arrives. When employed, the agent receives an exogenously specified fixed wage $w > 0$ from the principal.[3] The principal incurs a flow cost of $c > w$ which we interpret as the cost of carrying out experimentation and the wage paid to the agent. If G arrives, the game ends with the principal receiving $\mathit{\Gamma}$. If B arrives, the game ends with the principal receiving 0.

[2] Irreversible firing is not restrictive because in our equilibria, once the principal fires he will never hire again.

[3] This is w.l.o.g. in that we can allow the principal to choose any fixed wage as long as it is at least equal to w.

Both players discount future payoffs at rate r normalized to equal 1.[4]

The terminal payoffs are:

1. If principal fires the agent, both players receive 0.

2. If G obtains, the principal receives $\Gamma - R$ and the agent receives R.

3. If B obtains, the principal receives $-F$ and the agent receives F.

Letting τ denote the stochastic time at which either the agent is fired or conclusive news arrives, the agent's payoff is given by

$$u(p_0) = \mathbb{E}_{a,s}\left[\int_0^\tau e^{-u} w \mathrm{d}u + e^{-\tau}\left[\mathbb{1}_{y_\tau = G} R + \mathbb{1}_{y_\tau = B} F\right]\right],$$

and the principal's payoff is given by

$$v(p_0) = \mathbb{E}_{a,s}\left[\int_0^\tau e^{-u}(-c)\mathrm{d}u + e^{-\tau}\left[\mathbb{1}_{y_\tau = G}(\Gamma - R) + \mathbb{1}_{y_\tau = B}(-F)\right]\right].$$

By dividing both players' payoffs by w, we can set, without loss of generality, $w = 1$.[5] Lastly, we assume $R \geq 1$. That is, the amount the principal pays to the agent upon obtaining a G signal, is no less than the discounted value of the agent's wage. We interpret good news as the principal adopting the project and employing the agent to work on it. The agent should thus continue to receive at least the flow rents he receives during the experimentation stage.

1.2.1 Strategies and equilibrium

Let P_t be the posterior probability that $\theta = 1$ at time t conditional on the agent's allocation history and signal realizations. We restrict attention to Markov Perfect Equilibria (equilibria or MPE henceforth) using P_t as state variable.

A Markov strategy for the principal is a reward structure $(R, F) \in [1, \infty] \times [0, \infty]$ and a function $s : [0, 1] \to \{0, 1\}$ that specifies hiring ($s = 1$) and firing ($s = 0$) at each belief.

[4] This normalization amounts to merely calculating time in different units.

[5] By doing so, the agent's wage becomes 1, while his terminal payoff becomes S/w and F/w depending on the signal. For the principal, the flow cost is c/w and the terminal payoffs are $\Gamma/w - S/w$ and $-F/w$ depending on the signal.

A Markov strategy for the agent is a function $a : [0, 1] \to [0, 1]$ specifying an allocation at each belief.

The posterior belief P_t is a stochastic process that takes a value $1(0)$ for all $t > \tau$ such that $y_\tau = G(B)$. In the absence of a conclusive signal and when $a(\cdot)$ is continuous, P_t follows the law of motion given by, [6]

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = [(1 - a(P_t))\lambda_b - a(P_t)\lambda_g]P_t(1 - P_t). \tag{1.1}$$

We make the following assumptions on $a : [0, 1] \to [0, 1]$ to ensure that there exists a unique continuous function $P : [0, \infty) \to [0, 1]$ that satisfies (1.1) whenever $a(\cdot)$ is continuous.

Assumption 1. *1. The function $a(\cdot)$ is piecewise continuous.*[7]

2. Define

$$a^f = \frac{\lambda_b}{\lambda_b + \lambda_g}. \tag{1.2}$$

For any $\hat{p}$ where $a(\cdot)$ is discontinuous, if $\lim_{q \uparrow \hat{p}} a(q) \leq a^f$ and $\lim_{q \downarrow \hat{p}} a(q) \geq a^f$, then $a(\hat{p}) = a^f$.

Note that using (1.1) we can show that

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = 0,$$

when $a(P_t) = a^f$. That is, beliefs do not move in the absence of a conclusive signal if the agent allocates a^f to the good news technology. We call a^f as the freezing allocation and when agent chooses a^f at some belief p, we say that "the agent freezes beliefs at p".

Denote the space of Markov strategies for the agent by $\mathcal{A}$ and the space of hiring/firing Markov strategies of the principal by $\mathcal{S}$. Given $P_0 = p_0$ and $(a, s) \in \mathcal{A} \times \mathcal{S}$, define the induced stochastic process $\{P_t, A_t, S_t\}$ by setting $A_t = a(P_t), s_t = s(P_t)$ and letting $\{P_t\}_t$

[6] Since beliefs are a martingale, we have that $\lambda_g a(P_t)P_t\mathrm{d}t + (1 - [\lambda_g a(P_t)P_t + \lambda_b(1 - a(P_t))(1 - P_t)]\mathrm{d}t)(P_t + \dot{P}_t\mathrm{d}t) = P_t$. Dividing by $\mathrm{d}t$ we obtain (1.1).

[7] A piecewise continuous function is continuous except at a finite number of points in its domain.

follow (1.1).

The value function for the agent is

$$u(p|a,s,R,F) := \mathbb{E}_{a,s}\left[\int_0^\tau e^{-u} w \mathrm{d}u + e^{-\tau}\left[\mathbb{1}_{y_\tau=G} R + \mathbb{1}_{y_\tau=B} F\right] \Big| P_0 = p\right],$$

and for the principal is

$$v(p|a,s,R,F) := \mathbb{E}_{a,s}\left[\int_0^\tau e^{-u}(-c)\mathrm{d}u + e^{-\tau}\left[\mathbb{1}_{y_\tau=G}(\mathit{\Gamma} - R) + \mathbb{1}_{y_\tau=B}(-F)\right] \Big| P_0 = p\right].$$

Finally, we define the notion of equilibrium in our setting.

Definition 1. *An equilibrium is a collection* $(a,s,R,F) \in \mathcal{A} \times \mathcal{S} \times [1,\infty] \times [0,\infty]$ *such that:*

1. Agent optimality. *For each* $p \in (0,1)$,

$$a \in \operatorname{argmax}_{\hat{a}\in\mathcal{A}} u(p|\hat{a},s,R,F).$$

2. Principal optimality.

(a) Firing strategy s *is optimal at all beliefs* $p \in (0,1)$ *given* (R,F).

Given any (R,F), *for each* $p \in (0,1)$,

$$s \in \operatorname{argmax}_{\hat{s}\in\mathcal{S}} v(p|a,\hat{s},R,F).$$

(b) Define, $\mathcal{E}(R,F) := \{(a,s) : (a,s)$ *satisfies* 1 *and* 2a *for the given* $(R,F).\}$, *and* $v_*(p|R,F) := \sup_{(a,s)\in\mathcal{E}(R,F)} v(p|a,s,R,F)$.

The initial choice of (R,F) *must be optimal.*

$$(R,F) \in \operatorname{argmax}_{\hat{R},\hat{F}\in[1,\infty]\times[0,\infty]} v_*(p_0|\hat{R},\hat{F}).$$

We assume that the value the principal receives when a G signal arrives is sufficiently high relative to the cost of experimentation:

Assumption 2. $\dfrac{\lambda_b \lambda_g}{\lambda_b + \lambda_g}(\varGamma - 1) - c > 0.$

1.3 Results

Given a reward structure, we characterize the principal-optimal equilibrium as follows. First, we fix a stopping strategy for the principal and find the agent's best response. Thereafter, given the agent's best response we find the principal's optimal stopping strategy. For any reward structure, there is a unique principal-optimal equilibrium. Therefore, we compare the principal-optimal equilibria across various reward structures to obtain the optimal reward structure given the initial belief p_0.

1.3.1 Agent's best response

Suppose the principal hires on an interval $[\underline{p}, \bar{p}] \subseteq [0, 1]$, that is,

$$s(p) = \begin{cases} 1 & \text{if } p \in [\underline{p}, \bar{p}], \\ 0 & \text{if } p \notin [\underline{p}, \bar{p}]. \end{cases}$$

What would the agent do at each belief? He has two choices: look for good news or look for bad news. If he looks for bad news, the game ends with him receiving F if a B signal arrives, while the beliefs move up in the absence of news. If he looks for good news, the game ends with him receiving R if a G signal arrives, while the beliefs move down in the absence of news. The agent strictly prefers being employed over getting fired, while the principal wishes to hire the agent when sufficiently optimistic about project quality. If F is large enough (say $F = 1$) looking for bad news—even though costly to the principal, imposes no hazard for the agent. But what about when F is small, say $F = 0$ e.g.? In that case, looking for bad news imposes a hazard as a B signal would end the game with the agent receiving nothing. Alternatively, the agent can look for good news which entails a possibility of a reward R if a G signal is obtained, however, beliefs move down to the firing cutoff $\underline{p}$ in case no signal arrives. As Lemma 1 clarifies, the agent is heavily predisposed against looking for bad news when F is smaller than 1, the agent's wage. However, he makes use of the bad news technology by combining it with the good news technology to freeze beliefs at $\underline{p}$. The observability of the agent's allocation plays a critical role here. It allows the agent to continue employment by preventing the principal's beliefs from falling any further in the absence of a signal. Above $\underline{p}$, he only looks for good news. That is, the agent delays looking for bad news as much as possible. The remedy, should the principal want the agent to look for bad news, is to have $F \geq 1$. In that case, the agent looks for bad news below a cutoff belief p^f (defined below) and looks for good news above. As a consequence, if the beliefs reach p^f, they remain frozen there until the uncertainty is resolved. We discuss the intuition behind how p^f is calculated in the discussion following Lemma 1 which characterizes the best response of the agent. All the proofs are presented in the appendix.

$$p^f := \frac{\lambda_b(F-1)}{\lambda_b(F-1) + \lambda_g(R-1)}. \tag{1.3}$$

Lemma 1. *Suppose the principal hires the agent when $p \in [\underline{p}, \bar{p}] \subset [0,1]$ and fires otherwise. The best response of the agent, for any $p \in [\underline{p}, \bar{p}]$, is the following:*[8]

If $(R, F) \neq (1,1)$, and

1. *if $F < 1$, then*

$$a(p) = \begin{cases} a^f & \text{if } p = \underline{p}, \\ 1 & \text{if } p \in (\underline{p}, \bar{p}]; \end{cases}$$

2. *if $F \geq 1$ and $p^f \geq \bar{p}$, then*

$$a(p) = \begin{cases} a^f & \text{if } p = \bar{p}, \\ 0 & \text{if } p \in [\underline{p}, \bar{p}); \end{cases}$$

[8] Outside the hiring region $[\underline{p}, \bar{p}]$, the agent is indifferent across any allocation.

3. *if* $F \geq 1$ *and* $p^f \leq \underline{p}$*, then*

$$a(p) = \begin{cases} a^f & \text{if } p = \underline{p}, \\ 1 & \text{if } p \in (\underline{p}, \bar{p}]; \end{cases}$$

4. *if* $F \geq 1$ *and* $p^f \in (\underline{p}, \bar{p})$

$$a(p) = \begin{cases} 0 & \text{if } p \in (\underline{p}, p^f), \\ a^f & \text{if } p = p^f, \\ 1 & \text{if } p \in (p^f, \bar{p}]. \end{cases}$$

If $(R, F) = (1, 1)$*, the agent's best response is any* $a \in \mathcal{A}$ *such that* $a(\underline{p}) \leq a^f$ *and* $a(\bar{p}) \geq a^f$*.*

To gain some intuition assume the agent has only three choices at any belief: look for good news alone, bad news alone or freeze beliefs, and that the hiring interval is of the form $[\underline{p}, 1]$. Qualitatively, the agent's best response has two forms: one when $F < 1$ and another when $F \geq 1$.

We first consider $F < 1$. Suppose $R \geq 1$ and $F < 1$, and the hiring interval is of the form $[\underline{p}, 1]$. Notice that the agent can choose to freeze beliefs at any $p \in [\underline{p}, 1]$, giving us a lower bound on his value function. Starting at some $p_0 \geq \underline{p}$, consider the following two Markovian strategies:

1. Look for good news until beliefs reach $q \in (\underline{p}, p_0)$ in the absence of signal and freeze at q.

2. Look for good news until beliefs reach $q - \epsilon \in (\underline{p}, p_0)$ in the absence of signal and freeze at $q - \epsilon$, for a small $\epsilon > 0$.

It is easy to see that (2) performs strictly better than (1). If $\theta = 0$, a B signal will arrive after the agent switches to freezing and the agent receives 0 eventually with either policy. However, in (2), a B signal is delayed because the agent freezes beliefs later. If $\theta = 1$, the agent will receive 1 until a G signal arrives, in which case the agent receives $R \geq 1$. In (2), the agent spends strictly larger time looking for a G signal and therefore, is expected to receive R earlier. Therefore, regardless of the state, the agent does better.

The above argument suggests that if the agent switches from good news to freezing at some point, he would postpone it as much as possible. Therefore, freezing at any belief except at $\underline{p}$ cannot be optimal. At $\underline{p}$, the agent will not choose an allocation $a(\underline{p}) > a^f$ as, in the absence of a signal, the beliefs drift downward yielding a continuation payoff of 0. Therefore, the agent must either freeze beliefs at $\underline{p}$, or look for bad news at $\underline{p}$, which has an identical effect of freezing beliefs at $\underline{p}$, since the agent switches back to using the good news arm once beliefs are higher than $\underline{p}$. Also, It is easy to see that looking for bad news forever is worse than freezing beliefs for analogous reasoning as in the previous discussion.

Therefore, if the agent does in fact use $a = 0$ (look for bad news) then, he must eventually shift to looking for good news at some belief $p \in (\underline{p}, 1)$. However, at such a switching belief p—a belief such that to its left the agent looks for bad news and to its right the agent looks for good news—the beliefs remain frozen conditional on reaching there. As argued previously, freezing at $p \in (\underline{p}, 1)$ is strictly suboptimal. Therefore, we have a candidate for the optimal policy when $R \geq 1$ and $F < 1$—Look for good news on $(\underline{p}, 1)$ and freeze at $\underline{p}$. Its optimality is then established using the usual verification arguments, more importantly, by not imposing the restriction that $a(p) \in \{0, a^f, 1\}$.

The only way to incentivize the agent to look for bad news is by offering a reward $F \geq 1$. To this end, suppose $(R, F) \neq (1, 1)$ and $F \geq 1$. Notice that the agent can guarantee himself a payoff of at least 1 by freezing beliefs. Therefore, we could focus on the excess payoff the agent receives over 1. If G obtains, the excess payoff is $R - 1$ and for B it is $F - 1$. At some belief p, by looking for good news for a small time $\mathrm{d}t$, the agent's expected myopic payoff is $\lambda_g p(R - 1)\mathrm{d}t$, which is increasing in p. Similarly, the expected myopic payoff by looking for bad news is $\lambda_b(1 - p)(F - 1)\mathrm{d}t$, which is decreasing in p. At p^f, the switching belief such that the agent looks for bad news to its left and good news to its right, the two expected myopic payoffs are equal. We would like to emphasize though, that reasoning based on the myopic payoff comparison for the two kind of news is illustrative but incomplete. Dynamic considerations should play a role in deciding what news to look for. By looking for bad news, the beliefs move upwards and good news drives them downwards, and in amounts proportional to λ_b and λ_g. Therefore, the agent's choice is an outcome of both the myopic payoff comparisons and the curvature of the optimal value function that determines the spread of continuation values.

1.3.2 Principal's problem

In light of Lemma 1, we first argue that we have two cases to consider insofar as the principal's optimal choice of reward structure is concerned: $F = 0$ and $F \geq 1$. To see this, note that for any $F \in [0, 1)$, the agent's behavior is unchanged while the costs increase in F for the principal. Therefore, the principal would either set $F \geq 1$ to induce the agent to look for bad news on an interval or would set $F = 0$. Also, it is easy to see that in the principal's optimal equilibrium, the higher endpoint of the hiring interval $\bar{p}$, is equal to 1. Therefore, in the principal's optimal equilibrium, the principal's strategy is simple: fire if $p < \underline{p}$, and hire otherwise.

Suppose, in the principal's optimal equilibrium $R > 1$ and $F > 1$. Lemma 11 in the appendix shows that we could then lower R and F to keep the agent behavior unchanged while increasing the principal's payoffs. The logic is straightforward. If $(R, F) >> (1, 1)$, the agent looks for bad news when beliefs are below p^f and good news when the beliefs are above p^f, where p^f is given by (1.3). It is clear that for any $(R, F) >> (1, 1)$, we can choose (R', F') such that $1 < R' < R, 1 < F' < F$ to obtain the same p^f. Therefore $(R, F) > (1, 1)$ cannot be optimal for the principal. Lastly, it is easy to see that $R = 1, F > 1$ or $R > 1, F = 1$ is suboptimal for similar reasons: lowering F or R (whichever is larger than 1) keeps p^f the same (0 or 1) and so keeps the agent's behavior unchanged. Therefore, we have either $R = F = 1$ or $R = 1$ and $F = 0$. As a result, we can reduce the principal's problem of finding the optimal reward structure to simply choosing between $F = 1$ or $F = 0$. That is, the principal either does not reward bad news at all or rewards it the same as the good news. The above discussion motivates the following proposition whose detailed proof can be found in the appendix.

Proposition 1. ***Reduction of the Principal's Problem.*** *In a principal-optimal equilibrium, $R = 1$ and $F \in \{0, 1\}$.*

Finally we compare the principal's optimal values for two cases: $(R, F) = (1, 0)$ and $(R, F) = (1, 1)$. Notice that in case $(R, F) = (1, 1)$, the agent's best response is not unique.

In fact, any allocation policy of the agent that results in a non-negative drift of beliefs at the cutoff belief below which the principal fires i.e. $a(\underline{p}) \leq a^f$, is a best response for the agent. Therefore, we are left with the following questions:

1. Given the agent's indifference when $(R, F) = (1, 1)$, what is the principal's preferred behavior for the agent?

2. What is the optimal $\underline{p}$ when $(R, F) = (1, 0)$ and when $(R, F) = (1, 1)$?

3. What is the optimal reward structure for the principal? In particular, is it ever optimal to have $F = 1$? Conversely, is it ever optimal to have $F = 0$?

We answer (1) in the appendix, Section A.7. For an exogenously specified firing cutoff $\underline{p}$, we show that the principal's preferred behavior for the agent under the constraint that $a(\underline{p}) \leq a^f$ is simple: if $\underline{p} < p^s$ (defined in (1.4) below), look for bad news when $p < p^s$, look for good news when $p > p^s$ and freeze beliefs at p^s. If $\underline{p} \geq p^s$, look for good news when $p > \underline{p}$, and freeze beliefs at $\underline{p}$.

$$p^s = \frac{\lambda_b(c - F)}{\lambda_g(\mathit{\Gamma} - R + c) + \lambda_b(c - F)} = \frac{\lambda_b(c - 1)}{\lambda_g(\mathit{\Gamma} - 1 + c) + \lambda_b(c - 1)}. \tag{1.4}$$

In light of the reduction of the principal's problem, we are left with the following two candidate agent strategies for the principal hiring region of the form $[\underline{p}, 1]$. a_0^* is the strategy when $F = 0$ while a_1^* is the strategy when $F = 1$.

$$a_0^*(p) = \begin{cases} 1 & \text{if } p > \underline{p}, \\ a^f & \text{if } p = \underline{p}, \\ 0 & \text{if } p < \underline{p}. \end{cases} \tag{1.5}$$

$$a_1^*(p) = \begin{cases} 1 & \text{if } p > \max\{p^s, \underline{p}\}, \\ a^f & \text{if } p = \max\{p^s, \underline{p}\}, \\ 0 & \text{if } p < \max\{p^s, \underline{p}\}. \end{cases} \tag{1.6}$$

Now we turn to answer (2). For a given reward structure and the agent behavior, the principal chooses an optimal stopping belief $\underline{p}$. Let the optimal stopping beliefs for $F = 0$ and $F = 1$ be p_0^* and p_1^* respectively.[9] A natural candidate for p_0^* is the belief at which the principal's value is 0, given that the agent freezes beliefs at that belief. Lemma 14 in

[9]More details about these beliefs can be found in the appendix, (A.12) and (A.23).

Section A.6 shows that indeed such a belief is the optimal stopping belief. In the case when when $F = 1$ and $\underline{p} \geq p^s$, note that the agent's behavior is identical to the case when $F = 0$ for a given stopping belief. Hence, in this case p_1^* is calculated in an identical way as above. In the case when $F = 1$ and $\underline{p} < p^s$, the optimal stopping belief p_1^* is set to the belief at which the principal's value is 0, given that the agent is following the strategy given by (1.5).

We finally answer (3). Let the associated optimal value functions for the principal be given by $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ for the case $F = 0$ and $F = 1$ respectively. The explicit expressions are given in the appendix, (A.14) and (A.24) respectively.

Insofar as the principal's optimal reward structure is concerned, what remains now is to compare $v_*^{F=0}(p_0)$ and $v_*^{F=1}(p_0)$, where p_0 is the prior at time 0. We formally state this comparison in the proposition below. Let the principal's optimal value function in the game with $P_0 = p_0$ (recall that P_t is the belief at time t) be denoted by $v^{p_0}(\cdot)$.

Proposition 2. *For any initial prior p_0, the principal's optimal value function, $v^{p_0}(p)$ is the following:*

$$v^{p_0}(\cdot) = \begin{cases} v_*^{F=0}(\cdot) & \text{if } v_*^{F=0}(p_0) \geq v_*^{F=1}(p_0), \\ v_*^{F=1}(\cdot) & \text{if } v_*^{F=1}(p_0) > v_*^{F=0}(p_0). \end{cases}$$

Proposition 2 answers when to reward bad news by setting $F = 1$. The principal compares the two value function for $F = 0$ and $F = 1$ at time 0 and chooses the maximum. However, it does not answer the question of whether $F = 1$ ever obtains and, conversely, is $F = 0$ ever optimal? The following proposition answers these question. Before we state the proposition, define, $\Lambda := \frac{\lambda_b \lambda_g}{\lambda_b + \lambda_g}$.

Proposition 3. *Suppose*

$$c > (1 + \lambda_g). \tag{1.7}$$

There exists $\underline{\lambda}_b$ such that for all initial beliefs, $\lambda_b > \underline{\lambda}_b$ implies that $F = 1$ is optimal, and *in particular, strictly optimal when $v_*^{F=1}(\cdot) > 0$. On the other hand, if*

$$\frac{c+\Lambda}{\Lambda\Gamma} > \frac{\lambda_b(c-1)}{\lambda_g(\Gamma-1+c)+\lambda_b(c-1)} \ , \tag{1.8}$$

then $F = 0$ is optimal for all initial beliefs, and in particular, strictly optimal when $v_*^{F=0}(\cdot) > 0$.

The sufficient condition for optimality of $F = 1$ is derived by simply asking, when is it the case that $p_1^* < p_0^*$? In that case, since the stopping cutoff with $F = 1$ is strictly lower than the cutoff for $F = 0$, it is at least optimal to have $F = 1$ when $p \in (p_1^*, p_0^*)$. However, it turns out that whenever $p_1^* < p_0^*$, $v_*^{F=1}(p) > v_*^{F=0}(p)$ for all $p > p_1^*$. That is, it is optimal to set $F = 0$ for all prior beliefs when the principal does not fire the agent right away.

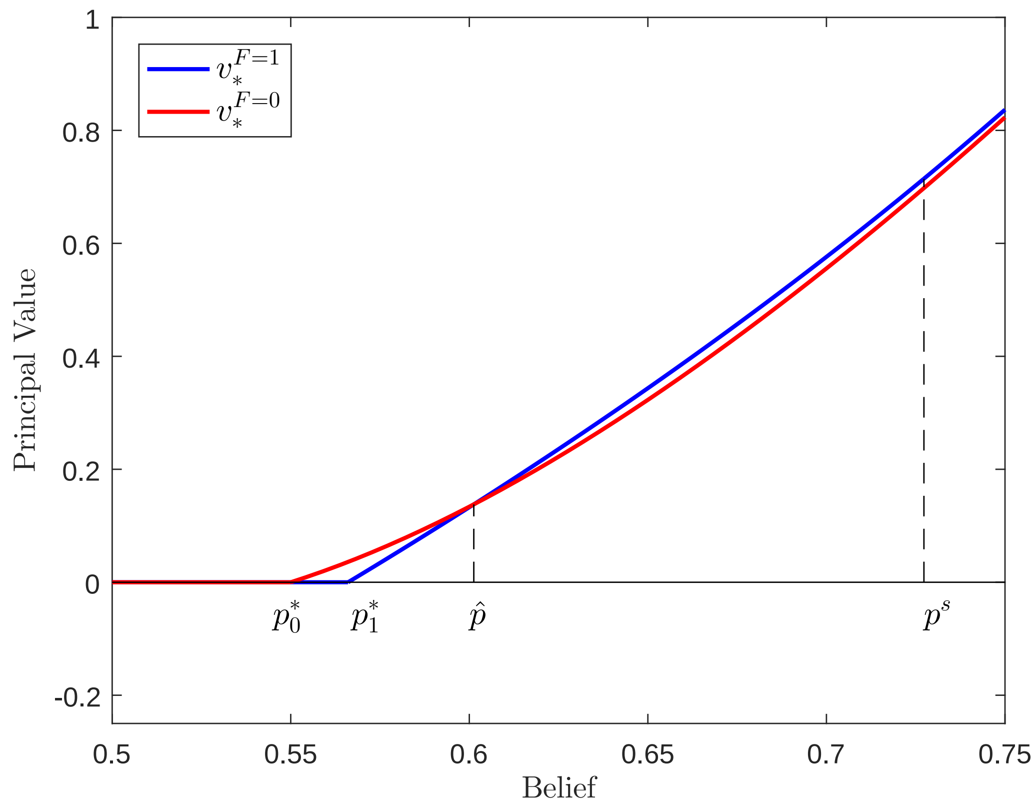
The sufficient condition for optimality of $F = 0$ is straightforward. Consider the case when $F = 1$ and set $\underline{p} < p^s$. Look at the principal's value function when the agent looks for bad news below p^s and good news above it. If this value is negative at p^s, then it will not be optimal for the principal to hire at p^s. In that case, we will have $p_1^* > p^s$, and the agent behavior (on path) would be to look for good news at all beliefs above p_1^*, and freeze beliefs at p_1^*. This behavior, qualitatively, is identical to the agent behavior with $F = 0$. Therefore, since the principal's costs are higher when $F = 1$, we will have $p_1^* > p_0^*$, and $v_*^{F=0}(p) > v_*^{F=1}(p)$ for all $p > p_0^*$.

While our sufficient conditions establish that we can have either $F = 0$ or $F = 1$ as the optimal reward for all initial priors, what happens when the two conditions are violated? It will still be the case that optimally $F \in \{0, 1\}$, but the answer will depend on the prior belief p_0 as well. As Figure 1.1 shows, it is possible to have $p_1^* > p_0^*$ and yet, $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ not being globally ranked. In fact, $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ cross at most once. Moreover, if they cross, they cross in a way that $v_*^{F=1}(\cdot)$ dominates $v_*^{F=0}(\cdot)$ above a certain $\hat{p}$, and the other way below it. That is, whenever it is optimal for the principal to reward bad news for some prior p_0, it is optimal to reward bad news for all priors larger than p_0. On the other hand, whenever it is optimal to not reward bad news for some prior p_0, it is optimal to do so for all the lower priors. This observation is summarized in Proposition 4 below. The proofs can be found in Lemma 31 and the lemmata before it in the appendix.

Proposition 4. *If $v_*^{F=0}(p_1^*) > 0$ and $v_*^{F=0}(p^s) < v_*^{F=1}(p^s)$ then $\exists\ \hat{p}$ such that, $v_*^{F=1}(p) > v_*^{F=0}(p)$ whenever $p \in (\hat{p}, 1)$ and $v_*^{F=1}(p) < v_*^{F=0}(p)$ whenever $p \in (p_0^*, \hat{p})$.*

The intuition behind this, perhaps counter-intuitive reward structure relates to the observation we made earlier: rewarding the agent for producing bad news is costly. If initial prior is high, then the expected revenue from the project is also high which implies that the principal is willing to bear the cost of rewarding for producing bad news, for she expects that to be a less likely outcome.

FIGURE 1.1. The case when value functions cross.



1.3.3 When and how to reward bad news

Observations from Propositions 2, 3, and 4, lead to two important takeaways. First, it may indeed be optimal to reward the bearer of bad news. But always rewarding for bad news is not necessarily optimal. Second, the answer can also depend on the initial prior. If the bad news technolgy is sufficiently informative ($\lambda_b > \underline{\lambda_b}$) and if experimentation is sufficiently costly (1.7), then it is optimal to reward bad news. This may explain why, in the technology sector, there is a growing push towards rewarding reporting bad news, since it is relatively easier to find bugs in softwares.[10] At the same time, if the good news technology is extremely informative ($\lambda_g \to \infty$ e.g.) then it is not optimal to provide incentives to look for bad news (Inequality (1.8) holds). Some explanations suggest that the employees fear the negative consequences of being the bearer of bad news. Hence, bad news is not transmitted efficiently to the management.[11] We view our explanation as an alternate one—it is perhaps not fear but rather the reward structure that disincentivizes the employees from acquiring such information. That is, it is not that the employees hide negative information, but rather that they choose not to acquire it. In a software company, this would mean that employees do not look for bugs in their products in the absence of adequate incentives. Widely adopted use of the "bug bounty programs" in the recent times is consistent with this explanation.

1.4 Beyond Markov Perfect Equilibria

The following strategy profile (σ^*) is an MPE: The agent always looks for bad news and the principal always fires. In σ^*, both players are simultaneously min-maxed, and therefore, σ^* is the worst threat to both players. Using σ^* as punishment, intuitively, we can implement any behavior from either players in a non-MPE, in particular the first best.[12] To see this, suppose the principal wants to implement any allocation $\tilde{a}(t)$ starting at time 0. The grim-trigger strategy profile where the principal hires as long as the agent follows the allocation policy $\tilde{a}(t)$ and reverts to σ^* following any deviation from the agent is an equilibrium. The agent is willing to follow $\tilde{a}(t)$ because it guarantees him wages and any deviation leads to firing.

This grim-trigger equilibrium relies heavily on the presence of a severe threat. We

[10]For example, Keil and Mähring (2010) and Tan, Smith, Keil, and Montealegre (2003) document the importance of rewarding bad news in project management in the technology sector.

[11]As documented by Smith and Keil (2003), and to quote Barry M. Staw and Jerry Ross, *"Because no one wants to be the conveyor of bad news, information is filtered as it goes up the hierarchy."* https://hbr.org/1987/03/knowing-when-to-pull-the-plug

[12]Formalizing this discussion involves handling the well known issues in continuous time games described in Simon and Stinchcombe (1989) and Bergin and MacLeod (1993). For our purposes, an illustration suffices.

impose a simple requirement of weak renegotiation-proofness due to Farrell and Maskin (1989), adapted to dynamic games by Bergemann and Hege (2005).[13]

Definition 2. *In the subgame after choosing a reward structure* (R, F)*, a subgame perfect equilibrium* $\{a, s\}$ *is* weakly renegotiation-proof *if there do not exist continuation equilibria at some* h_t *and* $h_{t'}$ *with* $P(h_t) = P(h_{t'})$ *and* $h_t \neq h_{t'}$ *such that* $u(h_t) \geq u(h_{t'})$ *and* $v(h_t) \geq v(h_{t'})$ *with at least one strict inequality.*

The above definition, viewed as an internal consistency requirement, requires that after any two histories such that the beliefs are the same after the two histories, the continuation play must not be Pareto ranked. It is easy to see that the grim-trigger equilibrium is not weakly renegotiation-proof. On the other hand, any MPE is weakly renegotiation-proof because, the continuation play is the same after any two histories such that the beliefs are the same. This straightforward observation is summarized below.

Proposition 5. *All MPEs are weakly renegotiation proof.*

Since we are interested in the principal-optimal equilibria, the question remaining is whether there are weakly renegotiation-proof equilibria that deliver higher payoff to the principal compared to the principal-optimal MPE. We are not aware of any such equilibrium and leave this issue to be resolved in future research.

1.5 Extensions

1.5.1 Unobservable allocation choice

A key feature of our model is that the agent's allocation is observable to the principal. However, it is natural to explore what would happen if the agent's actions were not observable to the principal. We show that main insights of our model remain unchanged in this environment.

We modify the model presented in Section 2.4 by assuming that the allocation choice of the agent is not observed by the principal, however any signal (good news or bad news) is

[13]Formal details on defining strategies and histories in our setting can be provided if needed.

observed by both parties. Because allocation choice is privately observed by the agent, the principal's strategy cannot depend on the belief about the state which is known only to the agent once the game commences. Hence the notion of MPE in beliefs is not applicable in this setting. The solution concept in this case is Perfect Bayesian Equilibrium.

Given a reward structure $(R.F)$, a pure strategy for the principal is a stopping time $T \in [0, \infty]$ such that the principal hires the agent when $t < T$ and fires the agent when $t \geq T$, in the absence of a conclusive signal. The agent's history at any point where no conclusive news has arrived is $h^t := (a_t)_t$ where a_t is the allocation at time t. Any private history maps into a belief $\widetilde{P}_t$ for the agent which evolves according to the law of motion given by (1.1). The set strategies for the agent is the set of functions $a : [0,1] \times [0, \infty) \to [0,1]$ which specifies an allocation at time t when belief is $\widetilde{P}_t$. Given a reward structure (R, F), an equilibrium is a tuple (T, a), such that each player best responds to the other.

We first consider the case when $R = 1$ and $F = 0$. In this case, the agent has no incentive to look for bad news. This is because a bad news leads to termination without any reward to the agent. Given that the agent looks for good news exclusively, the principal's optimal behavior is simply to fire the agent when beliefs drift down sufficiently in the absence of a signal. Suppose that the initial belief is p_0, and define

$$\hat{p}_0 = \frac{c}{\lambda_g(\mathit{\Gamma} - 1)}, \tag{1.9}$$

then the principal-optimal equilibrium is established by the following lemma.

Lemma 2. *When $R = 1, F = 0$, the principal-optimal equilibrium is given by (τ^*, a^*) such that*

1. *Agent's allocation: $a^*(p, t) = 1$ for all $t \leq \tau^*$*
2. *If $p_0 \leq \hat{p}_0$, then $\tau^* = 0$.*
3. *If $p_0 > \hat{p}_0$, then $\tau^* = \inf\{t : P_t^a = \hat{p}_0\}$, where P_t^a denotes the Principal's posterior probability that $\theta = 1$ calculated assuming $a_u = 1$ for all $u \leq t$.*

When $F < 1$, given any stopping time T of the principal, the agent prefers to delay the search for bad news as much as he can. The intuition is the following. Suppose the principal hires follows a finite stopping time policy, i.e. fires the agent if no signal arrives by time T. Take a strategy $(a_t)_t$ such that the agent devotes $T_g := \int_0^T a_t dt$ time to look for good news and the remaining for bad news. Now, define another strategy $(\hat{a}_t)_t$ where, the agent searches sets $\hat{a}_t = 1$ when $t \leq T_g$ and sets $\hat{a}_t = 0$ thereafter. Note that if $\theta = 1$, the payoff of the agent is the same under both strategies since a B signal never arrives. However, if $\theta = 0$, bad news arrives earlier in expectation under a compared to $\hat{a}$. Since, $F < 1$, the agent prefers $\hat{a}$ to a since he can collect a flow wage of 1 for longer in expectation. Therefore, in any best response, the agent will search for good news up to some time T_1 and may search for bad news thereafter. Note that the principal has a profitable deviation in case the agent searches for bad news. She can simply lower her stopping time to T_1 and be better off since she knows that after T_1, the agent can only produce a B signal that leads to the abandonment of the project and she can save the cost of experimentation by abandoning the project herself. Notice that given this equilibrium behavior of the agent, the principal is might as well set $R = 1$ and $F = 0$. Hence $R = 1$ and $F \in (0, 1)$ does not improve upon $R = 1$ and $F \in (0, 1)$. The above discussion is summarized the lemma below.

Lemma 3. *In the principal-optimal equilibrium, either $F = 0$ or $F \geq 1$.*

Next, we look at the case when $R = F = 1$. We ask the following question: If the principal could choose a policy for the agent ignoring the incentive constraints of the agent, what policy would she choose? Lemma 33 shows us that there are three possibilities.

1. G policy: Search for good news when $p \in [\hat{p}_0, 1]$.
2. $G - B - G$ policy: There exists a cutoff $\tilde{p}$ with $\hat{p}_0 < \tilde{p} < p^s$ such that
    - Search for good news when $p \in [\hat{p}_0, \tilde{p}] \cup [p^s, 1]$.
    - Search for bad news when $p \in (\tilde{p}, p^s)$.
3. $B - G$ policy:
    - Search for good news when $p \in [p^s, 1]$.

- Search for bad news when $p \in [p_1^*, p^s)$.

Note that desired behavior of the agent in case (1) can be implemented by setting $R = 1, F = 0$ as we have shown in preceding discussion. In case (2), when initial prior $p_0 < \hat{p}$ the desired agent behavior can be implemented similarly. When in case (2) with $p_0 > \hat{p}$ and case(3), the principal can implement the desired agent behavior as follows. Set the stopping time $T = \infty$, i.e. never fire the agent. The agent now is indifferent between any policy and in particular is willing to follow the policy desired by the principal. We have shown that we can implement principal's optimal policy when $R = F = 1$, which implies that the principal cannot do any better by setting higher rewards. The above discussion is summarized in the proposition below.

Proposition 6. *When allocation is unobservable, the optimal reward structure is either* $R = F = 1$ *or* $R = 1, F = 0$.

The question on when to reward bad news also carries over from the case when allocation is observable. In particular, the sufficient conditions shown in proposition 3 hold in the case of unobserved allocation as well. When the principal does not reward bad news, observe that the (implied) belief at which the principal fires the agent is $\hat{p}_0$, which is strictly lower than p_0^*, the firing belief in the optimal MPE when $F = 0$. A lower cutoff belief also results in the principal attaining a strictly higher value when the agent's actions are not observed. When the principal optimally chooses $F = 1$, the value of the principal is identical under both unobserved actions and MPE, since the behavior of the agent is identical on-path. We summarize this in the proposition below, letting $v_*^{obs}(\cdot)$ $(v_*^{unobs}(\cdot))$ stand for the principal-optimal value function when the agent's action is observable (not observable).

Proposition 7. $v_*^{uobs}(p) \geq v_*^{obs}(p)$ *for all* p*. Moreover, the inequality is strict whenever* $v_*^{unobs}(p) > 0$ *if* $F = 0$ *is optimal for the principal when the agent's action is observable to her.*

1.5.2 Signals not fully revealing

Throughout the paper, we have assumed that both the good news and the bad news arm are fully revealing. We now argue that even without fully revealing arms, our main result—

either do not reward the bad news at all or reward it as much as the good news—is preserved. We continue to maintain that the good news arm is fully revealing, i.e. the rate of arrival conditional on $a = 1$, are $\lambda_g \theta$. This is in line with our interpretation that good news apart from revealing that the state is high, results in a direct payoff i.e. the project can be implemented. However, we now assume that the bad news arm is not fully revealing. In particular, the rate of arrival of a B signal when $\theta = 1$ is $\lambda_b^1 < \lambda_b^0$, the rate of arrival for the bad signal when $\theta = 0$. As before, starting at a prior p, a G signal takes the posterior to 1, while a B signal takes the posterior to $p' \in (0, p)$. In particular, the law of motion is:

$$\mathrm{d}P_t = [(\lambda_b^0 - \lambda_b^1)(1 - a_t) - \lambda_g a_t] P_t (1 - P_t) \mathrm{d}t.$$

A natural analog of our (R, F) in this environment would be an amount R upon a conclusive G signal, while an amount F upon a B signal after which the principal terminates the relationship. An important observation is the following: Even in this environment, the agent can ensure that he does not get fired without a B signal. The idea is very intuitive. Suppose there is a cutoff belief $\underline{p}$ below which the principal fires the agent. As before, the agent can set $a = \frac{\lambda_b^0 - \lambda_b^1}{\lambda_g + \lambda_b^0 - \lambda_b^1}$ in order to ensure that the beliefs do not move from $\underline{p}$ in the absence of a signal. Therefore, for any $F \in [0, 1)$, the agent faces the same problem: delay the time at which the beliefs move below $\underline{p}$ as much as possible. This is because the agent receives his flow wage while employed, and can receive more if a G signal arrives, while his outside option is strictly inferior. Therefore, if the principal wishes to set $F \in (0, 1)$, she might as well set $F = 0$.

To see why, if $F \geq 1$, it is optimal to have $R, F = 1$, notice that the agent's payoff is the following.

$$v(p) = \sup_{a \in \mathcal{A}} \mathbb{E}^a \left[(1 - e^{-\tau}) 1 + e^{-\tau} [\mathbb{1}_{y_\tau = G} R + \mathbb{1}_{y_\tau = B} F] \right],$$

where τ is the smallest time such that either G obtains or where the posterior belief upon a B signal goes into the firing region. Subtracting 1 from either side, and dividing by $F - 1$, the agent's problem now, depends only on the ratio $\frac{R-1}{F-1}$, and its comparison to 1.

Therefore, for any $(R, F) >> (1, 1)$, we can choose a lower (R', F') as in Proposition 1, to keep the agent behavior unchanged while improving the principal profits. Therefore, if $F \geq 1$, it must be optimal to have $R = F = 1$.

1.6 Conclusion

In this paper, we studied a simple model of a principal-agent relationship with experimentation and limits to contractibility. The main focus of the paper was to determine whether and when the principal should reward the agent for bearing bad news, and how the optimal reward scheme should be structured. Our main takeaway is that either the principal should offer no reward to the agent for bearing bad news, or she should offer the same reward regardless of the type of news, good or bad. Given that rewarding bad news is costly, the sole reason for offering such a reward is to incentivize the agent to search for bad news, thereby potentially saving future experimentation costs. Prior to this paper, most research that prescribed rewarding bad news has focused on providing incentives to the agent to disclose bad news. In contrast, we show that even when such concerns are absent, i.e. the news is public, a fundamental source of conflict arises due to the agent's aversion to searching for bad news because its arrival triggers his termination. In addition, we also show that despite the simplicity of our framework, the above message also holds if the agent's action is not observed by the principal.

A key feature of our model—viewing experimentation as acquiring information from multiple sources—brings out novel dynamics. Our model predicts that rewarding for bad news may be more common in experimentation environments where the informativeness of the bad news source is high. Our results may also provide an alternative explanation to why bad news is not transmitted efficiently to management in organizations—it is not that the employees hide negative information, but rather that they choose not to acquire it when there is no reward for finding negative information.

Chapter 2

Supervising to Motivate

2.1 Introduction

In most organizations, supervisors have a role beyond the passive job of writing contracts with employees and disseminating the terms of the contract. They also play an active role in production, investing resources to augment the quality or quantity of the output produced. When working jointly with an employee on a project of uncertain quality, if a supervisor has private information (such as past output from the project), her investments in the project may transmit her private information to the employee.[1] This may not be desirable — it may be costly to induce effort from an employee who learns that the project is unlikely to produce output. In such situations, a supervisor faces the trade-off between investing optimally in the project based on her private information and transmitting information about quality of the project to the employee.[2] How should the supervisor manage this trade-off over time? In particular, should the supervisor's investment be sensitive to her private information? If yes, when and how?

I pursue these questions in a dynamic principal-agent model in which the principal learns privately about a project of uncertain quality (good or bad) over time. The principal invests costly resources into the project. The agent learns about the quality of the project through the investments made by the principal and exerts costly effort. The agent's mo-

[1] Models that allow private information with superviosrs have been studied in the literature on subjective and private evaluations, for example, MacLeod (2003) and Fuchs (2007).

[2] Here, optimal investment is the investment that maximizes the supervisor's value in the absence of information transmission concerns.

tivation (willingness to exert effort) increases in the principal's investment and the agent's belief (probability that project quality is good). The principal's willingness to invest increases with her belief. I solve for the principal's optimal equilibrium and find three main insights. First, if the relationship is in a optimistic stage (both players have high beliefs), the principal ignores information transmission concerns and invests optimally, i.e. the principal's investment is sensitive to her private information. Second, if the relationship moves to a pessimistic stage, or starts at a pessimistic stage, the principal ignores her private information when investing — to prevent any further information transmission and thereby preserving the motivation of the agent to exert effort. Third, if the relationship starts at a pessimistic stage, a principal who learns that the project is good, may still be forced to invest sub-optimally. However, such distortions are transient and disappear eventually.

More specifically, I study a continuous time principal-agent relationship where the principal seeks to induce effort from the agent on a project of unknown quality, either good or bad. Initially, both players have a common prior μ_0 about the quality of the project being good. At each instant, the principal chooses how much to invest (x_t) in the project at a cost that is convex in investment. In the same instant, but after observing the principal's investment, the agent chooses whether to exert effort at a cost or not. Not exerting effort is costless. The agent's effort is observable, but not contractible. Conditional on the quality of the project being good and the agent exerting effort, lump-sum output worth $1 + x_t$ arrives at the exponential rate of λ. Output is privately observed by the principal. The principal's investment does not affect the arrival rate of output, however it affects the value of the output if output arrives. That is, the principal can augment the output by investing, but cannot produce an output in the absence of effort from the agent. There is no output if either the project quality is bad or the agent does not exert effort. At the beginning of the relationship, both players agree to a non-negotiable sharing rule of output where the principal receives a share γ and the agent receives a share $1 - \gamma$ of the output. Both players discount future payoffs at the rate r.[3] The outside option of both players are normalized to zero. There are no transfers.

[3] Since the agent does not observe output during the course of the relationship, I interpret the rate of discounting r as the exogenous rate at which the game ends resulting in the agent realizing his share of the accumulated output.

At any point in the relationship the principal is one of two types. The "high" type, or type H principal is one who has observed an output and knows that the project is of good quality. The "low" type, or type L principal is one who has not observed an output. Since the investment cost of the principal is convex, in the absence of information transmission concerns, the principal's optimal investment is increasing in her belief that the project is of good quality. I study Markov Perfect Bayesian Equilibria, henceforth equilibria, of this game. The relevant state variable is the pair of beliefs (μ, μ^a) where μ is the belief of the type L principal that the project is of good quality and μ^a is the belief of the agent that she faces the type H principal.

The key question I ask is the following: How does the principal manage the trade-off between her current value and future value by controlling the information transmitted to the agent? I find that the answer to this question is critically dependent on the cost of effort of the agent, which captures the severity of misalignment of incentives of the players.

When the cost of effort is low (below some c^*), I show that the principal optimal equilibrium is fully separating at every belief of the type L principal (Proposition 9). That is, there is full information transmission to the agent. The type H principal always chooses her optimal investment and the type L principal chooses an investment strictly lower than the type H principal's optimal investment. At intermediate beliefs type L principal invests above her optimal level to induce effort from the agent and when beliefs are sufficiently low she quits the relationship. The misalignment in the incentives of players results in the type L principal investing above her optimal investment to motivate the agent to exert effort at intermediate beliefs. However, the misalignment is not sufficiently high for it to be optimal to stop the transmission of information about project quality to the agent in equilibrium.

When the cost of effort for the agent is high (above c^*), the misalignment in incentives of the players is enough that in the principal optimal equilibrium, in addition to investing above her optimal investment, the type L principal also stops the flow of information to the agent when her beliefs are sufficiently low. In this case, the nature of the principal optimal equilibrium is sensitive to the initial prior μ_0. When μ_0 is high (above some $\bar{\mu}$), the principal optimal equilibrium has three regions of interest in the space of type L principal's beliefs (Proposition 11). When beliefs are high, there is separation and transmission of information to the agent. However, once beliefs reach a cut-off ($\bar{\mu}$), the pooling phase begins. Both types of principal pool on the type H principal's optimal investment until beliefs reach another cutoff $\underline{\mu}$ at which point the type L principal quits the relationship. By stopping the flow of information to the agent at $\bar{\mu}$, the type L principal exploits the agent's uncertainty in the pooling phase and is able to reduce her investment while inducing effort from the agent.

When μ_0 is intermediate ($\mu_0 \in [\mu_g, \bar{\mu}]$), the separation region does not exist (Proposition 12). Both type of principals start with pooling until beliefs reach $\underline{\mu}$ at which point type L principal quits. Although qualitatively similar to the case with high initial prior, the nature of investment during pooling is different in this case. When starting with high initial prior, the type L principal optimally chooses when to start pooling ($\bar{\mu}$). However, when initial belief is lower than $\bar{\mu}$, such a luxury does not exist. Since the relationship begins with the agent being more pessimistic ($\mu_0 < \bar{\mu}$) the average investment needed during the pooling phase to induce effort from the agent is strictly higher compared to the case when $\mu_0 > \bar{\mu}$. This implies in this case the average pooling investment is strictly higher than type H principal's optimal investment. When $\mu_0 < \mu_g$, the average pooling investment needed is high enough that the relationship does not even start.

This leads to an important observation. When the cost of effort is high and the relationship starts at intermediate beliefs, type H principal invests sub-optimally during the pooling phase. This results from the inability of the type H principal to separate herself from type L principal. The sub-optimality of type H principal's investment is not perpetual. Once the type L principal quits the relationship, type H principal is revealed to the agent and continues to invest optimally thereafter. That is to say, the type H principal's investment may be sub-optimal in the short run, but in the long run, optimality is restored.

The above insights can be useful in understanding the optimal behavior of managers. In organizations, workers are given feedback through periodic evaluations and appraisals. These represent the costless (cheap talk) channel through which managers provide information to workers to motivate them to exert effort. However, how workers learn about the relevant aspects of the production environment (project's quality/ agent's ability) is not limited to these periodic chunks of information. The day to day actions of a manager, how much interest they show in a workers activities by investing their resources credibly transmits the private information of the manager to the workers affecting their motivation to work. Accounting for this channel of credible information transmission may enhance how managers motivate their workers and better achieve organizational goals.

The paper is organized as follows. I review the connection of this paper with literature in Section 2.2, followed by a one period example in Section 2.3. In Section 2.4, I present the model and then present results in Section 2.5. Extensions are discussed in Section 2.6. All proofs are relegated to the appendix.

2.2 Literature Review

Motivating agents to exert effort through costly signalling has been previously studied in static settings by Hermalin (1998) and Komai, Stegeman, and Hermalin (2007) who analyze models of moral hazard in teams where a leader endowed with private information about a project spends costly resources to signal the quality of the project to her followers in order to induce effort. In contrast, this paper analyzes a dynamic environment where the principal learns about the quality of a project over time and balances between investing optimally and transmitting information to the agent. Dong (2018) studies a similar problem where two players simultaneously, but separately experiment to learn about the state of the world, with one player possessing superior information from the outset, as opposed to my model of joint experimentation where players start the relationship with symmetric information. My results are qualitatively different in that at high beliefs about project quality, our model predicts full information transmission. Halac (2012) studies a game with transfers where a principal with persistent private information induces effort from the agent. She shows that a separating equilibrium does not exist, and in equilibrium, the principal discloses her type gradually. In contrast, my model is without transfers and the principal acquires private information over time which result in very different dynamics.

This paper is related to the literature on private learning. In particular, Bonatti and Hörner (2011) study a model of moral hazards in teams in an exponential bandit framework and show that the incentive to free-ride on other players' leads to reduction of effort. Augmenting the setting of Bonatti and Hörner (2011), Guo and Roesler (2016) analyze a model of collaboration and private learning with publicly observable exit decisions. However, in their setting players have access to a fully revealing bad signal and communicate their private information by exiting or the lack of it. Bimpikis, Drakopoulos, Ehsani, et al. (2018) study a strategic experimentation model of private learning where information can be credibly disclosed without a cost through commitment. Akcigit and Liu (2016) examine an innovation competition between two firms which decide whether to pursue a risky or a safe project. Since a firm benefits when its competitor works in a less rewarding direction, it never reveals dead-end finding.

More broadly, this paper is related to the literature on experimentation (See, for instance Bolton and Harris (1999), Keller et al. (2005), Keller and Rady (2015)), particularly in principal-agent settings with conflicting interests studied by Guo (2016), Garfagnini (2011) and Kuvalekar and Lipnowski (2018) among others.

2.3 A one period example

In this section I present a one period example that highlights the main forces I study in this paper. Both the principal (she) and the agent (he) share common prior $\mu \in [0,1]$ that the quality of a project is good. The principal moves first and invests $x \in [0, \bar{x}]$ at a cost of $\frac{1}{2}ax^2$. After observing the principal's investment, the agent can exert effort ($e = 1$) at a cost of c or not ($e = 0$). Not exerting effort is costless. If the project is of good quality and the agent exerts effort then an output $1 + x$ is produced with probability p. No output is produced if project quality is bad or the agent does not exert effort. The principal gets a share γ and the agent gets a share $1 - \gamma$ of the output.

Supposing that the agent exerts effort when the belief that project quality is good is μ, the principal's payoff when she invests x is given by:

$$\gamma\mu p(1 + x) - \frac{1}{2}ax^2.$$

This gives the principal's optimal choice of investment as a function of her belief

$$x^*(\mu) = \frac{\gamma\mu p}{a},$$

resulting in a value of

$$\gamma\mu p + \frac{(\gamma\mu p)^2}{2a}.$$

The agent's problem is to exert effort or not. Given a belief μ and the choice of investment x of the principal, the agent exerts effort if

$$(1-\gamma)\mu p(1+x) - c \geq 0.$$

This gives the minimum the principal needs to invest $x_a(\mu)$ at belief μ in order to induce effort from the agent,

$$x_a(\mu) = \frac{c}{(1-\gamma)\mu p} - 1.$$

Note that $x_a(\mu)$ is decreasing in μ. As the assessment about the quality of the project decreases, the agent demands higher investment from the principal to be willing to exert effort. That is, the more pessimistic the agent is about the prospect of the project, the more costly it is for the principal to induce effort from the agent.

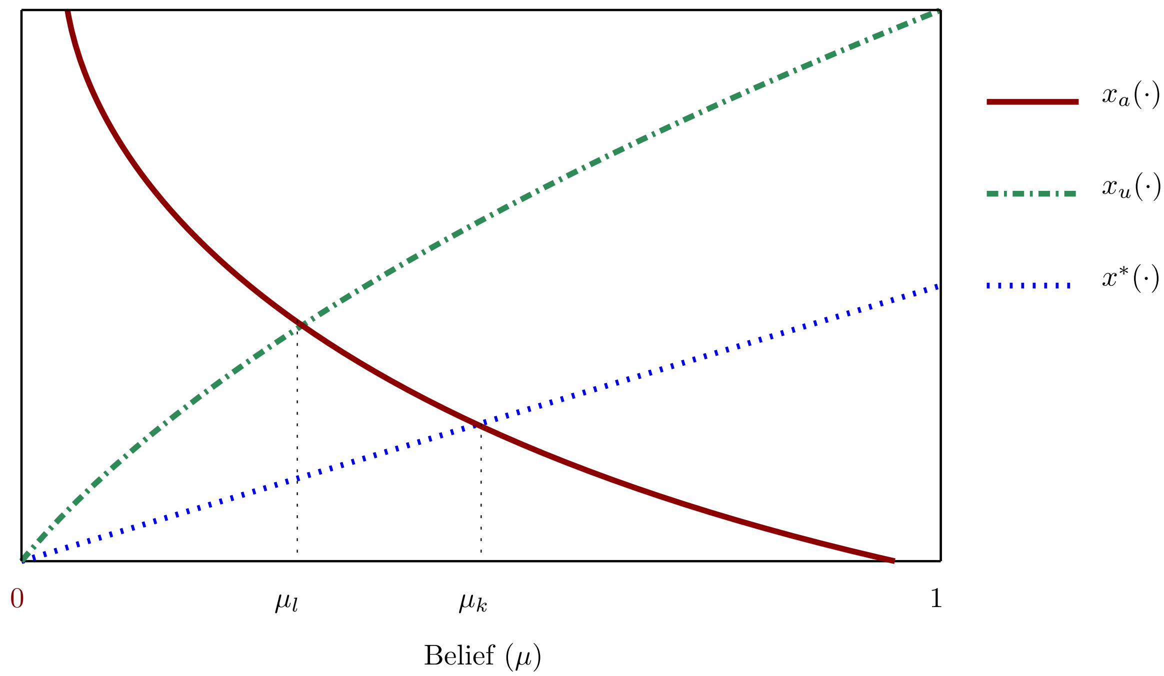
At any belief μ, there is an upper bound to what the principal is willing to invest to induce effort. This upper bound $x_u(\mu)$ solves

$$\gamma\mu p(1+x_u) - \frac{1}{2}ax_u^2 = 0.$$

Note that $x_u(\mu)$ is increasing in μ. Define μ_l as the belief at which $x_a(\cdot)$ and $x_u(\cdot)$ intersect. Also, define μ_k as the belief at which $x_a(\cdot)$ and $x^*(\cdot)$ intersect.

Note from Figure 2.2 that when beliefs are sufficiently high, i.e. in the region $[\mu_k, 1]$, there is no conflict between the principal and the agent. The principal's optimal investment is at least as much what the agent requires to exert effort. For intermediate beliefs, i.e. in the region $[\mu_l, \mu_k)$, the agent's participation constraint binds ($x_a(\cdot) > x^*(\cdot)$). In this region, the principal invests higher than her optimal level in order to appease the agent into exerting effort. For sufficiently low beliefs, i.e. in the region $[0, \mu_l)$, the cost of appeasing the agent is too high for the principal and the principal prefers to not induce effort from the agent. This observation is summarized in the proposition below.

FIGURE 2.1. One shot game



Proposition 8. *When belief is μ, the unique Subgame Perfect Nash Equilibrium of the one shot game is given by*

$$x = \begin{cases} \max\left\{\frac{c}{(1-\gamma)\mu p} - 1, \frac{\gamma\mu p}{a}\right\} \text{ if } \mu \geq \mu_l, \\ 0 \text{ if } \mu < \mu_l. \end{cases}$$

$$e(x) = \begin{cases} 1 \text{ if } x \geq \frac{c}{(1-\gamma)\mu p} - 1, \\ 0 \text{ otherwise }. \end{cases}$$

The one shot game highlights the difficulty in motivating an agent to exert effort when he is pessimistic about the project quality. In a dynamic environment where the principal learns about the project quality privately over time and makes investments into the project, the investment of the principal today not only affects the output today, but also affects how much the agent learns about project quality which in turn affects the continuation value of the principal. In the rest of the paper, I study how the principal optimally manages the trade-off between current value and her continuation value by influencing how the agent learns about project quality.

2.4 Model

Players: Time t is continuous and runs from 0 to ∞. A principal hires an agent to work on a project of unknown quality. The quality of the project is good ($\theta = 1$), or bad ($\theta = 0$). At time 0, both players have common prior μ_0 that project quality is good.

Actions: At each instant, the principal invest an amount $x \in [0, \bar{x}]$ of resource into the project at a cost of $\frac{ax^2}{2}$ where a is a constant. In the same instant, but after having observed the investment of the principal, the agent chooses either to exert effort ($e = 1$) in the project or not ($e = 0$). The agent incurs a cost of c if she exerts ($e = 1$) effort. Not exerting effort is costless. The agent's effort choice is observed by the principal but is not contractible.

Payoffs The agent's effort and the quality of the project affects the arrival rate of lump sum output which is exponentially distributed. The arrival rate of output is $\lambda\theta e$ where λ is the sensitivity of the production technology. Output is only produced if the project is of good quality and the agent exerts effort. Output is privately observed by the principal. Note that the first arrival of output resolves all the uncertainty for the principal: the realization of the first output results in the posterior of the principal taking the value $\mu = 1$. The output at time t, if it arrives is given by $1 + x_t$, where x_t is the principal's investment at time t. The principal's investment does not affect the arrival rate of output, however it affects the value of the output if output arrives. That is, the the principal can only augment the output by investing. I denote by $y_t \in \{\phi, (1 + x_t)\}$ the output at time t, where ϕ denotes no output.

At the beginning of the relationship, both players commit to a non-negotiable sharing rule of output where the principal receives a share γ and the agent receives a share $1 - \gamma$ of the output. Both players discount future payoffs at the rate r. Notice that the agent does not observe output during the course of the relationship.[4] I interpret the rate of discounting r as the exogenous rate at which the game ends resulting in the agent realizing his share

[4] This is an important feature since otherwise there is no information asymmetry between the players and the main tension in the model is absent.

of the accumulated output. There are no transfers. Letting $\mathcal{T}$ denote the set of stochastic times at which an output arrives, the principal's payoff is given by

$$U(\mu_0) = \mathbb{E}^{e,x}\left[\gamma \sum_{s\in\mathcal{T}} e^{-rs}y_s - \int_{t=0}^{\infty} re^{-rt}\frac{1}{2}ax_t^2 dt\right],$$

and the agent's payoff is given by

$$V(\mu_0) = \mathbb{E}^{e,x}\left[(1-\gamma)\sum_{s\in\mathcal{T}} e^{-rs}y_s - \int_{t=0}^{\infty} re^{-rt}cedt\right].$$

2.4.1 Strategies and Equilibrium

Since the principal observes the history of outputs privately, she has private information about project quality. Given the fully revealing nature of the production technology, the relevant private information of the principal at time t can be categorized into two classes: either the principal has observed an output or has not observed an output before t. Denote the private history of the principal at time t as h_t^p defined as $h_t^p = \{y_s\}_{s=0}^t$. h_t^p contains all information about realized outputs up to time t. Denote the set of all feasible private histories for the principal by H_t^p. The principal can one of two types at time t depending on her private history H_t^p. Type L is a principal who has not observed any output. Type H is a principal who has observed output at some point. I denote the belief (probability that project quality is good) of type $T \in \{L, H\}$ principal by μ^T. Given that the agent exerts effort ($e = 1$) at every instant before t, the belief of the type L principal at time t is given by

$$\mu^L = \frac{\mu_0}{\mu_0 + (1-\mu_0)\exp(\lambda t)}.$$

The belief of a type H principal at any time is $\mu^H = 1$. At the start of the game, by definition, the principal is type L. Also note that type H is an absorbing type.

The agent learns about project quality through the investment of the principal. In particular, the information available to the agent about project quality depends on the strategies of both types of principal. There are two relevant beliefs of the agent that we need to track throughout the game

1. μ^a: The agent's belief (probability) that the principal is type H.

2. $\bar{\mu}^a$: The agent's belief (probability) that the project quality is good.

Note that the agent's beliefs are known to the principal since the agent does not possess any private information. At the start of the game $\mu^a = 0$ and $\bar{\mu}^a = \mu_0$ since at the beginning of the game the principal is type L and there is no information asymmetry. At any time in the game, μ^a and $\bar{\mu}^a$ are related as follows

$$\begin{aligned}\bar{\mu}^a =& \mu^a \mu^H + (1 - \mu^a)\mu^L \\ =& \mu^a[1] + (1 - \mu^a)\left[\frac{\mu_0}{\mu_0 + (1 - \mu_0)\exp \lambda t}\right].\end{aligned}$$

Solution Concept: The solution concept used in this paper is Markov Perfect Bayesian Equilibrium. Strategies are Markov if they depend only on payoff-relevant past events that includes a player's own payoff-relevant private information. The payoff-relevant information for a principal at any point where she makes her investment decision, is her own private type $T \in \{L, H\}$, the belief of the type L principal μ^L and the public belief of the agent μ^a. The payoff-relevant information for the agent at any point where he makes his effort decision is his belief about the type of principal before investment μ^a, the belief of the type L principal μ^L and the investment of the principal. The state variable for this problem is the pair $\mathcal{S} = (\mu^L, \mu^a) \in [0, 1]^2$ which is observed by both players. At the beginning of the game, the principal is of type L and the agent knows he is facing type L principal. Hence the game starts with the state $\mathcal{S}_0 = (\mu_0, 0)$.

The principal's strategy depends on her type and the state $\mathcal{S}$. A pure strategy for type T principal is a function

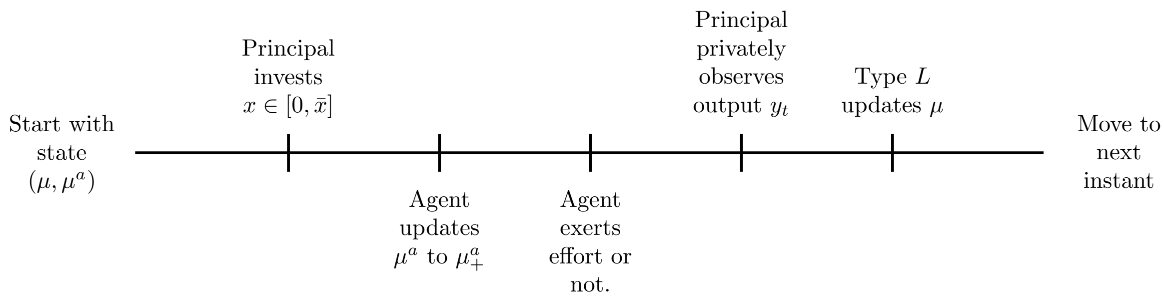
$$x_T : [0, 1]^2 \to [0, \bar{x}].$$

Two histories for the principal that lead to the same posterior for the principal and the agent, result in an identical investment choice by the principal. The agent's strategy depends on the state $\mathcal{S}$ and the observed investment $x \in [0, \bar{x}]$. A pure strategy for the agent is a function

$$e : [0,1]^2 \times [0,\bar{x}] \to \{0,1\}.$$

With slight abuse of notation, I denote the strategy of type T principal as $x_T(\mu^L, \mu^a)$ and the strategy of the agent as $e(\mu^L, \mu^a_+, x)$, where μ^a_+ is the updated belief of the agent after having observed investment x.

Figure 2.2. Timeline within an instant



Updating of beliefs off-path: An important part of the description of a Perfect Bayesian Equilibrium is the updating of beliefs after players observe off-path actions. I maintain consistency in the rule the agent uses to update his belief after observing investments that are off-path, across equilibria presented in this paper. I describe the rule below and refer to it when specifying an equilibrium.

Definition 3. ***(Updating of beliefs off-path)*** *Suppose X^σ is the set of all possible on path investments at state $\mathcal{S} = (\mu, \mu^a)$ under a strategy profile σ. Set $x^1 = \max X^\sigma$ and $x^2 = \min X^\sigma$. The updated belief of the agent after observing an investment $x \notin X^\sigma$ is given by*

$$\mu^a_+(S,x) = \begin{cases} \mu^a_+(S,x^1) \text{ if } x > x^1, \\ \mu^a_+(S,x^2) \text{ if } x \in (x^2,x^1), \\ 0 \text{ if } x < x^2, \end{cases}$$

where $\mu^a_+(S,x^1)$ and $\mu^a_+(S,x^2)$ are calculated using Bayes' rule.

Note that this definition applies even if $\mu^a = 1$. That is, even if the agent assigns probability 1 that he is facing type H principal before investment (and thereby $|X^\sigma| = 1$), the agent can update her belief if he observes off-path investment. This updating rule specifies that after observing an off-path investment, the probability the agent assigns to be facing the type H principal is equal to what he would have assigned after observing the highest on-path investment lower than the investment observed. If the off-path investment is lower than all on-path investments, the agent believes that he is facing type L. Note that if type H principal invests higher than type L principal (which will be the case in any equilibrium), this belief updating rule is monotonic in investment. This updating rule specifies the worst beliefs for the agent ($\mu^a_+ = 0$) following any downward deviation of the principal, which results in the strongest incentives for the principal to not deviate from the path of play.

I make an assumption on parameters so that there are no distortions in a relationship where it is common knowledge that the principal is type H. The following suffices.

Assumption 3. $(1-\gamma)\lambda - c > 0$.

Since I am interested in equilibria that maximize the principal's ex-ante payoff, the following definition is useful.

Definition 4. ***(Principal Optimal Equilibrium)*** *For any initial state $\mathcal{S}_0 = (\mu_0, 0)$, denote by $\Sigma_{\mathcal{S}_0}$ the set of all equilibria at $\mathcal{S}_0$.*[5] *For any $\sigma \in \Sigma_{\mathcal{S}_0}$ denote the value to the type L principal at $\mathcal{S}_0$ as $U^\sigma_L(\mathcal{S}_0)$. I call $\hat{\sigma} \in \Sigma_{\mathcal{S}_0}$ a principal optimal equilibrium if,*

$$U^{\hat{\sigma}}_L(\mathcal{S}_0) = \max_{\sigma' \in \Sigma_{\mathcal{S}_0}} U^{\sigma'}_L(\mathcal{S}_0).$$

2.5 Results

I first describe the equilibrium in the game where it is common knowledge among both players that project is of good quality ($\theta = 1$). Since the principal has no private information, and hence no concern for information transmission, in this case the principal chooses

[5] The set of all equilibria is the set of Markov Perfect Equilibria satisfying the belief updating rule given in Definition 3.

her optimal investment at every instant by maximizing her expected flow payoff given by

$$\gamma\lambda(1+x)-\frac{ax^2}{2}.$$

Her optimal investment is $\frac{\lambda\gamma}{a}$. The agent exerts effort at every instant and his payoff from the relationship is denoted as z^a given by

$$z^a=(1-\gamma)\lambda(1+\frac{\gamma\lambda}{a})-c, \tag{2.1}$$

which is strictly positive since Assumption 3 guarantees that the agent's payoff is strictly positive regardless of the investment of the principal. I denote the value to the principal in this relationship as z^p given by

$$z^p=\gamma\lambda(1+\frac{\gamma\lambda}{2a}). \tag{2.2}$$

The above discussion is summarized by the following lemma whose proof I omit.

Lemma 4. *Suppose both players know that project is of good quality. The unique equilibrium is given by*

$$x^*=\frac{\lambda\gamma}{a};\ \ e^*=1.$$

I next describe an autarkic equilibrium which is used as punishment off the path of play in the equilibria I construct.

Lemma 5. *Suppose* $\mu\leq\frac{c}{\lambda(1-\gamma)}$. *The strategy profile* $\sigma^n=(x^n_H,x^n_L,e^n)$ *described below is an equilibrium at state* $(\mu,0)$.

$$\begin{aligned}x^n_H(\cdot,\cdot)=&x^n_L(\cdot,\cdot)=0;\\ e^n(\cdot,\cdot,\cdot)&=0.\end{aligned}$$

Agent updates beliefs after observing off-path investment using the rule described in Definition 3.

The proof can be found in the appendix. Note that the autarkic equilibrium shown above exists at the state when type L principal's belief μ is less than $\dfrac{c}{\lambda(1-\gamma)}$ and the agent believes he is facing the type L principal. When the type L principal's belief is above $\dfrac{c}{\lambda(1-\gamma)}$, the agent's value is strictly positive from exerting effort even if the principal does not invest, hence autarky is not an equilibrium when $\mu > \dfrac{c}{\lambda(1-\gamma)}$. When players revert to autarky, we say that players have quit the relationship.

I begin the characterization of the principal's optimal equilibria by stating a necessary condition that must be satisfied in any equilibrium that specifies pooling at some point.

Lemma 6. *Fix an equilibrium that specifies pooling at some interval (p_1, p_2) of type L principal's beliefs and the agent exerts effort in this interval.*[6] *Denote by $x^P(p)$ the pooling investment function in the interval (p_1, p_2). Let $D \subseteq (p_1, p_2)$ be a Lebesgue measurable set of beliefs where $x^P(p) < \dfrac{\gamma\lambda}{a}$ when $p \in D$. Then, D has Lebesgue measure zero.*

Note that $\dfrac{\gamma\lambda}{a}$ is the type H type principal's optimal investment. The result above states that in any equilibrium which specifies pooling in an interval (p_1, p_2) of type L principal's beliefs, the pooling investment must be at least equal to $\dfrac{\gamma\lambda}{a}$ except at a measure zero subset of (p_1, p_2). The idea behind the result is as follows. The agent's belief updating rule given by Definition 3 specifies that during pooling, the agent's updated belief after observing an investment higher than the equilibrium investment is identical to the updated belief after observing equilibrium investment. If there is a subset of (p_1, p_2) of positive measure where the equilibrium investment function specifies an investment smaller than $\dfrac{\gamma\lambda}{a}$, the type H principal can simply deviate to $\dfrac{\gamma\lambda}{a}$ for all beliefs in this subset and improve her payoff since she will then invest optimally. Note that the agent's belief updating will be unchanged by this deviation which implies that continuation equilibrium when type L principal's belief reaches p_1 will be unchanged following the deviation.

Since the principal's behavior on a measure zero set of beliefs during pooling has no payoff consequence to either player, going forward, I will assume that in any equilibrium, the investment specified during pooling (if there is pooling) must be atleast equal to $\dfrac{\gamma\lambda}{a}$.

The above result leads to the following important observation. The amount type L principal is willing to invest in order to induce effort from the agent increases in her belief

[6] The dependence of the principal's strategy on the agent's belief has been suppressed for the ease of notation.

since her value of the relationship is lower at smaller beliefs. If her beliefs are sufficiently low, she will not invest $\frac{\gamma\lambda}{a}$ to induce effort. This implies that at sufficiently low beliefs of type L principal, there cannot be pooling in equilibrium. I define $\underline{\mu}$ as the cut-off belief at which type L principal is indifferent between quitting and inducing effort from the agent by investing $\frac{\gamma\lambda}{a}$.

Definition 5. *Define*

$$\underline{\mu} = \frac{r\lambda\gamma^2}{2[r\lambda\gamma^2 + a(r\gamma + z^p)]}.$$

Interpretation of $\underline{\mu}$: Suppose type L principal has a choice of investing $\frac{\gamma\lambda}{a}$ and inducing effort from the agent or quit the relationship. If the belief of the type L principal $\mu > \underline{\mu}$, she strictly prefers to induce effort. If $\mu < \underline{\mu}$, she strictly prefers to quit. If $\mu = \underline{\mu}$, she is indifferent between inducing effort and quitting.

In light of the above discussion, a natural question arises: Is the agent willing to exert effort at $\underline{\mu}$ following an investment of $\frac{\gamma\lambda}{a}$ if he knows he is facing the type L principal? The answer to this condition depends on whether condition $C1$ (defined below) is satisfied or not. If $C1$ is satisfied then the agent does not exert effort, and if $C1$ is violated, then the agent exerts effort.

Condition 1 ($C1$)**:**

$$r\left[\underline{\mu}\lambda(1-\gamma)\left(1+\frac{\lambda\gamma}{a}\right) - c\right] + \underline{\mu}\lambda z^a < 0.$$

I provide a heuristic intuition behind $C1$. Note that $r\left[\underline{\mu}\lambda(1-\gamma)\left(1+\frac{\lambda\gamma}{a}\right) - c\right]\mathrm{d}t$ is the expected flow value of the agent and $\underline{\mu}\lambda z^a\mathrm{d}t$ is the expected continuation value which is the product of the probability ($\underline{\mu}\lambda\mathrm{d}t$) that an output is produced (implying that the type L principal transitions to type H) and the value to the agent from continuing the relationship with the type H principal going forward (z^a). The sum of these two values gives the overall value to the agent from exerting effort at $\underline{\mu}$. Also note that fixing other parameters of the model, $C1$ is satisfied when the cost of effort of the agent is high and is violated otherwise. In other words, $C1$ can also be stated in terms of a cut-off on the cost of effort of the agent i.e. $C1$ is equivalent to

$$c > \underline{\mu}\lambda(1-\gamma)\left(1+\frac{\lambda\gamma}{a}\right) - \frac{\underline{\mu}\lambda z^a}{r} = c^*.$$

The nature of principal's optimal equilibrium critically depends on whether $C1$ is satisfied or not i.e. whether the agent's cost of effort is high or low. I present the principal optimal equilibrium for both cases in the following sections.

2.5.1 Low cost of effort

In this section I present the principal optimal equilibrium when $C1$ is violated i.e. agent's cost of effort is low in Proposition 9 presented below. Before we go to the proposition, I define some useful objects.

Definition 6. *Define μ^c as the unique value of μ that solves*

$$\frac{\mu\gamma\lambda}{a} = f(\mu),$$

and μ^s as the unique value of μ that solves

$$r\left[\mu\lambda\gamma(1+f(\mu)) - a\frac{(f(\mu))^2}{2}\right] + \mu\lambda z^p = 0,$$

where

$$f(\mu) = \frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1. \tag{2.3}$$

Also define type L principal's optimal investment as a function of her belief as

$$g(\mu) = \frac{\mu\gamma\lambda}{a}.$$

Proposition 9. *Suppose $C1$ does not hold, then, $\mu^s \leq \underline{\mu} < \mu^c$, and the principal optimal equilibrium is a fully separating equilibrium σ^* given by,*

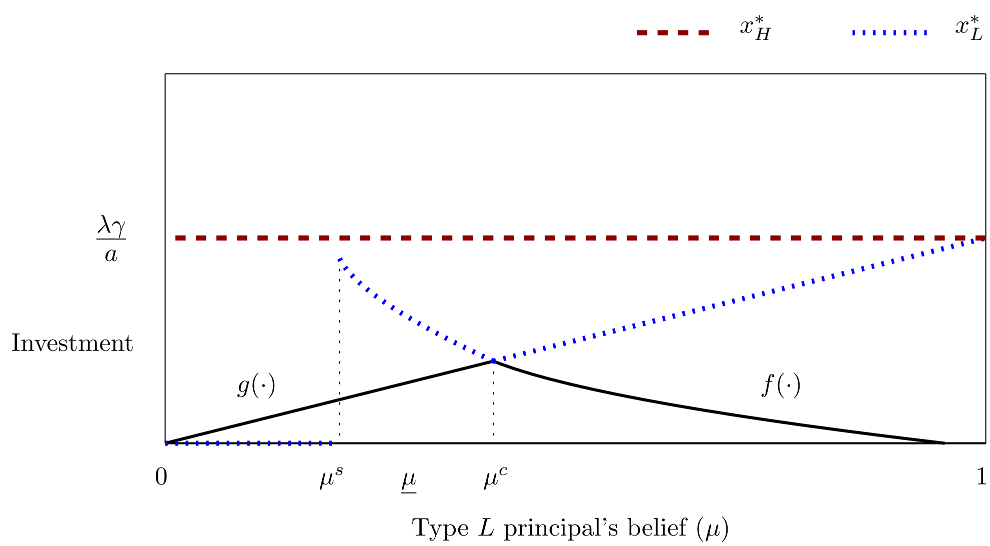
$$x_H^*(\mu, \mu^a) = \frac{\gamma\lambda}{a}; \quad x_L^*(\mu, \mu^a) = \begin{cases} \dfrac{\mu\gamma\lambda}{a} \textit{ if } \mu \in [\mu^c, 1], \\ f(\mu) \textit{ if } \mu \in [\mu^s, \mu^c), \\ 0 \textit{ if } \mu < \mu^s; \end{cases}$$

$$e^*(\mu, \mu_+^a = 1, x) = 1;$$

$$e^*(\mu, \mu_+^a = 0, x) = \begin{cases} 1 \textit{ if } \mu < \mu^c \textit{ and } x \geq f(\mu), \\ 1 \textit{ if } \mu \geq \mu^c, \\ 0 \textit{ otherwise.} \end{cases}$$

Agent updates beliefs using Bayes' rule on path and using the rule described in Definition 3 after observing off-path investment.

FIGURE 2.3. Equilibrium when cost of effort is low.



The principal's optimal equilibrium when $C1$ does not hold is a fully separating equilibrium, i.e. separation is specified for every belief of the type L principal. There is full information transmission, through the investments of the principal, the agent learns perfectly the type of principal she is facing. Type H principal invests optimally at every point in the relationship. Type L principal invests optimally when her beliefs are high ($\mu \in [\mu^c, 1]$), invests higher than her optimal ($g(\mu)$) at intermediate beliefs ($\mu \in [\mu^s, \mu^c)$) and quits the relationship at low beliefs ($\mu < \mu^s$). Note that this implies that if the initial prior μ_0 is less than μ^s, the principal never invests and the agent never exerts effort, i.e. there is no production.

Whenever the type L principal invests higher than her optimal investment, note that the value of the agent facing her must be zero, otherwise she can reduce her investment by a little and still induce effort and consequently be better off. This is indeed the case when $\mu < \mu^c$. Type L principal simply chooses the minimum investment needed to induce effort from the agent using the equation below,

$$\underbrace{r\left[\mu\lambda(1-\gamma)\left(1+x_L^*(\mu,\mu^a)\right)-c\right]\mathrm{d}t}_{\text{expected flow value}} + \underbrace{\mu\lambda z^a \mathrm{d}t}_{\text{expected continuation value}} = 0.$$

Note that the expected continuation value to the agent comes only from facing the type H principal in the next instant. This is because, if the agent faces the type L principal at the next instant, as argued above, the value of the agent is zero.

Type L principal's investment is always weakly less than the optimal investment of type H principal. To see this observe that since since $C1$ is violated,

$$r\left[\underline{\mu}\lambda(1-\gamma)\left(1+\frac{\lambda\gamma}{a}\right)-c\right]\mathrm{d}t + \underline{\mu}\lambda z^a \mathrm{d}t \geq 0,$$

which implies type L principal can induce effort from the agent at $\underline{\mu}$ by investing weakly lower than $\dfrac{\lambda\gamma}{a}$, and strictly so if the above inequality is strict. Also note that type L principal's investment can never be greater than $\dfrac{\lambda\gamma}{a}$ when $\mu < \underline{\mu}$ by the definition of $\underline{\mu}$. Since type L principal's equilibrium investment is strictly decreasing in μ when $\mu \in [\mu^s, \mu^c]$, her investment is always weakly less than $\dfrac{\lambda\gamma}{a}$.

When the agent's cost of effort is low, the misalignment in the incentives of the players results in the type L principal investing above her optimal investment to motivate the agent to exert effort at intermediate beliefs. However, the misalignment is not sufficiently high to stop the flow of information about the project quality to the agent. As we will see in the next section, if the cost of effort for the agent is high enough, in the principal optimal equilibrium, in addition to investing above her optimal investment, the type L principal behavior results in the stoppage of the flow of information to the agent when her beliefs are sufficiently low.

2.5.2 High cost of effort

When the agent's cost of effort is high, i.e. condition $C1$ holds, the principal's optimal equilibrium is sensitive to the initial prior μ_0 about the project quality. Before we proceed, I define a belief $\bar{\mu}$ which will be useful in our analysis.

Definition 7. *Suppose starting at state $(\mu, 0)$ with $\mu > \underline{\mu}$, type H principal invests $\frac{\lambda\gamma}{a}$ in perpetuity and type L principal invests $\frac{\lambda\gamma}{a}$ until beliefs drift down to $\underline{\mu}$ and then quits. Denote by $V(\mu)$ the value of the agent if he exerts effort following investment. Define $\bar{\mu}$ as the unique belief at which $V(\bar{\mu}) = 0$.*

As in the case with low cost of effort, if the initial prior is sufficiently low, the only equilibrium is autarky, i.e. the principal invests zero and the agent never exerts effort. I state this in the following proposition.

Proposition 10. *Suppose $C1$ holds. There exists $\mu_g \in (\underline{\mu}, \bar{\mu})$ such that if $\mu_0 < \mu_g$, the only equilibrium is autarky, i.e. players quit right away. If $\mu_0 \geq \mu_g$ an equilibrium with production exists.*

The proof is delivered by Lemma 43 which can be found in the appendix. The principal's optimal equilibrium is qualitatively different depending on whether the initial prior μ_0 is greater than $\bar{\mu}$ or not. I first consider the case when $\mu_0 > \bar{\mu}$. Recalling the definition of μ^c and $f(\mu)$ given in Definition 6 before Proposition 9, the principal's optimal equilibrium is given in the proposition below.

Proposition 11. *Suppose $C1$ holds. If $\mu_0 > \bar{\mu}$, the principal's optimal equilibrium σ^* is a three phase equilibrium given by,*

- ***Separating Phase:*** *$\mu > \bar{\mu}$:*

$$x_H^*(\mu, \mu^a) = \frac{\gamma\lambda}{a}; \quad x_L^*(\mu, \mu^a) = \begin{cases} \frac{\mu\gamma\lambda}{a} \ \textit{if } \mu \in (\max\{\bar{\mu}, \mu^c\}, 1], \\ f(\mu) \ \textit{if } \mu \in (\bar{\mu}, \max\{\bar{\mu}, \mu^c\}];^7 \end{cases}$$

$$e^*(\mu, \mu_+^a = 1, x) = 1;$$

$$e^*(\mu, \mu_+^a = 0, x) = \begin{cases} 1 \ \textit{if } \mu > \max\{\bar{\mu}, \mu^c\}, \\ 1 \ \textit{if } \mu \in (\bar{\mu}, \max\{\bar{\mu}, \mu^c\}) \ \textit{and } x \geq f(\mu), \\ 0 \ \textit{if } \mu \in (\bar{\mu}, \max\{\bar{\mu}, \mu^c\}) \ \textit{and } x < f(\mu); \end{cases}$$

- ***Pooling Phase:*** *$\mu \in [\underline{\mu}, \bar{\mu}]$:*

$$x_H^*(\mu, \mu^a) = x_L^*(\mu, \mu^a) = \frac{\gamma\lambda}{a};$$

$$e^*(\mu, \mu_+^a, x) = 1 \ \textit{if } \mu_+^a \neq 0;$$

  Players revert to autarky if at any point in this phase the agent assigns $\mu_+^a = 0$.

- ***Quitting Phase:*** *$\mu < \underline{\mu}$:*

$$x_H^*(\mu, \mu^a) = \frac{\gamma\lambda}{a}; \ x_L^*(\mu, \mu^a) = 0;$$

$$e^*(\mu, \mu_+^a = 1, x) = 1;$$

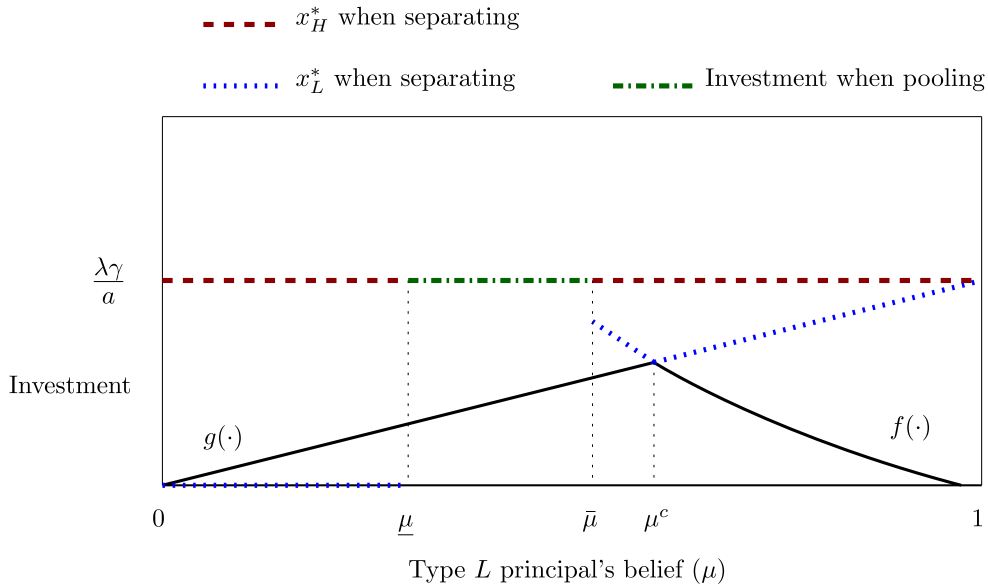
$$e^*(\mu, \mu_+^a = 0, x) = 0;$$

Agent updates beliefs using Bayes' rule on path and using the rule described in Definition 3 after observing off-path investment.

The principal optimal equilibrium exhibits three phases when initial prior $\mu_0 > \bar{\mu}$. When beliefs of the type L principal are high, separation occurs. Type L principal invests less than type H principal who invests optimally. In this region, the behavior of the principal is qualitatively similar to the fully separating equilibrium presented in Proposition 9. Within

[7] Intervals of the form $(a, a]$, (a, a) and $[a, a)$ are considered as empty sets.

FIGURE 2.4. Equilibrium when cost of effort is high and prior is high.



the separating region the type L principal invests optimally when $\mu > \max\{\bar{\mu}, \mu^c\}$. If $\mu^c > \bar{\mu}$, there can be a region $(\bar{\mu}, \mu^c)$ within the separating phase where type L principal invests above her optimal level since the agent's participation constraint begins to bind at beliefs lower than μ^c. Once beliefs reach $\bar{\mu}$, the pooling phase begins where both players pool on the type H principal's optimal investment until belief reach $\underline{\mu}$, at which point type L principal quits the relationship.

Even when cost of effort is high, if the relationship begins with a high belief that the project is of good quality, it is optimal for both types of principals to invest optimally and allow full information transmission to the agent, at least until beliefs drift down enough where the conflict between the type L principal and the agent begins. Anticipating the high costs (investment) of motivating an agent who is pessimistic (low belief that project is of good quality) to exert effort in future, type L principal has incentives to stop the flow of information to the agent, which leads to pooling at intermediate beliefs in equilibrium. A natural question arises then — When does pooling begin? By Definition 7, if pooling begins at $\bar{\mu}$ and both types of principal invest type H principal's optimal investment $(\dfrac{\gamma\lambda}{a})$ then the agent gets a value of zero at $\bar{\mu}$. If pooling is delayed, i.e. pooling begins at $\mu' < \bar{\mu}$, the average investment in the pooling phase must be strictly higher than $(\frac{\gamma\lambda}{a})$ to induce effort from the agent since the agent is pessimistic at μ' compared to $\bar{\mu}$. However note that in this case type L principal can deviate at belief $\mu' + \varepsilon < \bar{\mu}$. By choosing the type H principal's separating investment at $\mu' + \varepsilon$, type L principal can guarantee she doesn't have to invest any more than $\frac{\gamma\lambda}{a}$ in future to induce effort. If pooling begins earlier at $\mu' > \bar{\mu}$, note that the agent's value at μ' is strictly positive since the minimum possible investment during pooling is $\frac{\gamma\lambda}{a}$. By decreasing the pooling cutoff, the type L principal can separate and invest less than $\frac{\gamma\lambda}{a}$ for a little longer and hence is better off. Therefore in an ex-ante principal optimal equilibrium pooling must begin at $\bar{\mu}$. Note that during the pooling phase, after observing investment, the agent's belief that he faces the type H principal given by μ^a_+ is strictly positive (and increases as μ goes down over time) since there is a positive probability that the principal has observed output since the end of the separating phase. Any downward deviation by the principal in this phase leads to the agent updating μ^a_+ to 0 and players continue with autarky (quitting the relationship).

By stopping the flow of information to the agent at $\bar{\mu}$, the type L principal exploits the agent's uncertainty (μ^a_+ is increasing as μ goes down) in the pooling phase and is able to reduce her investment while inducing effort from the agent. In particular, she is able to induce effort from the agent when her belief $\mu \in [\underline{\mu}, \mu_g)$ which would not have been possible if the agent was aware that he is facing her.

Next, we turn to the principal's optimal equilibrium when the initial prior $\mu_0 \in [\mu_g, \bar{\mu}]$. Before I present it, the following definition will be useful.

Definition 8. *Denote by $\mathcal{X}(\mu_0)$ as set of pooling investment function defined from $[\underline{\mu}, \mu_0]$ to $[\frac{\lambda\gamma}{a}, \bar{x}]$ that give both the type L principal and the agent, a non negative value in the pooling region $[\underline{\mu}, \mu_0]$ when the relationship begins at the state $(\mu_0, 0)$. Define $x^P(\cdot; \mu_0) \in \mathcal{X}(\mu_0)$ as the function that maximizes the ex-ante value of the type L principal at the state $(\mu_0, 0)$.*

I state the proposition characterizing the principal's optimal equilibrium when $\mu_0 \in [\mu_g, \bar{\mu}]$ below.

Proposition 12. *Suppose $C1$ holds. If $\mu_0 \in [\mu_g, \bar{\mu}]$, the principal's optimal equilibrium σ^* is a two phase equilibrium given by,*

- ***Pooling Phase:*** *$\mu \in [\underline{\mu}, \bar{\mu}]$:*

$$
\begin{aligned}
x_H^*(\mu, \mu^a) = & x_L^*(\mu, \mu^a) = x^P(\mu; \mu_0); \\
e^*(\mu, \mu_+^a, x) = & 1 \ \textit{if} \ \mu_+^a \neq 0, \ \textit{or} \ \mu = \mu_0;
\end{aligned}
$$

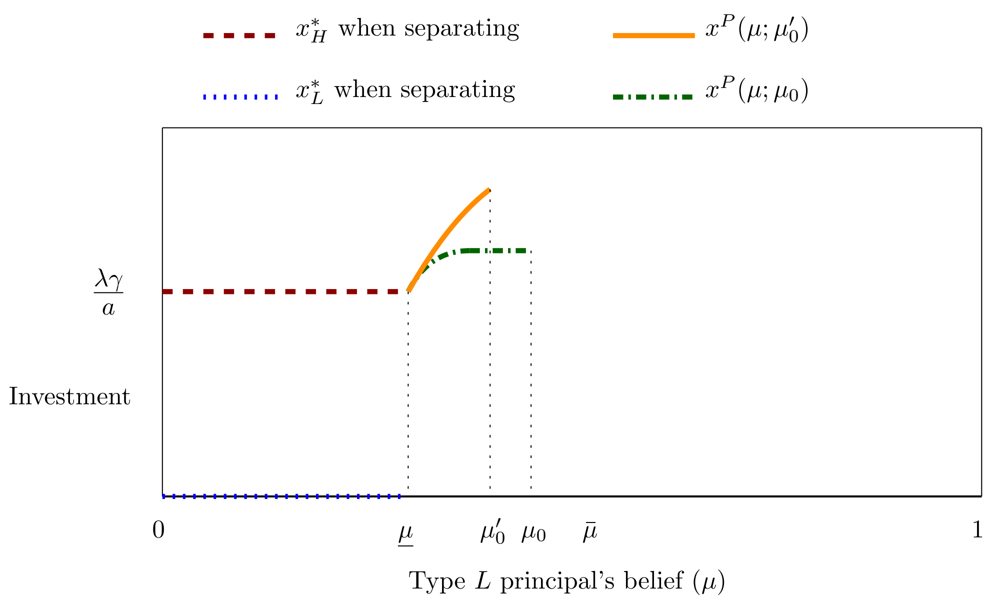
Players revert to autarky if at any point in this phase the agent assigns $\mu_+^a = 0$, except when $\mu = \mu_0$.

- ***Quitting Phase:*** *$\mu < \underline{\mu}$:*

$$
\begin{aligned}
& x_H^*(\mu, \mu^a) = \frac{\gamma\lambda}{a}; \ \ x_L^*(\mu, \mu^a) = 0; \\
& e^*(\mu, \mu_+^a = 1, x) = 1; \\
& e^*(\mu, \mu_+^a = 0, x) = 0;
\end{aligned}
$$

Agent updates beliefs using Bayes' rule on path and using the rule described in Definition 3 after observing off-path investment.

FIGURE 2.5. Equilibrium when cost of effort is high and prior is low



The principal's optimal equilibrium in this case is qualitatively similar to when $\mu_0 > \bar{\mu}$

except that the separation region does not exist and pooling starts at the outset. When $\mu_0 \in [\mu_g, \bar{\mu}]$ the agent is already pessimistic and demands to be appeased to exert effort. For the reasons argued in the preceding discussion, it is optimal to pool from the start of the relationship. However, the nature of pooling investment is different now. When pooling begins at a belief $\mu_0 < \bar{\mu}$, for the reasons pointed out in the preceding discussion, the average investment in the pooling phase must be higher than $\frac{\gamma\lambda}{a}$ in order to induce effort from the agent.[8] In particular, as μ_0 decreases, the average investment needed during pooling starting at μ_0 increases. Figure 2.5 shows that for lower initial prior $\mu_0' < \mu_0$, the average investment is higher in the pooling phase. Eventually at beliefs lower than $\mu_g > \underline{\mu}$, the investment required to induce effort is too high and the type L principal prefers not induce effort if $\mu_0 < \mu_g$. The principal's optimal pooling equilibrium investment function $x^P(\mu; \mu_0)$ is the pooling investment function that satisfies the participation constraint of both the type L principal and agent in the pooling region and maximizes the value of the type L principal at μ_0, the initial prior. As in the case with high μ_0, in the pooling phase, equilibrium is supported using the threat of autarky following a downward deviation of the principal.

2.5.3 Equilibrium distortions

Since the agent's cost of effort is strictly positive in this model, the conflict between the agent and the type L principal is inevitable. Once the belief of type L principal is low enough, the agent will demand investment higher than the optimal investment of type L in order to exert effort. As one would expect, this results in distortion i.e. sub-optimally high levels of investment of type L principal to appease the agent at low beliefs.

Perhaps interestingly, the type H principal's investment may also be distorted in equilibrium. As Proposition 12 shows, if the cost of effort is high enough and initial beliefs are in the intermediate range, in the pooling region, the type H principal invests strictly above her optimal investment on average. This results from the inability of the type H principal to separate herself from type L principal in any equilibrium. To see why, suppose that type H

[8] For any pooling investment function staring at the state $(\mu_0, 0)$, the average investment is the constant pooling investment that delivers the same value to the agent at state $(\mu_0, 0)$.

could separate herself at some belief in the pooling region, then, post separation she would invest at her optimal level and induce effort from the agent who knows that he is facing type H principal. But then, type L principal could simply mimic type H's investment, convince the agent that she is type H and induce effort by investing type H principal's optimal investment which on average is lower than the equilibrium pooling investment.

The sub-optimality of type H principal's investment during pooling is not perpetual. Once the type L principal's beliefs reach $\underline{\mu}$ and she quits the relationship, type H principal is revealed to the agent and continues to invest optimally thereafter. That is to say, the type H principal's investment may be sub-optimal in the short run, but in the long run, optimality is restored. This discussion is summarized in the proposition below.

Proposition 13. *If $\mu_0 \in [\mu_g, \bar{\mu}]$ and $c > \bar{c}$, then the type L principal invests sub-optimally in the pooling phase. However, once type L principal quits, type L principal invests optimally in perpetuity.*

Note that when cost of effort $c \leq \bar{c}$ or initial prior $\mu_0 > \bar{\mu}$, there is no distortion in type H principal's investment. When $c \leq \bar{c}$ the type L principal can induce effort by investing lower than type H principal's optimal investment whenever she wants which is why there is no distortion. When $c > \bar{c}$, and $\mu_0 > \bar{\mu}$, the intuition is as follows. Note that during the relationship, type L principal transitions to become type H if she observes output. That is, distortion in type H principal's investment also hurts the type L principal who anticipates that she could transition to become type H in future. In particular, this is indeed a concern for the type L principal at the initial belief μ_0. This concern for the type L principal at μ_0 results in no distortions for type H principal in the principal optimal equilibrium.

2.5.4 Discussion

In this model, the conflict between the principal and the agent has two sources: the cost of effort for the agent and the beliefs of the players about the quality of the project being good. Propositions 9, 11 and 12 highlight how these sources of conflict interact and shape the principal's optimal behavior in equilibrium. In particular, note that if the cost of effort for the agent is zero, there is no conflict: the agent does not need to be appeased to exert effort, and hence there is no distortion in the investment of both types of principal. When the cost of effort is low ($c < \bar{c}$), there is conflict when players are sufficiently pessimistic which results in type L principal investing sub-optimally and appeasing the agent to exert effort. However, the conflict is not enough to warrant stopping information transmission to the agent, and hence, the equilibrium is fully separating at every belief. When the cost of effort is high ($c > \bar{c}$), the conflict between players is high enough that equilibrium exhibits stoppage of information transmission at low beliefs. In this case the equilibrium exhibits information transmission at high beliefs if players are optimistic to begin with, but eventually, as the beliefs of players fall, the flow of information to the agent is stopped to preserve his motivation, anticipating the high investment costs of motivating the agent to exert effort if he becomes more pessimistic. Proposition 13 shows that if the cost of effort is high and relationship starts with low priors, even the type H principal's investment is distorted owing to her inability to credibly convince the agent of her type. However, such distortions are transient and go away in the long run.

In the model presented, the uncertainty in the environment is about the project quality. One can also interpret the uncertainty to be about the agent's ability, that is, it is the agent's ability that affects whether output is produced or not and, the principal, who is experienced and better skilled, learns about the agent's ability privately.

In organizations, workers are given feedback through periodic evaluations and appraisals. These represent a costless (cheap talk) channel through which managers provide information to workers to motivate them to exert effort. However, how workers learn about the relevant aspects of the production environment (project's quality/ agent's ability) is not limited to these periodic chunks of information. The day to day actions of a manager, how much interest they show in a workers activities by investing their resources credibly transmits the private information of the manager to the workers affecting their motivation to work. This novel channel of credible information transmission, which, if accounted for, may enhance how managers motivate their workers and better achieve organizational goals.

2.6 Extensions

2.6.1 Continuous effort choice

In this paper I assumed that the agent's effort choice is binary, either she exerts effort at a cost c or does not exert effort. I now show that if we allow the agent chose effort with a constant marginal cost, the results of this paper are unaffected. Suppose the agent can choose $e \in [0, 1]$ at a cost of ce. As before $e = 0$ is costless and $e = 1$ costs c. First note that effort has two effects, first it affects probability of arrival of output, and second it affects the speed of learning. In particular, with continuous choice of e, the rate of arrival of output $\lambda e \theta$ can be controlled more precisely by the agent. However, note that since marginal cost of effort is constant, and the decision problem of the agent at any instant is linear (due to exponential arrival rates), the optimal effort choice must be bang-bang. In other words, allowing for continuous effort with constant marginal cost has no effect on our results. However, if the marginal cost of effort is not constant, i.e. optimal effort allocation may be strictly interior, the results in this paper may not hold.

2.6.2 Breakdowns

In this paper, the observation of an output reveals to the principal that project is of good quality. That is, this is a model of fully revealing good news. Suppose instead of observing a fully revealing output, the principal privately observes breakdowns, i.e. fully revealing signals that confirm that project quality is bad. Output is hidden from both the principal and the agent. In this case note that the type L principal (the one who knows project is bad quality for sure) has no incentive to mimic the optimal investment of type H principal (the one who is optimistic since the start of the game on account of not seen a breakdown yet). Hence any equilibrium where the principal invests and the agent exerts effort must be fully separating at all beliefs of type H principal and both types invest their optimal investment. That is, there is no distortion in equilibrium.

2.7 Conclusion

This paper studies a simple principal-agent relationship with private learning where the principal faces a trade-off between optimally choosing her actions and transmitting information to the agent. There are three main takeaways. First, if the relationship is in a optimistic stage, i.e. both parties believe the relationship is likely to bear fruits in future, the principal should ignore information transmission concerns and invest optimally, i.e. the principal's actions should be sensitive to her private information. Second, if relationship moves to a pessimistic stage, or starts at a pessimistic stage, the principal should ignore her private information in choosing her actions, to prevent any information transmission and hence preserving the motivation of the agent to work. Third, if the relationship starts at a pessimistic stage, a principal who learns that the project is good, may still be forced to choose her actions sub-optimally in the short run, however, in the long run, she chooses her optimal actions.

The insights of this paper can be useful in understanding the optimal behavior of managers in organizations. In particular, feedback on project quality/agent's ability that affects the motivation of a worker is not limited to appraisals (periodic and costless information transmission), but also regularly through how much a manager actively invests in the employee (or the project he is working on) as the relationship progresses. This novel channel of credible information transmission, if accounted for, may enhance how managers motivate their workers and better achieve the objectives of their organization.

Appendix A

Appendix: When and How to Reward Bad News

A.1 Notation

Notation 1. *Let $u(\cdot|s,a)$ be the agent's value function given s and a. And $u(\cdot|s)$ be the optimal value function. Similarly, $v(\cdot|s,a)$ be the principal's value function and $v(\cdot|a)$ be his optimal value function given a. Lastly, let $v(\cdot)$ be the principal's optimal value function.*

Definition 9. *Let $\mathcal{A}$ denote the space of piecewise continuous functions from $[0,1]$ to $[0,1]$*

$\mathcal{A}$ is our space of admissible strategies for the agent.

Remark 1. *The optimal control a, for a fixed Markovian s is Markovian in p, since the evolution of state is Markovian.*

There are two Hamilton-Bellman-Jacobi (HJB) equations that underlie most of our analysis, one for the agent and one for the principal.

A.2 Agent's HJB equation and its solutions

The Agent's HJB equation is given by

$$\begin{aligned} u(p) =& 1 - \lambda_b(1-p)u(p) + \lambda_b p(1-p)u'(p) + \lambda_b(1-p)F \\ &+ \max_{a\in[0,1]} a\left\{\lambda_b(1-p)\left[u(p) - pu'(p)\right] + \lambda_g p\left[R - u(p) - (1-p)u'(p)\right] - \lambda_b(1-p)F\right\}. \end{aligned} \tag{A.1}$$

Define:

$$\Delta(p) := \lambda_b(1-p)\left[u - pu'\right] + \lambda_g p\left[R - u - (1-p)u'\right] - \lambda_b(1-p)F. \tag{A.2}$$

If the agent uses $a = 1$, the equation becomes,

$$u(p) = 1 - \lambda_g pu + \lambda_g pR - \lambda_g p(1-p)u'.$$

Its solution is,

$$u_1(p) = 1 + \frac{\lambda_g p(R-1)}{1+\lambda_g} + c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}. \tag{A.3}$$

If the agent uses $a = 0$, the equation becomes,

$$u(p) = 1 - \lambda_b(1-p)u + \lambda_b p(1-p)u' + \lambda_b(1-p)F.$$

Its solution is,

$$u_0(p) = \frac{1 + \lambda_b p + F\lambda_b(1-p)}{1+\lambda_b} + c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}. \tag{A.4}$$

A.3 Principal's HJB equation and its solutions

The principal's HJB equation is given by

$$\begin{aligned} v(p) =& -c - \lambda_b(1-p)F + \lambda_b p(1-p)v'(p) - \lambda_b(1-p)v(p) \\ &+ \max_{a\in[0,1]} a\left[\lambda_g p(\Gamma - R) + \lambda_b(1-p)F - (\lambda_b+\lambda_g)p(1-p)v'(p) - (\lambda_g p + \lambda_b(1-p))v(p)\right]. \end{aligned} \tag{A.5}$$

When $a = 1$, the solution is given by

$$v_1(p) = \frac{p\lambda_g(\Gamma - R + c)}{1+\lambda_g} - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}, \tag{A.6}$$

where C_1 is the constant of integration.

When $a = 0$,the solution is given by,

$$v_0(p) = -\frac{p\lambda_b c + c + \lambda_b F(1-p)}{\lambda_b + 1} + C_0 p \left[\frac{p}{1-p}\right]^{\frac{1}{\lambda_b}}, \tag{A.7}$$

where C_0 is the constant of integration.

Notation 2. *We denote by $u_0(\cdot; C)$ the solution to the agent HJB with $a = 0$ given by*(A.4) *with $C_0 = C$. Similar, notations apply for u_1, v_0 and v_1.*

Two other value functions that will prove to be useful is the value that the players receive (v^f for the principal and u^f for the agent) if the agent chooses $a = a^f$ (defined below) everywhere. When the agent chooses a^f, the beliefs do not move in the absence of news. We define

$$a^f = \frac{\lambda_b}{\lambda_b + \lambda_g}, \tag{A.8}$$

$$v^f(p) = \frac{-c + \mathit{\Lambda} p(\mathit{\Gamma} - R) - \mathit{\Lambda}(1-p)F}{1 + \mathit{\Lambda}}, \text{ and} \tag{A.9}$$

$$u^f(p) = \frac{\lambda_g + \lambda_b + \lambda_b \lambda_g [pR + (1-p)F]}{\lambda_g + \lambda_b + \lambda_b \lambda_g}, \tag{A.10}$$

where $\mathit{\Lambda} := \frac{\lambda_b \lambda_g}{\lambda_b + \lambda_g}$. Also, the following beliefs will turn out to be useful for later analysis. We define

$$p_m^R = \frac{c}{\lambda_g(\mathit{\Gamma} - R)}, \tag{A.11}$$

$$p_0^* = \frac{c}{\mathit{\Lambda}(\mathit{\Gamma} - R)}, \text{ and} \tag{A.12}$$

$$p_f^* = \frac{c + \mathit{\Lambda} F}{\mathit{\Lambda}[\mathit{\Gamma} - R + F]}. \tag{A.13}$$

p_m^R is the myopic experimentation cutoff, i.e., if the agent used $a = 1$ then the flow profit to the principal is 0 at p_m^R. p_0^* and p_f^* are the beliefs where $v_f(p)$ is zero with $F = 0$ and 1 respectively.

A.4 Agent's best response

Definition 10. *For any reward structure* $(R, F) \neq (1, 1)$*, define the following sets of beliefs:*

$$\bar{P} = \{p : \lambda_g p(R-1) > \lambda_b(1-p)(F-1)\}\,, \text{ and}$$
$$\underline{P} = \{p : \lambda_g p(R-1) < \lambda_b(1-p)(F-1)\}\,.$$

Using $\bar{P}$ *and* $\underline{P}$ *we then define*

$$p^f = \begin{cases} \inf \bar{P} = \sup \underline{P} & \text{if } \bar{P} \neq \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 1 & \text{if } \bar{P} = \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 0 & \text{if } \underline{P} = \emptyset \text{ and } \bar{P} \neq \emptyset. \end{cases}$$

Notice that if $p \in \bar{P} \Rightarrow [p, 1] \subset \bar{P}$ *and, if* $p \in \underline{P} \Rightarrow [0, p] \subset \underline{P}$*.*

Proof of Lemma 1: When $(R, F) \neq (1, 1)$, we show case by case.

1. If $p^f \leq \underline{p}$: By Lemma 7, for any value larger than the freezing value given by (A.10) at $\underline{p}$, the agent's best response is to use the good arm on the interior. So, the only question is can the agent receive a value strictly higher than the freezing value at $\underline{p}$. The drift of beliefs, conditional on no news must be non-negative at $\underline{p}$. But, since the agent is using the good arm to the right of $\underline{p}$, the only admissible policy with a non-negative drift is to use the freezing policy. Therefore, his value at $\underline{p}$ is $u^f(\underline{p})$ and his optimal policy at $\underline{p}$ is $a(\underline{p}) = \frac{\lambda_b}{\lambda_b + \lambda_b}$.

2. If $p^f \geq \bar{p}$: By Lemma 9, for any value larger than the freezing value given by (A.10) at $\bar{p}$, the agent's best response is to use the bad arm on the interior. So, the only question is can the agent receive a value strictly higher than the freezing value at $\bar{p}$. The drift of beliefs, conditional on no news must be non-positive at $\bar{p}$. But, since the agent is using the bad arm to the left of $\bar{p}$, the only admissible policy with a non-positive drift is to use the freezing policy. Therefore, his value at $\bar{p}$ is $u^f(\bar{p})$ and his optimal policy at $\bar{p}$ is $a(\bar{p}) = \frac{\lambda_b}{\lambda_b + \lambda_b}$.

3. If $p^f \in (\underline{p}, \bar{p})$: This case is a combination of the above two cases. We set $\underline{p} = p^f$ in case (1) and set $\bar{p} = p^f$ in case (2).

When $(R, F) = (1, 1)$, consider the class of policies for the agent where he has the following allocation

$$a(p) = \begin{cases} [0, \frac{\lambda_b}{\lambda_b + \lambda_g}] \text{ if } p = \underline{p}, \\ [\frac{\lambda_b}{\lambda_b + \lambda_g}, 1] \text{ if } p = \bar{p}. \end{cases}$$

Note that under any policy in this class, the drift of belief in the absence of a signal is non-negative when $p = \underline{p}$ and non-positive when $p = \bar{p}$. This implies that if the agent follows a policy in this class, she is never fired in the absence of a signal. $(R, F) = (1, 1)$ implies that the value of the agent under a policy in this class at any belief $p \in [\underline{p}, \bar{p}]$ is equal to 1, since the flow payoff is equal to the continuation payoff in case of a signal. Also observe that when hired, the value of the agent can never exceed 1, which implies that the agent cannot do better than any policy in this class. This establishes the best response when $(R, F) = (1, 1)$. □

Definition 11. *Suppose $p^f \leq \underline{p}$ and the agent is guaranteed to not be fired for $[\underline{p}, \bar{p}]$. And, at $\underline{p}$ he receives an exogenously specified value $u^* \geq u^f(\underline{p})$. We call the agent's problem as the auxiliary problem (a) and denote his best response as $\tilde{a}$.*

Lemma 7. *In the auxiliary problem (a), $\tilde{a}(p) = 1\ \forall p \in (\underline{p}, \bar{p})$.*

Proof. For any fixed $\underline{p} \geq p^f$, let $u^{\bar{p}}$ denote the value function for the agent for the auxiliary problem (a) $[\underline{p}, \bar{p}]$. It is obvious that $u^{\bar{p}_2} \geq u^{\bar{p}_1}$ pointwise if $\bar{p}_2 \geq \bar{p}_1$.[1] Therefore, we will first solve the auxiliary problem (a) $[\underline{p}, 1]$. Then, it is straightforward to see that for any auxiliary problem (a) $[\underline{p}, \bar{p}]$, $u^{\bar{p}} = u^1$ on $[\underline{p}, \bar{p}]$.

To this end, consider the auxiliary problem (a) $[\underline{p}, 1]$. The HJB equation for the agent

[1] This is trivially true for any $p \in (\bar{p}_1, \bar{p}_2)$. For other beliefs, a candidate policy for the agent is to use the good arm for any such belief until the beliefs hit $\bar{p}_1$, and thereafter follow the policy in the auxiliary problem (a) $[\underline{p}, \bar{p}_1)$.

is,

$$\begin{aligned}u(p) =& 1 - \lambda_b(1-p)u(p) + \lambda_b p(1-p)u'(p) \\ &+ \max_{a\in[0,1]} a\left[(\lambda_b(1-p) - \lambda_g p)u(p) - u'(p)p(1-p)[\lambda_b + \lambda_g] + \lambda_g pR - \lambda_b(1-p)F\right].\end{aligned}$$

Our candidate value function is obtained by using $a = 1$ on the interval $(\underline{p}, 1]$. So, the value function is,

$$u_1(p) = 1 + \frac{\lambda_g p(R-1)}{1+\lambda_g} + c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}},$$

where the constant is determined by the boundary condition $u(\underline{p}) = u^*$.

To prove that this is indeed the optimal value function, we need to prove that the above function satisfies the HJB equation. The key object that determines whether $a = 1$ or 0 is,

$$\Delta(p) = (\lambda_b(1-p) - \lambda_g p)u(p) - u'(p)p(1-p)[\lambda_b + \lambda_g] + \lambda_g pR - \lambda_b(1-p)F.$$

Lemma 8 establishes that if $\Delta(p) \geq 0$ for some $p \geq p^f$, for our candidate value function, then it is strictly positive for all higher beliefs. Therefore, we only need to prove that $\Delta(\underline{p}) \geq 0$. By our boundary condition,

$$\begin{aligned}u^* = u(\underline{p}) = 1 + \frac{\lambda_g \underline{p}(R-1)}{1+\lambda_g} + c_1(1-\underline{p})\left(\frac{1-\underline{p}}{\underline{p}}\right)^{\frac{1}{\lambda_g}} \\ \Rightarrow c_1(1-\underline{p})\left(\frac{1-\underline{p}}{\underline{p}}\right)^{1/\lambda_g} = u^* - 1 - \frac{\lambda_g \underline{p}(R-1)}{1+\lambda_g}.\end{aligned}$$

$$\begin{aligned}\Delta(p) &= \lambda_b(1-p)(1-F) + c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right] + \frac{\lambda_g p(R-1)}{1+\lambda_g}. \\ \Delta(\underline{p}) &= \lambda_b(1-\underline{p})(1-F) + \left[u^* - 1 - \frac{\lambda_g \underline{p}(R-1)}{1+\lambda_g}\right]\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right] + \frac{\lambda_g \underline{p}(R-1)}{1+\lambda_g} \\ &= 0 \quad \text{if } u^* = u^f(\underline{p}).\end{aligned}$$

Therefore, for any terminal $u^* \geq u^f(\underline{p})$, $\Delta \geq 0$, and therefore, it is strictly positive on

$(\underline{p}, 1)$. That is, our candidate value function satisfies the HJB equation and the boundary condition and, therefore, is the optimal value function for the problem $[\underline{p}, \bar{p}]$. Moreover, the constructed value function does not depend on $\bar{p}$, and therefore, is the optimal value function for all $\bar{p} > \underline{p}$. $\square$

Define,

$$\Delta_1(p) := \lambda_b(1-p)\left[u_1 - pu_1'\right] + \lambda_g p\left[R - u_1 - (1-p)u_1'\right] - \lambda_b(1-p)F,$$

where u_1 is defined in (A.3).

Lemma 8. *If $\hat{p} \in \bar{P}$ and $\Delta(\hat{p}) = \Delta_1(\hat{p}) = 0$ then $\Delta(p) > 0\ \forall p > \hat{p}$.*

Proof. Plugging in u_1 and u_1' we get

$$\begin{aligned}
\Delta_1(p) =& \lambda_b(1-p)\left[1 + \frac{\lambda_g p(R-1)}{1+\lambda_g} + c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}} - \frac{\lambda_g p(R-1)}{1+\lambda_g}\right] \\
&+ \lambda_b(1-p)c_1\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\frac{1+\lambda_g p}{\lambda_g} \\
&+ \lambda_g p\left[R - 1 - \frac{\lambda_g p(R-1)}{1+\lambda_g} - c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\right] \\
&- \lambda_g p(1-p)\left[\frac{\lambda_g(R-1)}{1+\lambda_g} - c_1\left[\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\frac{1+\lambda_g p}{\lambda_g p}\right]\right] - \lambda_b(1-p)F \\
=& \lambda_b(1-p)(1-F) + c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right] + \frac{\lambda_g p(R-1)}{1+\lambda_g}.
\end{aligned}$$

Therefore,

$$\Delta_1'(p) = -\lambda_b(1-F) - c_1\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right]\left[\frac{1+\lambda_g p}{\lambda_g p}\right] + \frac{\lambda_g(R-1)}{1+\lambda_g}.$$

Consider a $p_1 \in \bar{P}$ such that $\Delta_1(p_1) = 0$, Then we have

$$\begin{aligned}
\Delta_1(p_1) = 0 =& \lambda_b(1-p_1)(1-F) + c_1(1-p_1)\left(\frac{1-p_1}{p_1}\right)^{\frac{1}{\lambda_g}}\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right] + \frac{\lambda_g p_1(R-1)}{1+\lambda_g} \\
\Rightarrow & c_1\left(\frac{1-p_1}{p_1}\right)^{\frac{1}{\lambda_g}}\left[\frac{\lambda_b + \lambda_g + \lambda_b\lambda_g}{\lambda_g}\right] = -\lambda_b(1-F) - \frac{\lambda_g p_1(R-1)}{(1+\lambda_g)(1-p_1)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}\mathit{\Delta}_1'(p_1) &= -\lambda_b(1-F) + \left[\lambda_b(1-F) + \frac{\lambda_g p_1(R-1)}{(1+\lambda_g)(1-p_1)}\right]\left[\frac{1+\lambda_g p_1}{\lambda_g p_1}\right] + \frac{\lambda_g(R-1)}{1+\lambda_g} \\ &= \frac{\lambda_b(1-F)}{\lambda_g p_1} + \frac{R-1}{1-p_1} > 0 \text{ since } p_1 \in \bar{P}.\end{aligned}$$

This shows that $\mathit{\Delta}_1(p)$ is strictly increasing at p_1. By continuity of $\mathit{\Delta}_1'(p)$, we have that $\mathit{\Delta}_1(p)$ is strictly increasing in some ϵ neighborhood of p_1. Note that for all $p > p_1$, $p \in \bar{P}$, which concludes the proof. □

Definition 12. *Suppose $p^f \geq \bar{p}$ and the agent is guaranteed to not be fired for $[\underline{p}, \bar{p}]$. And, at $\bar{p}$ he receives an exogenously specified value $u^* \geq u^f(\bar{p})$. We call the agent's problem as the auxiliary problem (b) and denote his best response as $\tilde{a}$.*

Lemma 9. *In the auxiliary problem (b), $\tilde{a}(p) = 0$ $\forall p \in (\underline{p}, \bar{p})$.*

Proof. For any fixed $\bar{p} \leq p^f$, let $u^{\underline{p}}$ denote the value function for the agent for the auxiliary problem (b) $[\underline{p}, \bar{p}]$. It is obvious that $u^{\underline{p}_1} \geq u^{\underline{p}_2}$ pointwise if $\bar{p}_1 \leq \bar{p}_2$.[2] Therefore, we will first solve the auxiliary problem (b)$[0, \bar{p}]$. Then, it is straightforward to see that for any auxiliary problem (b)$[\underline{p}, \bar{p}]$, $u^{\underline{p}} = u^0$ on $[\underline{p}, \bar{p}]$.

To this end, consider the auxiliary problem (a) $[0, \bar{p}]$. The HJB equation for the agent is,

$$\begin{aligned}u(p) =& 1 - \lambda_b(1-p)u(p) + \lambda_b p(1-p)u'(p) \\ &+ \max_{a\in[0,1]} a\left[(\lambda_b(1-p) - \lambda_g p)u(p) - u'(p)p(1-p)[\lambda_b + \lambda_g] + \lambda_g pR - \lambda_b(1-p)F\right].\end{aligned}$$

Our candidate value function is obtained by using $a = 0$ on the interval $[0, \bar{p})$. So, the value function is,

$$u_0(p) = \frac{1 + \lambda_b p + F\lambda_b(1-p)}{1+\lambda_b} + c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}},$$

[2]This is trivially true for any $p \in (\underline{p}_1, \underline{p}_2)$. For other beliefs, a candidate policy for the agent is to use the bad arm for any such belief until the beliefs hit $\underline{p}_2$, and thereafter follow the policy in the auxiliary problem (b) $(\underline{p}_2, \bar{p}]$.

where the constant is determined by the boundary condition $u(\bar{p}) = u^*$.

To prove that this is indeed the optimal value function, we need to prove that the above function satisfies the HJB equation. The key object that determines whether $a = 1$ or 0 is,

$$\varDelta(p) = (\lambda_b(1-p) - \lambda_g p)u(p) - u'(p)p(1-p)[\lambda_b + \lambda_g] + \lambda_g pR - \lambda_b(1-p)F.$$

Lemma 10 establishes that if $\varDelta(p) \leq 0$ for some $p \leq p^f$, for our candidate value function, then it is strictly negative for all lower beliefs. Therefore, we only need to prove that $\varDelta(\bar{p}) \leq 0$. By our boundary condition,

$$\begin{aligned} u^* = u(\bar{p}) &= \frac{1 + \lambda_b \bar{p} + F\lambda_b(1-\bar{p})}{1+\lambda_b} + c_0 p \left(\frac{\bar{p}}{1-\bar{p}}\right)^{\frac{1}{\lambda_b}} \\ \Rightarrow c_0 p \left(\frac{\bar{p}}{1-\bar{p}}\right)^{\frac{1}{\lambda_b}} &= u^* - \frac{1 + \lambda_b \bar{p} + F\lambda_b(1-\bar{p})}{1+\lambda_b}. \end{aligned}$$

$$\begin{aligned} \varDelta(p) &= \frac{\lambda_b(1-p)(1-F)}{1+\lambda_b} + \lambda_g p(R-1) - c_0 p \left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}} \left[\frac{\lambda_g + \lambda_b + \lambda_g \lambda_b}{\lambda_b}\right]. \\ \varDelta(\bar{p}) &= \frac{\lambda_b(1-\bar{p})(1-F)}{1+\lambda_b} + \lambda_g p(R-1) - \left[u^* - \frac{1 + \lambda_b \bar{p} + F\lambda_b(1-\bar{p})}{1+\lambda_b}\right] \left[\frac{\lambda_b + \lambda_g + \lambda_b \lambda_g}{\lambda_g}\right] \\ &= 0 \quad \text{if } u^* = u^f(\bar{p}). \end{aligned}$$

Therefore, for any terminal $u^* \geq u^f(\bar{p})$, $\varDelta \leq 0$, and therefore, it is strictly negative on $(0, \bar{p})$.

That is, our candidate value function satisfies the HJB equation and the boundary condition and, therefore, is the optimal value function for the problem $[\underline{p}, \bar{p}]$. Moreover, the constructed value function does not depend on $\underline{p}$, and therefore, is the optimal value function for all $\underline{p} < \bar{p}$. $\square$

Define:

$$\varDelta_0(p) := \lambda_b(1-p)\left[u_0 - pu_0'\right] + \lambda_g p\left[R - u_0 - (1-p)u_0'\right] - \lambda_b(1-p)F,$$

where u_0 is defined in (A.4).

Lemma 10. *If $\hat{p} \in \underline{P}$ and $\Delta(\hat{p}) = \Delta_0(\hat{p}) = 0$ then $\Delta(p) < 0$ $\forall p < \hat{p}$.*

Proof. Plugging in u_0 and u_0' we get

$$\begin{aligned}\Delta_0(p) =& \lambda_b(1-p)\left[\frac{1+\lambda_b p + F\lambda_b(1-p)}{1+\lambda_b} + c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\right] \\ &- \lambda_b(1-p)p\left[\frac{\lambda_b(1-F)}{1+\lambda_b} - c_0\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\left[\frac{1+\lambda_b(1-p)}{\lambda_b(1-p)}\right]\right] \\ &+ \lambda_g p\left[R - \frac{1+\lambda_b p + F\lambda_b(1-p)}{1+\lambda_b} - c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\right] \\ &- \lambda_g p(1-p)\left[\frac{\lambda_b(1-F)}{1+\lambda_b} + c_0\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\left[\frac{1+\lambda_b(1-p)}{\lambda_b(1-p)}\right]\right] - \lambda_b(1-p)F.\end{aligned}$$

This can be simplified to

$$\Delta_0(p) = \frac{\lambda_b(1-p)(1-F)}{1+\lambda_b} + \lambda_g p(R-1) - c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\left[\frac{\lambda_g+\lambda_b+\lambda_g\lambda_b}{\lambda_b}\right].$$

And we differentiate $\Delta_0(p)$ to obtain

$$\Delta_0'(p) = -\frac{\lambda_b(1-F)}{1+\lambda_b} + \lambda_g(R-1) - c_0\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\left[\frac{\lambda_b+\lambda_g+\lambda_b\lambda_g}{\lambda_b}\right]\left[\frac{1+\lambda_b(1-p)}{\lambda_b(1-p)}\right].$$

Consider $p_0 \in \underline{P}$ such that $\Delta_0(p_0) = 0$, Then we have

$$\Delta_0(p_0) = 0 = \frac{\lambda_b(1-p_0)(1-F)}{1+\lambda_b} + \lambda_g p_0(R-1) - c_0 p_0\left(\frac{p_0}{1-p_0}\right)^{\frac{1}{\lambda_b}}\left[\frac{\lambda_g+\lambda_b+\lambda_g\lambda_b}{\lambda_b}\right],$$

which gives

$$c_0\left(\frac{p_0}{1-p_0}\right)^{\frac{1}{\lambda_b}}\left[\frac{\lambda_g+\lambda_b+\lambda_g\lambda_b}{\lambda_b}\right] = \lambda_g(R-1) + \frac{\lambda_b(1-p_0)(1-F)}{p_0(1+\lambda_b)}.$$

Now we evaluate $\varDelta_0'(p_0)$ by plugging in the above expression

$$\begin{aligned}\varDelta_0'(p_0) &= -\frac{\lambda_b(1-F)}{1+\lambda_b} + \lambda_g(R-1) - \left[\lambda_g(R-1) + \frac{\lambda_b(1-p_0)(1-F)}{p_0(1+\lambda_b)}\right]\left[\frac{1+\lambda_b(1-p_0)}{\lambda_b(1-p_0)}\right] \\ &= -\frac{\lambda_g}{\lambda_b}\left[\frac{\lambda_b(1-F)}{\lambda_g p} + \frac{R-1}{1-p}\right] > 0 \text{ since } p_0 \in \underline{P}.\end{aligned}$$

This shows that $\varDelta_0(p)$ is strictly increasing at p_0 which then by continuity of $\varDelta_0'(p)$ implies that $\varDelta_0(p)$ is strictly increasing in some ϵ neighborhood of p_0. Note that for all $p < p_0$, $p \in \underline{P}$, which concludes the proof.

□

A.5 Reduction of principal's problem

Now, we focus on the case where the agent receives a reward $R \geq 1$ upon producing a success and $F \geq 0$ upon producing a failure. For most of this part, we will describe the principal's policies by a triple $(\hat{p}, R, F)$ where $\hat{p}$ denotes the firing cutoff.[3] Let $v(\cdot|R, F)$ denote the principal's optimal value function for a fixed R, F. Lemma 1 provided the agent's best response in this case. Notice that, for any $R, F > 1$, qualitatively, the best response for the agent takes the following form: Bad news arm below a certain cutoff and good news arm above it. Recall (Definition 10) that the switching belief, p^f, was determined by the objects below.

$$\bar{P} := \{p : \lambda_g p(R-1) > \lambda_b(1-p)(F-1)\}.$$

$$\underline{P} := \{p : \lambda_g p(R-1) < \lambda_b(1-p)(F-1)\}.$$

$$p^f := \begin{cases} \inf \bar{P} = \sup \underline{P} & \text{if } \bar{P} \neq \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 1 & \text{if } \bar{P} = \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 0 & \text{if } \underline{P} = \emptyset \text{ and } \bar{P} \neq \emptyset. \end{cases}$$

[3] We can also consider more general hiring policies as in the previous section but we focus on the principal-optimal policies for the sake of exposition. It is straightforward to extended results like the connectedness of the hiring interval to this setting.

Notice that since $R \geq 1$, $\bar{P}$ is always non-empty. Moreover, if $F < 1$, then $\underline{P}$ is always empty. Therefore, the switching belief is given by,

$$p^f := \begin{cases} 0 & \text{if } F < 1, \\ \frac{\lambda_b(F-1)}{\lambda_g(R-1)+\lambda_b(F-1)} & \text{if } \frac{\lambda_b(F-1)}{\lambda_g(R-1)+\lambda_b(F-1)} \in (0,1). \end{cases}$$

A key observation from the agent's best response, therefore, is the following: For any R, F such that $(R, F > 1)$, we can choose an $1 < R' < R, 1 < F' < R$ while keeping the agent's behavior unchanged. Obviously, this increases the principal's profits for any belief p. This observation is summarized below.

Lemma 11. *For any principal policy $(\hat{p}, R, F)$ with $(R, F) > (1, 1)$, we can choose (R', F') such that*

(i) $(R, F) > (R', F') > (1, 1)$,

(ii) The agent's behavior is unchanged,

(iii) $v(p|R'F') > v(p|R, F)$ *whenever* $v(\cdot|R, F) > 0$. *Moreover,*

$$\{p : v(p|R, F) > 0\} \subset \{p : v(p|R', F') > 0\}.$$

Proof. It is obvious that we can choose (R', F') such that $(R, F) > (R', F') > (1, 1)$ such that

$$\frac{\lambda_b(F-1)}{\lambda_g(R-1)+\lambda_b(F-1)} = \frac{\lambda_b(F'-1)}{\lambda_g(R'-1)+\lambda_b(F'-1)}.$$

Therefore, preserving the agent's behavior by choosing a lower value of R and F is straightforward. Obviously, this unambiguously helps the principal (since both are costs for the principal) in increasing his profits and (iii) follows.

□

The only remaining case is when $F < 1$. Here, the agent's best response, by Lemma 1, is to use the good news arm at all beliefs except $\hat{p}$, where the agent uses a^f. By the exact same arguments as above, we can reduce F to 0 in particular, to keep the agent behavior unchanged while strictly improving upon the principal's profits. This observation is summarized in Lemma 12, whose proof we omit.

Lemma 12. *For any principal policy $(\hat{p}, R, F)$ such that $R \geq 1$, $F \in (0,1)$, we can choose an (R', F') such that,*

(i) $R' \leq R$, $0 = F' < F$,

(ii) The agent's behavior is unchanged.

(iii) $v(p|R'F') > v(p|R,F)$ *whenever* $v(\cdot|R,F) > 0$. *Moreover,*

$$\{p : v(p|R,F) > 0\} \subset \{p : v(p|R',F') > 0\}.$$

Proof of Proposition 1: Lemma 11 and lemma 12 together imply that in any principal-optimal equilibrium, we must have $R = 1$ and $F \in \{0,1\}$. □

A.6 Principal's problem when R=1, F=0

If $R = 1, F = 0$, then $p^f = 0$. We obtain the agent's best response as a corollary of Lemma 1.

Proposition 14. *Suppose the principal hires the agent when $p \in [\underline{p}, \bar{p}] \subset [0,1]$ and fires otherwise. Then the best response of the agent is given as:*

$$a(p) = \begin{cases} [0,1] \text{ if } p \notin [\underline{p}, \bar{p}], \\ \frac{\lambda_b}{\lambda_b + \lambda_g} \text{ if } p = \underline{p}, \\ 1 \text{ if } p \in (\underline{p}, \bar{p}]. \end{cases}$$

A.6.1 Characterization of equilibrium when R=1,F=0

Definition 13. *Given a strategy for the principal s, define H as the set of beliefs at which the principal hires, i.e. $H = \{p : s(p) = 0\}$.*

Lemma 13. *In any equilibrium, H is a connected set.*

Proof. Suppose $H = [\underline{p}_1, \bar{p}_1] \cup [\underline{p}_2, \bar{p}_2]$ with $\underline{p}_2 > \bar{p}_1$. Suppose that the best response of the agent is a. We know that

$$a(p) = \begin{cases} \frac{\lambda_b}{\lambda_b + \lambda_g} \text{ if } p \in \{\underline{p}_1, \underline{p}_2\}, \\ 1 \text{ if } p \in (\underline{p}_1, \bar{p}_1] \cup (\underline{p}_2, \bar{p}_2]. \end{cases}$$

However the best response of the agent in the interval $(\bar{p}_1, \underline{p}_2)$ can be any function satisfying Assumption 1. Also, note that the value of the principal from this strategy profile at $\bar{p}_1$ is $V(\bar{p}_1) > V_f(\bar{p}_1) > 0$, and the value of the principal from this strategy profile at $\underline{p}_2$ is $V(\underline{p}_2) = V_f(\underline{p}_2) > 0$.

We now show that for any behavior of the agent in the interval $(\bar{p}_1, \underline{p}_2)$, satisfying Assumption 1, the principal has an incentive to deviate. Notice that $V^f(p) > 0$ for all $p > \underline{p}_1$, where V^f is the value to the principal if the agent were to freeze beliefs everywhere. Therefore, for us to have an equilibrium where the principal hires on two disjoint intervals, i$a(p) \neq a^f$ for any $p \in (\bar{p}_1, \underline{p}_2)$. We will consider three cases regarding the limit of the strategy on the firing interval at $\bar{p}_1$ and $\underline{p}_2$.[4]

1. $\lim_{p \downarrow \bar{p}_1} a(p) < a^f$ and $\lim_{p \uparrow \underline{p}_2} a(p) > a^f$. In this case, the drift of beliefs is positive on $(\bar{p}_1, p)$ for some $p > \bar{p}_1$ and is negative $(p, \underline{p}_2)$ for some $p < \underline{p}_2$. Therefore, it must be 0 for some $\tilde{p} \in (\bar{p}_1, \underline{p}_2)$, i.e. $a(\tilde{p}) = a^f$. Since $\tilde{p} > \underline{p}_1$ where the principal obtains a non-negative value with agent using a^f, therefore, the principal obtains a strictly positive value at $\tilde{p}$ if he were to deviate not fire. A contradiction.

2. $\lim_{p \downarrow \bar{p}_1} a(p) > a^f$ or $\lim_{p \uparrow \underline{p}_2} a(p) < a^f$. We will argue only for the case $\hat{a} := \lim_{p \downarrow \bar{p}_1} a(p) > a^f$, as the argument for the other case is similar. Notice that $\hat{a} > a^f \Rightarrow p[\lambda_g a(p) - \lambda_b(1 - a(p))] > \epsilon$ for some $\epsilon > 0$ for $p \in (\bar{p}_1, \bar{p}_1 + \delta]$ for some $\delta > 0$. That is, the drift of beliefs, in the absence of news, is strictly negative and bounded from above by $-\epsilon$. Define, $\tau(p) := \inf\{t > 0 : P_t \in \{0, \bar{p}_1, 1\}\}$. Since the drift is bounded away from 0, $\lim_{p \downarrow \bar{p}_1} \tau(p) = 0$ a.s and $\mathbb{P}(P_{\tau(p)} = \bar{p}_1 | P_0 = p) = 1$. If the principal deviates to

[4]Since a is piecewie continuous, these limits exist.

continue until the beliefs hit either 0, 1 or $\bar{p}_1$, where he collects $V(\bar{p}_1)$, then his payoff from such a policy is,

$$V(p) = \mathbb{E}\left[(1-e^{-\tau(p)})(-c) + e^{-\tau(p)}\left(\mathbb{P}(P_{\tau(p)}=1)V(1) + \mathbb{P}(P_{\tau(p)}=\bar{p}_1)V(\bar{p}_1)\right)\right].$$

As $p \downarrow \bar{p}_1$, $V(p) \to V(\bar{p}_1) > 0$. Therefore, the principal would strictly prefer continuing and not firing just above $\bar{p}_1$, a contradiction.

3. $\lim_{p\downarrow\bar{p}_1} a(p) = a^f$ or $\lim_{p\uparrow\underline{p}_2} a(p) = a^f$ We will argue for the case $\lim_{p\uparrow\underline{p}_2} a(p) = a^f$, as the other case is straightforward. First of all, if $a(p) = a^f$ for any $p \in (\bar{p}_1, \underline{p}_2)$, we are done. So, assume wlog that $a(p) < a^f$ for all $p(\bar{p}_1, \underline{p}_2)$ with $\lim_{p\uparrow\underline{p}_2} a(p) = a^f$. Therefore, for any $\epsilon > 0$, $\exists\delta > 0$ such that, $a^f > a(p) > a^f - \epsilon$ for all $p \in (\underline{p}_2 - \delta, \underline{p})$. For any such p, if the principal were to deviate and not fire until the belief hits 0, 1 or $\underline{p}_2$ where he receives $V^f(\bar{p}_2) > 0$, his payoff is,

$$V(p) = \mathbb{E}\left[(1-e^{-\tau})(-c) + e^{-\tau}V(P_\tau)\right],$$

where $\tau := \inf\{t : P_t \in \{0, 1, \underline{p}_2\}\}$. Notice that, in the absence of a signal, the law of motion for beliefs is,

$$\begin{aligned}
&\mathrm{d}P_t = P_t(1-P_t)[\lambda_b(1-a_t) - \lambda_g a_t]\mathrm{d}t \\
&\Rightarrow \log\left(\frac{P_t}{1-P_t}\right) - \log\left(\frac{P_0}{1-P_0}\right) = \int_0^t [\lambda_b(1-a_t) - \lambda_g a_t]\mathrm{d}t.
\end{aligned}$$

Notice that $[\lambda_b(1-a_t) - \lambda_g a_t] > 0$. for all $t \leq \tau$. Therefore, there is a unique time t^* where the beliefs will reach $\underline{p}_2$ conditional on no signal. Moreover, $\tau \leq t^*$ almost surely. Therefore,

$$\mathbb{E}e^{-\tau}V(P_\tau) = \int_0^{t^*} \mathbb{P}(\tau = t \cap P_\tau = 1)e^{-t}V(1)\mathrm{d}t + \mathbb{P}(\tau = t^*)e^{-t^*}V(\underline{p}_2).$$

$$\mathbb{P}(\tau = t \cap P_\tau = 1) = p\lambda_g a_t \exp\left(-\int_0^t \lambda_g a_u \mathrm{d}u\right) \quad \text{if } \tau < t^*.$$
$$\mathbb{P}(\tau = t^*) = p\exp\left(-\int_0^{t^*} \lambda_g a_u \mathrm{d}u\right) + (1-p)\exp\left(-\int_0^{t^*} \lambda_b (1-a_u)\mathrm{d}u\right).$$

Since $|a_u - a^f| < \epsilon$ for all $u \leq \tau$, and since all the integrals are bounded, it is easy to see that all the quantities above are at most $K\epsilon$ away, for some positive constant K, (ignoring the higher order terms) from using $a_u = a^f$ for all u. Therefore, $|V(p) - V^f(p)| < M\epsilon$ for some constant M. For a small enough ϵ, this implies that $V(p) > 0$ since $V^f(p) > 0$ for all $p \in (\underline{p}_1, 1]$. Therefore, the principal would strictly prefer hiring for some $p \in (\underline{p}_2 - \delta, \underline{p}_2)$, a contradiction.

□

Now, we prove that p_0^* (defined in (A.12)) is the unique candidate for the lower cutoff in equilibrium.

Lemma 14. *In any equilibrium, $\underline{p} = p_0^*$.*

Proof. Suppose the principal hires on $[\underline{p}, \bar{p}]$ and the agent's best response to this hiring strategy is a. The agent chooses $a = a^f$ at $\underline{p}$. Therefore, $\underline{p} \geq p_0^*$. If not, the principal receives a strictly negative value at $\underline{p}$, a contradiction. Suppose $\underline{p} > p_0^*$. $\Rightarrow v(\underline{p}|a) > 0$. If $a(p) = a^f$ for any $p \in (p_0^*, \underline{p})$, the principal would strictly prefer hiring at such a p, contradicting that the hiring region is $[\underline{p}, \bar{p}]$. Since $a(\cdot)$ is piecewise continuous, it is continuous on $(\underline{p} - \epsilon, \underline{p})$ for some $\epsilon > 0$. Suppose $a(p) < a^f$ when $p \in (\underline{p} - \epsilon, \underline{p})$. By an argument analogous to Case 2 and 3 in Lemma 13, the principal would strictly benefit by lowering the firing cutoff from $\underline{p}$. Therefore, it cannot be an equilibrium. On the other hand, suppose $a(p) > a^f$ when $p \in (\underline{p} - \epsilon, \underline{p})$. Choose a $p \in (\underline{p} - \epsilon, \underline{p})$ such that, $a(q) > a^f + \delta$ for some $\delta > 0$ for all q in the neighborhood of p. Suppose the principal deviates to hire on a small interval $(p - \eta, p]$. Given the agent's strategy, the principal's payoff is,

$$v(p|a) = -c(1 - \exp(-\tau)) + \exp(-\tau))p(1 - \exp\left(-\int_0^\tau \lambda_g a(P_t)\mathrm{d}t\right)(\mathit{\Gamma} - R),$$

where $\tau = \inf\{t : P_t^p \notin (p-\eta, p)\}$ where P_t^p denotes the stochastic process with the initial state as p. As $\eta \to 0$, $\tau \to 0$ a.s. and, we have,

$$v(p|a) \approx -c\tau + \lambda_g p a(p)(\mathit{\Gamma} - R)\tau > 0 \quad \text{when } p > p_0^* \text{ and } a(p) > a^f.$$

Therefore, the principal would prefer hiring on $(p-\eta, p)$, a contradiction. □

Lemma 15. *All equilibria where the agent uses $a = 0$ in the firing region are characterized by a belief $\bar{p}$, $\bar{p} \in [p_0^*, 1]$. The players' strategies are given by:*

$$a(p) = \begin{cases} 0 \ \textit{if } p \in [0, p_0^*) \cup (\bar{p}, 1], \\ a^f \ \ \textit{if } p = p_0^*, \\ 1, \ \textit{if } p \in (p_0^*, \bar{p}]. \end{cases} \qquad s(p) = \begin{cases} 1, \ \textit{if } p \in [0, p_0^*) \cup (\bar{p}, 1], \\ 0, \ \textit{if } p \in [p_0^*, \bar{p}]. \end{cases}$$

When $\lambda_b < \hat{\lambda}_b$ there does not exist an equilibrium where the agent is hired at any interior belief.

Proof. In any equilibrium the hiring region is of the form $[\underline{p}, \bar{p}]$ by Lemma 13. Agent's best response is given by Proposition 14. For the lower firing cutoff, $\underline{p} \geq p_0^*$ because otherwise the principal receives a strictly negative value at $\underline{p}$. Strict inequality is not possible because the agent uses the bad arm below $\underline{p}$, and therefore, the principal would like to lower the cutoff if $\underline{p} > p_0^*$. That the principal would continue hiring above p_0^* is immediate from Lemma 16 and that $v(p)$ is increasing. □

Proposition 15. *The principal-optimal equilibrium, which features the same on path behavior, is the following:*

$$a(p) = \begin{cases} [0, a^f] & \textit{if } p \in [0, p_0^*), \\ a^f & \textit{if } p = p_0^*, \\ 1 & \textit{if } p \in (p_0^*, 1]. \end{cases} \qquad s(p) = \begin{cases} 1 & \textit{if } p \in [0, p_0^*), \\ 0 & \textit{if } p \in [p_0^*, 1]. \end{cases}$$

Proof. By Lemma 14, the lower cutoff is uniquely pinned down. Therefore, the principal-optimal equilibrium, and also the unique Pareto optimal equilibrium, would be one with the largest hiring region, i.e. $\bar{p} = 1$. Therefore, all we need to prove is that the above strategy is in fact an equilibrium. Proposition 14 shows that the agent's strategy in the candidate equilibrium is indeed the best response to the principal's strategy. We need to show that principal's strategy is the best response to the agent's strategy in the candidate equilibrium. We split the proof in two cases:

Suppose, $p \in (p_0^*, 1]$: In this region the agent uses the good news arm exclusively. The principal's value function is given by:

$$v(p) = \lambda_g \gamma p - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}.$$

Where C_1 is the constant of integration that is determined using $v(p_0^*) = 0$.

$$\lambda_g \gamma p_0^* - c + C_1(1-p_0^*)\left[\frac{1-p_0^*}{p_0^*}\right]^{\frac{1}{\lambda_g}} = 0.$$

Lemma 16 tells us that $v'_+(p_0^*) > 0$ and $v'(p) > 0$ when $p \in (p_0^*, 1]$. This implies that $v(p) > 0$ for all $p > p_0^*$. This establishes that $s(p) = 0$ is the best response for all $p > p_0^*$.

□

Lemma 16. *Denote by $v'_+(p_0^*)$ the right hand derivative of $v(p)$ at p_0^*. Then $v'_+(p_0^*) > 0$ and $v'(p) > 0$ when $p \in (p_0^*, \bar{p}]$.*

Proof. When the good news arm is used ($a = 1$), the differential equation governing the value of the principal is given by

$$v_1(p) = \lambda_g p(\mathit{\Gamma} - 1) - c - \lambda_g p v_1(p) - \lambda_g p(1-p)v'_1(p).$$

Since the boundary condition at p_0^* dictates that $v(p_0^*) = 0$, we have $v(p) = v_1(p; C_1^*)$ when $p \in [p_0^*, 1]$ where C_1^* is determined by setting $v_1(p_0^*) = 0$.[5] Next, note that the right hand derivative $v'_+(p_0^*)$ is given by

[5]The function $v_1(p; C)$ denotes the value function v_1 with $C_1 = C$, as defined in Notation 2.

$$v'_+(p_0^*) = v'_1(p_0^*) = \frac{\lambda_g p_0^*(\mathit{\Gamma}-1)-c}{\lambda_g p_0^*(1-p_0^*)}.$$

Note that

$$\lambda_g p_0^*(\mathit{\Gamma}-1)-c = \frac{\lambda_g c(\mathit{\Gamma}-1)}{\mathit{\Lambda}(\mathit{\Gamma}-1)} - c = c\left[\frac{\lambda_g}{\mathit{\Lambda}}-1\right] > 0,$$

which implies that $v'_+(p_0^*) > 0$. The principal's value function when $p \in [p_0^*, \bar{p}]$ is given by (A.6) is as follows

$$v(p) = v_1(p) = \frac{p\lambda_g(\mathit{\Gamma}-R+c)}{1+\lambda_g} - c + C_1^*(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}.$$

We differentiate once to get

$$v'_1(p) = \frac{\lambda_g(\mathit{\Gamma}-1+c)}{1+\lambda_g} - C_1^*\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}\left[\frac{\frac{1}{\lambda_g}+p}{p}\right],$$

and twice to get

$$v''_1(p) = C_1^*\left[\frac{1+\lambda_g}{\lambda_g^2 p^2(1-p)}\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\right].$$

Now suppose $C_1^* \leq 0$, clearly this means that $v'_1(p) > 0$ since $\mathit{\Gamma}-1+c > 0$. On the contrary, if $C_1^* > 0$, we know that $v''_1(p) > 0$, and since $v'_1(p_0^*) > 0$, $v'_1(p) > 0$ for $p \in (p_0^*, 1]$. Since $v(p) = v_1(p)$ for $p \in (p_0^*, \bar{p}]$, we have $v'(p) > 0$ when $p \in (p_0^*, \bar{p}]$ and $v'_+(p_0^*) > 0$. □

Going forward, to keep track of the principal value function for the case of $F = 0$, we will denote it by $v_*^{F=0}(\cdot)$. That is,

$$v_*^{F=0}(p) = \lambda_g \gamma p - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}, \tag{A.14}$$

where C_1 is the constant of integration that is determined using $v_*^{F=0}(p_0^*) = 0$.

A.7 Principal's problem when R=F=1

When $R = F = 1$, the agent is indifferent across all policies with the only restriction being that at the left (right) endpoint of the hiring interval the drift of beliefs must be non-negative (non-positive). So, supposing that the hiring interval is of the form $[\hat{p}, 1]$, we want to find the optimal firing cutoff and the agent policy for the principal satisfying the following two:

1. $a(p)$ is piecewise continuous.
2. $a(\hat{p}) \leq a^f$.

Principal's problem is,

$$v_*(p) = \sup_{\{a,\hat{p}\}} \mathbb{E}^a \left[(1 - e^{-\tau})(-c) + e^{-\tau} v(P_\tau) | P_0 = p\right],$$
$$\text{subject to } a(\hat{p}) \leq a^f \text{ and } a \text{ is piecewise continuous,}$$

such that, $\tau := \inf\{t : P_t \notin [\hat{p}, 1)\}$, $v(0) = -F = -1, v(1) = \mathit{\Gamma} - R = \mathit{\Gamma} - 1$ and $v(\hat{p}) = 0$.

Let the optimal stopping belief be $\hat{p}_*$ and the associated optimal policy in the continue region be denoted by a^*. We will first conjecture that the optimal policy is to use the bad news arm on $[\hat{p}_*, p^s]$[6] and good news arm for higher beliefs.

To this end, let us suppose that the firing cutoff is exogenously specified to be some $\hat{p}$ and the associated optimal policy be denoted by $a^{\hat{p}}$ and the associated value function be $v^{\hat{p}}$. We conjecture that the optimal policy is to use the bad news arm below some p^s and good news arm above it. We will find the optimal p^s within such policies and then argue that it is indeed optimal across all the policies. The optimal p^s for the principal is calculated using the value matching and smoothpasting conditions for v_f and v_1 (or equivalently v_f and v_0).

So, the conjectured value function is (where $R = F = 1$),

$$v(p) = \begin{cases} v_0(p) = -\frac{p\lambda_b c + c + \lambda_b F(1-p)}{\lambda_b + 1} + C_0^s p \left[\frac{p}{1-p}\right]^{\frac{1}{\lambda_b}} & \text{if } p \in [\hat{p}, p^s], \\ v_f(p) = \frac{-c - \mathit{\Lambda} F + p\lambda[\mathit{\Gamma} - R + F]}{1 + \mathit{\Lambda}} & \text{if } p = p^s, \\ v_1(p) = \frac{p\lambda_g(\mathit{\Gamma} - R + c)}{1 + \lambda_g} - c + C_1^s (1-p) \left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}} & \text{if } p > p^s, \end{cases}$$

where $v_0(p), v_f(p)$ and $v_1(p)$ are obtained from (A.7), (A.9) and (A.3) respectively.

[6] We deal with the case when $\hat{p}_* > p^s$ in the proof of proposition 3 given in section A.8.

Conjecturing, continuity and smooth pasting, we have the following equations: $v_0(p^s) = v_f(p^s) = v_1(p^s)$ and $v_0'(p^s) = v_f'(p^s) = v_1'(p^s)$. We pin down C_0^s, C_1^s and p^s given by,

$$p^s = \frac{\lambda_b(c-F)}{\lambda_g(\Gamma - R + c) + \lambda_b(c-F)}, \tag{A.15}$$

$$C_0^s = \left[\frac{\lambda_g(\Gamma - R + c)}{\lambda_b(c-F)}\right]^{\frac{1}{\lambda_b}} \left[\frac{\lambda_b}{1+\lambda_b}\right] \left[\frac{\Lambda}{1+\Lambda}\right] (\Gamma - R + c), \tag{A.16}$$

$$C_1^s = \left[\frac{\lambda_b(c-F)}{\lambda_g(\Gamma - R + c)}\right]^{\frac{1}{\lambda_g}} \left[\frac{\lambda_g}{1+\lambda_g}\right] \left[\frac{\Lambda}{1+\Lambda}\right] (c-F). \tag{A.17}$$

Note that we need $c \geq F$ for p^s to be interior and well defined. This is also the interesting case since if $c < F$, the principal would rather incur the costs of experimentation forever than give out a reward for bringing in bad news. Hence we assume $c \geq F$ for this section. Let us denote the value functions obtained by using the above constants as $v_0(p^s; C_0^s)$ and $v_1(p; C_1^s)$. Our conjectured optimal policy for any exogenusly specified firing cutoff $\hat{p} \leq p^s$[7] and the conjectured optimal value function of the principal are,

$$a^*(p) = \begin{cases} 0 & \text{if } p \in [\hat{p}, p^s), \\ a^f & \text{if } p = p^s, \\ 1 & \text{if } p > p^s, \end{cases}$$

and,

$$v_*^{\hat{p}}(p) = \begin{cases} v_0(p; C_0^s) & \text{if } p \in [\hat{p}, p^s), \\ v_f(p^s) = v_0(p^s; C_0^s) = v_1(p^s; C_1^s) & \text{if } p = p^s, \\ v_1(p; C_1^s) & \text{if } p > p^s. \end{cases} \tag{A.18}$$

We now prove the optimality of the above policy in steps. In Lemma 17 we prove that for any piecewise continuous control, the principal value function is differentiable. Lemma 18, 19 and 20 combine to reduce the candidate policies $a' \succ a^*$ to those where we can have

[7] As we will see, the only situation when $\hat{p} > p^s$ would be optimal is when $v(p^s) < 0$, in which case it is optimal to use $F = 0$.

at most one switch from the bad news arm to the good news arm. Optimality of a^* is proved in Lemma 22.

Definition 14. *Given $a, a' \in \mathcal{A}$, we say that $a' \succ a$ iff $v(p|a') \geq v(p|a)$ $\forall p$ with the inequality being strict for some p.*

Definition 15. *Define,*

$$\eta(a) := \lambda_b(1-a) - \lambda_g a.$$

Lemma 17. *If $a(p)$ is continuous on $[p_1, p_2]$ such that $\eta(a(p)) \neq 0$ on $[p_1, p_2]$, then $v(\cdot|a)$, denoted by $v(\cdot)$ in this lemma, is C^1 on (p_1, p_2), right differentiable at p_1 and left continuous at p_2. Moreover, on (p_1, p_2),*

$$v'(p) = \frac{c + [1 + p\lambda_g a(p) + (1-p)\lambda_b(1-a(p))]v(p) - p\lambda_g a(p)(\mathit{\Gamma} - 1) + (1-p)\lambda_b(1-a(p))}{\eta(a(p))p(1-p)}. \tag{A.19}$$

Proof. There are two cases: $\eta(a(p)) > 0$ on $[p_1, p_2]$ and $\eta(a(p)) < 0$ on $[p_1, p_2]$. We will assume that $\eta(a(p)) > 0$ on $[p_1, p_2]$ and leave the other case to the reader.

For any $p \in (p_1, p_2)$, notice that, for any $\delta > 0$,

$$v(p) = \mathbb{E}(1 - e^{-\tau})(-c) + e^{-\tau}v(p_\tau),$$

where $\tau := \inf\{t : P_t \notin (p, p+\delta)\}$. Since $(1-p)\lambda_b(1-a(p)) - p\lambda_g a(p) > 0$, $\tau \to 0$ a.s. as $\delta \to 0$ and $p_\tau \to p$ *a.s.* Therefore, for any sequence $p_n \downarrow p$,

$$v(p) - v(p_n) = \mathbb{E}(1 - e^{-\tau_n})(-c) + e^{-\tau_n}v(p_{\tau_n}) - v(p_n) \to 0$$

as $n \to \infty$. For $p_n \uparrow p$, define $\tau_n := \inf\{t : P_t \notin (p_n, p)\}$ and repeat the argument above.

Therefore, v is continuous on (p_1, p_2), right(left) continuous at $p_1(p_2)$. For right differentiability, we need to show,

$$\lim_{h \downarrow 0} \frac{v(p+h) - v(p)}{h}$$

exists. For a small h, define τ as before and recall, due to the continuity of $a(\cdot)$,

$$\int_p^{p+h} \frac{1}{\eta(a(p))p(1-p)} \mathrm{d}p = \Delta.$$

By mean value theorem and continuity of a, we have,

$$h \approx \eta(a(p))p(1-p)\Delta$$

for a small Δ. ignoring the second order terms. Therefore, $\tau \leq \hat{t} := \frac{h}{\eta(a(p))p(1-p)}$. We know that,

$$\begin{aligned}
v(p) &= \mathbb{E}(1-e^{-\tau})(-c) + e^{-\tau}v(P_\tau). \\
\mathbb{E}e^{-\tau} &= \int_0^{\hat{t}} [p\lambda_g a(P_t)e^{-\int_o^t \lambda_g a(P_u)\mathrm{d}u} + (1-p)\lambda_b(1-a(P_t))e^{-\int_o^t \lambda_b(1-a(P_u))\mathrm{d}u}]e^{-t}\mathrm{d}t \\
&\quad + [pe^{-\int_0^{\hat{t}} \lambda_g a(P_u)\mathrm{d}u} + (1-p)e^{-\int_0^{\hat{t}} \lambda_b(1-a(P_u))\mathrm{d}u}]e^{-\hat{t}}.
\end{aligned}$$

For a sufficiently small h, using continuity of a and first order approximations, we get,

$$\mathbb{E}e^{-\tau} = 1 - \hat{t}.$$

Similar calculations show that,

$$\begin{aligned}
\mathbb{E}e^{-\tau}v(P_\tau) = (1-\hat{t})(1 - [p\lambda_g a(p) + (1-p)\lambda_b(1-a(p)]\hat{t})v(p+h) \\
+ p\lambda_g a(p)\hat{t}(\Gamma - 1) + (1-p)\lambda_b(1-a(p)\hat{t}(-1).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{v(p+h)-v(p)}{h} =& \frac{c\hat{t} + (1 + p\lambda_g a(p) + (1-p)\lambda_b(1-p)v(p+h)\hat{t}}{h} \\
&+ \frac{-p\lambda_g a(p)(\Gamma-1)\hat{t} + (1-p)\lambda_b(1-a(p))\hat{t}}{h}.
\end{aligned}$$

Therefore, it is easy to see that $\lim_{h\downarrow 0}\frac{v(p+h)-v(p)}{h}$ exists and is equal to

$$\frac{c+[1+p\lambda_g a(p)+(1-p)\lambda_b(1-a(p))]v(p)-p\lambda_g a(p)(\Gamma-1)+(1-p)\lambda_b(1-a(p))}{\eta(a(p))p(1-p)}.$$

Therefore, v is right differentiable on (p_1,p_2), and its right derivative is continuous, and bounded for any interval $[p_1,p_2]$. Standard results in analysis show that if a continuous function has continuous right derivatives at each point in an interval, and the right derivatives are continuous, then the function is differentiable on the interval.[8] Therefore, v is differentiable on (p_1,p_2) with the derivative given by (A.19).

□

First, for any continuous a such that $a\neq a^f$ on the interval, we also have $\eta(a,p)>0$. By Lemma 17, we know that $v(\cdot|a)$ is differentiable and satisfies the following differential equation:

$$\begin{aligned}v(p)&=\lambda_g pa(p)(\Gamma-1)+\lambda_b(1-p)(1-a(p))(-1)-c\\&\quad+[\lambda_b(1-a(p))-\lambda_g a(p)]p(1-p)v'(p)-[\lambda_g pa(p)+\lambda_b(1-p)(1-a(p))]v(p),\end{aligned}$$

where, with some abuse of notation, we denote $v(\cdot|a)$ by $v(\cdot)$. Rearranging the above,

$$v(p)=-\lambda_b(1-p)-c+\lambda_b p(1-p)v'(p)-\lambda_b(1-p)v(p)+a(p)H(p,v(p),v'(p)),\tag{A.20}$$

$$\text{where } H(x,y,z):=\lambda_g x(\Gamma-1)+\lambda_b(1-x)-(\lambda_b+\lambda_g)x(1-x)z-(\lambda_g x-\lambda_b(1-x)y).\tag{A.21}$$

Notice that H is continuous in each of its argument.

Lemma 18. *Suppose, $a\in\mathcal{A}$ is continuous on $[p_1,p_2]\subset(\hat{p},1)$. Suppose, for some $p\in(p_1,p_2)$, $a(p)\notin\{0,a^f,1\}$ and $H(p,v(p),v'(p)\neq 0$. Then, there is an $a'\in\mathcal{A}$ and an $\epsilon>0$ such that, either, $a'(q)\in\{0,1\}$ for all $q\in[p-\epsilon,p]$ or for all $q\in[p,p+\epsilon]$ and $v(\cdot|a')\geq v(\cdot|a)$.*

Proof. There are two possible cases:

[8] For example, https://math.stackexchange.com/questions/418737/continuous-right-derivative-implies-differentiability

1. $H(p, v, v') > 0$,

2. $H(p, v, v') < 0$.

1. Suppose, $H(p, v(p), v'(p)) > 0$.

   By continuity of a and H, $a(q) \notin \{0, 1\}$ and $H(q, v(q), v'(q) > 0$ for all $q \in B_\epsilon(p)$ for some $\epsilon > 0$. Consider an alternative control a' such that $a'(q) = 1$ for $q \in (p, p+\epsilon]$. Moreover, we assume that as soon as the beliefs hit either p, we switch to a, i.e., the control is non-markovian. However, due the markovian structure of the underlying problem, if a' outperforms a, then there is a Markovian control that also outperforms a. Therefore, for any $q \in (p, p+\epsilon)$, the value function $v(\cdot|a')$ (denoted by $\tilde{v}$ henceforth) satisfies the following differential equation:

$$\begin{aligned}\tilde{v}(q) =& -\lambda_b(1-q) - c + \lambda_b q(1-q)\tilde{v}'(q) - \lambda_b(1-q)\tilde{v}(q) + H(q, \tilde{v}(q), \tilde{v}'(q)) \\ \text{and } v(q) =& -\lambda_b(1-q) - c + \lambda_b q(1-q)v'(q) - \lambda_b(1-q)v(q) + H(q, v(q), v'(q)).\end{aligned}$$

   Moreover, $\tilde{v}(p) = v(p)$. Therefore,

$$\begin{aligned}\tilde{v}(q) - v(q) =& \lambda_b q(1-q)(\tilde{v}'(q) - v'(q)) - \lambda_b(1-q)(\tilde{v}(q) - v(q) \\ &+ H(q, \tilde{v}(q), \tilde{v}'(q)) - a(q)H(q, v(q), v'(q)),\end{aligned}$$

   which implies

$$\begin{aligned}\tilde{v}(q) - v(q) - (1-a)H(q, v(q), v'(q)) =& \lambda_b q(1-q)(\tilde{v}'(q) - v'(q)) \\ &+ [H(q, \tilde{v}(q), \tilde{v}'(q)) - H(q, v(q), v'(q))] \\ &- \lambda_b(1-q)(\tilde{v}(q) - v(q)), \\ =& -\lambda_g q(1-q)[\tilde{v}'(q) - v'(q)] - \lambda_g q(\tilde{v}(q) - v(q)).\end{aligned}$$

   where the last equality uses the expression for H defined in (A.21). Notice that as $q \downarrow p$, $\tilde{v}(q) - v(q) \to 0$. However, $H(q, v(q), v'(q)) > 0$. Therefore, $-\lambda_g q(1-q)[\tilde{v}'(q) - v'(q)] < 0$ as $q \downarrow p$. Therefore, in the neighborhood of p, $\tilde{v}'(q) > v'(p)$, i.e. $\tilde{v}(q) > v(q)$

for all $q \in (p, p + \epsilon_1)$ for some $\epsilon_1 > 0$.

2. Suppose, $H(p, v(p), v'(p)) < 0$. The argument is exactly as above with $a' = 0$ on some interval $(p - \epsilon, p)$ and following the policy at p thereafter. Similar calculations as before yield, for any $q \in (p - \epsilon, p)$,

$$\tilde{v}(q) - v(q) + aH(q, v(q), v'(q)) = \lambda_b q(1 - q)[\tilde{v}'(q) - v'(q)] + \lambda_b(1 - q)(\tilde{v}(q) - v(q)).$$

   Again, taking limits as $q \uparrow p$, and since $H(q, v(q), v'(q)) < 0$, we must have $\tilde{v}'(q) < v'(q)$. Therefore, some ϵ_1 $\tilde{v}(q) > v(q)$ for all $q \in (p - \epsilon_1, p)$.

Notice that even though a' maybe non-Markovian, due the Markovian structure of the problem, if a non-Markovian control does strictly better than a, there exists a Markovian control that does strictly better than a. Therefore, in general, there exists a Markovian control $a' \succ a$ and, has $a'(q) = 1(0)$ for all $q \in (p, p + \epsilon_1)$ if $H(p, v(p), v'(p)) > (<)0$. □

Lemma 19. *Suppose, $a \in \mathcal{A}$ is continuous on $[p_1, p_2] \subset (\hat{p}, 1)$. Suppose, for some $p \in (p_1, p_2)$, $a(p) \notin \{0, a^f, 1\}$ and $H(p, v(p), v'(p)) = 0$. Then, at least one of the following hold:*

1. *There is an $a' \in \mathcal{A}$ and an $\epsilon > 0$ such that, either, $a'(q) \in \{0, 1\}$ for all $q \in [p - \epsilon, p]$ or for all $q \in [p, p + \epsilon]$ and $v(\cdot|a') \geq v(\cdot|a)$.*

2. *$\exists q$ such that $v(q|a^*) > v(q|a)$.*

Proof. There are four cases:

1. For some $\epsilon > 0$, $H(q, v(q), v'(q)) = 0$ for all $q \in [p - \epsilon, p]$ or $[p, p + \epsilon]$. Notice that v must satisfy (A.20) and (A.21) where we set $H(q, v(q)v'(q)) = 0$. It is easy to check that the two imply that v must be linear. However, It is straightforward to see that there is no v of the form $K_1 p + K_2$ for some constants K_1 and K_2 that satisfies both the equations.

2. $\exists \epsilon > 0$ such that. $\{q : H(q, v(q), v'(q) = 0, q \in B_\epsilon(p)\} = \{p\}$. By Lemma 18, since H is signed on $(p, p + \epsilon)$, we can find a control a' that is valued in $\{0, 1\}$ that does strictly better than a. Moreover, setting $a'(p) = 1$ if $a(p) > a^f$ and 0 otherwise leaves the value at p unchanged.

3. $\exists \epsilon > 0$ such that both of the following hold:

   (a) $H(q, v(q), v'(q)) \geq 0$ or $H(q, v(q), v'(q)) \leq 0$ for all $q \in (p, p + \epsilon)$.

   (b) $H(q, v(q), v'(q)) \geq 0$ or $H(q, v(q), v'(q)) \leq 0$ for all $q \in (p - \epsilon, p)$.

   That is, $H(q, v(q), v'(q))$ does not change sign on either side of p for some open interval. Here, again, we can set $a'(q) = 1$ or 0 depending on whether $H \geq 0$ or ≤ 0, for any $q \in B_\epsilon(p) \backslash p$. At p, we can set $a'(p) = 1$ if $a(p) > a^f$ and 0 otherwise as before.

4. At least one of the following holds:

   (a) For any $\epsilon > 0$, $\exists q_1, q_2 \in (p - \epsilon, p)$ such that $H(q_1, v(q_1), v'(q_1)) > 0$ and $H(q_2, v(q_2), v'(q_2)) < 0$.

   (b) For any $\epsilon > 0$, $\exists q_1, q_2 \in (p - \epsilon, p)$ such that $H(q_1, v(q_1), v'(q_1)) > 0$ and $H(q_2, v(q_2), v'(q_2)) < 0$.

   We will argue only for case 4a. For any $\epsilon > 0$, $\exists q \in (p, p+\epsilon)$ such that $H(q, v(q), v'(q)) > 0$. Define,

$$\begin{aligned} e :=& \sup\{w \in (p, q) : H(w, v(w), v'(w)) < 0\} \\ =& \inf\{w \in (p, q) : H(x, v(x), v'(x)) \geq 0 \forall x \in (w, q)\}. \end{aligned}$$

   The equality is obvious and, it is also easy to see, due to continuity, that $e \in (p, q)$. Therefore, we can define a control, a'' that takes the value 1 on (e, q) such that $a'' \succ a$ by Lemma 18. Moreover, by definition of e, $\exists$ a sequence $q_n \uparrow e$ such that $H(q_n, v(q_n), v'(q_n)) < 0$. By continuity, for every such q_n, $\exists$ an interval (p_n, q_n) such that $H < 0$ on the entire interval. Therefore, we can define a control a' modifying a'' by setting $a(w) = -1$ on (p_n, q_n) such that $a' \succ a'' \succ a$.[9] Lastly, notice that $a(q_n) = 0$ and $q_n \to e$ and $a(w) = 1$ for $w \in (e, q)$. Therefore, $v(e|a') = v^f(e)$, value by freezing at e. We can repeat this construction to obtain another point, say e' in (p, e) where the value obtained is the value by freezing. By Lemma 21, $v(\cdot|a^*) > v^f(\cdot)$ for all p except p^s. Therefore, $\exists$ a $w \in \{e, e'\}$ such that $v(w|a^*) > v(w|a)$.

[9] Notice that this control may not be piecewise is continuous but, the argument goes through by choosing a finite number of intervals (p_n, q_n) close to e.

□

Definition 16. *We say, H is* ***signed on an interval*** (x, y) *if, for all $z \in (x, y)$ either $H(z, v(z), v'(z) \geq 0$ or ≤ 0. If $H(z, v(z), v'(z) \geq 0$ we say H is $+$ and if $H(z, v(z), v'(z) \leq 0$ we say H is $-$. If for some point z, $H(z, v(z), v'(z)$ has different signs on either side of z, we say that H changes sign at z.*

Lemma 20. *If $a \in \mathcal{A}$ such that $a \succ a^*$, then H changes sign at most once.*

Proof. By Lemma 19, we know that if $a \succ a^*$, for every p such that $H(p, v(p|a), v'(p|a)) = 0$, we are in Case 2 or 3 in Lemma 19. That is, for every p such that H is 0, $\exists \epsilon(p)$ such that H is signed on for $q \in [p, p + \epsilon(p)]$ and on $[p - \epsilon(p), p]$. Let I be the set of points where H changes sign. If $|I| > 1$, we must have at least one point, say p, where the sign changes from $-$ to $+$. Replacing a by 0 to the left of p and 1 to its right, we know that we achieve a strictly higher payoff. Moreover, at p, the value is equal to $v^f(p)$. Therefore, if $p \neq p^s$, $a \succ a^*$ is not possible, as $v(p|a^*) > v^f(p)$.

Therefore, if there is a $p > p^s$ where H changes sign, it must be from $+$ to $-$, and, more importantly, H stays negative on $(p, 1)$. In that case, we can set $a'(q) = 0$ on $(p, 1)$ so that $a' \succ a \succ a^*$. It is straightforward to see that using the bad news arm on $(1 - \epsilon, 1)$ is strictly dominated by using the good news arm on $(1 - \epsilon, 1)$ for a sufficiently small $\epsilon > 0$. Therefore, $a' \succ a^*$ is not possible.

Hence, H cannot change its sign at $p > p^s$, must be $+$ on $(p^s, 1)$ and a sign change from $-$ to $+$, if present, must occur only at p^s. Therefore, the only possibilities are, H stays $+$ throughout or changes from $-$ to $+$ at p^s.

□

Lemma 21. *$v(p^s|a^*) = v^f(p^s)$ and $v(p|a^*) > v^f(p)$ for all $p \neq p^s$.*

Proof. Recall, by definition of a^*, $v(p^s|a^*) = v^f(p^s)$ and $v'(p^s|a^*) = v^{f\prime}(p^s)$. It is easy to see that $v'(p|a^*)$ is increasing. Since $v^f(\cdot)$ is linear and $v'(p|a^*)$ is increasing, $v(p|a^*) > v^f(p)$ for all $p \neq p^s$. □

Lemma 22. *Suppose the principal hires in the interval $[\hat{p}, 1]$ such that $\hat{p} \leq p^s$, then the policy a^* described in equation A.18 is optimal for the principal.*

Proof. Lemma 20 tells us that there can be at most one switch from bad news arm to good news arm in $[\hat{p}, 1]$. Suppose the belief at which switching occurs is $\tilde{p} > p^s$. Then in this case, the value function of the principal is given by

$$\tilde{v}^{\hat{p}}(p) := \begin{cases} v_0(p; \tilde{C}_0) & \text{if } p \in [\hat{p}, \tilde{p}), \\ v_f(\tilde{p}) = v_0(\tilde{p}; \tilde{C}_0) = v_1(\tilde{p}; \hat{C}_1) & \text{if } p = \tilde{p}, \\ v_1(p; \hat{C}_1) & \text{if } p \in (\tilde{p}, 1], \end{cases} \tag{A.22}$$

where $\tilde{C}_0$ and $\tilde{C}_1$ are computed by invoking continuity at $\tilde{p}$. Note that since $v_1(p; C_1^s)$ is convex, tangent to $v_f(\cdot)$ at p_s, and $v_1(p; C_1^s) \geq v_f(p)$, it must be the case that $v_1(\tilde{p}; C_1^s) > v_f(\tilde{p}) = v_1(\tilde{p}; \tilde{C}_1)$. This implies that $C_1^s > \tilde{C}_1$ and consequently $v_1(p; C_1^s) > v_1(p; \tilde{C}_1)$ for all $p \in [\tilde{p}, 1]$. This results in $v_*^{\hat{p}}(p) > \tilde{v}^{\hat{p}}(p)$ for all $p \in [\tilde{p}, 1]$.

We now claim that $\tilde{C}_0 > 0$. To see this note that $v_f(\tilde{p}) = v_0(\tilde{p}; \tilde{C}_0) > 0$. Equation A.7 implies that $v_0(\tilde{p}; \tilde{C}_0) > 0 \implies \tilde{C}_0 > 0$, which implies that $v_0(p; \tilde{C}_0)$ is convex. Next note that since $v_0(p; C_0^s)$ is convex, tangent to $v_f(\cdot)$ at p_s and $v_0(p; C_0^s) \geq v_f(p)$, it must be the case that $v_0(\tilde{p}; C_1^s) > v_f(\tilde{p}) = v_0(\tilde{p}; \tilde{C}_0)$. The fact that $v_0(p; C_0^s)$ and $v_0(p; \tilde{C}_0)$ do not intersect, together with the convexity of $v_0(p; \tilde{C}_0)$ implies that $v_0(p_1; \tilde{C}_0) = v_f(p_1)$ for some $p_1 < p^s$. This implies that it must be the case that $v_0(p; \tilde{C}_0) < v_f(p)$ for $p \in (p_1, \tilde{p})$, implying that $v_*^{\hat{p}}(p) > \tilde{v}^{\hat{p}}(p)$ for all $p \in (p_1, \tilde{p})$. Next note that since $v_0(p; C_1^s) > v_0(p; \tilde{C}_0)$ we have $v_*^{\hat{p}}(p) > \tilde{v}^{\hat{p}}(p)$ for all $p \in [\hat{p}, p_1]$ This shows that $v_*^{\hat{p}}(p) > \tilde{v}^{\hat{p}}(p)$ for all $p \in [\hat{p}, 1]$. With a mirror argument as above (left to the reader) we can show that we have the same result as above when $\tilde{p} < p^s$. This implies that p^s is the optimal switching belief. Hence a^* is the optimal policy for the principal. □

Since a^* is optimal for any exogenously specified stopping cutoff $\hat{p}$, the optimal stopping cutoff is the following:

Definition 17. *The* ***optimal stopping cutoff*** *is given by*

$$p_1^* := \{p : v(p|a^*) = 0\}. \tag{A.23}$$

Here, we have extended a^* on $[0, 1]$ by assuming that $a^*(p) = 0 \ \forall p < p^s$.

Notation 3. *As before, to keep track of the principal value function for the case of $F = 1$, we will denote it by $v_*^{F=1}(\cdot)$.*

Define,

$$v_*^{F=1}(p) := \begin{cases} 0 & \text{if } p \leq p_1^*, \\ v_*^{p_1^*}(p) & \text{otherwise}, \end{cases} \tag{A.24}$$

where $v_*^{p_1^*}(p)$ is defined in Equation A.18.

A.8 Comparison between F=1 and F=0 cases

Proof of Proposition 3: Lemma 26 establishes that if $c > F(1 + \lambda_g)$ then there exists $\underline{\lambda_b}$ such that when $\lambda_b > \underline{\lambda_b}$ we have $p_1^*(\lambda_b) < p_0^*(\lambda_b)$. Lemma 27 shows that if $p_1^* < p_0^*$ then $R = 1, F = 1$ dominates $R = 1, F = 0$ for all prior belief $p_0 \in (0, 1)$ which concludes the proof of the first part.

Lemma 23 shows that in the case when $R = 1, F = 1$,

$$p_f^* > p^s \iff \frac{c + \Lambda}{\Lambda(\Gamma - R + 1)} > \frac{\lambda_b(c-1)}{\lambda_g(\Gamma - R + c) + \lambda_b(c-1)},$$

where p_f^* is the belief at which the value of the principal equals 0 when the agent uses the policy $a = a^f$. This implies that when $R = 1, F = 1$ and $p_f^* > p^s$ the principal's value at p^s is negative under the policy a^* as defined in (A.18). We claim that in this case the optimal hiring region of the principal is $[p_f^*, 1]$ and the principal-optimal policy of the agent is given by

$$\tilde{a}(p) = \begin{cases} 1 & \text{if } p \in [p_f^*, 1], \\ a^f & \text{if } p = p_f^*. \end{cases}$$

To see this, recall that the principal-optimal policy of the agent features at most one switch from the bad news arm to the good news arm (Lemma 20). Fixing the lower end of the hiring interval at p_f^*, in the candidate policy $\tilde{a}$ the agent uses the good news arm everywhere above p_f^* and freezes the belief at p_f^*. It is easy to see that $\tilde{a}$ dominates the policy where the agent uses $a = 0$ for all $p \geq p_f^*$ since the value of the principal is strictly negative under such a policy. All we need to show is that $\tilde{a}$ dominates any other policy where there is a switch from the bad news arm to the good news arm at a belief $\tilde{p} > p_f^*$. From lemma 22 we know that since $\tilde{p} > p^s$, lowering the belief at which switching occurs, improves the value of the principal. This implies that in the optimal policy given the lowering firing cutoff p_f^*, switching must occur at p_f^*, which implies that in the optimal policy, the agent must be freezing belief at p_f^*. To see that p_f^* is an optimal choice of lower end of hiring interval, simply note that $v_f(p_f^*) = 0$. If the principal chooses a cutoff $\hat{p}$ strictly lower than p_f^*, principal's value at $\hat{p}$ is strictly negative and hence cannot be an equilibrium. If the principal chooses a cutoff $\hat{p}$ strictly higher than p_f^*, principal's value is strictly lower at all beliefs in $(p_f^*, 1]$ compared to $\tilde{a}$. Hence p_f^* is the optimal choice of lower end of the hiring interval.

Note that in this case the principal can achieve an identical behavior from the agent by setting $F = 0$ instead. Additionally, this improves the payoff of the principal since the agent does not have to be paid anything if a bad news is obtained. Hence, in this case the optimal reward structure is to set $R = 1, F = 0$. □

In comparing the optimality of $F = 1$ vs $F = 0$, an obvious case when $F = 0$ would dominate $F = 1$ is one when $p_1^* > p^s$ (or equivalently $p_f^* > p^s$). The following lemma gives the condition when that can happen.

Lemma 23.

$$p_f^* > p^s \iff \frac{c + \mathit{\Lambda} F}{\mathit{\Lambda}(\mathit{\Gamma} - R + F)} > \frac{\lambda_b(c - F)}{\lambda_g(\mathit{\Gamma} - R + c) + \lambda_b(c - F)}.$$

Proof. Proof follows from the definitions of p_f^* and p^s defined in (A.13) and (A.15) respectively. □

Similarly, a natural case where $F = 1$ would dominate $F = 0$ (at least for some prior beliefs) is one where $p_1^* < p_0^*$ where p_0^* is defined in (A.12).

Lemma 24. *p_1^* is decreasing in λ_b*[10]*.*

Proof. Suppose the agent uses the bad news arm in the interval $[p_0, p_1]$ with $p_0 \neq p_1$ and at p_1, the principal gets specified value $\bar{V}(p_1) \geq 0$ at p_1 . We calculate the value of the principal under the given experimentation strategy of the agent for a given value of λ_b. We call this value as $\bar{V}(p_0; \lambda_b)$. To calculate this value we first define $\bar{t}$ as the time it takes for beliefs to drift from p_0 to p_1 in the absence of a signal. Note that

$$\bar{t} = \frac{1}{\lambda_b} \int_{p_0}^{p_1} \frac{\mathrm{d}p}{p(1-p)} = \frac{1}{\lambda_b} \ln \left[\frac{p_1}{1-p_1} \frac{1-p_0}{p_0} \right].$$

We can write

$$\begin{aligned} \bar{V}(p_0; \lambda_b) = (1-p_0) \int_0^{\bar{t}} \lambda_b e^{-\lambda_b s} \underbrace{\left[-c(1-e^{-s}) - Fe^{-s}\right]}_{\text{Value when signal arrives at } s} \mathrm{d}s \\ + \underbrace{\left[p_0 + (1-p_0)e^{-\lambda_b \bar{t}}\right]}_{\text{Prob. of no signal until } \bar{t}} \left[e^{-\bar{t}} \bar{V}(p_1) - c[1-e^{-\bar{t}}]\right]. \end{aligned}$$

The integral can be evaluated to give:

$$\begin{aligned} \bar{V}(p_0; \lambda_b) = (1-p_0) \left[-c(1-K) + (c-F) \frac{\lambda_b}{1+\lambda_b}(1-K^{\frac{1+\lambda_b}{\lambda_b}})\right] \\ + [p_0 + (1-p_0)K] \left[K^{\frac{1}{\lambda_b}}(\bar{V}(p_1) + c) - c\right]. \end{aligned}$$

where K is given by

$$K = \left[\frac{p_0}{1-p_0} \frac{1-p_1}{p_1}\right] < 1.$$

We then claim that $\bar{V}(p_0; \lambda_b)$ is strictly increasing in λ_b. To see this first note that the second term term $[p_0 + (1-p_0)K] \left[K^{\frac{1}{\lambda_b}}(\bar{V}(p_1) + c) - c\right]$ is strictly increasing in λ_b since $K < 1$. We just need to check for the first term. Taking the derivative we get

$$\frac{\mathrm{d}}{d\lambda_b} \left[\frac{\lambda_b}{1+\lambda_b}(1-K^{\frac{1+\lambda_b}{\lambda_b}})\right] = \frac{1}{(1+\lambda_b)^2} \left[1 + K^{\frac{1+\lambda_b}{\lambda_b}} \left[\frac{1+\lambda_b}{\lambda_b} \ln K - 1\right]\right].$$

[10] p_1^* is a function on several variables including λ_b. Here we study the behavior of p_1^* as a function of λ_b ceterus paribus.

Define the function $f(m) = 1+K^m(m \ln K - 1)$, where $m = \frac{1+\lambda_b}{\lambda_b}$. Note that m is decreasing in λ_b and when $\lambda_b = 0, m = \infty$ and when $\lambda_b = \infty, m = 1$. Note that

$$f'(m) = mK^m(\ln K)^2 > 0 \text{ for all } K > 0,$$

and

$$f(1) = 1 - K + K \ln K \geq 0 \text{ for all } K > 0.$$

This implies that $\bar{V}(p_0; \lambda_b)$ is strictly increasing in λ_b.

Now, assume $\lambda_b' > \lambda_b''$ and consider $V_0(p; \lambda_b'')$. Call $p^\dagger$ the optimal belief at which switch from bad arm to good arm happens. Note that we must have $V_0(p^\dagger; \lambda_b'') \geq 0$. Suppose the principal chooses the same cut off $p^\dagger$ to switch from bad arm to good arm when the arrival rate of the bad news arm is λ_b'. Let's call the value function under this policy to the left of $p^\dagger$ as $\tilde{V}_0(p; \lambda_b')$. Consider a belief $p_1 \leq p^\dagger$. We now show that $\tilde{V}_0(p_1; \lambda_b') > V_0(p_1; \lambda_b'')$. First we observe that $\tilde{V}_0(p^\dagger; \lambda_b') > V_0(p^\dagger; \lambda_b'')$. This is because

$$\tilde{V}_0(p^\dagger; \lambda_b') = V^f(p^\dagger; \lambda_b') > V^f(p^\dagger; \lambda_b'') = V_0(p^\dagger; \lambda_b''),$$

since $V^f(p; \lambda_b)$ is increasing in λ_b. Now, we can write

$$\begin{aligned}
V_0(p_1; \lambda_b'') &= (1-p_1)\left[-c(1-K) + (c-F)\frac{\lambda_b''}{1+\lambda_b''}(1 - K^{\frac{1+\lambda_b''}{\lambda_b''}})\right] \\
&\quad + [p_1 + (1-p_1)K]\left[K^{\frac{1}{\lambda_b''}}(V_0(p^\dagger; \lambda_b'') + c) - c\right] \text{ and} \\
\tilde{V}_0(p_1; \lambda_b') &= (1-p_1)\left[-c(1-K) + (c-F)\frac{\lambda_b'}{1+\lambda_b'}(1 - K^{\frac{1+\lambda_b'}{\lambda_b'}})\right] \\
&\quad + [p_1 + (1-p_1)K]\left[K^{\frac{1}{\lambda_b'}}(\bar{V}_0(p^\dagger; \lambda_b'') + c) - c\right],
\end{aligned}$$

where $K = \left[\frac{p_1}{1-p_1}\frac{1-p^\dagger}{p^\dagger}\right] < 1$. Since we have shown above that $\bar{V}(p_0; \lambda_b)$ is strictly increasing in λ_b, we have $\tilde{V}_0(p^f; \lambda_b') > V_0(p; \lambda_b'')$. Next note that $p^\dagger$ may or may not be the optimal cutoff choice when the arrival rate is λ_b'. Hence we must have $V_0(p; \lambda_b') > V_0(p; \lambda_b)$ when $p < p^\dagger$ which in turn implies that $p_1^*(\lambda_b') < p_1^*(\lambda_b'')$ since $V_0'(p, \lambda_b) > 0$ when $p \geq p_1^*(\lambda_b)$.

$\square$

Lemma 25. $\lim_{\lambda_b \to \infty} p_1^*(\lambda_b) = \dfrac{F(1+\lambda_g)}{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c}$ *and* $\lim_{\lambda_b \to \infty} p_0^*(\lambda_b) = \dfrac{c}{\lambda_g(\mathit{\Gamma} - R)}$.

Proof. Taking limit of A.16 as $\lambda_b \to \infty$ gives us

$$C_0^f = \frac{\lambda_g}{1+\lambda_g}(\mathit{\Gamma} - R + c).$$

This gives us

$$\begin{aligned}\lim_{\lambda_b \to \infty} V_0(p;\lambda_b) &= -pc - (1-p)F + \frac{\lambda_g}{1+\lambda_g}(\mathit{\Gamma} - R + c)p \\ &= -F + p\left[\frac{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c}{1+\lambda_g}\right].\end{aligned}$$

which yields

$$\lim_{\lambda_b \to \infty} p_1^*(\lambda_b) = \frac{F(1+\lambda_g)}{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c}.$$

Also note that

$$\lim_{\lambda_b \to \infty} p_0^*(\lambda_b) = \lim_{\lambda_b \to \infty} \frac{c}{\mathit{\Lambda}(\mathit{\Gamma} - R)} = \frac{c}{\lambda_g(\mathit{\Gamma} - R)}.$$

$\square$

Lemma 26. *If* $c > F(1+\lambda_g)$ *then there exists* $\underline{\lambda}_b$ *such that when* $\lambda_b > \underline{\lambda}_b$ *we have* $p_1^*(\lambda_b) < p_0^*(\lambda_b)$.

Proof. First we note that

$$c > F(1+\lambda_g) \iff \frac{F(1+\lambda_g)}{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c} < \frac{c}{\lambda_g(\mathit{\Gamma} - R)}.$$

Next, note that $p_1^*(\lambda_b)$ is continuous in λ_b. Lemma 24 establishes that $p_1^*(\lambda_b)$ is decreasing in λ_b. Also lemma 25 establishes that $\lim_{\lambda_b \to \infty} p_1^*(\lambda_b) = \dfrac{F(1+\lambda_g)}{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c}$ and $\lim_{\lambda_b \to \infty} p_0^*(\lambda_b) = \dfrac{c}{\lambda_g(\mathit{\Gamma} - R)}$. Therefore, since

$$\lim_{\lambda_b \to \infty} p_1^*(\lambda_b) = \frac{F(1+\lambda_g)}{F(1+\lambda_g) + \lambda_g(\mathit{\Gamma} - R) - c} < \frac{c}{\lambda_g(\mathit{\Gamma} - R)} = \lim_{\lambda_b \to \infty} p_0^*(\lambda_b),$$

there must exist $\underline{\lambda}_b$ such that $p_1^*(\lambda_b) < \dfrac{c}{\lambda_g(\Gamma - R)}$ when $\lambda_b > \underline{\lambda}_b$. Note that $p_0^* \geq \dfrac{c}{\lambda_g(\Gamma - R)}$ for all λ_b, which implies that $p_1^*(\lambda_b) < p_0^*(\lambda_b)$ when $\lambda_b > \underline{\lambda}_b$. $\square$

Lemma 27. *If $p_1^* < p_0^*$ then $R = 1, F = 1$ dominates $R = 1, F = 0$ for all prior belief $p_0 \in (0, 1)$.*

Proof. Recall the principal's value function when $R = F = 1$ given by (A.18)

$$v_*^{F=1}(p) := \begin{cases} 0 & \text{if } p \in [0, p_1^*), \\ v_0(p; C_0^s) & \text{if } p \in [p_1^*, p^s), \\ v_f(p^s) = v_0(p^f; C_0^s) = v_1(p^f; C_1^s) & \text{if } p = p^s, \\ v_1(p; C_1^s) & \text{if } p > p^s, \end{cases}$$

and when $R = 1, F = 0$ is given by

$$v_*^{F=0}(p) := \begin{cases} 0 & \text{if } p \in [0, p_0^*), \\ v_0(p; C_0^*) & \text{if } p \in [p_0^*, 1]. \end{cases}$$

From lemma 29, we know that if $p_1^* < p_0^*$ then $v_1(p; C_1^s) > v_1(p; C_1^*)$ for all $p \in [p^s, 1]$ which implies that $v_*^{F=1}(p) > v_*^{F=0}(p)$ for all $p \in [p^s, 1]$. We are left to show that $v_*^{F=1}(p) > v_*^{F=0}(p)$ for all $p \in [p_0^*, p^s]$. To that end, first note that $v_*^{F=1}(p_0^*) = v_0(p_0^*; C_0^s) > v_1(p_0^*; C_1^*) = 0$. From lemma 30 we know that $v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked and $v_1'(p^s; C_1^*) - v_0'(p^s; C_0^s) > 0$, which implies that $v_1(p; C_1^*) - v_0(p; C_0^s)$ attains it's maximum value in $[p_0^*, p^s]$ at either p_0^* or p^s. Note that $v_0(p^s; C_0^s) = v_1(p^s; C_1^s) > v_1(p^s; C_1^*)$. Hence $v_1(p; C_1^*) - v_0(p; C_0^s) < 0$ when $p \in [p_0^*, p^s]$ and we are done. $\square$

Lemma 28. *$v_0(p; C_0^s) < v_1(p; C_1^s)$ if $p \in [0, p^s)$ and $v_0(p; C_0^s) > v_1(p; C_1^s)$ if $p \in (p^s, 1]$.*

Proof. We note that C_1^s and C_0^s are both positive. We evaluate $v_1''(p; C_1^s) - v_0''(p; C_0^s)$ which is given by

$$
\begin{aligned}
&C_1^s\left[\frac{1+\lambda_g}{\lambda_g^2 p^2(1-p)}\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}\right]-C_0^s\left[\frac{1+\lambda_b}{\lambda_b^2 p(1-p)^2}\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\right]\\
=&\frac{1}{p(1-p)}\underbrace{\left[\frac{C_1^f(1+\lambda_g)}{\lambda_g^2 p}\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}-\frac{C_0^f(1+\lambda_b)}{\lambda_b^2(1-p)}\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}\right]}_{=\Phi(p)}.
\end{aligned}
$$

Since C_1^s and C_0^s are both positive, $\Phi(p)$ is strictly decreasing. Also note that $\lim_{p\downarrow 0}\Phi(p) = \infty = -\lim_{p\uparrow 1}\Phi(p)$ which implies that there exists a unique p^d such that $\Phi(p^d) = 0$. We next show that $v_1''(p^s; C_1^s) - v_0''(p^s; C_0^s) = 0$ by plugging in p^s as defined in (A.15) (algebra left to the reader) implying that $p^d = p^s$. Therefore $v_1''(p; C_1^s) - v_0''(p; C_0^s) > 0$ when $p < p^s$ and $v_1''(p; C_1^s) - v_0''(p; C_0^s) < 0$ when $p > p^s$. Now we know that at p^s, $v_1(p^s; C_1^s) = v_0(p^s; C_0^s)$ and $v_1'(p^s; C_1^s) = v_0'(p^s; C_0^s)$. Hence we get $v_0(p; C_0^s) < v_1(p; C_1^s)$ if $p \in [0, p^s)$ and $v_0(p; C_0^s) > v_1(p; C_1^s)$ if $p \in (p^s, 1]$. □

Lemma 29. *If $p_1^* < p_0^*$, then $v_1(p; C_1^s) > v_1(p; C_1^*)$ for all p.*

Proof. Recall that $v_1(p; C_1^*)$ is the value function of the principal when $p \in [p_0^*, 1]$ in the case when $F = 0$. Since $p_1^* < p_0^*$, we have $v_*^{F=1}(p_0^*) > v_*^{F=0}(p_0^*) = 0$. If $p_0^* < p^s$, then $v_0(p_0^*; C_0^s) > v_1(p_0^*; C_1^*) = 0$. We know from lemma 28 that $v_1(p_0^*; C_1^s) > v_0(p_0^*; C_0^s)$ which implies that $v_1(p_0^*; C_1^s) > v_1(p_0^*; C_1^*)$. If $p_0^* \geq p^f$, then $v_1(p_0^*; C_0^s) > v_1(p_0^*; C_1^*) = 0$, since $p^s > p_1^*$.

This implies that $v_1(p; C_1^s) > v_1(p; C_1^*)$ for all p, since if $C \neq C'$ then $v_1(p; C) \neq v_1(p; C')$ for any p. □

Lemma 30. *$v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked and moreover if $C_1^f > C_1^*$ then $v_1'(p^s; C_1^*) - v_0'(p^s; C_0^s) > 0$.*

Proof. To see that $v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked, following identical steps in lemma 28 we can show that $v_1''(p; C_1^*) - v_0''(p; C_0^s)$ is strictly decreasing and is positive below a cutoff and negative above it, establishing that $v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked.

Note that $v_0'(p^s; C_0^s) = v_1'(p^s; C_1^s)$ by the definition of p^s. Also note from lemma 29 that $C_1^s > C_1^*$. $v_1'(p^s; C_1^*) - v_0'(p^s; C_0^s)$is given by

$$\left[\frac{\lambda_g(\varGamma - R + c)}{1+\lambda_g} - C_1^*\left[\frac{1-p^s}{p^s}\right]^{\frac{1}{\lambda_g}}\left[\frac{\frac{1}{\lambda_g}+p^s}{p^s}\right]\right] - \left[-\frac{\lambda_b(c-F)}{\lambda_b+1} + C_0^s\left[\frac{p^s}{1-p^s}\right]^{\frac{1}{\lambda_b}}\left[\frac{\frac{1}{\lambda_b}+1-p^s}{1-p^s}\right]\right].$$

Since $C_1^s > C_1^*$ we must have $v_0'(p^s; C_0^s) - v_1'(p^s; C_1^*) > 0$.

$\square$

Lemma 31. *Suppose $p_1^* \geq p_0^*$.*

1. *If $v_*^{F=0}(p^s) > v_*^{F=1}(p^s)$, then $v_*^{F=0}(p) > v_*^{F=1}(p)$ for all $p \in [p_0^*, 1]$.*
2. *If $v_*^{F=0}(p^s) < v_*^{F=1}(p^s)$, then there exists $\hat{p} \in [p_1^*, p^s]$ such that $v_*^{F=1}(p) > v_*^{F=0}(p)$ when $p > \hat{p}$, $v_*^{F=1}(p) < v_*^{F=0}(p)$ when $p \in (p_0^*, \hat{p})$, and $v_*^{F=1}(\hat{p}) = v_*^{F=0}(\hat{p})$.*
3. *If $v_*^{F=0}(p^s) = v_*^{F=1}(p^s)$, then $v_*^{F=1}(p) = v_*^{F=0}(p)$ when $p \geq p^s$, and $v_*^{F=1}(p) < v_*^{F=0}(p)$ when $p \in (p_0^*, p^s)$.*

In particular, if $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ cross, they cross exactly once.

Proof. We prove case by case.

1. $v_*^{F=0}(p^s) > v_*^{F=1}(p^s)$: In this case $C_1^* > C_1^s$. This implies that $v_*^{F=0}(p) > v_*^{F=1}(p)$ when $p \geq p^s$. Also know from lemma 28 that $v_*^{F=1}(p) = v_0(p; C_0^s) < v_1(p; C_1^s) = v_*^{F=0}(p)$ when $p < p^s$. This implies that $v_*^{F=1}(\cdot)$ and $v_*^{F=0}(\cdot)$ never cross.
2. $v_*^{F=0}(p^s) < v_*^{F=1}(p^s)$: In this case we have $C_1^* < C_1^s$. We know from lemma 30 that $v_1'(p^s; C_1^*) - v_0'(p^s; C_0^s) > 0$ and $v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked. It is easy to see that $\lim_{p\downarrow 0} v_1'(p; C_1^*) - v_0'(p; C_0^s) = -\infty$. Also note that $v_1(p^s; C_1^*) - v_0(p^s; C_0^s) < 0$ and $v_1(p_1^*; C_1^*) - v_0(p_1^*; C_0^s) > 0$. This implies that $v_1(p; C_1^*) - v_0(p; C_0^s)$ must be decreasing in some subset of $[p_1^*, p^s]$. Since $v_1'(p; C_1^*) - v_0'(p; C_0^s)$ is single peaked, $\lim_{p\downarrow 0} v_1'(p; C_1^*) - v_0'(p; C_0^s) = -\infty$ and $v_1'(p^s; C_1^*) - v_0'(p^s; C_0^s) > 0$, it must be the case that $v_1(p; C_1^*) - v_0(p; C_0^s)$ is decreasing in $[p_1^*, p^\dagger]$ where $p^\dagger < p^s$. This implies that $v_1(p; C_1^*)$ and $v_0(p; C_0^s)$ cross exactly once at $\hat{p} \in [p_1^*, p^s]$ implying that $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ cross exactly once in $[p_1^*, p^s]$. It is easy to see that $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ do not cross in $[p^s, 1]$.

3. $v_*^{F=0}(p^s) = v_*^{F=1}(p^s)$: In this case we have $C_1^* = C_1^s$, hence $v_*^{F=1}(p) = v_*^{F=0}(p)$ when $p \geq p^s$. Also from we know from lemma 28 that $v_*^{F=1}(p) = v_0(p; C_0^s) < v_1(p; C_1^s) = v_*^{F=0}(p)$ when $p < p^s$. Hence $v_*^{F=1}(p) < v_*^{F=0}(p)$ when $p \in (p_0^*, p^s)$.

□

The following corollary to the above lemma is useful.

Corollary 1. *If $p_1^* \geq p_0^*$ and $v_*^{F=0}(p^s) < v_*^{F=1}(p^s)$, then $v_*^{F=0}(\cdot)$ and $v_*^{F=1}(\cdot)$ cross exactly once.*

A.9 Unobserved allocation

Proof of lemma 2: Since $R = 1$ and $F = 0$, the agent has no incentives to look for bad news on and off path and hence, $a(p,t) = 1$ is optimal.

Given that $a(p,t) = 1$, note that the principal's belief drifts down in the absence of a signal. The principal's value function at any belief p is given by (A.6) restated below.

$$v(p) = \frac{p\lambda_g(\mathit{\Gamma} - 1 + c)}{1 + \lambda_g} - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}},$$

C_1 and the stopping belief $\hat{p}_0$ (belief at which principal's value is zero) are jointly determined by imposing smooth-pasting and value matching with the function $f(p) = 0$ at $\hat{p}_0$. After some algebra we find that $\hat{p}_0 = \frac{c}{\lambda_g(\mathit{\Gamma}-1)}$. The principal's optimal stopping time is simply the time t at which belief reach $\hat{p}_0$. If initial prior is less than or equal to $\hat{p}_0$, then it is optimal to choose $T^* = 0$.

□

Proof of lemma 3: Suppose $R = 1$ and $F \in (0,1)$ and fix a stopping time for the principal T. We know from lemma 32 that if the agent searches for bad news at some time $\bar{t} < T$, he must be looking exclusively for bad news when $t \in [\bar{t}, T]$. Note that in this case the principal has a profitable deviation. She can simply lower her stopping time to $\bar{t}$ and be better off since she knows that after $\bar{t}$, the agent can only bring her bad news that leads to the abandonment of the project and she can save the cost of experimentation by abandoning the project herself at $\bar{t}$. Therefore in any equilibrium when $R = 1$ and $F \in (0, 1)$, it must be the case that the agent using the good news arm solely before the stopping time is reached. Notice that given this equilibrium behavior of the agent, the principal is equally better off by setting $R = 1$ and $F = 0$. Hence $R = 1$ and $F \in (0, 1)$ does not improve upon $R = 1$ and $F \in (0, 1)$. $\square$

Lemma 32. *Suppose $R = 1$ and $F \in (0, 1)$. Fix a stopping time for the principal T. Fix a strategy of the agent a such that $T_0 = \int_0^T (1 - a(p, t))\,dt > 0$. Consider another strategy for the agent $\bar{a}$ such that $\bar{a}(p, t) = 1$ when $t \in [0, T - T_0]$ and $\bar{a}(p, t) = 0$ when $t \in (T_0, T]$. The agent strictly prefers $\bar{a}$ to a.*

Proof. When $\theta = 1$. B signal never arrives under both strategies. The ex ante probability of arrival of G signal by T is the same under both strategies since the amount of time allocated to search for good news is $T - T_0$ under both strategies. If G signal arrives, the value of the agent is 1 under either strategy. If no signal arrives by T, then the value of the agent is equal to $1 - e^{-T}$ under both strategies. Hence when $\theta = 1$, both policies yield the same ex ante payoff to the agent.

When $\theta = 0$. G signal never arrives under both strategies. If no signal arrives by T, then the value of the agent is equal to $1 - e^{-T}$ under both strategies. The ex ante probability of arrival of a B signal by T is the same under both strategies since the amount of time allocated to search for bad news is T_0 under both strategies. However, note that B signal arrives later in expectation in strategy $\bar{a}$ compared to a since the agent has delayed the use of bad news arm in $\bar{a}$. Since $F < 1$, the agent strictly prefers $\bar{a}$ to a, since the agent can collect a flow wage of 1 for longer in expectation under $\bar{a}$.

$\square$

Lemma 33. *Suppose $R = F = 1$, then the principals optimal policy is one of the following*

1. *G policy: Search for good news when $p \in [\hat{p}_0, 1]$.*

2. *$G - B - G$ policy: There exists cutoff $\tilde{p}$ with $\hat{p}_0 < \tilde{p} < p^s$ such that*

   - *Search for good news when $p \in [\hat{p}_0, \tilde{p}] \cup [p^s, 1]$.*

- *Search for bad news when* $p \in (\tilde{p}, p^s)$.

3. *$B - G$ policy:*

   - *Search for good news when* $p \in [p^s, 1]$.
   - *Search for bad news when* $p \in [p_1^*, p^s)$.

Proof. Lemma 20 establishes that in the optimal policy there can be at most one switch from bad new arm to good news arm in any experimentation region $[p^\dagger, 1]$. Suppose there is no switch, then it is easy to see that the optimal policy must be to use the good news arm in the hiring region and thereby, the optimal lower cutoff of experimentation must be $\hat{p}_0$ as defined in lemma 2. If there is a switch then we have cases (2) and (3) as possibilities. Using a similar argument as in lemma 22 we can show that if there is a switch from bad news arm it must be at p^s defined in (1.4). In the $B - G$ policy, the optimal stopping belief is equal to p_1^* as in the case of observed allocation. In the $G - B - G$ policy, the optimal stopping belief is $\hat{p}_0$ by the same argument as in lemma 2. If $p_1^* < \hat{p}_0$ then the optimal policy is the $B - G$ policy. The proof is delivered by by following identical steps as in lemma 27. If $p_1^* < \hat{p}_0$ then the optimal policy must be $G - B - G$ since there is switch from bad news arm to good news arm at p^s. We denote the belief at which optimal policy switches from good news arm to bad news arm in the $G - B - G$ policy as $\tilde{p}$. $\square$

Appendix B

Appendix: Supervising to Motivate

B.1 Preliminaries

Definition 18. *I define two classes of strategy profiles as follows.*

1. ***Three phase strategy profile:*** *A three phase strategy profile $\sigma = (x_H^\sigma, x_L^\sigma, e^\sigma)$ is a strategy profile that satisfies the following on path. Given cutoffs $\bar{\mu}$ and $\underline{\mu}$, with $\underline{\mu} > \bar{\mu}$*

   - ***Separating Phase:*** *When $\mu^L > \bar{\mu}$: $x_H^\sigma(\mu^L, \mu^a) \neq x_L^\sigma(\mu^L, \mu^a)$.*
   - ***Pooling Phase:*** *When $\mu^L \in [\underline{\mu}, \bar{\mu}]$: $x_H^\sigma(\mu^L, \mu^a) = x_L^\sigma(\mu^L, \mu^a) = x^p(\mu^L)$.*
   - ***Quitting Phase:*** *When $\mu^L < \underline{\mu}$: $x_H^\sigma(\mu^L, \mu^a) = \dfrac{\gamma\lambda}{a}$; $x_L^\sigma(\mu^L, \mu^a) = 0$.*

2. ***Two phase strategy profile:*** *A two phase strategy profile $\sigma = (x_H^\sigma, x_L^\sigma, e^\sigma)$ is a strategy profile that satisfies the following on path.*

   - ***Pooling Phase:*** *When $\mu^L \geq \underline{\mu}$: $x_H^\sigma(\mu^L, \mu^a) = x_L^\sigma(\mu^L, \mu^a) = x^p(\mu^L)$.*
   - ***Quitting Phase:*** *When $\mu^L < \underline{\mu}$: $x_H^\sigma(\mu^L, \mu^a) = \dfrac{\gamma\lambda}{a}$; $x_L^\sigma(\mu^L, \mu^a) = 0$.*

Proof of Lemma 5: Note that starting at a state $(\mu, 0)$, the strategy profile prescribes that the agent never exerts effort and the principal invests 0. Suppose the agent deviates and exerts effort for a small time interval Δ, then the agent's payoff is given by

$$\begin{aligned} &\mu\lambda\Delta r(1-\gamma) - rc\Delta, \\ =&\Delta[\mu\lambda r(1-\gamma) - rc]. \end{aligned}$$

For any positive Δ, the payoff to the agent from this deviation is negative if $\lambda r(1-\gamma)-rc<0$ or $\mu \leq \dfrac{c}{\lambda(1-\gamma)}$. Hence the agent has no incentive to deviate. Given the agent's behavior, it is easy to see that the behavior of both types of principal are optimal. $\square$

Proof of Lemma 6 : Suppose $\mathcal{L}(D) > 0$. Consider the deviation of type H principal to an investment level of $\dfrac{\gamma\lambda}{a}$ for all $p \in D$. By Definition 3, everywhere in the interval (p_1, p_2) the agent's belief is equal to what it would have been in the absence of deviation. Hence, the continuation equilibrium at p_1 is unchanged given the agent exerts effort everywhere in (p_1, p_2). The agent's flow utility is higher when the principal invests $\dfrac{\gamma\lambda}{a}$ compared to the on-path investment level $x^P(p)$ when $p \in D$, hence, given that continuation play is identical at p_1, the agent continues to exert effort in-spite of the deviation. Since continuation play does not change at p_1, and $\mathcal{L}(D) > 0$, the utility of types H is strictly higher by deviating to $\dfrac{\gamma\lambda}{a}$ for $p \in D$, since type H is now choosing her optimal level of investment. $\square$

B.2 Low cost of effort

In this section, I provide the proofs of the case when $C1$ is violated, i.e. cost of effort is low.

Proof of Proposition 9: We first show that the strategy profile is an equilibrium, we then argue that it is indeed the principal's optimal equilibrium

First we observe that since the strategy profile is fully separating, on equilibrium path the agent knows perfectly the type of principal he is facing when making his effort decision. I now show that no player has any incentive to deviate from the candidate equilibrium strategy profile.

- Type H principal: Under our candidate equilibrium strategy profile, type H principal chooses her flow optimal investment $\dfrac{\lambda\gamma}{a}$ after any history while inducing effort from the agent. Clearly, type H principal cannot achieve a higher payoff by deviating to any other strategy.

- Type L principal: First we observe that the optimal level of investment of type L with belief μ is equal to $\frac{\mu\lambda\gamma}{a}$ since it maximizes the flow value of the type L principal given by
$$\mu\lambda\gamma(1+x) - a\frac{x^2}{2}.$$
Note that $\frac{\mu\gamma\lambda}{a}$ is linear and increasing in μ. Also, $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$ is decreasing in μ, taking a value of ∞ as μ goes to 0 and is less than $\frac{\gamma\lambda}{a}$ when $\mu = 1$ since $z^a = \lambda(1-\gamma)(1+\frac{\gamma\lambda}{a}) - c > 0$ by definition. Hence, there must exist a unique μ^c such that
$$\frac{\mu^c\gamma\lambda}{a} = \frac{c}{\mu^c\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1.$$
Next, define μ_1 such that
$$f(\mu_1) = \frac{c}{\mu_1\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1 = \frac{\lambda\gamma}{a}.$$
Note that since ***C1*** does not hold, $\mu_1 \leq \underline{\mu}$. Also $f(\mu) > \frac{\lambda\gamma}{a}$ for all $\mu < \mu_1$. Recalling the definition of $\underline{\mu}$, we can say that there exists a $\mu^s \in (\mu_1, \underline{\mu}]$ such that if $\mu < \mu^s$ the type L principal strictly prefers to quit the relationship, if $\mu > \mu^s$ the type L principal strictly prefers to induce effort by investing $f(\mu)$, and is indifferent when $\mu = \mu^s$. Also, it is easy to see that if $\mu_1 < \underline{\mu}$ then $f(\mu^s) < \frac{\lambda\gamma}{a}$. Since the principal is indifferent between hiring and firing at μ^s, the value of the type L principal at μ^s is given by
$$U_L^*(\mu^s, \mu^a = 0, x_L^*(\mu^s)) = 0 = \mathrm{d}t\left[r\left(\mu^s\lambda\gamma(1+x_L^*(\mu^s)) - a\frac{(x_L^*(\mu^s)^2)}{2}\right) + \mu^s\lambda z^p\right],$$
which implies that
$$r\left(\mu^s\lambda\gamma(1+x_L^*(\mu^s)) - a\frac{(x_L^*(\mu^s)^2)}{2}\right) + \mu^s\lambda z^p = 0.$$

Now observe that $r\left(\mu\lambda\gamma(1+x_L^*(\mu))-a\frac{(x_L^*(\mu)^2)}{2}\right)+\mu\lambda z^p$ is increasing in μ, which implies that the value of the principal is strictly positive when $\mu \in (\mu^s, \mu^c]$. Therefore, the type L has no incentive to deviate when $\mu^L \in [\mu^s, \mu^c]$ since by deviating to a lower investment than $x_L^*(\mu^L)$ type L principal only pauses the relationship (since the agent does not exert effort) and by deviating to a higher investment the principal does no better because $x_L^*(\mu) < x_H^*(\mu)$ for all $\mu \in [\mu^s, \mu^c]$. Note that when $\mu \in [\mu^c, 1)$, type L principal is investing at her myopic optimal level and cannot do better by deviating to any other investment level. Hence I have shown that type L principal has no incentive to deviate from her candidate equilibrium strategy.

- Agent: I first show that if $\mu^L \leq \mu^c$, then the value to the agent facing a type L principal is 0 in the candidate equilibrium. The value of an agent facing type L principal can be written as

$$V^*(\mu^L, \mu_+^a = 0, x_L^*) = r\mathrm{d}t[\mu^L\lambda(1-\gamma)(1+x_L^*) - c]$$
$$+ (1 - r\mathrm{d}t)[\mu^L\lambda \mathrm{d}t z^a + (1-\mu^L\lambda\mathrm{d}t)V^*(\mu^L + \mathrm{d}\mu^L, \mu_+^a = 0, x_L^*)],$$

where

$$\mathrm{d}\mu^L = -\lambda\mu^L(1-\mu^L)\mathrm{d}t.$$

The above expression can be rearranged to give

$$V^*(\mu^L, \mu_+^a = 0, x_L^*) = \underbrace{\mathrm{d}t[r(\mu^L\lambda(1-\gamma)(1+x_L^*) - c) + \mu^L\lambda z^a]}_{A(\mu^L)}$$
$$+ (1 - r\mathrm{d}t)(1-\mu^L\lambda\mathrm{d}t)\underbrace{V^*(\mu^L + \mathrm{d}\mu^L, \mu_+^a = 0, x_L^*)}_{B}.$$

The value of the agent facing a type L principal can be decomposed into an pseudo flow term A and a continuation value B adjusted for discounting and the probability that agent will face a type L principal after $\mathrm{d}t$. Plugging in for x_L^*, it is straightforward to see that $A(\mu^L) = 0$. Also note that we can decompose the continuation value after B in a similar manner. Note that for any belief of type L principal $\mu \leq \mu^c$, we have

$A(\mu) = 0$. $V^*(\mu^L, \mu^a_+ = 0, x^*_L)$ is just a weighted aggregate of $A(\mu)$ for $\mu \in [\mu^s, \mu^L]$ and hence must be equal to 0. Therefore, the agent has no incentive to deviate and terminate the relationship when $\mu^L \in (\mu^s, \mu^c]$.

Next, I show that the value of the agent is strictly positive when $\mu^L \in (\mu^c, 1)$. To see this, simply note that $r(\mu^L \lambda(1-\gamma)(1+x^*_L) - c) + \mu^L \lambda z^a > 0$ when $\mu^L \in (\mu^c, 1)$. Therefore, the agent has no incentive to deviate and terminate the relationship when $\mu^L \in (\mu^c, 1)$.

This establishes that the strategy profile is indeed an equilibrium. I now show that this is also the principal optimal equilibrium. First we observe that the type L principal always invests strictly lower than type H principal's investment who invests optimally. Hence, from Lemma 6 we know that any equilibrium with pooling can be improved upon. This implies that the principal optimal equilibrium must be a fully separating equilibrium. Note that in our candidate equilibrium, the type L principal invests above her optimal investment only when the value of the agent is 0, i.e. the agent's participation constraint binds. Hence, the type L principal cannot decrease her investment when the agent's IR binds and be better off. This establishes that the equilibrium is indeed optimal for the principal.

□

Lemma 34. *Consider any equilibrium σ with separation at (μ^L, μ^a) and the agent facing type L principal exerts effort following investment $x_L(\mu^L, \mu^a) > \dfrac{\mu^L \lambda \gamma}{a}$. Define $V^\sigma(\mu^L, \mu^a_+ = 0, x_L(\mu^L, \mu^a))$ as the value of the agent facing type L principal under σ. Then $V^\sigma(\mu^L, \mu^a_+ = 0, x_L(\mu^L, \mu^a)) = 0$*

Proof. We start by observing that $\dfrac{\mu^L \gamma \lambda}{a}$ is the flow optimal level of investment of type L principal at (μ^L, μ^a). Suppose $V^\sigma(\mu^L, \mu^a_+ = 0, x_L(\mu^L, \mu^a)) > 0$, then consider the deviation of type L at (μ^L, μ^a) to $x' = \dfrac{\mu^L \gamma \lambda}{a}$. We observe that any downward deviation of the type L principal does not change the continuation play as long as the agent continues to exert effort at (μ^L, μ^a). This is because after a downward deviation the agent continues to believe that he is facing the type L principal and the principal's deviation at (μ^L, μ^a) does not affect her belief going forward given that the agent continues to exert effort. There are two possible cases

1. The value of the agent after the deviation, $V^{\sigma}(\mu^L, \mu_+^a = 0, x') > 0$. In this case the agent continues to exert effort and the type L principal is better off because she improves her instantaneous payoff while keeping the continuation unchanged.

2. The value of the agent after the deviation, $V^{\sigma}(\mu^L, \mu_+^a = 0, x') \leq 0$. In this case consider an alternate deviation from type L principal $x'' = x_L(\mu^L, \mu^a) - \epsilon > \frac{\mu^L \gamma \lambda}{a}$ where ϵ is chosen such that $V^{\sigma}(\mu^L, \mu_+^a = 0, x'') > 0$. Under this alternate deviation the agent continues to exert effort and the type L principal is better off because she improves her instantaneous payoff while keeping the continuation unchanged.

Hence $V^{\sigma}(\mu^L, \mu_+^a = 0, x_L(\mu^L, \mu^a)) = 0$. □

B.3 High cost of effort

In this section, I provide the proofs of the case when $C1$ is satisfied , i.e. cost of effort is high.

Definition 19. *Suppose $\mu_0 \leq \bar{\mu}$. Consider a two phase strategy profile $\bar{\sigma}$ such that*

$$\bar{x}_H(\mu^L, \mu^a) = \bar{x}_L(\mu^L, \mu^a) = \bar{x}(\mu^L), \text{ such that } \bar{x}(\mu_L) \in [\frac{\lambda\gamma}{a}, \bar{x}] \text{ if } \mu^L \in [\underline{\mu}, \mu_0];$$

$$\bar{x}_H(\mu^L, \mu^a) = \frac{\gamma\lambda}{a} \text{ if } \mu^L < \underline{\mu};$$

$$\bar{x}_L(\mu^L, \mu^a) = 0 \text{ if } \mu^L < \underline{\mu};$$

$$\bar{e}(\mu^L, \mu_+^a, x) = \begin{cases} 0 \text{ if } \mu_+^a = 0 \text{ and } \mu^L < \mu_0, \\ 1 \text{ otherwise }. \end{cases}$$

*We call $\bar{\sigma}$ **feasible** if it satisfies the individual rationality (IR) constraints of type L principal (see lemma 35) and the agent (see lemma 36) for all $\mu^L \in [\underline{\mu}, \mu_0]$. Note that if $\bar{\sigma}$ is IR for the type L principal then it is IR for type H principal.*

Lemma 35. *The IR constraint of type L principal under $\bar{\sigma}$ when $\mu^L = \mu \in [\underline{\mu}, \mu_0]$ is given by*

$$\int_{\underline{\mu}}^{\mu}\left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu\lambda\gamma(1+\bar{x}(\phi))-\frac{1}{2}a\bar{x}^2(\phi)]\frac{d\phi}{\lambda\phi(1-\phi)}+\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{\mu-\underline{\mu}}{1-\underline{\mu}}z^p\geq 0. \tag{IC_{μ}^{L}}$$

Proof. Starting at $\mu^L=\mu\in[\underline{\mu},\mu_0]$, the probability that a type L principal observes a output and transitions to become a type H principal by the time $\mu^L=\underline{\mu}$ is given by $\mu(1-e^{-\lambda(t(\underline{\mu})-t(\mu))})$ where $t(\underline{\mu})-t(\mu)$ is the time it takes for μ^L to drift down from μ to $\underline{\mu}$. Recall that in the absence of observing a output

$$\frac{\mathrm{d}\mu^L}{\mathrm{d}t}=-\mu^L(1-\mu^L)\lambda.$$

Rearranging and integrating both sides I get

$$\int_{\mu}^{\underline{\mu}}\frac{\mathrm{d}\phi}{\phi(1-\phi)}=-\lambda\int_{t(\mu)}^{t(\underline{\mu})}\mathrm{d}t,$$

which gives us

$$e^{-\lambda(t(\underline{\mu})-t(\mu))}=\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu},$$

leading to

$$\mu(1-e^{-\lambda(t(\underline{\mu})-t(\mu))})=\frac{\mu-\underline{\mu}}{1-\underline{\mu}}.$$

Note that if the principal is type H at $\mu^L=\underline{\mu}$, then under $\bar{\sigma}$ her value is z^p. The value of type L principal under $\bar{\sigma}$ when $\mu^L=\mu$ is then given by

$$U_L^{\bar{\sigma}}(\mu)=\int_{t(\mu)}^{t(\underline{\mu})} re^{-r(s-t(\mu))}[\mu(s)\lambda\gamma(1+\bar{x}(\mu(s)))-\frac{1}{2}a\bar{x}^2(\mu(s))]ds+e^{-r(t(\underline{\mu})-t(\mu))}\frac{\mu-\underline{\mu}}{1-\underline{\mu}}z^p.$$

Changing the variable of integration from time to type L principal's belief we get

$$U_L^{\bar{\sigma}}(\mu)=\int_{\underline{\mu}}^{\mu}\left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu\lambda\gamma(1+\bar{x}(\phi))-\frac{1}{2}a\bar{x}^2(\phi)]\frac{d\phi}{\lambda\phi(1-\phi)}+\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{\mu-\underline{\mu}}{1-\underline{\mu}}z^p.$$

$\square$

Lemma 36. *The IR constraint of the agent under $\bar{\sigma}$ when $\mu^L = \mu \in [\underline{\mu}, \mu_0]$ is given by*

$$\int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_0\lambda(1-\gamma)(1+\bar{x}(\phi)) - c]\frac{d\phi}{\lambda\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}} \frac{\mu_0 - \underline{\mu}}{1-\underline{\mu}} z^a \geq 0. \tag{IC_μ^A}$$

Proof. Starting at $\mu^L = \mu_0$, the probability that the agent will face type H principal is given by $\mu_0(1 - e^{-\lambda(t(\underline{\mu}) - t(\mu_0))})$ where $t(\underline{\mu}) - t(\mu_0)$ is the time it takes for μ^L to drift down from μ_0 to $\underline{\mu}$. Following the steps in the proof of lemma 35, we get

$$\mu_0(1 - e^{-\lambda(t(\underline{\mu}) - t(\mu_0))}) = \frac{\mu_0 - \underline{\mu}}{1 - \underline{\mu}}.$$

Note that if the principal is type H at $\mu^L = \underline{\mu}$, then under $\bar{\sigma}$ the value to the agent at $\mu^L = \underline{\mu}$ is z^a. The value of the agent under $\bar{\sigma}$ when $\mu^L = \mu$ is then given by

$$V^{\bar{\sigma}}(\mu) = \int_{t(\mu)}^{t(\underline{\mu})} re^{-r(s-t(\mu))}[\mu_0\lambda(1-\gamma)(1+\bar{x}(\mu(s))) - c]ds + e^{-r(t(\underline{\mu})-t(\mu))}\frac{\mu_0 - \underline{\mu}}{1-\underline{\mu}} z^a.$$

Changing the variable of integration from time to type L principal's belief we get

$$V^{\bar{\sigma}}(\mu) = \int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_0\lambda(1-\gamma)(1+\bar{x}(\phi)) - c]\frac{d\phi}{\lambda\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}} \frac{\mu_0 - \underline{\mu}}{1-\underline{\mu}} z^a.$$

$\square$

Definition 20. *We call the two phase strategy profile $\bar{\sigma}$ defined in 19* ***just feasible*** *if under $\bar{\sigma}$ the IR constraint for type L principal binds for all $\mu \in [\underline{\mu}, \mu_0]$. I denote a just feasible strategy profile by $\hat{\sigma}$ and the corresponding pooling investment function by $\hat{x} : [\underline{\mu}, \mu_0] \to [\frac{\lambda\gamma}{a}, \bar{x}]$.*

Lemma 37. *Suppose $\hat{\sigma}$ is a just feasible strategy profile. The just feasible pooling investment* *function $\hat{x}$ is the unique solution to the first order non linear differential equation given by:*

$$\hat{x}'(\mu) = \frac{\mu\lambda\gamma(1+\hat{x}(\mu))(1+\frac{r}{\lambda}) - \frac{a\hat{x}(\mu)^2}{2}(\frac{r}{\lambda}+2\mu-1)}{\mu(1-\mu)[a\hat{x}(\mu)-\mu\lambda\gamma]}.$$

Proof. Denote by $U_T^{\hat{\sigma}}(\mu)$, the value of a type $T \in \{L, H\}$ principal under $\hat{\sigma}$ at $\mu^L = \mu \leq \mu_0$. Then by definition of $\hat{\sigma}$, we must have $U_L^{\hat{\sigma}}(\mu) = 0$ for all $\mu \in [\underline{\mu}, \mu_0]$. This means

$$\begin{aligned} U_L^{\hat{\sigma}}(\mu) = r\mathrm{d}t &\left[\mu\lambda\gamma(1+\hat{x}(\mu)) - \frac{a\hat{x}^2(\mu)}{2}\right] \\ &+ (1-r\mathrm{d}t)[\mu\lambda\mathrm{d}t U_H^{\hat{\sigma}}(\mu+\mathrm{d}\mu) + (1-\mu\lambda\mathrm{d}t)U_L^{\hat{\sigma}}(\mu+\mathrm{d}\mu)] = 0. \end{aligned}$$

Since by definition of $\hat{\sigma}$, $U_L^{\hat{\sigma}}(\mu+\mathrm{d}\mu) = 0$ we get

$$U_L^{\hat{\sigma}}(\mu) = r\mathrm{d}t\left[\mu\lambda\gamma(1+\hat{x}(\mu)) - \frac{a\hat{x}^2(\mu)}{2}\right] + (1-r\mathrm{d}t)\mu\lambda\mathrm{d}t U_H^{\hat{\sigma}}(\mu+\mathrm{d}\mu) = 0.$$

Since $\hat{x}(\mu) \in [\frac{\lambda\gamma}{a}, \bar{x}]$ and $0 \leq U_H^{\hat{\sigma}}(\mu) \leq z^p$ for all $\mu \in [\underline{\mu}, \mu_0]$, $U_H^{\hat{\sigma}}(\mu)$ is continuous on $[\underline{\mu}, \mu_0]$ (Perhaps a lemma for this). This implies that $\lim_{\Delta\mu\to 0} U_H^{\hat{\sigma}}(\mu+\Delta\mu) = U_H^{\hat{\sigma}}(\mu)$. Which implies that we can write

$$U_L^{\hat{\sigma}}(\mu) = \mathrm{d}t\left[r\left[\mu\lambda\gamma(1+\hat{x}(\mu)) - \frac{a\hat{x}^2(\mu)}{2}\right] + \mu\lambda U_H^{\hat{\sigma}}(\mu)\right] = 0.$$

Since by definition of $\hat{\sigma}$, $U_L^{\hat{\sigma}}(\mu) = 0$, which implies that

$$r\left[\mu\lambda\gamma(1+\hat{x}(\mu)) - \frac{a\hat{x}^2(\mu)}{2}\right] + \mu\lambda U_H^{\hat{\sigma}}(\mu) = 0. \tag{B.1}$$

for all $\mu \in [\underline{\mu}, \mu_0]$ except for any set $D \subset [\underline{\mu}, \mu_0]$ of measure zero. Further simplification of equation B.1 yields

$$\hat{x}(\mu) = \frac{r\mu\lambda\gamma + \sqrt{r^2\mu^2\lambda^2\gamma^2 + 2ar\mu\lambda(r\gamma + U_H^{\hat{\sigma}}(\mu))}}{ar}.$$

Observe that $U_H^{\hat{\sigma}}(\underline{\mu}) = z^p$ which implies that

$$\hat{x}(\underline{\mu}) = \frac{r\underline{\mu}\lambda\gamma + \sqrt{r^2\underline{\mu}^2\lambda^2\gamma^2 + 2ar\underline{\mu}\lambda(r\gamma + z^p)}}{ar} = \frac{\lambda\gamma}{a}. \tag{B.2}$$

Recall from lemma 35 that $U_L^{\hat{\sigma}}(\mu)$ is given by

$$U_L^{\hat{\sigma}}(\mu) = \int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu\lambda\gamma(1+\hat{x}(\phi)) - \frac{1}{2}a\hat{x}^2(\phi)]\frac{d\phi}{\lambda\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{\mu-\underline{\mu}}{1-\underline{\mu}}z^p.$$

Similarly $U_H^{\hat{\sigma}}(\mu)$ is given by

$$U_H^{\hat{\sigma}}(\mu) = \int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\lambda\gamma(1+\hat{x}(\phi)) - \frac{1}{2}a\hat{x}^2(\phi)]\frac{d\phi}{\lambda\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}} z^p.$$

Evaluating $U_H^{\hat{\sigma}}(\mu) - U_L^{\hat{\sigma}}(\mu)$ we get:

$$U_H^{\hat{\sigma}}(\mu) - U_L^{\hat{\sigma}}(\mu) = (1-\mu)\int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{1-\mu}{1-\underline{\mu}}z^p.$$

Under $\hat{\sigma}$, $U_L^{\hat{\sigma}}(\mu) = 0$, hence we get

$$\begin{aligned}
U_H^{\hat{\sigma}}(\mu) &= (1-\mu)\int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{1-\mu}{1-\underline{\mu}}z^p, \\
&= (1-\mu)\left[\int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{z^p}{1-\underline{\mu}}\right], \\
&= (1-\mu)\left[\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\left[\int_{\underline{\mu}}^{\mu} \left[\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\right]^{\frac{r}{\lambda}}\frac{z^p}{1-\underline{\mu}}\right].
\end{aligned}$$

Plugging in $U_H^{\hat{\sigma}}(\mu)$ in equation B.1 we get

$$\begin{aligned}
0 = {} & r[\mu\lambda\gamma(1+\hat{x}(\mu)) - \frac{1}{2}a\hat{x}(\mu)^2] \\
& + \mu\lambda(1-\mu)\left[\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\left[\int_{\underline{\mu}}^{\mu} \left[\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\right]^{\frac{r}{\lambda}}\frac{z^p}{1-\underline{\mu}}\right],
\end{aligned}$$

which on rearrangement yields

$$r[\frac{1}{2}a\hat{x}(\mu)^2 - \mu\lambda\gamma(1+\hat{x}(\mu))]\frac{1}{\mu\lambda(1-\mu)}\left[\frac{\mu}{1-\mu}\right]^{\frac{r}{\lambda}}$$
$$= \int_{\underline{\mu}}^{\mu}\left[\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\gamma(1+\hat{x}(\phi))]\frac{d\phi}{\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\right]^{\frac{r}{\lambda}}\frac{z^p}{1-\underline{\mu}}.$$

Differentiating the LHS above equation with respect to μ we get

$$\frac{\mathrm{d}LHS}{\mathrm{d}\mu} = r\left[\hat{x}'(\mu)(ax(\mu) - \mu\lambda\gamma) - \lambda\gamma(1+\hat{x}(\mu))\right]\frac{1}{\mu\lambda(1-\mu)}\left[\frac{\mu}{1-\mu}\right]^{\frac{r}{\lambda}}$$
$$+ r[\frac{1}{2}a\hat{x}(\mu)^2 - \mu\lambda\gamma(1+\hat{x}(\mu))]\frac{\frac{r}{\lambda}+2\mu-1}{\lambda\mu^2(1-\mu)^2}\left[\frac{\mu}{1-\mu}\right]^{\frac{r}{\lambda}}.$$

Differentiating the RHS above equation with respect to μ we get

$$\frac{\mathrm{d}RHS}{\mathrm{d}\mu} = \left[\frac{\mu}{1-\mu}\right]^{\frac{r}{\lambda}}\frac{r[\gamma(1+\hat{x}(\mu))]}{\mu(1-\mu)}.$$

On equating $\frac{\mathrm{d}LHS}{\mathrm{d}\mu}$ to $\frac{\mathrm{d}RHS}{\mathrm{d}\mu}$ and and simplifying we get

$$\hat{x}'(\mu) = \frac{\mu\lambda\gamma(1+\hat{x}(\mu))(1+\frac{r}{\lambda}) - \frac{a\hat{x}(\mu)^2}{2}(\frac{r}{\lambda}+2\mu-1)}{\mu(1-\mu)[a\hat{x}(\mu) - \mu\lambda\gamma]}.$$

Note that the above expression is a first order non linear differential equation and standard results in the theory of differential equation guarantee existence and uniqueness of the solution. The precise solution $\hat{x}(\mu)$ is pinned down by using the boundary condition given by equation B.2.

□

Definition 21. *For any pooling investment function $x(\cdot)$ defined for $\mu \in [\underline{\mu}, \mu_0]$. Define the principal average investment under $x(\cdot)$ at μ_1, $A(x,\mu_1)$ as the unique solution of*

$$\int_{t(\mu_1)}^{t(\underline{\mu})} re^{-r(s-t(\mu_1))}x(\mu(s))ds = \int_{t(\mu_1)}^{t(\underline{\mu})} re^{-r(s-t(\mu_1))}A(x,\mu_1)ds.$$

Note that $A(x,\mu_1)$ is a constant for a given pooling investment function x and belief μ_1.

Lemma 38. *$A(\hat{x}, \mu)$ is increasing in μ.*

Proof. Recall from lemma 37 that

$$\hat{x}'(\mu) = \frac{\mu\lambda\gamma(1+\hat{x}(\mu))(1+\frac{r}{\lambda}) - \frac{a\hat{x}(\mu)^2}{2}(\frac{r}{\lambda}+2\mu-1)}{\mu(1-\mu)[a\hat{x}(\mu)-\mu\lambda\gamma]}.$$

It is easy to see that the denominator $\mu(1-\mu)[a\hat{x}(\mu)-\mu\lambda\gamma]$ is always positive. We focus on the numerator. In particular I define $\tilde{x}(\mu)$ such that the numerator is equal to 0. That is

$$\mu\lambda\gamma(1+\tilde{x}(\mu))(1+\frac{r}{\lambda}) - \frac{a\tilde{x}(\mu)^2}{2}(\frac{r}{\lambda}+2\mu-1) = 0.$$

On solving the quadratic equation we have

$$\tilde{x}(\mu) = \beta(\mu) + \sqrt{\beta(\mu)^2 + 2\beta(\mu)},$$

where

$$\beta(\mu) = \frac{\mu\lambda\gamma(1+\frac{r}{\lambda})}{a(\frac{r}{\lambda}+2\mu-1)}.$$

It can be shown that

$$\beta'(\mu) \text{ is } \begin{cases} \geq 0 \text{ if } \frac{r}{\lambda} \geq 1, \\ < 0 \text{ if } \frac{r}{\lambda} < 1. \end{cases}$$

This implies that

$$\tilde{x}'(\mu) \text{ is } \begin{cases} \geq 0 \text{ if } \frac{r}{\lambda} \geq 1, \\ < 0 \text{ if } \frac{r}{\lambda} < 1. \end{cases}$$

Note that

$$\hat{x}'(\mu) > 0 \text{ if } \hat{x}(\mu) < \tilde{x}(\mu),$$

$$\hat{x}'(\mu) = 0 \text{ if } \hat{x}(\mu) = \tilde{x}(\mu),$$

$$\hat{x}'(\mu) < 0 \text{ if } \hat{x}(\mu) > \tilde{x}(\mu).$$

It can be shown that

$$\lim_{\mu \uparrow 1} \tilde{x}(\mu) = \lim_{\mu \uparrow 1} \hat{x}(\mu) = \frac{\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 2a\lambda\gamma}}{a}.$$

and $\hat{x}'(\underline{\mu}) > 0$, implying that $\hat{x}(\underline{\mu}) < \tilde{x}(\underline{\mu})$.

Claim 1. *If $\frac{r}{\lambda} \geq 1$ then $\hat{x}(\mu)$ is increasing in $[\underline{\mu}, 1]$.*

Proof. We have $\hat{x}(\underline{\mu}) < \tilde{x}(\underline{\mu})$ and $\lim_{\mu \uparrow 1} \tilde{x}(\mu) = \lim_{\mu \uparrow 1} \hat{x}(\mu)$. Suppose $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ intersect in $[\underline{\mu}, 1)$. In particular, suppose that the smallest belief at which they intersect is $\mu_1 \in [\underline{\mu}, 1)$. That is $\hat{x}(\mu_1) = \tilde{x}(\mu_1)$. Then we know that $\hat{x}'(\mu_1) = 0$. We also know that $\tilde{x}'(\mu_1) \geq 0$ since $\frac{r}{\lambda} \geq 1$. This implies that $\hat{x}(\mu_1 - \epsilon) \geq \tilde{x}(\mu_1 - \epsilon)$ for some $\epsilon > 0$, which contradicts that μ_1 is the smallest belief at which $\hat{x}(\mu) = \tilde{x}(\mu)$. Therefore I have shown that $\hat{x}(\underline{\mu})$ and $\tilde{x}(\underline{\mu})$ never intersect when $\mu < 1$. This implies that $\hat{x}(\mu)$ is increasing in $[\underline{\mu}, 1]$. □

Claim 2. *If $\frac{r}{\lambda} < 1$ then $\hat{x}(\mu)$ is either increasing or single peaked in $[\underline{\mu}, 1]$.*

Proof. I first show that if $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ intersect, then they intersect exactly once. To that end suppose that $\hat{x}(\mu_2) = \tilde{x}(\mu_2)$. Note that since $\frac{r}{\lambda} < 1$, $\tilde{x}'(\mu_2) < 0$. Also $\hat{x}'(\mu_2) = 0$, which implies that $\hat{x}(\mu) > \tilde{x}(\mu)$ when $\mu \in (\mu_2, \mu_2 + \epsilon)$ for some $\epsilon > 0$. Using an argument identical to claim 1 we can say that $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ never intersect in $(\mu_2, 1)$. This implies that if $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ intersect, then they intersect exactly once. Note that when $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ intersect, $\hat{x}(\mu)$ is single peaked (increasing before intersection and decreasing after). If $\hat{x}(\mu)$ and $\tilde{x}(\mu)$ never intersect then $\hat{x}(\mu)$ is increasing in $[\underline{\mu}, 1]$ because $\hat{x}(\mu)$ ¡ $\tilde{x}(\mu)$ when $\mu \in [\underline{\mu}, 1]$. □

From Claim 1 and Claim 2 there are two possible cases.

1. $\hat{x}(\mu)$ is increasing in $[\underline{\mu}, 1]$: In this case it is easy to see that $A(\hat{x}, \mu)$ is increasing in μ.

2. $\hat{x}(\mu)$ is single peaked in $[\underline{\mu}, 1]$: First we observe that $A(\hat{x}, 1) = \frac{\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 2a\lambda\gamma}}{a}$ because when $\mu = 1$, the the principal is essentially of type H and the investment of a type L principal such that her value is equal to 0 is given by $\frac{\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 2a\lambda\gamma}}{a}$. Next, $A(\hat{x}, \mu) < A(\hat{x}, 1)$ when $\mu < 1$ since the type L principal's average investment under $\hat{x}$ is always

less than that of the type H principal. Suppose $\hat{x}(\mu)$ is increasing in $[\underline{\mu}, \mu_1)$ and decreasing in $(\mu_1, 1)$. As shown in case(1), $A(\hat{x}, \mu)$ is increasing when $\mu \in [\underline{\mu}, \mu_1)$. Since $\hat{x}(\mu)$ is decreasing in $(\mu_1, 1)$, it must be the case that $\hat{x}(\mu) > \lim_{\mu \uparrow 1} \hat{x}(\mu) = \frac{\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 2a\lambda\gamma}}{a}$ when $\mu \in [\mu_1, 1)$. In particular note that $\hat{x}(\mu) > \frac{\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 2a\lambda\gamma}}{a} > A(\hat{x}, \mu)$ when $\mu \in [\mu_1, 1)$. This implies that $A(\hat{x}, \mu)$ is increasing when $\mu \in [\mu_1, 1)$. Hence $A(\hat{x}, \mu)$ is increasing in μ.

$\square$

Definition 22. *Define the agent minimum average investment at $\mu^\dagger > \underline{\mu}$, $B(\mu^\dagger)$ as the unique solution of*

$$\int_{t(\mu^\dagger)}^{t(\underline{\mu})} re^{-r(s - t(\mu^\dagger))} \left[\mu^\dagger \lambda(1-\gamma)(1 + B(\mu^\dagger)) - c\right] ds + e^{-r(t(\underline{\mu}) - t(\mu^\dagger))} \frac{\mu^\dagger - \underline{\mu}}{1 - \underline{\mu}} z^a = 0.$$

Lemma 39. *$B(\mu^\dagger)$ is decreasing in $\mu^\dagger$.*

Proof. I integrate the expression

$$\int_{t(\mu^\dagger)}^{t(\underline{\mu})} re^{-r(s - t(\mu^\dagger))} \left[\mu^\dagger \lambda(1-\gamma)(1 + B(\mu^\dagger)) - c\right] ds + e^{-r(t(\underline{\mu}) - t(\mu^\dagger))} \frac{\mu^\dagger - \underline{\mu}}{1 - \underline{\mu}} z^a = 0.$$

to get

$$\left[1 - \left[\frac{\underline{\mu}}{1-\underline{\mu}} \frac{1-\mu^\dagger}{\mu^\dagger}\right]^{\frac{r}{\lambda}}\right] \left[\mu^\dagger \lambda(1-\gamma)(1 + B(\mu^\dagger)) - c\right] + \left[\frac{\underline{\mu}}{1-\underline{\mu}} \frac{1-\mu^\dagger}{\mu^\dagger}\right]^{\frac{r}{\lambda}} \frac{\mu^\dagger - \underline{\mu}}{1 - \underline{\mu}} z^a = 0.$$

Which can be rewritten as

$$\begin{aligned} 0 = &\left[1 - \left[\frac{\underline{\mu}}{1-\underline{\mu}} \frac{1-\mu^\dagger}{\mu^\dagger}\right]^{\frac{r}{\lambda}}\right] \left[\mu^\dagger \lambda(1-\gamma)(1 + \frac{\lambda\gamma}{a}) - c\right] \\ &+ \left[\frac{\underline{\mu}}{1-\underline{\mu}} \frac{1-\mu^\dagger}{\mu^\dagger}\right]^{\frac{r}{\lambda}} \frac{\mu^\dagger - \underline{\mu}}{1 - \underline{\mu}} z^a + \left[1 - \left[\frac{\underline{\mu}}{1-\underline{\mu}} \frac{1-\mu^\dagger}{\mu^\dagger}\right]^{\frac{r}{\lambda}}\right] \left[\mu^\dagger \lambda(1-\gamma) \left(B(\mu^\dagger) - \frac{\lambda\gamma}{a}\right)\right]. \end{aligned}$$

I define $\bar{V}(\mu;\bar{B})$ as

$$\bar{V}(\mu;\bar{B}) = \left[1-\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\right]\left[\mu\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c\right]$$
$$+\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{\mu-\underline{\mu}}{1-\underline{\mu}}z^a+\left[1-\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\right]\left[\mu\lambda(1-\gamma)\left(\bar{B}-\frac{\lambda\gamma}{a}\right)\right].$$

It is easy to see that if $\bar{B}_1 > \bar{B}_2$ then $\bar{V}(\mu;\bar{B}_1) > \bar{V}(\mu;\bar{B}_2)$ for all μ. Therefore, if $\bar{V}(\mu;\bar{B}) = 0$, then $\bar{B}$ must be unique, which in turn implies that as $\bar{B}$ goes up the belief μ that solves $\bar{V}(\mu;\bar{B}) = 0$, decreases.

□

Definition 23. *Define $\bar{\mu}$ as the belief at which*

$$B(\bar{\mu}) = \frac{\lambda\gamma}{a}.$$

Lemma 40. *Autarkic equilibrium exists at state $(\mu,0)$ when $\mu \in [\underline{\mu},\bar{\mu}]$.*

Proof. Consider a strategy profile σ^m, where

$$x_H^m(\mu,\mu^a) = \frac{\lambda\gamma}{a};\ x_L^m(\mu,\mu^a) = \begin{cases} \frac{\lambda\gamma}{a} \text{ if } \mu \geq \underline{\mu} \\ 0 \text{ if } \mu < \underline{\mu} \end{cases}$$

$$e^m(\mu,\mu_+^a,x) = \begin{cases} 0 \text{ if } \mu_+^a = 0, \\ 1 \text{ if } \mu_+^a = 1. \end{cases}$$

At the state $(\bar{\mu},0)$, note that by the definition of $\bar{\mu}$, the value of the agent $V^m(\bar{\mu})$ under σ^m is equal to 0. Note that the autarkic strategy profile (defined in the proof of Lemma 5) starting at state $(\bar{\mu},0)$ specifies s strictly lower level of investment for type L principal in when $\mu^L \in [\underline{\mu},\bar{\mu}]$ compared to σ^m. Hence, the value of the agent under the autarkic strategy profile is strictly negative if the agent chooses to deviate and exert effort at any point. Hence, an autarkic equilibrium exists at state $(\bar{\mu},0)$. It is easy to see that the above argument holds for all states $(\mu,0)$ such that $\mu \in [\underline{\mu},\bar{\mu}]$ and hence , an autarkic equilibrium exists at state $(\mu, 0)$ such that $\mu \in [\underline{\mu}, \bar{\mu}]$. $\square$

Lemma 41. *Suppose there exists a feasible two phase strategy profile σ at state $\mathcal{S} = (\mu_1, 0)$ with $\mu_1 < \bar{\mu}$ and $V^{\sigma}(\mathcal{S}) > 0$. Then, there exists another feasible two phase strategy profile σ_0 at $\mathcal{S}$ such that $V^{\sigma_0}(\mathcal{S}) = 0$.*

Proof. Suppose x^{σ} is the pooling investment function in the pooling region $[\underline{\mu}, \mu_1]$ under σ. First observe that the agent's value at any $\mu_L = \mu \in [\underline{\mu}, \mu_1]$ under σ is given by

$$V^{\sigma}(\mu) = \int_{\underline{\mu}}^{\mu} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_1\lambda(1-\gamma)(1+x^{\sigma}(\phi))-c]\frac{d\phi}{\lambda\phi(1-\phi)} + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}} \frac{\mu_1 - \underline{\mu}}{1-\underline{\mu}} z^a.$$

Note that $V^{\sigma}(\mu)$ is continuous in $[\underline{\mu}, \mu_1]$. Since $V^{\sigma}(\mu_1) > 0$, we have two cases.

Case 1: $V^{\sigma}(\mu) > 0$ for all $\mu \in [\underline{\mu}, \mu_1]$: Consider an alternate two phase strategy profile σ' with the following investment function in the pooling phase:

$$x^{\sigma'}(\mu) = \begin{cases} x^{\sigma}(\mu) \text{ if } \mu \in [\underline{\mu}, \hat{\mu}], \\ \frac{\lambda\gamma}{a} \text{ if } \mu \in (\hat{\mu}, \mu_1]. \end{cases}$$

where $\hat{\mu} \in [\underline{\mu}, \mu_1]$. The value of the agent at $\mu_L = \mu_1$ under σ' is given by

$$\begin{aligned} V^{\sigma'}(\mu_1; \hat{\mu}) = &\int_{\underline{\mu}}^{\hat{\mu}} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_1\lambda(1-\gamma)(1+x^{\sigma}(\phi))-c]\frac{d\phi}{\lambda\phi(1-\phi)} \\ &+ \int_{\hat{\mu}}^{\mu_1} \left[\frac{1-\mu}{\mu}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_1\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c]\frac{d\phi}{\lambda\phi(1-\phi)} \\ &+ \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}} \frac{\mu_1 - \underline{\mu}}{1-\underline{\mu}} z^a. \end{aligned}$$

Next, note that $V^{\sigma'}(\mu_1; \hat{\mu} = \underline{\mu}) < 0$. To see this note that from lemma 39 we know that $B(\mu_1) > \frac{\lambda\gamma}{a}$ since $B(\bar{\mu}) = \frac{\lambda\gamma}{a}$ and $\mu_1 < \bar{\mu}$. Since $B(\mu_1) > \frac{\lambda\gamma}{a}$, we must have

$$\int_{\underline{\mu}}^{\mu_1}\left[\frac{1-\mu_1}{\mu_1}\frac{\phi}{1-\phi}\right]^{\frac{r}{\lambda}} r[\mu_1\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c]\frac{d\phi}{\lambda\phi(1-\phi)}+\left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu_1}{\mu_1}\right]^{\frac{r}{\lambda}}\frac{\mu_1-\underline{\mu}}{1-\underline{\mu}}z^a<0.$$

Also note that $V^{\sigma'}(\mu_1;\hat{\mu}=\mu_1)=V^{\sigma}(\mu_1)>0$ by assumption. It is easy to see that $V^{\sigma'}(\mu_1;\hat{\mu})$ is continuous in $\hat{\mu}$ which implies that there must exist a $\tilde{\mu}$ such that $V^{\sigma'}(\mu_1;\hat{\mu}=\tilde{\mu})=0$. I now show that the strategy profile σ' with $\hat{\mu}=\tilde{\mu}$ is feasible. I first show that σ' satisfies the IR constraints of the type L principal for all $\mu\in[\underline{\mu},\mu_1]$. Since σ is feasible, the IR constraints of the the type L principal for all $\mu\in[\underline{\mu},\tilde{\mu}]$ are satisfied. Note that $x^{\sigma}(\mu)\geq x^{\sigma'}(\mu)=\frac{\lambda\gamma}{a}$ for all $\mu\in(\tilde{\mu},\bar{\mu}]$ which implies that the IR constraints of the the type L principal for all $\mu\in(\tilde{\mu},\bar{\mu}]$ are satisfied.

Next I show that σ' satisfies the IR constraints of the agent for all $\mu\in[\underline{\mu},\mu_1]$. On the lines argued above, it is easy to see that the IR constraints of the the agent for all $\mu\in[\underline{\mu},\tilde{\mu}]$ are satisfied. Next note that

$$\mu_1\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c<0,$$

which implies that the value of the agent under σ' at $\mu_L=\mu$, $V^{\sigma'}(\mu)$ is increasing in μ when $\mu\in(\tilde{\mu},\mu_1]$. Since $V^{\sigma'}(\mu_1)=0$, it must be the case that $V^{\sigma'}(\mu_1)>0$ for all $\mu\in(\tilde{\mu},\mu_1]$, which establishes that which implies that the IR constraints of the the agent for all $\mu\in(\tilde{\mu},\bar{\mu}]$ are satisfied.

Since σ' with $\hat{\mu}=\tilde{\mu}$ satisfies the IR constraints of both type L principal and the agent when $\mu\in[\underline{\mu},\mu_1]$, σ' with $\hat{\mu}=\tilde{\mu}$ is feasible and $V^{\sigma'}(\mu_1)=0$ and hence proof is complete.

Case 2: There exists a set $\mathcal{M}^{\sigma}\subset[\underline{\mu},\mu_1]$ such that $V^{\sigma}(\mu)=0$ for $\mu\in\mathcal{M}^{\sigma}$: In this case, define $\tilde{\mu}=\sup\mathcal{M}^{\sigma}$. Consider an alternate two phase strategy profile σ' with the following investment function in the pooling phase:

$$x^{\sigma'}(\mu) = \begin{cases} x^{\sigma}(\mu) \text{ if } \mu \in [\underline{\mu}, \tilde{\mu}], \\ \frac{c}{\mu_1 \lambda(1-\gamma)} - 1 \text{ if } \mu \in (\tilde{\mu}, \mu_1]. \end{cases}$$

Note that $\mu_1 \lambda(1-\gamma)(1 + x^{\sigma'}(\mu)) - c = 0$ when $\mu \in (\tilde{\mu}, \mu_1]$, this implies that $V^{\sigma'}(\mu) = 0$ when $\mu \in (\tilde{\mu}, \mu_1]$. This implies that the agent's IR constraint's are satisfied for $\mu \in [\underline{\mu}, \mu_1]$ under σ'. Next observe that $V^{\sigma}(\mu) > V^{\sigma'}(\mu) = 0$ when $\mu \in (\tilde{\mu}, \mu_1]$. This implies that if σ' is IR for type L principal, since σ' specifies a weakly lower level of investment compared to σ when $\mu \in (\tilde{\mu}, \mu_1]$. Since σ' satisfies the IR constraints of both type L principal and the agent when $\mu \in [\underline{\mu}, \mu_1]$, σ' is feasible and $V^{\sigma'}(\mu_1) = 0$ and hence proof is complete.

□

Lemma 42. *Suppose there exists a feasible two phase strategy profile at state $\mathcal{S} = (\mu_1, 0)$ where $\mu_1 < \bar{\mu}$. Then, there exists a feasible two phase strategy profile at state $\mathcal{S}' = (\mu_1', 0)$ where $\mu_1 < \mu_1' \leq \bar{\mu}$.*

Proof. Recall from Lemma 39 that the agent's minimum average investment at μ_1' is smaller than the agent's minimum average investment at μ_1. Also note from lemma 38 we know that the type L principal's average maximum investment is increasing in μ. Note that since a feasible two phase strategy profile exists at state $(\mu_1, 0)$ and the minimum average investment demanded by the agent is smaller at $(\mu_1', 0)$ and the maximum average investment that type L is willing to invest is higher at $(\mu_1', 0)$, it must be the case that a feasible two phase exists at $(\mu_1, 0)$. □

Lemma 43. *There exists a belief $\mu_g \in (\underline{\mu}, \bar{\mu})$ such that,*

1. *if $\mu \geq \mu_g$, then there exist a two feasible two phase strategy profile at state $(\mu, 0)$;*
2. *if $\mu < \mu_g$, then there does not exist a feasible two phase strategy profile at state $(\mu, 0)$.*

Proof. By definition of $\underline{\mu}$, the type L principal does not invest any higher than $\frac{\gamma\lambda}{a}$ at $\underline{\mu}$. From lemma 39 we know that $B(\mu)$ is decreasing in μ. We also know that $B(\bar{\mu}) = \frac{\gamma\lambda}{a}$. Since $\bar{\mu} > \underline{\mu}$, we know that $B(\underline{\mu}) > \frac{\gamma\lambda}{a}$. Hence there does not exist a two feasible two phase strategy profile at state $(\underline{\mu}, 0)$. By continuity, there exists an $\epsilon > 0$ such that there does not exist a two feasible two phase strategy profile at state $(\mu, 0)$ when $\mu \in [\underline{\mu}, \underline{\mu} + \epsilon)$. Suppose μ_g is the smallest belief such that a feasible two phase strategy profile exists at state $(\mu_g, 0)$, then we know from lemma 42 that a feasible two phase strategy profile exists at all states $(\mu, 0)$ such that $\mu \geq \mu_g$.

□

Lemma 44. *Suppose $\mu_1 \in [\mu_g, \bar{\mu}]$. Then, there exists a two phase equilibrium at state $(\mu_1, 0)$.*

Proof. We know from Lemma 43 that there exists a feasible two phase strategy profile at $(\mu_1, 0)$. Let us denote a feasible two phase strategy profile by $\hat{\sigma} = (\hat{x}_H, \hat{x}_L, \hat{e})$ given by

$$\hat{x}_H(\mu, \mu^a) = \hat{x}_L(\mu, \mu^a) = \hat{x}(\mu) \textit{ such that } \hat{x}(\mu) \in [\frac{\lambda\gamma}{a}, \bar{x}] \textit{ for } \mu \in [\underline{\mu}, \mu_1];$$

$$\hat{x}_H(\mu, \mu^a) = \frac{\gamma\lambda}{a};\ \hat{x}(\mu^L, \mu^a) = 0 \textit{ for } \mu < \underline{\mu};$$

$$\hat{e}(\mu, \mu^a_+, x) = \begin{cases} 1 \textit{ if } \mu^a_+ = 0 \textit{ and } \mu = \mu_1; \\ 0 \textit{ if } \mu^a_+ = 0 \textit{ and } \mu \neq \mu_1, \\ 1 \textit{ otherwise }. \end{cases}$$

I now show that the above two phase strategy profile can be supported as an equilibrium using the threat of autarky following any downward deviation by the principal when $\mu^L \in [\underline{\mu}, \mu_1]$. Lemma 40 tells us that an autarkic equilibrium exists at state $(\mu, 0)$ when $\mu \in [\underline{\mu}, \bar{\mu}]$ which implies that an autarkic equilibrium exists at state $(\mu, 0)$ when $\mu \in [\underline{\mu}, \mu_1]$. I fist show that both types of principal have no incentive to deviate in the pooling region ($\mu^L \in [\underline{\mu}, \bar{\mu}]$). Note that the value of both types of principal is non negative at any point of the relationship starting at state $(\mu_1, 0)$ since $\hat{\sigma}$ is feasible. Any upward deviation principal during the pooling phase is unprofitable because it only reduces the principal's value without affecting the agent's behavior. Any downward deviation by the principal during the pooling phase leads to autarky that yields 0 to the principal and hence is unprofitable. In the firing phase, by deviating to atleast $\frac{\gamma\lambda}{a}$, the type L principal can induce effort from the agent, but this deviation leaves the type L principal strictly worse off. Type H principal has no incentive to deviate in the firing phase since she is investing her optimal level of investment and agent is exerting effort.

The agent has no incentive to deviate in the pooling phase since $\hat{\sigma}$ is feasible and any deviation leads to autarky which cannot improve the agent's payoff. Hence $\hat{\sigma}$ can be supported as an equilibrium by using the threat of autarky. $\square$

Lemma 45. *Suppose $\mu \in (\underline{\mu}, \bar{\mu}]$. There does not exist an equilibrium at state $(\mu, 0)$ that specifies separation at $\mu^L \in (\underline{\mu}, \mu]$.*

Proof. Suppose there exists an equilibrium σ at the state $(\mu, 0)$ that specifies separation at $\mu^L = \mu_1 \in (\underline{\mu}, \mu]$. Note that post separation, type H principal invests $\frac{\lambda\gamma}{a}$ going forward since the agent now knows the he is facing the type H principal and hence willing to exert effort. Note that for σ to be feasible, it must specify average investments are higher than $\frac{\lambda\gamma}{a}$ starting at state $(\mu_1, 0)$. This is because from lemma 39 we know that the agent minimum average investment at state $(\mu, 0)$ given by $B(\mu)$ is decreasing in μ and $B(\bar{\mu}) = \frac{\lambda\gamma}{a}$. Note that this means that the type L principal has a deviation. She can mimic type H principal when $\mu^L = \mu_1$ and induce effort from the agent by investing $\frac{\lambda\gamma}{a}$ until beliefs reach $\underline{\mu}$ at which point she quits. Note that by mimicking type H, she can guarantee herself an average investment that is strictly lower. Also note that her investment does not vary with time post her deviation. Since the principal's investment cost is convex, this guarantees that type L principal improves her payoff by deviating. Hence there cannot be separation before $\mu^L = \underline{\mu}$ in any equilibrium starting at state $(\mu, 0)$ if $\mu \in (\underline{\mu}, \bar{\mu}]$.

$\square$

Proof of Proposition 12. We know from lemma 44 that a two phase equilibrium exist at state $(\mu, 0)$ where $\mu \in [\mu_g, \bar{\mu}]$. Moreover, Lemma 45 tells us that any equilibrium where the agent exerts effort at state $(\mu_0, 0)$ with $\mu_0 \in [\underline{\mu}, \bar{\mu}]$ cannot exhibit separation at any point. This implies that a principal optimal equilibrium, if it exists must be a two phase equilibrium. I now argue that a principal optimal equilibrium, does in fact exist. To that end, note that every two phase equilibrium at $(\mu_0, 0)$ is characterized by the corresponding pooling investment function that maps $[\underline{\mu}, \mu_0]$ to $[\frac{\lambda\gamma}{a}, \bar{x}]$. Denote by by $\mathcal{X}(\mu_0)$ the set of all equilibrium pooling investment functions at state $(\mu_0, 0)$. Note that each of these equilibrium pooling functions is bounded an satisfies the type L principal's IR (IC_μ^L) and agent's IR (IC_μ^A) at every $\mu \in [\underline{\mu}, \mu_0]$. I show that the set $\mathcal{X}(\mu_0)$ is compact which implies that there exists a pooling equilibrium function that maximizes the principal's value over $\mathcal{X}(\mu_0)$.

□

Proof of Proposition 11 . I first show that the strategy profile is an equilibrium. I then show that it is indeed the principal's optimal equilibrium.

We first observe that the type H principal has no incentive to deviate at any stage since she is investing at her optimal level $(\frac{\lambda\gamma}{a})$ and inducing effort from the agent. Next I consider type L principal's behavior. We start with the quitting phase. Note that in this phase, the agent will exert effort if the type L principal invests $(\frac{\lambda\gamma}{a})$ from the beginning of the phase. But note that by the definition of $\underline{\mu}$, the type L principal strictly prefers to not induce effort by investing $(\frac{\lambda\gamma}{a})$ when her belief about project quality is strictly smaller than $\underline{\mu}$, which implies that the type L principal is behaving optimally by not investing.

In the pooling phase, I first show that under this strategy profile, the value of the type L principal is strictly positive when $\mu \in (\underline{\mu}, \bar{\mu}]$. To see this consider a specific pooling investment function $\hat{x}$ defined when $\mu \in [\underline{\mu}, \bar{\mu}]$ with the property that the value of type L principal is 0 in the pooling phase, i.e. the individual rationality constraint of type L principal binds at all beliefs in the pooling phase. Lemma 37 and lemma 38 together show that $\hat{x}(\mu) > \frac{\lambda\gamma}{a}$ for all $\mu \in (\underline{\mu}, \bar{\mu}]$. Note that in our candidate equilibrium, the investment during pooling phase is $\frac{\lambda\gamma}{a}$, i.e. strictly lower than $\hat{x}$. This implies that the value of type L principal is strictly positive in the pooling phase when $\mu \in (\underline{\mu}, \bar{\mu}]$. Note that any upward deviation by type L principal in the pooling phase does not alter the agent's behavior but decreases the payoff of the type L principal. Hence there is no incentive to deviate and invest higher. However, any downward deviation leads to the commencement of autarky which yields 0 to type L principal. Hence there is no incentive to deviate downward in the pooling region. This establishes the optimality of type L principal's behavior in the pooling region.

Now we turn to the separating region. First note that when $\mu \geq max\{\bar{\mu}, \mu^c\}$, the type L principal invests her optimal investment $\frac{\mu\lambda\gamma}{a}$ and induces effort from the agent and hence has no incentive to deviate. When $\mu \in (\bar{\mu}, max\{\bar{\mu}, \mu^c\})$, the type L's investment $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$ is strictly smaller than $\frac{\lambda\gamma}{a}$. Since the agent exerts effort following this investment, the type L principal has no incentive to deviate and invest higher. If type L principal invests lower, then the continuation equilibrium is autarky which yields 0 to type L principal and hence there is no inventive to deviate and invest lower than the specified investment. This completes the argument to show that type L principal has no incentive to deviate.

Lastly, I show that the agent's behavior is a best response. We start with the firing region. In the firing region, it is optimal for the agent to exert effort if he knows if she is facing a type L principal only if the investment of type L principal is atleast $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$. This is because the agent's flow value is equal to 0 when the type L principal invests $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$. When the agent knows he is facing type H principal, he is willing to exert effort regardless of the investment since his value is strictly positive even if the type H principal does not invest.

In the pooling region, by Definition 23 we know that the agent's value at the beginning of pooling (when type L principal's belief is $\bar{\mu}$), the agent's value is 0. Note that the agent's value in the pooling region is strictly positive, except at $\bar{\mu}$. This is because the agent's value at $\bar{\mu}$ in the the pooling region is given by

$$\left[1 - \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\bar{\mu}}{\bar{\mu}}\right]^{\frac{r}{\lambda}}\right]\left[\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a}) - c\right] + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\bar{\mu}}{\bar{\mu}}\right]^{\frac{r}{\lambda}}\frac{\bar{\mu}-\underline{\mu}}{1-\underline{\mu}}z^a = 0.$$

This implies that

$$\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a}) - c < 0.$$

Note that the agent's value at $\mu < \bar{\mu}$ in the the pooling region is given by

$$\left[1 - \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\right]\left[\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a}) - c\right] + \left[\frac{\underline{\mu}}{1-\underline{\mu}}\frac{1-\mu}{\mu}\right]^{\frac{r}{\lambda}}\frac{\bar{\mu}-\underline{\mu}}{1-\underline{\mu}}z^a.$$

The agent value at any belief μ in the pooling region is a linear combination of $\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c$ and $\dfrac{\bar{\mu}-\mu}{1-\underline{\mu}}z^a$. Note that the weight on $\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c$ is decreasing in μ. Since $\bar{\mu}\lambda(1-\gamma)(1+\frac{\lambda\gamma}{a})-c<0$, the agent's value must be strictly decreasing in μ in the pooling region. Since value at $\bar{\mu}$ is 0, the value of the agent must be strictly positive when $\mu \in [\underline{\mu}, \bar{\mu})$. This implies that the agent has no incentive to deviate and not exert effort in the pooling region since not exerting effort leads to autarky which we know exists when $\mu \leq \bar{\mu}$.

In the separating region, when $\mu \in (\bar{\mu}, \max\{\mu^c, \bar{\mu}\})$, the agent's value is 0. To see this note that the type L principal's investment in this region is $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$ which implies that the flow value of the agent is 0 in this region. Note that eventually beliefs reach $\bar{\mu}$ at which point pooling begins. We know from the previous paragraph that the agent's value is equal to 0 at $\bar{\mu}$, hence the value of the agent is equal to zero when $\mu \in (\bar{\mu}, \max\{\mu^c, \bar{\mu}\})$. Hence the agent cannot do any better by deviating since any deviation leads to autarky and a value of 0 for the agent.

Lastly, in the region where $\mu > \max\{\mu^c, \bar{\mu}\}$, the agent's flow is strictly positive since the type L principal invests $\frac{\mu\lambda\gamma}{a}$ which is strictly above $\frac{c}{\mu\lambda(1-\gamma)} - \frac{z^a}{r(1-\gamma)} - 1$ in this region. Hence the agent gets a strictly positive value in this region. Note that by deviating the agent only delays the value that he will get and hence it is suboptimal for the agent to deviate. This completes the proof that the specified strategy profile is indeed an equilibrium.

Now I show that σ^* is indeed the principal optimal equilibrium. We know from Lemma 6 that in the pooling phase the pooling investment level must be atleast $\frac{\lambda\gamma}{a}$. We also know from Lemma 45 that the continuation equilibrium at $\bar{\mu}$ must specify pooling when $\mu \in [\underline{\mu}, \bar{\mu}]$. Note that in the pooling phase under σ^*, the investment is the least admissible for pooling. Hence no other pooling continuation equilibrium at $\bar{\mu}$ can improve the principal's payoff. Next, given that pooling begins at $\bar{\mu}$, I show that the behavior in the separating phase is optimal. I will show this through two cases.

Case 1: $\mu^c \leq \bar{\mu}$. In this case both types of principal invest their optimal investment in the separating phase. Hence no other separating behavior in this phase can improve the ex ante payoff of the principal.

Case 1: $\mu^c > \bar{\mu}$. In this case the the type L principal invests optimally when $\mu \geq \mu^c$ and above her optimal investment when $\mu \in (\bar{\mu}, \mu^c)$. Note that in any principal optimal equilibrium it must be the case that type L principal invests optimally when $\mu \in (\bar{\mu}, \mu^c)$, otherwise the principal's payoff can be improved. Note that when $\mu \in (\bar{\mu}, \mu^c)$, type L principal's investment is above her optimal and the agent's value is equal to 0, which implies that type L principal cannot reduce her investment and improve her payoff since that will result to the violation of the agent's IR constraint and the agent will stop exerting effort.

The only question remaining to be answered is that can the cutoff at which pooling begins $\bar{\mu}$ be increased while improving the ex ante payoff of the principal? I show that it is not the case. Suppose the pooling phase begins at $\mu_1 > \bar{\mu}$. From Lemma 39 we know that $B(\mu_1) < B(\bar{\mu}) = \frac{\lambda\gamma}{a}$. We know that in the pooling phase the optimal investment must be $\frac{\lambda\gamma}{a}$, this implies that the value of the agent at the beginning of pooling $(\bar{\mu})$ must be strictly positive. Consider a belief $\mu_1 - \epsilon > \bar{\mu}$ for some $\epsilon > 0$. Suppose the pooling phase begins at $\mu_1 - \varepsilon$ with the same pooling investment level. Note that now the type L can invest strictly lower than $\frac{\lambda\gamma}{a}$ and induce effort from the agent. Note that ex ante, the principal is better off in this case because the type L principal invests strictly lower in the region $(\mu_1 - \epsilon, \mu_1)$ and still induces effort from the agent. This establishes that the optimal belief at which pooling phase begins must be $\bar{\mu}$. Therefore σ^* is the principal optimal equilibrium. $\square$

Bibliography

U. Akcigit and Q. Liu. The role of information in innovation and competition. *Journal of the European Economic Association*, 14(4):828–870, 2016.

D. Bergemann and U. Hege. The financing of innovation: Learning and stopping. *RAND Journal of Economics*, pages 719–752, 2005.

J. Bergin and W. B. MacLeod. Continuous time repeated games. *International Economic Review*, 34(1):21–37, 1993.

S. Bernstein, X. Giroud, and R. R. Townsend. The impact of venture capital monitoring. *The Journal of Finance*, 71(4):1591–1622, 2016.

K. Bimpikis, K. Drakopoulos, S. Ehsani, et al. Disclosing information in strategic experimentation. Technical report, 2018.

D. Blackwell. Equivalent comparison of experiments. *Annals of Mathematics and Statistics*, 32:265–272, 1953.

P. Bolton and C. Harris. Strategic experimentation. *Econometrica*, 1999.

A. Bonatti and J. Hörner. Collaborating. *American Economic Review*, 101(2):632–63, April 2011.

H. Chade and N. Kovrijnykh. Delegated information acquisition with moral hazard. *Journal of Economic Theory*, 162:55 – 92, 2016. ISSN 0022-0531.

Y.-K. Che and K. Mierendorff. Optimal sequential decision with limited attention. *unpublished, Columbia University*, 2016.

E. Damiano, H. Li, and W. Suen. Learning while experimenting. *University of British Columbia working paper*, 2017.

M. Dong. Strategic experimentation with asymmetric information. *Working paper*, 2018.

J. Farrell and E. Maskin. Renegotiation in repeated games. *Games and Economic Behavior*, 1(4):327–360, 1989.

W. Fuchs. Contracting with repeated moral hazard and private evaluations. *The American Economic Review*, 97(4):1432–1448, 2007.

D. Fudenberg, P. Strack, and T. Strzalecki. Stochastic choice and optimal sequential sampling. 2017.

U. Garfagnini. Delegated experimentation. 2011.

Y. Guo. Dynamic delegation of experimentation. *American Economic Review*, 106(8): 1969–2008, 2016.

Y. Guo and A.-K. Roesler. Private learning and exit decisions in collaboration. *Northwestern WP*, 2016.

M. Halac. Relational contracts and the value of relationships. *The American Economic Review*, 102(2):750–779, 2012.

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

B. E. Hermalin. Toward an economic theory of leadership: Leading by example. *The American Economic Review*, 88(5):1188–1206, 1998.

S. Hidir. Contracting for experimentation and the value of bad news. *o Unpublished*, 2017.

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

M. Keil and M. Mähring. Is your project turning into a black hole? *California Management Review*, 53(1):6–31, 2010.

G. Keller and S. Rady. Breakdowns. *Theoretical Economics*, 10(1):175–202, 2015.

G. Keller, S. Rady, and M. Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

M. Komai, M. Stegeman, and B. E. Hermalin. Leadership and information. *American Economic Review*, 97(3):944–947, June 2007. doi: 10.1257/aer.97.3.944. URL http://www.aeaweb.org/articles?id=10.1257/aer.97.3.944.

A. V. Kuvalekar and E. Lipnowski. Job insecurity. *SSRN working paper*, 2018.

J. Lerner and J. Tåg. Institutions and venture capital. *Industrial and Corporate Change*, 22(1):153–182, 2013.

S. D. Levitt and C. M. Snyder. Is no news bad news? information transmission and the role of "early warning" in the principal-agent model. *The RAND Journal of Economics*, 28(4):641–661, 1997.

A. Liang and X. Mu. Overabundant information and learning traps. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 71–72. ACM, 2018.

A. Liang, X. Mu, and V. Syrgkanis. Optimal learning from multiple information sources. *arXiv preprint arXiv:1703.06367*, 2017.

W. B. MacLeod. Optimal contracting with subjective evaluation. *The American Economic Review*, 93(1):216–240, 2003.

G. Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.

T. Mayskaya. Dynamic choice of information sources. 2017.

L. K. Simon and M. B. Stinchcombe. Extensive form games in continuous time: Pure strategies. *Econometrica: Journal of the Econometric Society*, pages 1171–1214, 1989.

H. J. Smith and M. Keil. The reluctance to report bad news on troubled software projects: a theoretical model. *Information Systems Journal*, 13(1):69–95, 2003.

V. Stefflre. Organizational obstacles to innovation: A formulation of the problem. *Journal of Product Innovation Management*, 2(1):3–11, 1985.

B. C. Tan, H. J. Smith, M. Keil, and R. Montealegre. Reporting bad news about software projects: Impact of organizational climate and information asymmetry in an individualistic and a collectivistic culture. *IEEE Transactions on Engineering Management*, 50(1): 64–77, 2003.