WHEN AND HOW TO REWARD BAD NEWS

ADITYA KUVALEKAR† AND NISHANT RAVI ‡

(Job Market Paper)

Click here for the most recent version.

ABSTRACT

We examine 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 observes the agent's allocation and news and makes a firing decision. We show that the principal's optimal Markov perfect 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.

Keywords: dynamic agency, experimentation

JEL codes: C73, D83, D86, M51.

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

†DEPARTMENT OF ECONOMICS, UNIVERSIDAD CARLOS III DE MADRID, AKUVALEK@ECO.UC3M.ES
‡ UNIVERSITY OF PENNSYLVANIA, NISHANTR@SAS.UPENN.EDU
NOVEMBER 10, 2018
Nishant Ravi is deeply indebted to his advisers George Mailath, Steve Matthews, and Mallesh Pai for constant guidance and encouragement. We would like to thank Aislinn Bohren, Joyee Deb, Rahul Deb, William Fuchs, Marina Halac, Rohit Lamba, Annie Liang, Elliot Lipnowski, Andrew Postlewaite, Maher Said, Johannes Schneider, Rakesh Vohra, Yuichi Yamamoto and audience members at University of Carlos III de Madrid and University of Pennsylvania,. Aditya Kuvalekar gratefully acknowledges support from the Ministerio Economia y Competitividad, Maria de Maeztu grant (MDM 2014-0431), and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444).

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 can help 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.

[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.

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 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 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 et al. (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 we present the model and then present results in Section 3. Lastly we discuss extensions in Section 5. All proofs are relegated to the appendix.

Related Literature: 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 et al. (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 et al. (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 ranked in the sense of Blackwell (1953), the dynamics are richer in our environment. Halac et al. (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 et al. (2017), Liang and Mu (2018), Fudenberg et al. (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.

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 Γ. If B arrives, the game ends with the principal receiving 0. 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

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

---

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

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.

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. 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]

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

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) whenever $a(\cdot)$ is continuous.

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

(2) Define

$$a^f = \frac{\lambda_b}{\lambda_b + \lambda_g}. \tag{2}$$

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) we can show that

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

[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.

[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).

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

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 $\{P_t\}_t$ follow (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

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

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, $\mathcal{E}(R,F) := \{(a,s) : (a,s)$ *satisfy* 1 *and* 2*a respectively 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. $\dfrac{\lambda_b \lambda_g}{\lambda_b + \lambda_g}(\Gamma - 1) - c > 0.$

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.

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.

$$p^f := \frac{\lambda_b(F-1)}{\lambda_b(F-1)+\lambda_g(R-1)}. \tag{3}$$

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

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

(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.

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

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

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.

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 8 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 (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 (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 $\underline{p}$.

$$p^s = \frac{\lambda_b(c-F)}{\lambda_g(\Gamma - R + c) + \lambda_b(c-F)} = \frac{\lambda_b(c-1)}{\lambda_g(\Gamma - 1 + c) + \lambda_b(c-1)}. \tag{4}$$

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 $F = 1$.

$$a_0^*(p) = \begin{cases} 1 & \text{if } p > \underline{p}, \\ a^f & \text{if } p = \underline{p}, \\ 0 & \text{if } p < \underline{p}. \end{cases} \tag{5}$$

$$a_1^*(p) = \begin{cases} 1 & \text{if } p > \max\{p^s, \underline{p}\}, \\ a^f & \text{if } p = \max\{p^s, \underline{p}\}, \\ 0 & \text{if } p < \max\{p^s, \underline{p}\}. \end{cases} \tag{6}$$

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 11 in 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 (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, (23) and (33) 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 $v^{p_0}(\cdot)$.

[9] More details about these beliefs can be found in the appendix, (21) and (32).

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*

$$c > (1 + \lambda_g). \tag{7}$$

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

$$\frac{c + \Lambda}{\Lambda \Gamma} > \frac{\lambda_b (c-1)}{\lambda_g (\Gamma - 1 + c) + \lambda_b (c-1)}\,, \tag{8}$$

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 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 28 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.

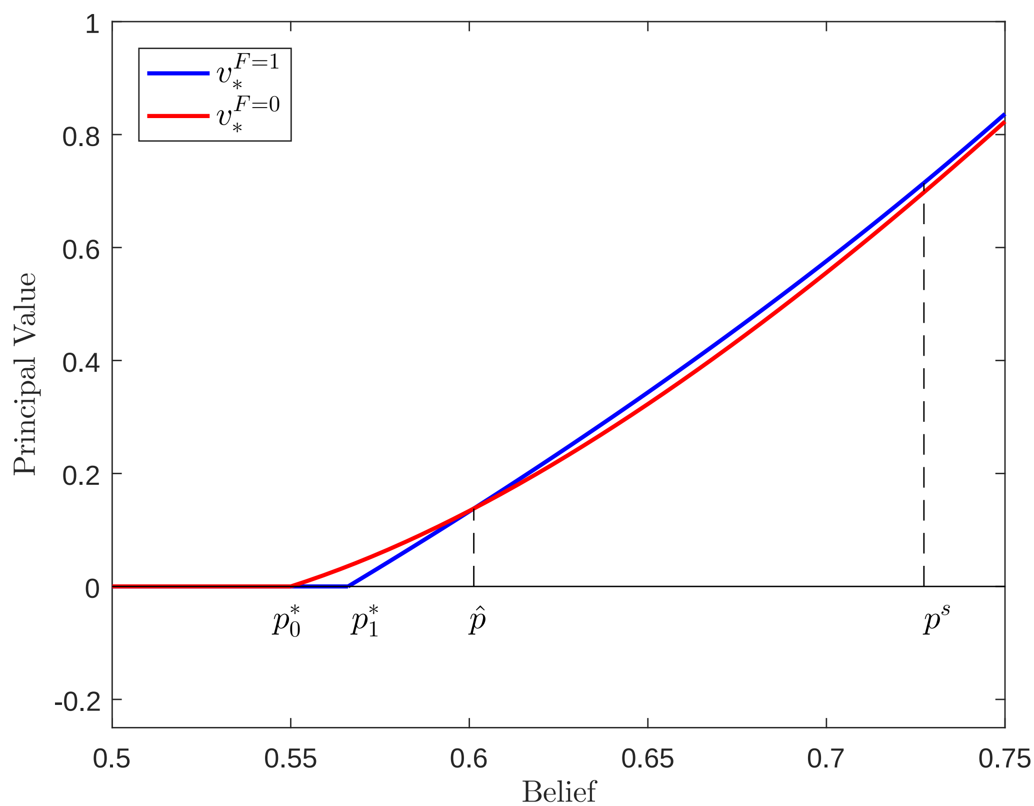

Figure 1. When $v_*^{F=1}$ and $v_*^{F=0}$ cross. Plot is for parameter values: $\lambda_g = 1$, $\lambda_b = 20$, $\Gamma = 11$ and $c = 5$.

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 (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 (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.

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.

[10] For example, Keil and Mähring (2010) and Tan et al. (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.

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

5. EXTENSIONS

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 by assuming that the allocation choice of the agent is not observed by the principal, however any signal (good news or bad news) is 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.

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

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

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). 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

$$\hat{p}_0 = \frac{c}{\lambda_g(\Gamma - 1)}, \tag{9}$$

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 $T_g := \int_0^T a_t \mathrm{d}t$ 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 30 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.*

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 $R = F = 1$.

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.

A. APPENDIX

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 3. *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

$$\begin{aligned} u(p) =& 1-\lambda_b(1-p)u(p)+\lambda_b p(1-p)u'(p)+\lambda_b(1-p)F \\ &+\max_{a\in[0,1]} a\left\{\lambda_b(1-p)\left[u(p)-pu'(p)\right]+\lambda_g p\left[R-u(p)-(1-p)u'(p)\right]-\lambda_b(1-p)F\right\}. \end{aligned} \tag{10}$$

Define:

$$\Delta(p) :=\lambda_b(1-p)\left[u-pu'\right]+\lambda_g p\left[R-u-(1-p)u'\right]-\lambda_b(1-p)F. \tag{11}$$

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,

$$u_1(p)=1+\frac{\lambda_g p(R-1)}{1+\lambda_g}+c_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{1}{\lambda_g}}. \tag{12}$$

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,

$$u_0(p)=\frac{1+\lambda_b p+F\lambda_b(1-p)}{1+\lambda_b}+c_0 p\left(\frac{p}{1-p}\right)^{\frac{1}{\lambda_b}}. \tag{13}$$

A.3. PRINCIPAL'S HJB EQUATION AND ITS SOLUTIONS

The principal's HJB equation is given by

$$\begin{aligned} v(p) = &- c - \lambda_b(1-p)F + \lambda_b p(1-p)v'(p) - \lambda_b(1-p)v(p) \\ &+ \max_{a\in[0,1]} a\left[\lambda_g p(\Gamma - R) + \lambda_b(1-p)F - (\lambda_b+\lambda_g)p(1-p)v'(p) - (\lambda_g p + \lambda_b(1-p))v(p)\right]. \end{aligned} \tag{14}$$

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

$$v_1(p) = \frac{p\lambda_g(\Gamma - R + c)}{1+\lambda_g} - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}, \tag{15}$$

where C_1 is the constant of integration.

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

$$v_0(p) = -\frac{p\lambda_b c + c + \lambda_b F(1-p)}{\lambda_b + 1} + C_0 p\left[\frac{p}{1-p}\right]^{\frac{1}{\lambda_b}}, \tag{16}$$

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*(13) *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

$$a^f = \frac{\lambda_b}{\lambda_b + \lambda_g}, \tag{17}$$

$$v^f(p) = \frac{-c + \Lambda p(\Gamma - R) - \Lambda(1-p)F}{1+\Lambda}, \text{ and} \tag{18}$$

$$u^f(p) = \frac{\lambda_g + \lambda_b + \lambda_b\lambda_g[pR + (1-p)F]}{\lambda_g + \lambda_b + \lambda_b\lambda_g}, \tag{19}$$

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

$$p_m^R = \frac{c}{\lambda_g(\Gamma - R)}, \tag{20}$$

$$p_0^* = \frac{c}{\Lambda(\Gamma - R)}, \text{ and} \tag{21}$$

$$p_f^* = \frac{c + \Lambda F}{\Lambda[\Gamma - R + F]}. \tag{22}$$

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

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

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 4, for any value larger than the freezing value given by (19) 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 6, for any value larger than the freezing value given by (19) 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 5. *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 4. *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 $\bar{p}_2 \geq \bar{p}_1$.[14] 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 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 5 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

[14] 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) $[\underline{p},\bar{p}_1)$.

$\Delta(\underline{p}) \geq 0$. By our boundary condition,

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

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

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 $\bar{p} > \underline{p}$. ■

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 (12).

Lemma 5. *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

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

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,

$$\begin{aligned}\Delta_1'(p_1) &= -\lambda_b(1-F) + \left[\lambda_b(1-F) + \frac{\lambda_g p_1(R-1)}{(1+\lambda_g)(1-p_1)}\right]\left[\frac{1+\lambda_g p_1}{\lambda_g p_1}\right] + \frac{\lambda_g(R-1)}{1+\lambda_g}\\ &= \frac{\lambda_b(1-F)}{\lambda_g p_1} + \frac{R-1}{1-p_1} > 0 \text{ since } p_1 \in \bar{P}.\end{aligned}$$

This shows that $\Delta_1(p)$ is strictly increasing at p_1. By continuity of $\Delta_1'(p)$, we have that $\Delta_1(p)$ 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 6. *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 6. *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 $\bar{p}_1 \leq \bar{p}_2$.[15] 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,

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

[15]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) $(\underline{p}_2,\bar{p}]$.

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

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

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,

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

Lemma 7 establishes that if $\Delta(p) \leq 0$ 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 $\Delta(\bar{p}) \leq 0$. By our boundary condition,

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

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

Therefore, for any terminal $u^* \geq u^f(\bar{p})$, $\Delta \leq 0$, 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 $\underline{p} < \bar{p}$. ∎

Define:

$$\Delta_0(p) := \lambda_b(1-p)\left[u_0 - pu_0'\right] + \lambda_g p\left[R - u_0 - (1-p)u_0'\right] - \lambda_b(1-p)F,$$

where u_0 is defined in (13).

Lemma 7. *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 $\Delta_0'(p_0)$ by plugging in the above expression

$$\begin{aligned}\Delta_0'(p_0)&=-\frac{\lambda_b(1-F)}{1+\lambda_b}+\lambda_g(R-1)-\left[\lambda_g(R-1)+\frac{\lambda_b(1-p_0)(1-F)}{p_0(1+\lambda_b)}\right]\left[\frac{1+\lambda_b(1-p_0)}{\lambda_b(1-p_0)}\right]\\&=-\frac{\lambda_g}{\lambda_b}\left[\frac{\lambda_b(1-F)}{\lambda_g p}+\frac{R-1}{1-p}\right]>0 \text{ since } p_0\in\underline{P}.\end{aligned}$$

This shows that $\Delta_0(p)$ is strictly increasing at p_0 which then by continuity of $\Delta_0'(p)$ implies that $\Delta_0(p)$ 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.[16] 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 4) that the switching belief, p^f, was determined by the objects below.

$$\bar{P} := \left\{ p : \lambda_g p(R-1) > \lambda_b(1-p)(F-1) \right\}.$$
$$\underline{P} := \left\{ p : \lambda_g p(R-1) < \lambda_b(1-p)(F-1) \right\}.$$
$$p^f := \begin{cases} \inf \bar{P} = \sup \underline{P} & \text{if } \bar{P} \neq \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 1 & \text{if } \bar{P} = \emptyset \text{ and } \underline{P} \neq \emptyset, \\ 0 & \text{if } \underline{P} = \emptyset \text{ and } \bar{P} \neq \emptyset. \end{cases}$$

Notice that 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 8. *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,*

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

[16] 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.

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 9, whose proof we omit.

LEMMA 9. *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 8 and lemma 9 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 8. *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 7. *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 10. *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 $\underline{p}_2$.[17]

Case 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.

Case 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 $\hat{a} > a^f \Rightarrow p[\lambda_g a(p) - \lambda_b(1 - a(p))] > \varepsilon$ for some $\varepsilon > 0$ 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 $-\varepsilon$. 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 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.

---

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

Case 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 $\varepsilon > 0$, $\exists\delta > 0$ such that, $a^f > a(p) > a^f - \varepsilon$ 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 $t \le \tau$. Therefore, there is a unique time t^* where the beliefs will reach $\underline{p}_2$ conditional on no signal. Moreover, $\tau \le t^*$ almost surely. Therefore,

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

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

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

Since $|a_u - a^f| < \varepsilon$ for all $u \le \tau$, and since all the integrals are bounded, it is easy to see that all the quantities above are at most $K\varepsilon$ away, for some positive constant K, (ignoring the higher order terms) from using $a_u = a^f$ for all u. Therefore, $|V(p) - V^f(p)| < M\varepsilon$ 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 (21)) is the unique candidate for the lower cutoff in equilibrium.

LEMMA 11. *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, $\underline{p} \ge p_0^*$. 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 $(\underline{p} - \varepsilon, \underline{p})$ for some $\varepsilon > 0$. Suppose $a(p) < a^f$ when $p \in (\underline{p} - \varepsilon, \underline{p})$. By an argument analogous to Case 2 and 3 in Lemma 10, 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 $p \in (\underline{p} - \varepsilon, \underline{p})$. Choose a $p \in (\underline{p} - \varepsilon, \underline{p})$ 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,

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

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,

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

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

LEMMA 12. *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:*

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

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 10. Agent's best response is given by Proposition 8. 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 13 and that $v(p)$ is increasing. ■

PROPOSITION 9. *The principal-optimal equilibrium, which features the same on path behavior, is the following:*

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

PROOF. By Lemma 11, 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 8 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 13 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 13. *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

$$v_1(p) = \lambda_g p(\Gamma - 1) - c - \lambda_g p v_1(p) - \lambda_g p(1-p)v_1'(p).$$

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 $v_1(p_0^*) = 0$.[18] Next, note that the right hand derivative $v'_+(p_0^*)$ is given by

$$v'_+(p_0^*) = v_1'(p_0^*) = \frac{\lambda_g p_0^*(\Gamma - 1) - c}{\lambda_g p_0^*(1-p_0^*)}.$$

Note that

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

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

which implies that $v'_+(p_0^*) > 0$. The principal's value function when $p \in [p_0^*,\bar{p}]$ is given by (15) is as follows

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

We differentiate once to get

$$v_1'(p) = \frac{\lambda_g(\Gamma - 1 + c)}{1+\lambda_g} - C_1^*\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}\left[\frac{\frac{1}{\lambda_g}+p}{p}\right],$$

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 $\Gamma - 1 + c > 0$. 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,

$$v_*^{F=0}(p) = \lambda_g \gamma p - c + C_1(1-p)\left[\frac{1-p}{p}\right]^{\frac{1}{\lambda_g}}, \tag{23}$$

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, $\tau := \inf\{t : P_t \notin [\hat{p},1)\}$, $v(0) = -F = -1, v(1) = \Gamma - R = \Gamma - 1$ 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 $[\hat{p}_*, p^s]$[19] 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 $R = F = 1$),

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

where $v_0(p), v_f(p)$ and $v_1(p)$ are obtained from (16), (18) and (12) respectively.

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,

$$p^s = \frac{\lambda_b(c - F)}{\lambda_g(\Gamma - R + c) + \lambda_b(c - F)}, \tag{24}$$

$$C_0^s = \left[\frac{\lambda_g(\Gamma - R + c)}{\lambda_b(c - F)}\right]^{\frac{1}{\lambda_b}} \left[\frac{\lambda_b}{1+\lambda_b}\right] \left[\frac{\Lambda}{1+\Lambda}\right] (\Gamma - R + c), \tag{25}$$

$$C_1^s = \left[\frac{\lambda_b(c - F)}{\lambda_g(\Gamma - R + c)}\right]^{\frac{1}{\lambda_g}} \left[\frac{\lambda_g}{1+\lambda_g}\right] \left[\frac{\Lambda}{1+\Lambda}\right] (c - F). \tag{26}$$

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 $\hat{p} \leq p^s$[20] and the conjectured optimal value function of the principal are,

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

[20] 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 $F = 0$.

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

and,

$$v_*^{\hat{p}}(p) = \begin{cases} v_0(p; C_0^s) & \text{if } p \in [\hat{p}, p^s), \\ v_f(p^s) = v_0(p^s; C_0^s) = v_1(p^s; C_1^s) & \text{if } p = p^s, \\ v_1(p; C_1^s) & \text{if } p > p^s. \end{cases} \tag{27}$$

We now prove the optimality of the above policy in steps. In Lemma 14 we prove that for any piecewise continuous control, the principal value function is differentiable. Lemma 15, 16 and 17 combine to reduce the candidate policies $a' \succ a^*$ to those where we can have at most one switch from the bad news arm to the good news arm. Optimality of a^* is proved in Lemma 19.

DEFINITION 8. *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 9. *Define,*

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

LEMMA 14. *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 $(p_1$,$p_2)$,*

$$v'(p) = \frac{c + [1 + p\lambda_g a(p) + (1-p)\lambda_b(1-a(p))]v(p) - p\lambda_g a(p)(\Gamma - 1) + (1-p)\lambda_b(1-a(p))}{\eta(a(p))p(1-p)}. \tag{28}$$

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.[21] Therefore, v is differentiable on (p_1,p_2) with the derivative given by (28).

■

First, for any continuous a such that $a \neq a^f$ on the interval, we also have $\eta(a,p) > 0$. By Lemma 14, 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,

$$v(p) = -\lambda_b(1-p) - c + \lambda_b p(1-p)v'(p) - \lambda_b(1-p)v(p) + a(p)H(p,v(p),v'(p)), \quad (29)$$

$$\text{where } H(x,y,z) := \lambda_g x(\Gamma-1) + \lambda_b(1-x) - (\lambda_b+\lambda_g)x(1-x)z - (\lambda_g x - \lambda_b(1-x)y). \quad (30)$$

Notice that H is continuous in each of its argument.

LEMMA 15. *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 $\varepsilon > 0$ such that, either, $a'(q) \in \{0,1\}$ for all $q \in [p-\varepsilon,p]$ or for all $q \in [p,p+\varepsilon]$ and $v(\cdot|a') \geq v(\cdot|a)$.*

PROOF. There are two possible cases:

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

Case 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 $q \in B_\varepsilon(p)$ for some $\varepsilon > 0$. Consider an alternative control a' such that $a'(q) = 1$ for $q \in (p,p+\varepsilon]$. 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 $q \in (p,p+\varepsilon)$, the value function $v(\cdot|a')$ (denoted by $\tilde{v}$ henceforth) satisfies the following differential equation:

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

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

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 (30). 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 $q \in (p, p+\varepsilon_1)$ for some $\varepsilon_1 > 0$.

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

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

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 $q \in (p-\varepsilon_1,p)$.

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 $q \in (p, p+\varepsilon_1)$ if $H(p,v(p),v'(p)) > (<)0$. ■

LEMMA 16. *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 $\varepsilon > 0$ such that, either, $a'(q) \in \{0,1\}$ for all $q \in [p-\varepsilon,p]$ or for all $q \in [p,p+\varepsilon]$ and $v(\cdot|a') \geq v(\cdot|a)$.

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

PROOF. There are four cases:

Case 1: For some $\varepsilon > 0$, $H(q,v(q),v'(q)) = 0$ for all $q \in [p-\varepsilon,p]$ or $[p,p+\varepsilon]$. Notice that v must satisfy (29) and (30) 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.

Case 2: $\exists \varepsilon > 0$ such that. $\{q : H(q,v(q),v'(q) = 0, q \in B_\varepsilon(p)\} = \{p\}$. By Lemma 15, since H is signed on $(p, p+\varepsilon)$, 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.

Case 3: $\exists \varepsilon > 0$ 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 $q \in (p, p+\varepsilon)$.
(b) $H(q,v(q),v'(q)) \geq 0$ or $H(q,v(q),v'(q)) \leq 0$ for all $q \in (p-\varepsilon, p)$.

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 $q \in B_\varepsilon(p) \backslash p$. At p, we can set $a'(p) = 1$ if $a(p) > a^f$ and 0 otherwise as before.

Case 4: At least one of the following holds:
(a) For any $\varepsilon > 0$, $\exists q_1,q_2 \in (p-\varepsilon,p)$ 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 $\varepsilon > 0$, $\exists q_1,q_2 \in (p-\varepsilon,p)$ 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 $\varepsilon > 0$, $\exists q \in (p,p+\varepsilon)$ such that $H(q,v(q),v'(q)) > 0$. Define,

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

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 15. 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 $a' \succ a'' \succ a$.[22] 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 18, $v(\cdot|a^*) > v^f(\cdot)$ for all p except p^s. Therefore, $\exists$ a $w \in \{e,e'\}$ such that $v(w|a^*) > v(w|a)$.

■

DEFINITION 10. *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.*

[22]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.

LEMMA 17. *If $a \in \mathcal{A}$ such that $a \succ a^*$, then H changes sign at most once.*

PROOF. By Lemma 16, 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 16. That is, for every p such that H is 0, $\exists\varepsilon(p)$ such that H is signed on for $q \in [p, p+\varepsilon(p)]$ and on $[p-\varepsilon(p),p]$. 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 $(1-\varepsilon,1)$ is strictly dominated by using the good news arm on $(1-\varepsilon,1)$ for a sufficiently small $\varepsilon > 0$. 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 18. *$v(p^s|a^*) = v^f(p^s)$ 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 19. *Suppose the principal hires in the interval $[\hat{p},1]$ such that $\hat{p} \leq p^s$, then the policy a^* described in equation 27 is optimal for the principal.*

PROOF. Lemma 17 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

$$\tilde{v}^{\hat{p}}(p) := \begin{cases} v_0(p;\tilde{C}_0) & \text{if } p \in [\hat{p},\tilde{p}), \\ v_f(\tilde{p}) = v_0(\tilde{p};\tilde{C}_0) = v_1(\tilde{p};\hat{C}_1) & \text{if } p = \tilde{p}, \\ v_1(p;\hat{C}_1) & \text{if } p \in (\tilde{p},1], \end{cases} \tag{31}$$

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 16 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 11. *The **optimal stopping cutoff** is given by*

$$p_1^* := \{p : v(p|a^*) = 0\}. \tag{32}$$

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,

$$v_*^{F=1}(p) := \begin{cases} 0 & \text{if } p \leq p_1^*, \\ v_*^{p_1^*}(p) & \text{otherwise,} \end{cases} \tag{33}$$

where $v_*^{p_1^*}(p)$ is defined in Equation 27.

A.8. COMPARISON BETWEEN F=1 AND F=0 CASES

PROOF OF PROPOSITION 3: Lemma 23 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 24 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 20 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 (27). 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 17). 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 19 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 20.

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

PROOF. Proof follows from the definitions of p_f^* and p^s defined in (22) and (24) 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 (21).

Lemma 21. *p_1^* is decreasing in λ_b*[23].

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

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

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.

[23] 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.

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 $p \geq p_1^*(\lambda_b)$.

■

LEMMA 22. $\lim_{\lambda_b\to\infty} p_1^*(\lambda_b) = \dfrac{F(1+\lambda_g)}{F(1+\lambda_g)+\lambda_g(\Gamma-R)-c}$ *and* $\lim_{\lambda_b\to\infty} p_0^*(\lambda_b) = \dfrac{c}{\lambda_g(\Gamma-R)}$.

PROOF. Taking limit of 25 as $\lambda_b \to \infty$ gives us

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

This gives us

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

which yields

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

Also note that

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

■

LEMMA 23. *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

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

Next, note that $p_1^*(\lambda_b)$ is continuous in λ_b. Lemma 21 establishes that $p_1^*(\lambda_b)$ is decreasing in λ_b. Also lemma 22 establishes that $\lim_{\lambda_b \to \infty} p_1^*(\lambda_b) = \dfrac{F(1+\lambda_g)}{F(1+\lambda_g)+\lambda_g(\Gamma-R)-c}$ and $\lim_{\lambda_b \to \infty} p_0^*(\lambda_b) = \dfrac{c}{\lambda_g(\Gamma-R)}$. Therefore, since

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

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 $\lambda_b > \underline{\lambda}_b$. ■

LEMMA 24. *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 (27)

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

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 26, 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 27 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. ■

LEMMA 25. *$v_0(p;C_0^s) < v_1(p;C_1^s)$ 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

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

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 (24) (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 26. *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 25 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 27. *$v_1'(p;C_1^*) - v_0'(p;C_0^s)$ 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 25 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 26 that $C_1^s > C_1^*$. $v_1'(p^s;C_1^*) - v_0'(p^s;C_0^s)$is given by

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

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

■

LEMMA 28. *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 25 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 27 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 25 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 (15) restated below.

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

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 $\hat{p}_0 = \frac{c}{\lambda_g(\Gamma-1)}$. 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 29 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 $F \in (0,1)$. ■

LEMMA 29. *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 $\bar{a}$.

■

LEMMA 30. *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 17 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 19 we can show that if there is a switch from bad news arm it must be at p^s defined in (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 24. 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 $\tilde{p}$. ■

REFERENCES

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, 719–752.

Bergin, J. and W. B. MacLeod (1993). Continuous time repeated games. *International Economic Review 34*(1), 21–37.

Bernstein, S., X. Giroud, and R. R. Townsend (2016). The impact of venture capital monitoring. *The Journal of Finance 71*(4), 1591–1622.

Blackwell, D. (1953). Equivalent comparison of experiments. *Annals of Mathematics and Statistics 32*, 265–272.

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

Che, Y.-K. and K. Mierendorff (2016). Optimal sequential decision with limited attention. *unpublished, Columbia University*.

Damiano, E., H. Li, and W. Suen (2017). Learning while experimenting. *University of British Columbia working paper*.

Farrell, J. and E. Maskin (1989). Renegotiation in repeated games. *Games and Economic Behavior 1*(4), 327–360.

Fudenberg, D., P. Strack, and T. Strzalecki (2017). Stochastic choice and optimal sequential sampling.

Garfagnini, U. (2011). Delegated experimentation.

Guo, Y. (2016). Dynamic delegation of experimentation. *American Economic Review 106*(8), 1969–2008.

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

Hidir, S. (2017). Contracting for experimentation and the value of bad news. *o Unpublished*.

Hörner, J. and L. Samuelson (2013). Incentives for experimenting agents. *The RAND Journal of Economics 44*(4), 632–663.

Keil, M. and M. Mähring (2010). Is your project turning into a black hole? *California Management Review 53*(1), 6–31.

Keller, G. and S. Rady (2015). Breakdowns. *Theoretical Economics 10*(1), 175–202.

Keller, G., S. Rady, and M. Cripps (2005). Strategic experimentation with exponential bandits. *Econometrica 73*(1), 39–68.

Kuvalekar, A. V. and E. Lipnowski (2018). Job insecurity. *SSRN working paper*.

Lerner, J. and J. Tåg (2013). Institutions and venture capital. *Industrial and Corporate Change 22*(1), 153–182.

Levitt, S. D. and C. M. Snyder (1997). 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.

Liang, A. and X. Mu (2018). Overabundant information and learning traps. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pp. 71–72. ACM.

Liang, A., X. Mu, and V. Syrgkanis (2017). Optimal learning from multiple information sources. *arXiv preprint arXiv:1703.06367*.

Manso, G. (2011). Motivating innovation. *The Journal of Finance 66*(5), 1823–1860.

Mayskaya, T. (2017). Dynamic choice of information sources.

Simon, L. K. and M. B. Stinchcombe (1989). Extensive form games in continuous time: Pure strategies. *Econometrica: Journal of the Econometric Society*, 1171–1214.

Smith, H. J. and M. Keil (2003). The reluctance to report bad news on troubled software projects: a theoretical model. *Information Systems Journal 13*(1), 69–95.

Stefflre, V. (1985). Organizational obstacles to innovation: A formulation of the problem. *Journal of Product Innovation Management 2*(1), 3–11.

Tan, B. C., H. J. Smith, M. Keil, and R. Montealegre (2003). 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.