

# Waiting For Fake News\*

Raphael Boleslavsky  
Business Economics and Public Policy  
Indiana University  
rabole@iu.edu

July 20, 2023

## Abstract

This paper studies a dynamic model of information acquisition, in which information might be secretly manipulated. A principal must choose between a safe action with known payoff and a risky action with uncertain payoff, favoring the safe action under the prior belief. She may delay her decision to acquire additional news that reveals the risky action's payoff, without knowing exactly when such news will arrive. An uninformed agent with a misaligned preference may have the capability to generate a false arrival of news, which is indistinguishable from a real one, distorting the information content of news and the principal's search. The analysis characterizes the positive and normative distortions in the search for news arising from such manipulation, and it considers three remedies that increase the principal's payoff: a commitment to naive search, transfer of authority to the agent, and delegation to an intermediary who is biased in the agent's favor.

\*The author thanks Deepal Basak, Hector Chade, Navin Kartik, Ayça Kaya, Aaron Kolb, Marco Ottaviani, Marilyn Pease, Manuel Santos, Mehdi Shadmehr, Curtis Taylor, Galina Vereshchagina, Dong Wei, Mark Whitmeyer, Kun Zhang, and seminar audiences at Arizona State, and University of Miami for helpful comments.

# 1. Introduction

In the U.S. legal system, the decision to file charges against a suspect is usually made by the local prosecutor's office, based on an investigation conducted by the police department.<sup>1</sup> Imagine that an extensive police investigation has uncovered circumstantial evidence that implicates a suspect, but it has produced no direct evidence of the suspect's guilt or innocence. The prosecutor agrees with the police officer's interpretation of the circumstantial evidence, and both the prosecutor and the officer would like to see justice done. Nevertheless, the prosecutor might be less inclined to pursue a case based on circumstantial evidence than the officer would like: in an overloaded system, the prosecutor must consider the opportunity cost, which is not borne by the officer. Furthermore, because a significant investigation has failed to uncover direct evidence, the case may soon be closed. In this situation, a police officer who believes that the suspect is probably guilty based on the circumstantial evidence might be tempted to fabricate direct evidence against him.<sup>2</sup> When deciding whether to pursue the case, the prosecutor must therefore consider the possibility that seemingly conclusive evidence may actually be fabricated, impacting the entire investigation.

Such issues are not confined to criminal investigations. A startup firm may not fully internalize a venture capitalist's opportunity cost. Thus, the startup may believe its product to be sufficiently viable to justify additional investment, but the VC may not be willing to invest further given the information currently available. If additional information about product viability (e.g., a working prototype) takes too long to produce, the venture capitalist may withdraw her investment. The startup may therefore be tempted to fabricate such evidence, rather than lose its funding.<sup>3</sup> When deciding whether to withdraw funding, the venture capitalist must consider the possibility that evidence of a product's viability might be fake. A similar scenario might unfold when a firm decides whether to launch a new project or expand a division, or when a regulator decides whether to approve a product coming to market.

In this paper I study a dynamic principal-agent model of information acquisition, in which information can be faked strategically. A principal faces a single, irreversible choice between a safe action with a known payoff and a risky action with an uncertain payoff. The principal is free to choose her action at any time, and the game continues until she makes this choice.

---

<sup>1</sup> This is true of the vast majority of cases [Gross, Possley, Roll, and Stephens \(2020\)](#). The main exception is federal cases, in which the decision to bring charges is made by a grand jury.

<sup>2</sup> [Gross, Possley, Roll, and Stephens \(2020\)](#) document various types of police misconduct that distort evidence in criminal cases, including but not limited to planting physical evidence, coercing a false confession, misrepresenting the results of forensic analysis, and procuring false testimony or identifications from witnesses.

<sup>3</sup> An example is the medical device company Theranos, whose "pathbreaking" product was fake.

The payoff of the risky action might be higher or lower than the payoff of the safe action, but under the prior belief, the principal prefers safe. The agent has no private information about the risky payoff, and he obtains the same payoff as the principal if the risky action is chosen. But, if the safe action is chosen, then the agent's payoff is smaller than the principal's. The principal and agent agree on the best action in the absence of uncertainty. However, the agent's safe payoff is small enough that he prefers the risky action under the prior belief, generating a conflict of interest between the parties that can only be resolved with additional information. The agent has a privately known type, *manipulative* or *normal*, which is described in more detail below.

The principal can delay her choice in order to acquire additional information about the risky payoff. Information arrives through a news process, which generates a single arrival of news at some time. A *real news* arrival accurately reports the payoff of the risky action, either supporting the risky action or opposing it. The arrival time of real news is uncertain, and it is drawn from a continuous prior distribution that is independent of the risky payoff. If this arrival time is reached, then real news is instantly produced. *Fake news* is generated by the manipulative agent in order to influence the principal's decision. At any time he likes, the manipulative agent can fabricate a news arrival that supports the risky action. Though it is fake, such an arrival *appears real*. Thus, when news arrives, the principal cannot directly observe whether it is real or fake; she must infer this from the manipulative agent's strategy. The normal agent does not have the ability to generate fake news, and he waits passively for the game to end.

The conflict of interest distorts the search for information. In each player's "first best," he or she has full control over the timing and action and only real news arrives. Each player decides how long to wait for real news before stopping the search and selecting the action that he or she prefers under the prior belief. Because the arrival of real news could reveal that each player's favorite action under the prior is actually wrong, news is valuable for both players. Provided its arrival rate is sufficiently high, each player waits for some time before terminating the search and taking an action. Furthermore, when the preference misalignment is not too large, the agent's first best search duration is longer than the principal's—the paper focuses on this case.<sup>4</sup> In the game, the principal must consider the possibility that a news arrival that conveys favorable information about the risky action is actually fake. She may therefore treat such news with skepticism, selecting the safe action instead of risky. Furthermore, the possibility that future news is manipulated undermines its informativeness and reduces its

---

<sup>4</sup>In the context of the introductory examples, the prosecutor's and VC's opportunity costs are not too high.

value. Consequently, the principal may terminate her search early, selecting the safe action before news arrives. Compared to the agent's first best, these distortions in the principal's search strategy make it more difficult for him to implement the risky action, which he prefers under the prior. Thus, if the principal's skepticism is currently low and the chance of early termination is high, then the manipulative agent may try to implement the risky action by faking favorable news.

When the probability of the manipulative type is not too large, the game has a unique equilibrium with some intriguing features. The equilibrium incorporates two deadlines, an early "soft deadline" and a later "hard deadline." Before the soft deadline, the agent does not fabricate news. The principal follows the advice implied by any news arrival and only selects an action if news arrives. In other words, before the soft deadline, equilibrium search is identical to each player's first best. Once the soft deadline is reached, the search for news becomes distorted. At each time between the soft and hard deadlines, the agent randomly fakes a news arrival, and the principal randomly terminates the search by choosing the safe action. Despite the positive likelihood of fabrication by the manipulative type, news that supports the risky action is sufficiently informative to overturn the principal's prior in favor of safe, convincing her to select the risky action. A manipulative agent fakes a news arrival with probability one before the hard deadline. Thus, reaching the hard deadline with no arrival reveals that the agent is normal and all future arrivals are real. Nevertheless, the principal stops the search at the hard deadline, which coincides with her first best search duration.

Though equilibrium search lasts until the hard deadline with positive probability, from a normative perspective, it is identical to a first best search that is terminated at the soft deadline, too soon for either player's liking. Before the soft deadline, equilibrium search is undistorted, but once the soft deadline is reached, each player randomizes. Each player is therefore indifferent over all times between the soft and hard deadline. From a normative perspective, the principal may as well choose safe at the soft deadline, and the agent may as well fake, inducing risky. Thus, each player's equilibrium payoff is as if he or she waits until the soft deadline and then selects his or her favorite action under the prior if no real news has arrived.

The equilibrium can be interpreted as a "stochastic unraveling" of the first best. Imagine that the principal searches "naively," simply following her strategy from the first best benchmark without considering the possibility of fake news: she acts on all news that arrives before (or at) the hard deadline, picking the safe action if no news comes. In this case, the manipulative agent would not fake before the hard deadline. Before it is reached, the agent is always able to induce the risky action by faking a bit later, and waiting allows more

time for valuable real news to arrive. At the hard deadline, however, waiting is no longer an option: the agent must fake an arrival immediately, in order to preempt the principal's imminent choice of safe, resulting in masses of faking and search termination concentrated at the hard deadline.<sup>5</sup> In the unique equilibrium of the game, these probability masses are spread backwards in time, resulting in an interval with continuous mixing, starting from the soft deadline and ending at the hard one. In this sense, the naive search stochastically unravels into the unique equilibrium. Furthermore, a higher prior probability of a manipulative agent intensifies the agent's preemption motive, producing a longer unraveling and an earlier soft deadline. If the probability of the manipulative type exceeds a threshold, then the initial phase of first best search collapses. In this case a multiplicity of equilibria arises. In all of them, the principal gains nothing from search: her payoff is as if she implements safe immediately.

Motivated by these normative features, I consider three simple forms of commitment that mitigate the agent's preemption motive. First, I allow the principal to commit to implement the "naive search" strategy described above. By guaranteeing that she will select risky if favorable news arrives at the hard deadline, the principal eliminates the agent's preemption motive and the resulting unraveling. The manipulative agent simply waits until the hard deadline to fake without worrying that he "tips his hand" by being too predictable. From the principal's perspective, such a commitment mimics the first best, except for one scenario: when no real news comes before the hard deadline and the agent is manipulative, she selects risky instead of safe. When the probability of a manipulative agent is not too high, the principal benefits from a commitment to naive search. Thus, when the probability of manipulated news is relatively small, the principal might want to pretend that it does not exist. The principal might therefore wish to introduce institutional features that allow her to "plausibly deny" the existence of fake news.

A more straightforward commitment by the principal transfers authority over the search and action to the agent. The principal need not pretend, she can simply step back and allow the agent to do as he sees fit. In this case, there is clearly no preemption motive for the agent. The principal's search is nevertheless distorted, because the agent selects the wrong action if he terminates the search without news, and he does so at the wrong time. The cost of these distortions may be small; if so, the principal would like to transfer all authority to the agent. Thus, if delegation is feasible, then the credible threat of news manipulation allows the agent to command outsized influence within the organization, despite having no private

---

<sup>5</sup> These masses are inconsistent with equilibrium. Anticipating such a faking strategy, the principal would ignore news that arrives at the deadline, incentivizing the agent to fake a bit sooner.

information that is relevant to the choice of action.

Finally, I consider a more nuanced form of delegation in which the principal allocates the decision authority to an intermediary, whose payoff from the safe action is between hers and the agent's. On one hand, the conflict of interest between intermediary and agent is smaller, which shifts the soft deadline toward a later time and extends the duration of the initial "first best" phase. On the other hand, once the soft deadline is reached, the intermediary is indifferent between continuing the search and selecting the safe action with no news. The principal's safe payoff is larger than the intermediary's, and therefore the principal would like to terminate the search at all times after the soft deadline. Thus, delegation is harmful if the search extends for too long and is (weakly) helpful otherwise. When the equilibrium of the main model has beneficial search, such delegation benefits the principal. For example, the prosecutor benefits when the decision to indict a suspect is made by a grand jury that places a bit less weight on the opportunity cost of trial.

**Related Literature.** In my paper, information acquisition is modeled as a stopping problem in the spirit of [Wald \(1947\)](#) and [Arrow, Blackwell, and Girshick \(1949\)](#). While the classic literature focuses on a single searcher,<sup>6</sup> [Brocas and Carrillo \(2007\)](#) and [Henry and Ottaviani \(2019\)](#) study Wald stopping in settings where authority over the final decision and search duration is split between a principal and agent. These authors show that even without the ability to manipulate the realization of the news process, controlling the stopping decision can benefit the agent. In my analysis, the agent has the ability to stop the search: by faking, he shuts off the news process, which forces the principal to choose an action immediately. However, the agent does not benefit from this ability. In particular, if the agent cannot fabricate a news arrival but can take an action that stops the search, in equilibrium he would never do so.

A number of related papers consider stopping problems with *observable* manipulation of the news process ([Aghion and Jackson, 2016](#); [Kuvalekar and Lipnowski, 2020](#); [Orlov, Skrzypacz, and Zryumov, 2020](#)). In this strand, [Che, Kim, and Mierendorff \(2022\)](#) (CKM) is closest to my analysis. These authors consider a dynamic persuasion model in which an uninformed sender designs a flow experiment at each instant but cannot commit to future experiments. Similar to my model, the receiver's expectation of future informativeness affects her stopping strategy, which itself affects sender's incentive to manipulate the flow experiment. Unlike my model, which often has a unique equilibrium, in CKM a range of mutual expectations can be sustained, resulting in multiple equilibria. More broadly, in my model the agent's manipu-

---

<sup>6</sup> Recent contributions by [Che and Mierendorff \(2019\)](#) and [Liang, Mu, and Syrgkanis \(2022\)](#) endogenize a single searcher's information by allowing her to choose which source(s) to sample among a set of available signals. In my analysis the information structure is endogenous, but it arises from a strategic interaction.

lation of the news manipulation is unobservable. This difference is crucial: if manipulation were observable in my model, then the agent would not manipulate in equilibrium.<sup>7</sup>

I am aware of only one other paper that studies a similar stopping problem with unobservable news manipulation.<sup>8</sup> [Ekmekci, Gorno, Maestri, Sun, and Wei \(2022\)](#) consider a drift-diffusion model in which an informed sender would like to persuade the receiver to wait by secretly inflating an informative performance measure, showing that observed performance tends to spike just before receiver stops. This model differs from mine in two key respects: information arrives continuously, and the agent is informed about his type. These modeling differences imply a distinct set of applications and generate different results.<sup>9</sup>

Finally, my paper is related to [Boleslavsky and Taylor \(2023\)](#) (BT), which studies fraud in the context of production. With respect to this paper, there are a number of substantive differences in focus, model, and results. First, in BT the agent’s incentive to cheat comes from impatience—he wants to obtain a reward quickly. In contrast, in my paper the agent fakes news in order to influence the principal’s action.<sup>10</sup> Second, in BT the principal does not have an outside option, and thus, the stopping decision at the heart of my analysis is absent. Third, the equilibrium structure differs significantly. In BT, the agent fakes a project randomly in an early phase, and the principal mixes between approving and rejecting such a project when it arrives. In the late phase, all distortions vanish. In my paper, this structure is reversed: the search for news is distorted in the late phase of the equilibrium. Furthermore, the principal’s response to faking is also different. In BT the principal randomizes between approval and rejection after a project arrives. In my paper, principal randomizes between stopping and waiting before news arrives; if it arrives, she always “approves” it by selecting the recommended action.

---

<sup>7</sup> My analysis also connects to an emerging literature on dynamic information design, whereby a sender commits to a long run communication protocol in order to influence a receiver’s decision(s) over time ([Ely, 2017](#); [Halac, Kartik, and Liu, 2017](#); [Renault, Solan, and Vieille, 2017](#); [Ely and Szydlowski, 2020](#)). The pioneering contribution of [Kamenica and Gentzkow \(2011\)](#), who study information design in a static setting, is also related. Juxtaposed with this literature, the agent’s lack of commitment in my analysis is even more pronounced.

<sup>8</sup> [Dilme \(2019\)](#) studies a dynamic signaling model in which an informed sender’s unobserved effort affects a public news process. Unlike my analysis, the sender decides when to stop; see also [Kolb \(2015, 2019\)](#).

<sup>9</sup> For example, in the context of venture capital, the setting of [Ekmekci, Gorno, Maestri, Sun, and Wei \(2022\)](#) suggests that the firm is relatively established. Its current operations generate a continuous performance measure, and its past has endowed management with private information about its potential for success. In contrast, my setting suggests that the firm is just getting started: a breakthrough (in R&D, perhaps) is still needed for the firm to be viable, and it has not operated for long enough to acquire private information about its ability.

<sup>10</sup> In my model even a perfectly patient agent would manipulate news in equilibrium, provided the players prefer different actions under the prior belief.



## 2. Model

An organization consists of a principal and an agent, who play a game in continuous time. At any time she likes, the principal makes a single irreversible decision, choosing either a safe or risky action. When the principal chooses an action, the game ends and the players' payoffs are realized. The safe action ( $S$ ) delivers known lump-sum payoffs  $\theta \in (0, 1)$  and  $\beta \in (0, \theta)$  to principal and agent respectively, while the risky action ( $R$ ) delivers an uncertain lump-sum payoff  $\omega \in \{0, 1\}$ , identical for both players. Thus, the preference misalignment stems from the safe action, which gives the agent a smaller payoff than the principal.<sup>11</sup> Although the players would select the same action if the payoff of the risky action were known, uncertainty about it creates disagreement. In particular, the players have a common prior belief  $\mu \equiv \Pr(\omega = 1)$ , where  $\beta < \mu < \theta$ . In this case, the principal prefers the safe action under the prior, but the agent prefers the risky action. The agent and principal both discount future payoffs at rate  $\rho$  and have zero flow payoffs. In addition, the agent has a privately known type, either *manipulative* or *normal*, whose significance is described below. At the outset, the principal believes that the agent is manipulative with probability  $\sigma \in (0, 1)$ .

Though she prefers the safe action given the prior belief, the principal may delay her decision in order to acquire additional information about  $\omega$ . Both players observe a news process which generates a single arrival of news—either real or fake—at some time. A *real news* arrival accurately reports the payoff of the risky action. Thus, real news comes in two varieties “type-0” and “type-1,” and a real news arrival is type- $\omega$  if and only if the risky payoff is  $\omega \in \{0, 1\}$ . The arrival time of real news is uncertain, and both players believe it is distributed according to continuous cumulative distribution function  $G(\cdot)$ , with associated density  $g(\cdot)$ , and decreasing hazard rate  $H_R(\cdot)$ , supported on  $\mathbb{R}_+$ .<sup>12</sup> If the arrival time of real news is reached, then real news is instantly produced. Note that the arrival time of real news is statistically independent of the payoff of the risky action (and the agent's type). Thus, the content of real news is informative about the risky payoff, but its arrival time is not.

*Fake news* is produced strategically in order to influence the principal's decision. In particular, the manipulative type of agent has the ability to produce an instant arrival of fake type-1 news, which is indistinguishable from real type-1 news that could have arrived at the same

---

<sup>11</sup> A similar preference structure, featuring a misaligned payoff with respect to a single action, also appears in the cheap talk model of [Che, Dessein, and Kartik \(2013\)](#).

<sup>12</sup> The decreasing hazard rate implies that pessimism about the likelihood of a real news arrival grows over time, strengthening the incentive to stop. The analysis is unchanged if there is a positive probability that no news arrives,  $\lim_{t \rightarrow \infty} G(t) < 1$ . One example is an exponential bandit: if real news exists, then its arrival time is exponentially distributed; otherwise, it never arrives.



moment.<sup>13</sup> The normal type of agent does not have the ability to fabricate news and waits passively for the principal’s decision. In light of this, “the agent” should be understood as the manipulative type.

Three observations streamline the subsequent exposition and analysis. First, recall that the news process delivers only one arrival. Thus, the principal chooses between safe and risky immediately when news arrives. Second, recall that type-0 news is never faked.<sup>14</sup> Therefore, type-0 news is fully revealing, and the principal immediately selects the safe action when it arrives. Third, the non-arrival of news before a given time conveys no information about the payoff of the risky action. In particular, the arrival time of real news is drawn from  $G(\cdot)$ , and this arrival time is statistically independent of  $\omega$ . Furthermore, because the agent has no private information about  $\omega$ , the arrival time of fake news is also statistically independent of  $\omega$ , regardless of the agent’s strategy. The arrival time of news is the smaller of the real and fake arrival times. Because both of these times are statistically independent of the risky payoff, their minimum is also. Thus, players’ beliefs about the risky payoff do not change in the absence of news (see also footnote 19). By implication, at all times before news arrives, the principal prefers the safe action and the agent risky.

**First-Best Benchmarks.** To clarify the broader implications of the conflict of interest between the players (and in anticipation of the analysis to follow), it is helpful to understand how each player would search for real news in isolation. We therefore consider two “first-best” benchmarks, one for the principal and one for the agent. In both benchmarks, the relevant player has full authority over the decision and search, and the news is always real. Thus, in either benchmark, the player selects safe following a type-0 arrival and risky following a type-1. Furthermore, as described above, the passage of time does not affect players belief about the risky payoff. Thus, if acting without news, the principal defaults to safe, and the agent risky.

Therefore, the crucial decision for each player in the first-best benchmark is the search duration—how long to wait for news before selecting the default action. For  $i \in \{A, P\}$ , let  $u_i^{FB}(t)$  be player  $i$ ’s expected payoff of selecting search duration  $t$  in the first-best benchmark.

---

<sup>13</sup> I focus the case where generating a fake is costless. A small faking cost can be incorporated into the analysis with minimal changes to the results. In particular, only the principal’s equilibrium strategy adjusts when a small cost is included.

<sup>14</sup> If the agent had the ability to fake or suppress an arrival of type-0 news, he would not choose to do so in equilibrium. Because the agent is uninformed and a type-0 arrival might be real, type-0 news is always weakly bad news about the risky action and induces safe. Thus, he would never fake a type-0 arrival, since the agent prefers risky without a real type-0 arrival. If a real type-0 arrival does occur, however, then the agent prefers safe, and he would want it to be implemented immediately. Unlike [Shadmehr and Bernhardt \(2015\)](#) and [Sun \(2023\)](#), the agent has no incentive to suppress or censor type-0 news.

In addition, let

$$\phi_P \equiv \frac{\rho\theta}{\mu(1-\theta)} \quad \phi_A \equiv \frac{\rho\mu}{\beta(1-\mu)}.$$

The numerator of  $\phi_i$  represents player  $i$ 's marginal cost of delaying his or her default action by an instant, while the denominator represents the gains from selecting the correct action (instead of the default action) if news arrives. I introduce two assumptions on these parameters that streamline subsequent analysis,

$$\text{A1: } H_R(0) > \phi_P > H_R(\infty) \quad \text{A2: } \phi_P > \phi_A, \text{ or equivalently, } \beta > \underline{\beta} \equiv \mu\left(\frac{\mu}{\theta}\right)\left(\frac{1-\theta}{1-\mu}\right).$$

The significance and interpretation of (A1) and (A2) will be discussed below. For now, note that  $\mu < \theta \Rightarrow \underline{\beta} < \mu$ . Thus, the set of  $\beta$  that is consistent with (A2) is non-empty.

The following result characterizes each player's optimal search duration in his or her first best benchmark.

**Lemma 1.** (*First-Best Benchmarks*). *In player  $i \in \{A, P\}$  first-best benchmark,*

- (i)  $u_i^{FB}(\cdot)$  is a continuous, differentiable, single-peaked function.
- (ii) if (A1), then principal's optimal search duration  $\tau_P$  is strictly positive and finite, and it is the unique value such that  $H_R(\tau_P) = \phi_P$ .
- (iii) if (A2), then agent's optimal search duration  $\tau_A$  is larger than the principal's, and it is the unique value such that  $H_R(\tau_A) = \phi_A$ .

In the first-best benchmark, player  $i \in \{A, P\}$  waits for news until time  $\tau_i$ , at which the hazard rate of real news arrivals reaches  $\phi_i$ .<sup>15</sup> Despite their disagreement over the default action, both players may wish to wait for news, since the arrival of real news could reveal that their default action was actually the wrong choice. However, because the players disagree over the default action, they may also disagree over the duration of first-best search. A1 ensures that the principal selects a positive and finite duration in her first best benchmark. A2 ensures that the agent's first best search duration is longer than the principal's. This assumption softens (but does not eliminate) the disagreement over search duration, streamlining the analysis and exposition.<sup>16</sup> It holds whenever the preference misalignment is not too large.

---

<sup>15</sup> These are essentially the first order conditions for each player. By marginally extending her search, the principal allows a bit of extra time for news to arrive, but imposes costly delay if it does not.

<sup>16</sup> Suppose the agent prefers a longer first best duration and the principal follows her first best strategy. If

### 3. Equilibrium Analysis

#### 3.1. Preliminaries.

**Strategies.** A pure strategy for the manipulative agent is a “faking time,”  $t \in \{\mathbb{R}^+ \cup \infty\}$ , at which the agent will fake type-1 news if the game has not yet ended. A mixed strategy is a probability measure over “faking times,”  $t \in \{\mathbb{R}^+ \cup \infty\}$  represented by cumulative distribution function  $F_A(\cdot)$ .

The principal’s strategy consists of two parts. The first component of the principal’s strategy describes the search duration. The principal selects a stopping time,  $t \in \{\mathbb{R}^+ \cup \infty\}$ . If the principal’s stopping time is reached with no news, then she stops the search and selects the safe action (which is optimal in the absence of news). A mixed stopping strategy specifies a probability measure over stopping times, represented by cumulative distribution function  $F_P(\cdot)$ . The second component of the principal’s strategy specifies her decision when news arrives. Because type-0 news is always real, its arrival reveals that  $\omega = 0$ , and the principal selects safe. If type-1 news arrives at  $t \in \mathbb{R}^+$ , then the principal selects the risky action with probability  $a(t)$ .<sup>17</sup> If  $a(t) = 1$ , then the principal “acts on” or “follows” the news at time  $t$ .

The analysis focuses on the case in which the players’ mixed strategies  $(F_A(\cdot), F_P(\cdot))$  have an absolutely continuous component and a discrete component (with no continuous singular component). In particular, for  $i \in \{A, P\}$ ,

$$F_i(t) = F_i^*(t) + \sum_{m \in M} \mathcal{I}(t \geq t_m) f_m,$$

where  $M$  is a countable set indexing the set of atoms,  $t_m$  is the time of atom  $m$ ,  $f_m$  is the probability assigned to atom  $m$ ,  $\mathcal{I}(\cdot)$  is the standard indicator function, and  $F_i^*(\cdot)$  is absolutely

---

no news comes by the principal’s chosen deadline, the agent would like the principal to extend the search, but he has no action that would allow him to do so. Thus, he fakes at or just before the deadline only to preempt the principal’s imminent choice of safe. In contrast, if the agent prefers a shorter duration and the principal follows her first best strategy, then he would fake when his preferred duration is reached, which also induces his preferred action. In this case, the equilibrium strategy is shaped both by the agent’s incentive to preempt the safe action and to terminate search early. This case is substantially more complicated and is left for future study.

<sup>17</sup> An ambiguity arises if type-1 news arrives at the exact same time that the principal intends to stop the search and select safe—does she select risky according the news response profile, or does she select safe? The analysis assumes that the principal selects safe in this case. In other words, in the small period  $[t, t + dt]$  the principal decides whether to terminate the search before observing whether news arrives. This assumption can be adjusted with minimal impact on the results, though some of the notation becomes more cumbersome.

continuous. In the standard way,

$$F_i(t) = \int_0^t f_i^*(s)ds + \sum_{m \in M} \mathcal{I}(t \geq t_m) f_m,$$

where  $f_i^*(\cdot)$  is the derivative  $F_i^*(\cdot)$  almost everywhere. To economize on notation, it is helpful to represent the atoms using the Dirac  $\delta(\cdot)$ . In particular, let  $f_i(s) \equiv f_i^*(s) + \sum_{m \in M} f_m \delta(s - t_m)$ . Using the notation convention for the Dirac  $\delta(\cdot)$ ,<sup>18</sup>

$$F_i(t) = \int_0^t f_i(s)ds.$$

In the text,  $f_i(t)$  is sometimes loosely referred to as the “probability mass” on  $t$ .

**Payoffs.** From the perspective of each player, the other’s strategy distorts the first best. These distortions are described below and are used to derive each player’s payoff function.

*Principal.* From the principal’s perspective the agent’s faking strategy distorts the news process. In particular, when the payoff of the risky arm is low, the agent may fake type-1 news before real (type-0) news arrives. This has implications for the arrival times of type-0 and type-1 news, as well as the information content of type-1 news.

Suppose that the principal intends to wait for news until time  $t$  and then select the safe action. Her search for news can end in three ways—an arrival of type-0 news before  $t$ , type-1 news before  $t$ , or no news before  $t$ .

Because the agent never fakes type-0 news, it can only arrive from the real news process. Thus, such an arrival occurs at some time  $s < t$  if the payoff of the risky arm is low, real news arrives at  $s$ , and fake news does not arrive before  $s$ . In turn, fake news takes longer than  $s$  unless the agent is manipulative and his chosen faking time is smaller than  $s$ . Combining these observations, the probability mass of a type-0 arrival at  $s$  is

$$w_0^P(s) = (1 - \mu)g(s)(1 - \sigma F_A(s)).$$

In contrast, type-1 news can be faked by the manipulative agent, and it can therefore arrive even when the payoff of the risky action is zero. In particular, an arrival of type-1 news at time  $s < t$  can occur in three mutually exclusive ways: (1)  $\omega = 1$ , real news arrives at  $s$ , fake news

---

<sup>18</sup> In particular,  $\int_{t_0}^{t_1} \delta(s)ds = \mathcal{I}\{t_0 \leq 0 \leq t_1\}$ .

takes longer, (2)  $\omega = 1$ , fake news arrives at  $s$ , real news takes longer, (3)  $\omega = 0$ , fake news arrives at  $s$ , real news takes longer. The probability mass of a type-1 news arrival at  $s$  is

$$\begin{aligned} w_1^P(s) &= \underbrace{\mu g(s)(1 - \sigma F_A(s))}_{(1)} + \underbrace{\mu \sigma f_A(s)(1 - G(s))}_{(2)} + \underbrace{(1 - \mu) \sigma f_A(s)(1 - G(s))}_{(3)} \\ &= \mu g(s)(1 - \sigma F_A(s)) + \sigma f_A(s)(1 - G(s)). \end{aligned}$$

No news arrives at or before time  $t$  if and only if the arrival times of both real news and fake news exceed  $t$ ,<sup>19</sup>

$$W_\phi^P(t) \equiv \Pr(\text{no news at or before } t) = (1 - G(t))(1 - \sigma F_A(t)).$$

The information content of type-1 news is also affected by the agent's faking—it is not necessarily fully revealing. In particular, of the three ways that type-1 news could arrive at time  $s$  (described above), only the first two have  $\omega = 1$ . Thus, the principal's belief that  $\omega = 1$ , given an arrival of type-1 news at time  $s$  is

$$\mu_1(s) \equiv \Pr(\omega = 1 | \text{type-1 news at } s) = \frac{\mu g(s)(1 - \sigma F_A(s)) + \mu \sigma f_A(s)(1 - G(s))}{\mu g(s)(1 - \sigma F_A(s)) + \sigma f_A(s)(1 - G(s))}. \quad (1)$$

Note that the denominator of  $\mu_1(s)$  is always strictly positive, and therefore this belief can be derived using Bayes' rule at all times  $s \geq 0$ .

Building on these observations, if the principal waits for news until time  $t$  and acts on news according to  $a(\cdot)$ , her expected payoff is

$$\begin{aligned} u_P(t) &= \int_0^t \exp(-\rho s) \{w_0^P(s)\theta + w_1^P(s)((1 - a(s))\theta + a(s)\mu_1(s))\} ds \\ &\quad + \exp(-\rho t) W_\phi^P(t)\theta. \end{aligned} \quad (2)$$

*Agent.* Compared to his single-player benchmark, the principal's stopping strategy and action profile restrict the agent's ability to implement the risky action. In his single-player problem, the agent could stop waiting for real news and select risky at any time. However, in the game with the principal, he faces two distortions: the principal may stop the search and select the

---

<sup>19</sup>Note that the joint probability  $\Pr(\text{no news at or before } t | \omega = 1) = \Pr(\text{no news at or before } t | \omega = 0) = W_\phi^P(t)$ . That is, conditional on  $\omega$ , no news arrives by  $t$  if and only if real news takes longer than  $t$  (probability  $1 - G(t)$ ) and fake news takes longer than  $t$  (probability  $1 - \sigma F_A(t)$ ). By implication, the posterior belief that  $\omega = 1$  given no news arrival by  $t$  is the prior,  $\mu$ . Similar logic applies for the agent.

safe action before news arrives, and even if type-1 news arrives at  $t$  the principal selects the risky action with probability  $a(t)$ , which may be less than 1.

Suppose that the agent intends to fake at some time  $t$  if it reached. The search for news can end in three ways. First, there could be an arrival of real news at time  $s < t$ , which occurs before the principal stops her search. The probability mass of such an arrival is

$$\begin{aligned} w_0^A(s) &= (1 - \mu)g(s)(1 - F_P(s)) \\ w_1^A(s) &= \mu g(s)(1 - F_P(s)), \end{aligned}$$

where the subscript denotes the type of news arrival. Second, the principal could stop her search at  $s < t$ , before news arrives. The probability mass of such an arrival is

$$w_S^A(s) = f_P(s)(1 - G(s)).$$

Finally, the agent's intended faking time  $t$  could be reached without a news arrival. In other words, real news takes longer than  $t$ , and the principal's stopping time exceeds  $t$ , which occurs with probability

$$W_\phi^A(t) = (1 - G(t))(1 - F_P(t)).$$

Thus, if the agent waits until time  $t$  to generate a fake, his expected payoff is

$$\begin{aligned} u_A(t) &= \int_0^t \exp(-\rho s) \{w_0^A(s)\beta + w_1^A(s)((1 - a(s))\beta + a(s)) + w_S^A(s)\beta\} ds \\ &\quad + \exp(-\rho t) W_\phi^A(t) ((1 - a(t))\beta + \mu a(t)). \end{aligned} \quad (3)$$

In the previous expression, if the search ends with the arrival of type-1 news, the agent is confident that  $\omega = 1$ , and his payoff is 1 if the principal selects the risky action. In contrast, if the search ends at  $t$  with no news, then the agent's belief that  $\omega = 1$  is still the prior  $\mu$ . Thus, if the risky action is chosen at  $t$  (in response to the fake type-1 arrival), the agent's expected payoff is  $\mu$ .

**Equilibrium.** A profile of strategies  $(a(\cdot), F_P(\cdot), F_A(\cdot))$  is an equilibrium if and only if

- (i) The agent's faking strategy  $F_A(\cdot)$  is optimal given the principal's strategy  $(a(\cdot), F_P(\cdot))$ ,
- (ii) For all times  $t \geq 0$ , the principal's response to type-1 news,  $a(t)$  is optimal given  $\mu_1(t)$ , her posterior belief that  $\omega = 1$  following the arrival of type-1 news at  $t$  (derived in (1)).

(iii) The principal's stopping strategy  $F_P(\cdot)$  is optimal given the agent's strategy  $F_A(\cdot)$  and her response to type-1 news,  $a(\cdot)$ .

Condition (i) holds if and only if

$$f_A(t) > 0 \Rightarrow u_A(t) \geq u_A(t') \quad \text{for all } t, t'. \quad (4)$$

That is, the agent's mixed strategy assigns positive mass (or probability) only to faking times that generate the highest possible payoff.

Condition (ii) holds if and only if

$$a(t) = \begin{cases} 1 & \text{if } \mu_1(t) > \theta \\ [0, 1] & \text{if } \mu_1(t) = \theta \\ 0 & \text{if } \mu_1(t) < \theta, \end{cases} \quad (5)$$

for all  $t \geq 0$ . By selecting safe, the principal obtains guaranteed payoff  $\theta$ , but following type-1 news at time  $t$  her expected payoff of risky is  $\mu_1(t)$ , her posterior belief that  $\omega = 1$  given type-1 news at  $t$ , which is derived from Bayes' rule at all times (see (1)). The principal selects the action with the larger expected payoff, mixing only if the payoffs are equal.

Condition (iii) holds if and only if

$$f_P(t) > 0 \Rightarrow u_P(t) \geq u_P(t') \quad \text{for all } t, t'. \quad (6)$$

In other words, optimal stopping for the principal requires that her strategy assigns positive mass (or probability) only to times that generate the highest possible payoff.<sup>20</sup>

### 3.2. Equilibrium With Beneficial Search

This section characterizes equilibria in which the principal benefits from the ability to search for news despite the agent's manipulation, which are the main focus of the analysis. Equilibria in which the principal does not benefit from the ability to search are studied (briefly) in Section 3.3.

In an equilibrium with beneficial search, the principal's equilibrium payoff exceeds  $\theta$ , her payoff from making a decision immediately based only on the prior belief. To develop intu-

---

<sup>20</sup> Each player's mixed strategy can be optimal even if it assigns positive mass to a set of  $t$  at which  $u_i(t)$  is suboptimal, as long as the probability of a realization in this set is zero. However, if the realization of the player's faking time happened to be in this set, the player would be tempted to redraw it.



ition for the structure of such an equilibrium, first consider a “naive search” by the principal. In particular, suppose that the principal adopts her strategy from the first-best benchmark: wait for news until stopping time  $\tau_P$  and then select safe, while acting on all news that arrives ( $a(\cdot) = 1$ ). As long as the principal’s stopping time  $\tau_P$  has not passed, the agent’s payoff is identical to his first-best benchmark—the distortions associated with the principal’s stopping are absent. Therefore, the agent’s payoff is increasing in his faking time if  $t < \tau_P$  (his first-best payoff has a single peak at  $\tau_A > \tau_P$ ). Thus, the agent would not want to fake at such times, which reinforces the principal’s incentive to wait for news (principal’s first-best payoff has a single peak at  $\tau_P$ ). However, if time  $\tau_P$  passes without a real news arrival, the principal stops on the safe action, which the agent dislikes under the prior belief. Thus, the agent would prefer to preempt the principal’s stopping by generating a fake arrival of type-1 news just before  $\tau_P$ , thereby inducing the principal to select risky. Anticipating that it is fake, principal would ignore the type-1 arrival at time  $\tau_P$ , and the agent would like to preempt earlier, “unraveling” the search for news backwards from time  $\tau_P$ .

To explore this unraveling in more detail, it is helpful to consider incentives at two small adjacent time periods, an early period and a late period, both of which precede  $\tau_P$ . For ease of discussion, suppose that within-period, the news arrival occurs before the principal’s stopping decision.<sup>21</sup> As described above, unraveling is triggered by the agent’s expectation that the principal will stop at the end of the late period in the absence of news, which incentivizes the agent to fake at the beginning of the late period in order to preempt the principal’s imminent choice of safe. Anticipating such a faking strategy, the principal would be skeptical about such a convenient arrival of type-1 news. Indeed, with this faking strategy the probability of a real arrival in the late period is low (the period is small), but the probability of a fake is high. Thus, the principal identifies the arrival of type-1 news in the late period as a likely fake and ignores it, selecting safe in the late period whether or not type-1 news arrives. But if she always selects safe in the late period, why wait? It is better for the principal to avoid the cost of discounting by selecting safe at the end of the early period. Of course, if the agent anticipates such early stopping by the principal, he would want to preempt it by faking in the early period. However, if the manipulative agent always fakes early, then the principal does *not* want to stop her search at the end of the early period. Indeed, if no news comes in the early period, the principal concludes that the agent is *not* the manipulative type and that any future news is real. She therefore wants to wait for news until  $\tau_P$ . Thus, the agent wants to fake in the same period that the principal stops, but the principal wants to stop in

---

<sup>21</sup> In other words, if the principal plans to stop her search in a period, and type-1 news arrives in the period, then she selects risky according to her strategy  $a(\cdot)$ .

the other period. This cyclical structure of best responses suggests that the search unravels stochastically: the principal mixes between stopping in the early period and the late period, and the agent mixes between faking early and late.

Extended to continuous time, the preceding intuition suggests that in equilibrium, the atom of stopping and faking that would be concentrated at  $\tau_P$  under a “naive search” unravels into an interval of times with smooth, stochastic stopping and faking. But, if search is beneficial for the principal, the stochastic unraveling cannot extend all the way to time 0. If it does, then the principal’s payoff is the same as if she immediately terminates the search for news by selecting safe, i.e.,  $\theta$ . Furthermore, the unraveling cannot leave an atom of fakes at  $\tau_P$ ; the existence of such an atom triggered the unraveling in the first place. Provided there is no atom of fakes at  $\tau_P$ , an atom of stopping by the principal could occur at this time, since the agent’s faking strategy preempts the entire atom for certain.

The following proposition formalizes this intuition by characterizing the structure of an equilibrium with beneficial search.

**Proposition 1.** (*Structure of Beneficial Search Equilibrium.*) *If an equilibrium with beneficial search exists, then it has the following structure. There exists  $\tau_M > 0$  such that*

- (i) *the agent’s faking time is drawn from a mixed strategy with no mass points or gaps, supported on interval  $[\tau_M, \tau_P]$ .*
- (ii) *the principal’s stopping time is drawn from a mixed strategy with no gaps, supported on interval  $[\tau_M, \tau_P]$ ;  $\tau_P$  is the only possible mass point in the principal’s stopping strategy.*
- (iii) *the principal selects risky following type-1 news at all times,  $a(t) = 1$  for all  $t \geq 0$ .*

Consistent with the “stochastic unraveling” intuition described above, an equilibrium with beneficial search time features two deadlines: a soft deadline  $\tau_M$ , and a hard deadline  $\tau_P$ . If the hard deadline is reached with no news, the principal stops the search and selects safe. Before the soft deadline, the players incentives to wait for news are mutually reinforcing: during this phase there is no stopping by the principal or faking by the agent, and equilibrium search is identical to each player’s first best. However, past  $\tau_M$ , the search becomes distorted. At the soft deadline, the principal becomes “barely engaged” in search—she acts on a news arrival if one occurs, but she may terminate the search at any moment if one does not. The manipulative agent becomes “anxious” about the arrival rate of real news (which has decreased) and about the principal abandoning the search by selecting the safe action before news arrives. Consequently, he begins to fake randomly, which undermines the value of

information and reinforces the principal's minimal engagement in the search. Although type-1 news is randomly fabricated during this phase, it remains sufficiently informative that the principal prefers to select the risky action when such news arrives. Maintaining the principal's "bare engagement" with the search requires that a future type-1 news arrival is expected to be sufficiently informative to offset her cost of delay. By implication, such an arrival must also be informative enough to induce the risky action.

The equilibrium structure also reveals an interesting normative feature. The support of each player's mixed strategy begins at  $\tau_M$ , the soft deadline. Below  $\tau_M$ , there is no stopping or faking, and each player's payoff coincides with the first best. Furthermore, optimality of player strategies (4, 6) imply that each player's equilibrium payoff is equal to the payoff of stopping the search or faking at  $\tau_M$ . Because there are no atoms at  $\tau_M$ , this payoff is simply the first-best payoff evaluated at  $\tau_M$ . Thus, each player's payoff in the beneficial search equilibrium equal to the payoff he or she would receive in the first-best benchmark, but with a premature stopping time,  $\tau_M < \tau_P < \tau_A$ . In other words, if real news arrives in equilibrium before  $\tau_M$ , then both players get their first-best payoffs. But if the equilibrium search extends past  $\tau_M$ , players' payoffs are as if they simultaneously implemented their default actions at  $\tau_M$ , even though these default actions are different.

The following proposition completes the characterization of equilibria with beneficial search, deriving necessary and sufficient conditions for existence, establishing uniqueness, and providing closed-form expressions for the players' strategies.

**Proposition 2.** (*Equilibrium With Beneficial Search*). *An equilibrium with beneficial search exists if and only if  $\sigma < \bar{\sigma}$ , where*

$$\bar{\sigma} \equiv 1 - \exp\left\{-\int_0^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\}.$$

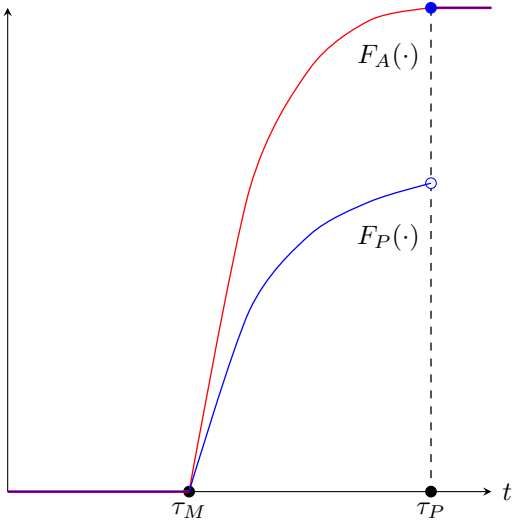
*In such an equilibrium, then the players' faking and stopping strategies are*

$$F_A(t) = \frac{1}{\sigma} \left(1 - \exp\left\{-\int_{\tau_M}^t \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\}\right),$$

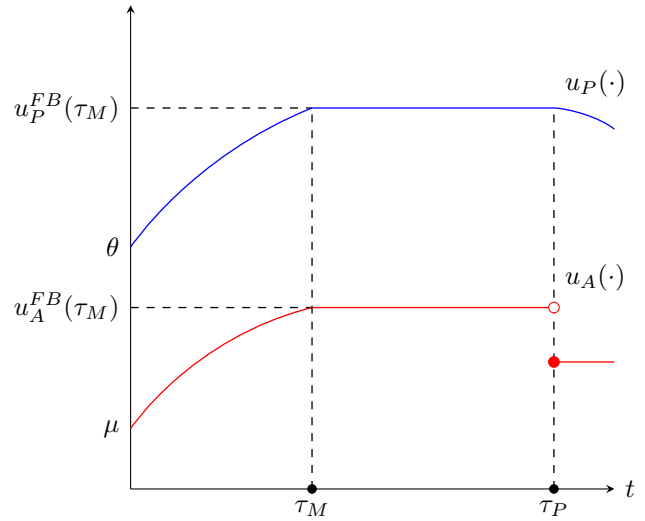
$$F_P(t) = 1 - \exp\left\{-\int_{\tau_M}^t \frac{\beta(1-\mu)H_R(s) - \rho\mu}{\mu - \beta} ds\right\} \quad \text{for } t \in [\tau_M, \tau_P) \quad \text{and} \quad F_P(\tau_P) = 1,$$

*both supported on  $[\tau_M, \tau_P]$ , where  $\tau_M$  is the unique value for which*

$$1 - \exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = \sigma.$$



**Figure 1** – Equilibrium Strategies  $F_A(\cdot), F_P(\cdot)$



**Figure 2** – Equilibrium payoffs  $u_A(\cdot), u_P(\cdot)$

The equilibrium payoff of player  $i \in \{A, P\}$  is  $u_i^{FB}(\tau_M)$ . No other equilibrium with beneficial search exists.

The equilibrium is illustrated graphically in Figures 1 and 2, which depict the equilibrium strategies and payoff functions of the players. As illustrated in Figure 1, the agent smoothly mixes on the equilibrium support  $[\tau_M, \tau_P]$ , but the principal’s strategy has an atom at  $\tau_P$ . Because of this atom, the agent’s payoff of waiting to fake until  $\tau_P$  has a downward jump discontinuity. The jump in the agent’s payoff does not disrupt the equilibrium, because the manipulative agent fakes the news with probability 1 before time  $\tau_P$  under his strategy. In other words, time  $\tau_P$  is only ever reached if the agent is *not* the manipulative type; therefore, the principal’s atom of stopping at this time does not disrupt the incentives of the manipulative agent. This atom has another interesting implication. Together with  $a(\cdot) = 1$  (see Proposition 1), the atom’s existence implies that the principal engages in a “naive search” with positive probability in equilibrium. To an outside observer, it may appear that the principal is completely unaware that the agent may have the ability to produce fake news, even though her equilibrium strategy accounts fully for this possibility.

Next, consider the informativeness of news in a beneficial search equilibrium. Although the underlying real news process is fully revealing, the equilibrium news observed by the principal is distorted by the agent’s faking. Consistent with the intuition underlying stochastic unraveling and Proposition 1, type-1 news is always sufficiently informative to induce the principal to select the risky action when it arrives. However, the informativeness of type-1 news is non-monotone in the arrival time. At early times, before the soft deadline, type-1

arrivals are fully revealing. At the soft deadline,  $\mu_1(t)$  drops discontinuously, and it increases gradually over the equilibrium support. At the hard deadline, type-1 news is fully revealing once again. The principal nevertheless prefers to exit at the hard deadline—even though a future news arrival reveals the risky payoff, it comes too slowly to justify additional delay. Thus, the informativeness of type-1 news exhibits a pattern of collapse and recovery, as formalized in the following corollary.

**Corollary 1.** (*Beneficial Search, Beliefs*). *In an equilibrium with beneficial search,*

(i) *the posterior belief that  $\omega = 1$  given type-1 news at  $t$  is*

$$\mu_1(t) = \begin{cases} 1 & \text{if } t \in [0, \tau_M) \\ \theta + (\theta - \mu) \frac{\rho\theta}{H_R(t)\mu(1-\mu) - \rho\theta} & \text{if } t \in [\tau_M, \tau_P] \\ 1 & \text{if } t \in [\tau_P, \infty). \end{cases}$$

(ii) *belief  $\mu_1(t)$  has a downward jump at  $\tau_M$ , is strictly increasing on  $[\tau_M, \tau_P)$ , and is continuous at  $\tau_P$ . In addition, an  $\varepsilon > 0$  exists such that  $\mu_1(t) \geq \theta + \varepsilon$  at all  $t \geq 0$ .*

(iii) *type-0 news reveals that  $\omega = 0$  at all times.*

The informativeness of the equilibrium news process is undermined most significantly near the soft deadline, at the earliest times in the support of the agent’s distribution of faking times. This might be surprising, given that the manipulative agent “panics” as the hard deadline approaches and the hazard rate of faking approaches infinity at the hard deadline. This effect arises from a “feedback loop” in the principal’s beliefs about the agent’s type. Reaching a time near the hard deadline without a type-1 arrival is, in and of itself, a strong signal that the agent is not manipulative, suggesting that a type-1 arrival is likely real. The manipulative agent exploits the principal’s optimism by faking aggressively. If type-1 news does not arrive, it is an even stronger signal that the agent is not manipulative, further reinforcing the manipulative type’s incentive to fake an instant later, and so on. The informativeness of type-1 news increases over time because the agent is less likely to be manipulative as time passes, which counteracts the increased aggressiveness of the manipulative agent’s faking.

The comparative statics with respect to  $\{\sigma, \theta\}$  are helpful in subsequent analysis. Intuitively, an increase in  $\theta$  intensifies the preference misalignment between players, as well as the disagreement over default action and search duration.<sup>22</sup> Thus, we might expect that the

---

<sup>22</sup>When the safe payoff is higher, the principal prefers safe more strongly under the prior, and in addition, prefers an even shorter duration under the first best, shifting it further away from the agent’s.

equilibrium distortions might be stronger, shifting the soft and hard deadlines, as well as the players' strategies toward earlier times. Meanwhile,  $\sigma$  determines the salience of the misalignment: if it is low, it is unlikely that the agent can affect the principal's action, even if he would gain more by doing so. Effectively, increases in  $\sigma$  intensify the conflict of interest and reductions weaken it. Thus, we might expect increases in  $\sigma$  to shift the soft deadline and the player's strategies toward earlier times, resulting in a stronger distortion from the first best.

**Corollary 2.** (*Beneficial Search, Comparative Statics*). *In an equilibrium with beneficial search, changes in model parameters have the following effects:*

- (i) *An increase in  $\sigma$  reduces  $\tau_M$ , has no effect on  $\tau_P$ , and generates a first order stochastic dominance shift in the stopping and faking strategies toward earlier times.*
- (ii) *An increase in  $\theta$  reduces  $\tau_M$  and  $\tau_P$ , and generates a first order stochastic dominance shift in the stopping and faking strategies toward earlier times.*

The normative properties of the equilibrium are also worth discussing further. As the probability of the agent being the faker increases, the stochastic unraveling intensifies and the soft deadline  $\tau_M$  shifts toward zero, thereby reducing the value of search and the equilibrium payoffs of the principal and manipulative agent. Furthermore, if the probability of the manipulative type is too large, then the unraveling completely undermines the value of search for the principal: no equilibrium with beneficial search exists. Perhaps the most striking normative feature of Proposition 2 is that the stochastic unraveling can completely undermine the value of search, even if the degree of preference misalignment between principal and agent is arbitrarily small. Indeed, provided that the players disagree over the default action ( $\mu \in (\beta, \theta)$ ), (A2) holds, and the agent is sufficiently likely to be manipulative, search has no value for the principal in equilibrium regardless of the degree of preference misalignment.

### 3.3. Equilibrium Without Beneficial Search

The preceding analysis implies that equilibrium with beneficial search exists only for a subset of parameters, in particular, when the prior probability of the manipulative agent is not too high. In the remaining situations, if an equilibrium exists, it must be such that search is not beneficial. Does such an equilibrium indeed exist when  $\sigma > \bar{\sigma}$ ? Furthermore, the logic of strategic unraveling described above suggests a strategic complementarity between the principal's stopping and the agent's faking. Could it be that, because of this strategic complementarity, multiple equilibria exist for some values of  $\sigma < \bar{\sigma}$ , one with beneficial search and another (or others) without?

This section answers the preceding questions, providing a partial characterization of equilibria in which the principal does not benefit from search. In particular, Proposition 3 establishes that an equilibrium without beneficial search exists if and only if  $\sigma$  is sufficiently large. Intuitively, if it is too unlikely that the agent can manipulate the news process, then he cannot “jam” the real news process enough to undermine its value completely. Furthermore, such an equilibrium exists whenever the equilibrium with beneficial search does not ( $\sigma > \bar{\sigma}$ ).

**Proposition 3.** (*Equilibrium Without Beneficial Search*). *An equilibrium without beneficial search exists if and only if  $\sigma \geq \underline{\sigma}$ , for some  $\underline{\sigma} \leq \bar{\sigma} < 1$ .*

The structure of such equilibria are worth discussing further. Note that if the principal does not benefit from search in equilibrium, then her equilibrium payoff cannot be higher than  $\theta$ , her payoff of choosing an action based on the prior alone. Simultaneously,  $u_P(0) = \theta$ , regardless of the agent’s strategy. Therefore, in such an equilibrium the principal’s equilibrium payoff must be  $\theta$ . Thus, if such an equilibrium exists, then choosing safe immediately is a best response to the agent’s equilibrium strategy. Furthermore, if the principal chooses safe immediately, then the agent is indifferent over all possible faking times (his payoff is always  $\beta$ ). In particular, the agent’s faking strategy continues to be optimal. By implication, immediate stopping for the principal, coupled with the agent’s original strategy is also an equilibrium. Thus, an equilibrium without beneficial search exists if and only if a faking strategy exists, such that immediate stopping is the principal’s best response. Thus, it is possible that the agent selects an adversarial strategy in equilibrium which minimizes the principal’s payoff and induces her to stop immediately. However, if the principal were to search, the agent would like to deviate. Without further restriction on the agent’s strategies, for  $\sigma \in (\underline{\sigma}, \bar{\sigma})$  it is possible that an equilibrium with beneficial search and without beneficial search coexist.

The equilibria without beneficial search can be further refined. One approach is to introduce a refinement similar to subgame perfection. Another approach considers a perturbed game in which there is a small probability that principal is a behavioral type who follows the naive search strategy. Either approach ensures that equilibria without beneficial search exist if and only if  $\sigma > \bar{\sigma}$ , eliminating the possible multiplicity for  $\sigma \in (\underline{\sigma}, \bar{\sigma})$ .<sup>23</sup>

## 4. Normative Remedies

The fundamental incentive problem in this game comes from the agent’s preemption motive, which triggers the stochastic unraveling of the search for news. Here, I describe three

---

<sup>23</sup>To streamline the exposition, I omit the argument. Additional details available upon request.



simple remedies that mitigate or eliminate the agent's preemption motive, deriving conditions under which they improve the principal's payoff. To streamline the analysis, I focus on the beneficial search equilibrium for  $\sigma \in (\underline{\sigma}, \bar{\sigma})$ .

**Commitment to Naive Search.** Consider a commitment by the principal to carry out a naive search, in which she waits until  $\tau_P$  to stop and acts on all news that arrives at or before this time. Effectively, the principal blinds herself to the possibility that the agent might be the manipulative type and behaves as if all news is real. Such a commitment eliminates the agent's preemption motive. In particular, because the principal always waits at  $t < \tau_P$ , and the agent can induce the risky action by faking type-1 news, the search is identical to the agent's first best at  $t < \tau_P$ . Thus, the agent has no incentive to fake at such times since he prefers a longer search duration in a first best search, i.e.,  $t < \tau_P < \tau_A$ . Given that the principal acts on it, the agent would therefore wait until  $\tau_P$  to fake a type-1 arrival. Without a binding commitment to act naively, the principal would recognize the fake and select the safe action if type-1 news arrives at  $\tau_P$ , triggering the agent to preempt. In contrast, with this binding commitment, the agent has no incentive to preempt: he can fake at  $\tau_P$  and still induce the risky action.

Though it eliminates the agent's preemption motive and the stochastic unraveling, such a commitment does not fully restore the principal's first-best search. In particular, if  $\tau_P$  is reached without a real news arrival, in the first-best the principal would like to stop and select the safe action. However, if she commits to a naive search, she may be forced to select risky. Indeed a manipulative agent will fake type-1 news at  $\tau_P$ . Following through on her commitment, the principal must select risky in this instance. In particular, with a commitment to naive search, the principal's expected payoff is

$$\begin{aligned} u_P^N &= \int_0^{\tau_P} \exp(-\rho s) g(s) (\mu + (1 - \mu)\theta) ds + \exp(-\rho \tau_P) (1 - G(\tau_P)) ((1 - \sigma)\theta + \sigma\mu) \\ &= u_P^{FB}(\tau_P) - \exp(-\rho \tau_P) (1 - G(\tau_P)) \sigma (\theta - \mu). \end{aligned}$$

The next proposition shows that such a commitment is always valuable for the principal if the equilibrium (without commitment) has beneficial search. More broadly, such a commitment is valuable unless the agent is very likely to be manipulative, so that the probability of selecting risky at the deadline is high.

**Proposition 4.** (*Naive Search*). *If the prior probability of the manipulative agent is not too high, then the principal's payoff with a commitment to naive search is higher than her equilibrium payoff in the*

main model. If  $\sigma < \bar{\sigma}$ , then commitment to naive search delivers a higher payoff to principal than the unique equilibrium without commitment, which has beneficial search. If  $\sigma \in (\bar{\sigma}, \tilde{\sigma}_N)$  for some  $\bar{\sigma} < \tilde{\sigma}_N \leq 1$ , then commitment to naive search delivers a higher payoff to the principal than any equilibrium without commitment, all of which feature non-beneficial search.

When the probability of manipulation is relatively low, the principal would like to pretend that manipulation is impossible. To support this pretense, the principal could develop an institution that allows her to plausibly deny the possibility of fabricated news. As awareness of such manipulation grows, it becomes more difficult for the principal to maintain this facade, threatening her commitment to naive search. To the extent that the principal is aligned with society more broadly, increased awareness of fabricated or manipulated information can therefore be counterproductive, unless it is also accompanied by increased accountability.

**Delegating Authority to the Agent.** Next, suppose that the principal can transfer the authority over the search and action directly to the agent. Doing so completely eliminates the agent's preemption motive, but it introduces an additional distortion in the decision process. In particular, if the agent controls the search, then he waits for news until  $\tau_A > \tau_P$  and if this time is reached without real news, then he selects risky: compared to the principal's first best, the agent waits too long for news and selects the wrong default action. Furthermore, unlike the naive search, where the risky action was selected at the deadline only if the agent is manipulative, with delegation the risky action is always chosen if the agent's deadline is reached. Thus, the principal's payoff of delegating authority to the agent is

$$\begin{aligned} u_P^D &= \int_0^{\tau_A} \exp(-\rho s) g(s) (\mu + (1 - \mu)\theta) ds + \exp(-\rho\tau_A) (1 - G(\tau_A)) \mu \\ &= u_P^{FB}(\tau_A) - \exp(-\rho\tau_A) (1 - G(\tau_A)) (\theta - \mu). \end{aligned}$$

As a first step, consider a simple choice for the principal: either she delegates the search to the agent, or she terminates the search immediately and selects the safe action. If the principal would rather terminate than delegate ( $u_P^D < \theta$ ), then the conflict of interest between the players is *severe*; otherwise it is *mild*. Obviously, with a severe conflict of interest, delegation could never be optimal—the principal's equilibrium payoff always exceeds  $\theta$  when she retains authority. Note that  $\theta$  and  $\beta$  can be arbitrarily close without violating any parametric assumption of the model, and in this case,  $u_P^D \approx u_P^{FB}(\tau_P) > \theta$ .<sup>24</sup> In other words, when the  $\theta$  and  $\beta$  are close,

---

<sup>24</sup>  $H_R(0) > \rho/(1 - \theta)$  guarantees (A1). Furthermore, as discussed in footnote 2, (A2) holds if  $\beta \in (\underline{\beta}, \mu)$ . Thus,  $\theta$  and  $\beta$  can be arbitrarily close to each other. It follows that  $\theta \approx \mu$  and  $\tau_P \approx \tau_A$ .

the conflict of interest is mild. Next, note that as long as  $\mu > \theta$ , we have  $\bar{\sigma} < 1$ . By implication, there is a range of large  $\sigma$  where the principal's payoff in the equilibrium is close to (or equal to)  $\theta$ . Continuity of the principal's equilibrium payoff as a function of  $\sigma$  delivers the following result.

**Proposition 5.** *(Delegating to Agent). If the conflict of interest is mild and the prior probability of the manipulative agent is sufficiently high, then the principal's payoff from delegating authority to the agent is higher than her equilibrium payoff in the main model. With a mild conflict and  $\sigma > \bar{\sigma}$ , the principal's payoff with delegation is higher than her payoff in any equilibrium of the main model. In addition, if  $\sigma \in (\tilde{\sigma}_D, \bar{\sigma})$  for some  $0 < \tilde{\sigma}_D < \bar{\sigma}$ , then the principal's payoff with delegation is higher than her payoff in the unique equilibrium of the main model, which features beneficial search.*

Proposition 5 suggests that the credible threat of manipulation of the news process allows the agent to gain outsized influence over the decision. Indeed, if the conflict of interest is relatively mild and the agent can probably manipulate the news, the principal would rather give the agent full authority over the search and decision than initiate a (possibly) manipulated search. Even though the agent has no private information about the best action, he is nevertheless granted complete authority over the decision. This contrasts with the classic logic supporting delegation, which is based on the agent's ability to access or produce superior payoff-relevant information (Aghion and Tirole, 1997; Dessein, 2002).

**Delegating Authority to an Intermediary.** In this section, a milder form of delegation is analyzed: the principal delegates the decision authority to a third party whose payoff from the safe action is different than her own. This section focuses on the more interesting case  $\sigma < \bar{\sigma}$ , which implies that the unique equilibrium has beneficial search. In addition, we focus on the principal's incentive to delegate "locally," so that the intermediary's payoff from the safe of option,  $\theta_I$ , is not too far from  $\theta$ . In this way, we can ensure that (A1) and (A2) are satisfied by the intermediary and that the equilibrium between the intermediary and agent features beneficial search.

By delegating to an intermediary, principal alters the equilibrium search for news: the incentives of the agent and intermediary determine the equilibrium, rather than the principal. Thus, with delegation the equilibrium is characterized by Proposition 2 with the intermediary standing in for the principal. Building on the comparative statics derived in Corollary 1, consider the effects of delegating authority to an intermediary with a smaller payoff from the safe action. On one hand, such an intermediary has a smaller preference conflict with the agent.

By delegating to such an intermediary, the principal weakens the incentives for stochastic unraveling, shifting the soft deadline to a later time. Thus by delegating, the principal extends the duration of the first best search, which increases her expected payoff. On the other hand, delegating authority to an intermediary also affect the agent's faking strategy. In equilibrium, the intermediary is indifferent between continuing the search and stopping at all times between the soft and hard deadlines if there has been no news. But the principal's payoff of stopping the search and selecting safe is higher, and she therefore prefers to stop at all such times. In other words, when she delegates, the principal's payoff function is decreasing between the soft and hard deadlines. Although her payoff is higher at the soft deadline, if the search extends for too long beyond it, the principal is harmed. This cost is particularly salient: by delegating, the principal shifts both the intermediary's stopping strategy and the agent faking strategy toward later times, increasing the likely duration of the search.

To make this discussion precise, consider the equilibrium of Proposition 2, with intermediary playing the role of principal. To distinguish the equilibrium with delegation, the intermediary's preference parameter is included explicitly in the notation. Thus,  $\tau_P(\theta_I)$  is the intermediary's first best search duration (and the hard deadline when it has the authority),  $\tau_M(\theta_I)$  is the soft deadline,  $F_A(\cdot|\theta_I)$ ,  $F_P(\cdot|\theta_I)$  are the equilibrium strategies of the agent and intermediary when it plays the role of principal, and  $u_P(t|\theta_I, \theta)$  is the principal's equilibrium expected payoff conditional on the intermediary's stopping time  $t$ . Note that the equilibrium strategies and the hard and soft deadline depend only on the intermediary's safe payoff  $\theta_I$ , but the principal's safe payoff enters explicitly into  $u_P(t|\theta_I, \theta)$ . Indeed, whenever the intermediary selects the safe action, the principal's payoff is  $\theta$ . In particular, we have

$$u_P(t|\theta_I, \theta) = u_P^{FB}(\tau_M(\theta_I)) + \int_{\tau_M(\theta_I)}^t \exp(-\rho s) \{w_0^P(s|\theta_I)\theta + w_1^P(s|\theta_I)\mu_1(s|\theta_I)\} ds + \exp(-\rho t)W_\phi^P(t|\theta_I)\theta,$$

where use has been made of the fact that the intermediary always selects the risky action following type-1 news in equilibrium (see Proposition 1). Recalling that the intermediary's equilibrium strategy has a mass point on  $\tau_P(\theta_I)$  but is otherwise differentiable, we have that the principal's expected payoff of delegating to intermediary with preference parameter  $\theta_I$  is

$$\begin{aligned} U(\theta_I|\theta) &= \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} f_P(s|\theta_I)u_P(s|\theta_I, \theta)ds + (1 - F_P(\tau_P^-(\theta_I)|\theta_I))u_P(\tau_P(\theta_I)|\theta_I, \theta) \\ &= \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} (1 - F_P(s|\theta_I))u_P'(s|\theta_I, \theta)ds + u_P^{FB}(\tau_M(\theta_I)), \end{aligned}$$

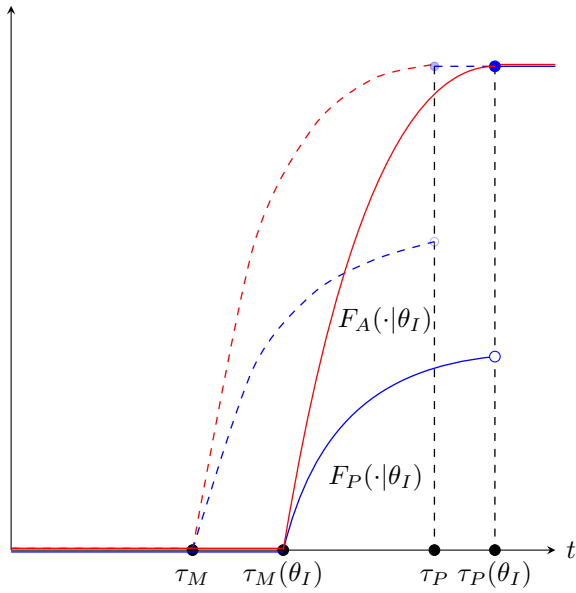


Figure 3 – Strategies, Delegation  $F_A(\cdot|\theta_I)$ ,  $F_P(\cdot|\theta_I)$

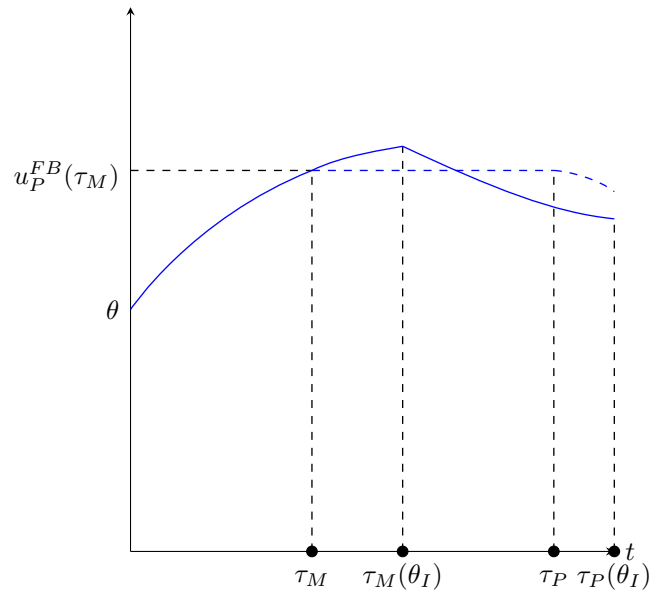


Figure 4 – Delegation payoff  $u_P(\cdot|\theta_I, \theta)$

where the last line follows from an integration by parts.<sup>25</sup> In this form, tradeoff is apparent: by delegating to an intermediary with a smaller payoff of the safe action, the principal extends the soft deadline and the duration of first best search, but if the soft deadline is reached with no news continued search is strictly harmful.

The tradeoff of delegating is illustrated graphically in Figures 3 and 4. Figure 3 illustrates the shift toward later stopping and faking times induced by delegation to an intermediary with a smaller payoff from the safe option. The dashed curves represent the equilibrium strategies when the principal keeps authority for herself, while the solid curves are the strategies under delegation. Figure 4 illustrates the principal's payoff when she delegates to an intermediary with safe payoff  $\theta_I$ , and when the intermediary select stopping time  $t$ . The solid blue curve is this payoff,  $u_P(\cdot|\theta_I, \theta)$ , while the dashed blue curve is the principal's payoff from keeping authority herself.

The following proposition shows that the benefits of extending the first best search outweigh the harm.

**Proposition 6.** *(Delegating to an Intermediary.) If  $\sigma < \bar{\sigma}$ , then the principal benefits by delegating decision authority to an intermediary whose payoff from the safe option is smaller than her own. In particular, a  $\theta_I^* < \theta$  exists, such that the principal benefits from delegation if  $\theta_I \in (\theta_I^*, \theta)$ . Furthermore, if the principal delegates authority to such an intermediary, then in equilibrium she is never tempted*

<sup>25</sup> See Proof of Proposition 6.

*to overrule the intermediary's chosen action.*

The second part of Proposition 6 highlights another difference between delegating to an intermediary and the agent. When delegating authority to the agent, the principal is tempted to overrule the agent if he tries to select risky at time  $\tau_A$ . When delegating to an intermediary with a small bias, the principal will not be tempted to overrule the intermediary's choice of risky action. As shown in Corollary 1, the posterior belief that  $\omega = 1$  given an arrival of type-1 news exceeds  $\theta_I$ , and furthermore, the gap is strictly positive. If  $\theta_I$  and  $\theta$  are close, then the same is also true of  $\theta$ .

## 5. Conclusion

I have studied a model of information acquisition, in which information can be faked strategically. A principal must choose between a safe action with a known payoff and a risky action with an uncertain payoff. An arrival of real news reveals the payoff of the risky action, arriving at an uncertain future time. The principal prefers the safe action under the prior, but she can wait for news to arrive. An agent has a smaller payoff of the safe action and prefers risky under the prior. The agent may have the ability to fabricate a news arrival that is indistinguishable from a real one.

The disagreement over the best action under the prior belief incentivizes preemption by the agent. whenever the principal stops before news arrives, she selects risky. Thus, if the agent expects that the principal is about to stop, he has an incentive to fake a news arrival, in order to induce the risky action, which he prefers under the prior belief. The agent's faking undermines the value of future news, which creates an incentive for her to stop without news.

From a positive perspective, two distortions arise in equilibrium. When the probability that the agent can manipulate news is not too high, the equilibrium features an early phase of undistorted search, followed by a phase in which agent randomly fakes news and the principal randomly stops before news arrives. Though it is sometimes manipulated, all news is sufficiently informative for the principal to follow. From a normative perspective, the possibility of faking undermines the value of information and reduces the principal's payoff. If the probability that the agent can fabricate news is sufficiently high, the benefits of news are undermined completely. I study three straightforward remedies that mitigate the agent's preemption motive and increase the value of search.

The model can be extended in a number of interesting directions. The current version studies only a single dynamic interaction between the players. It might be interesting to

consider incentives of an agent who interacts with principals repeatedly. In this case, the agent may be concerned about revealing his ability to fake to future principals, which may temper his incentive to manipulate. Another possibility is to consider stronger commitment power for the principal. While I have focused on relatively simple forms of commitment that dampen the agent's preemption motive, it would be interesting to know how the agent's ability to fake news allows him to obtain rents, even when the principal can commit to her entire strategy. I plan to address these and other related questions in future work.

## References

- Aghion, P., and M. O. Jackson, 2016, "Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits," *American Economic Journal: Microeconomics*, 8, 1–38.
- Aghion, P., and J. Tirole, 1997, "Formal and real authority in organizations," *Journal of political economy*, 105, 1–29.
- Arrow, K., D. Blackwell, and M. Girshick, 1949, "Bayes and Minimax Solutions of Sequential Decision Problems," *Econometrica*, 17, 213–244.
- Boleslavsky, R., and C. Taylor, 2023, "Make It 'Til You Fake It," *Mimeo*.
- Brocas, I., and J. D. Carrillo, 2007, "Influence Through Ignorance," *RAND Journal of Economics*, 38, 931–947.
- Che, Y.-K., W. Dessein, and N. Kartik, 2013, "Pandering to Persuade," *American Economic Review*, 103, 47–79.
- Che, Y.-K., K. Kim, and K. Mierendorff, 2022, "Keeping the Listener Engaged: a Dynamic Model of Bayesian Persuasion," *Journal of Political Economy*, forthcoming.
- Che, Y.-K., and K. Mierendorff, 2019, "Optimal Dynamic Allocation of Attention," *The American Economic Review*, 109, 2993–3029.
- Dessein, W., 2002, "Authority and Communication in Organizations," *Review of Economic Studies*, 69, 811–838.
- Dilme, F., 2019, "Dynamic quality signaling with hidden actions," *Games and Economic Behavior*, 113, 116–136.
- Ekmekci, M., L. Gorno, L. Maestri, J. Sun, and D. Wei, 2022, "Learning from Manipulable Signals," *American Economic Review*, 112, 3995–4040.



- Ely, J. C., 2017, "Beeps," *American Economic Review*, 107, 31–53.
- Ely, J. C., and M. Szydlowski, 2020, "Moving the Goalposts," *Journal of Political Economy*, 128, 468–506.
- Gross, S., M. Possley, K. Roll, and K. Stephens, 2020, "Government Misconduct and Convicting the Innocent, The Role of Prosecutors, Police and Other Law Enforcement," *SSRN Electronic Journal*.
- Halac, M., N. Kartik, and Q. Liu, 2017, "Contests for Experimentation," *Journal of Political Economy*, 125, 1523–1569.
- Henry, E., and M. Ottaviani, 2019, "Research and the Approval Process: The Organization of Persuasion," *American Economic Review*, 109, 911–955.
- Kamenica, E., and M. Gentzkow, 2011, "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.
- Kolb, A. M., 2015, "Optimal entry timing," *Journal of Economic Theory*, 157, 973–1000.
- Kolb, A. M., 2019, "Strategic real options," *Journal of Economic Theory*, 183, 344–383.
- Kuvalekar, A., and E. Lipnowski, 2020, "Job Insecurity," *American Economic Journal: Microeconomics*, 12, 188–229.
- Liang, A., X. Mu, and V. Syrgkanis, 2022, "Dynamically Aggregating Diverse Information," *Econometrica*, 90, 47–80.
- Orlov, D., A. Skrzypacz, and P. Zryumov, 2020, "Persuading the Principal to Wait," *Journal of Political Economy*, 128, 2542–2578.
- Renault, J., E. Solan, and N. Vieille, 2017, "Optimal dynamic information provision," *Games and Economic Behavior*, 104, 329–349.
- Shadmehr, M., and D. Bernhardt, 2015, "State censorship," *American Economic Journal: Microeconomics*, 7, 280–307.
- Sun, Y., 2023, "A dynamic model of censorship," *Theoretical Economics*, forthcoming.
- Wald, A., 1947, "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, pp. 279–313.

## 6. Appendix

**Proof of Lemma 1.** Consider the first-best benchmark for player  $i$ . Let  $(d_i, s_i)$  be player  $i$ 's expected payoff of the default action, and safe action respectively. That is  $(d_P, s_P) = (\theta, \theta)$  and  $(d_A, s_A) = (\mu, \beta)$ . We have

$$u_i^{FB}(t) = \int_0^t \exp(-\rho x) g(x) (\mu + (1 - \mu) s_i) dx + \exp(-\rho t) (1 - G(t)) d_i.$$

$G(\cdot)$  is differentiable, and therefore

$$\begin{aligned} \frac{d}{dt} u_i^{FB}(t) &= \exp(-\rho t) g(t) (\mu + (1 - \mu) s_i) - d_i \exp(-\rho t) (\rho (1 - G(t)) + g(t)) \\ &= \exp(-\rho t) (1 - G(t)) (\mu + (1 - \mu) s_i - d_i) \left\{ H_R(t) - \frac{d_i \rho}{\mu + (1 - \mu) s_i - d_i} \right\}. \end{aligned}$$

Note that  $d_i < \mu + (1 - \mu) s_i - d_i$  for  $i \in \{A, P\}$ . It follows that the first three terms are positive. The sign of the derivative is therefore determined by the sign of the fourth term. Because  $H_R(\cdot)$  is decreasing, the derivative of  $u_i^{FB}(\cdot)$  crosses zero at most once, and if it exists, this crossing is from above. It follows that  $u_i^{FB}(\cdot)$  is single-peaked. Under A1 and A2, the peak is at a strictly positive  $t$ . Substituting  $(d_i, s_i)$ , for  $i \in \{A, P\}$ , we find that the crossing occurs at  $\tau_i$  such that  $H_R(\tau_i) = \phi_i$ . A1 and A2 imply  $\tau_A > \tau_P > 0$ .  $\square$

### 6.1. Proof of Proposition 1

The proof makes use of a number of lemmas, stated in the next section. To simplify the exposition, the proofs of the lemmas are presented in the Supplemental Appendix.

### 6.2. Lemmas

**Observation 1.** Note that

$$\int_0^t w_0^P(s) + w_1^P(s) ds + W_\phi^P(t) = 1.$$

Indeed, the integral is the probability that a type-0 or type-1 arrival occurs at times less than or equal to  $t$ , while  $W_\phi^P(t)$  is the probability that an arrival of news takes longer than  $t$ . Rearranging, for  $t_1 \leq t_2$  we have

$$\int_{t_1}^{t_2} w_0^P(s) + w_1^P(s) ds = W_\phi^P(t_1) - W_\phi^P(t_2).$$

Similarly, for the agent we have,

$$\int_{t_1}^{t_2} w_0^A(s) + w_1^A(s) + w_S^A(s) ds = W_\phi^A(t_1) - W_\phi^A(t_2).$$

**Lemma 2.** *If (5), then inequalities (i), (ii), (iii) hold,*

- (i)  $u_P(t') - u_P(t) \leq (1 - \sigma F_A(t))\{u_P^{FB}(t') - u_P^{FB}(t)\}$  for  $t < t'$ .
- (ii)  $u_P(t) < u_P(\tau_P)$  for  $\tau_P < t \leq \infty$ .
- (iii)  $u_P(t') - u_P(t) \geq -(\exp(-\rho t) - \exp(-\rho t'))W_\phi^P(t)\theta$  for  $t < t'$ .
- (iv) Part (i) holds as an equality if  $F_A(t) = F_A(t')$ .

**Lemma 3.** *If (5), then  $u_P(\cdot)$  is continuous, regardless of the agent's strategy.*

**Lemma 4.** *If (5) and (6), then  $F_P(\tau_P) = 1$ .*

**Definition:** Let

$$v(t) \equiv \int_0^t \exp(-\rho x)g(x)(\mu + (1 - \mu)\beta)dx + \exp(-\rho t)(1 - G(t))\beta.$$

**Lemma 5.** *Inequalities (i), (ii) hold for  $t < t' \leq \tau_P$ , regardless of the principal's strategy.*

- (i)  $u_A(t') - u_A(t) \leq (1 - F_P(t))\{v(t') - v(t)\} + \exp(-\rho t')W_\phi^A(t')(\mu - \beta)a(t') - \exp(-\rho t)W_\phi^A(t)(\mu - \beta)a(t)$ .
- (ii)  $u_A(t') - u_A(t) \geq \{\exp(-\rho t')W_\phi^A(t')a(t') - \exp(-\rho t)W_\phi^A(t)a(t)\}(\mu - \beta) - \{\exp(-\rho t) - \exp(-\rho t')\}W_\phi^A(t)\beta$

(iii): *if  $F_P(t) = F_P(t')$ ,  $a(t) = a(t') = 1$ , and  $a(s) = 1$  for almost all  $s \in (t, t')$ , then*

$$u_A(t') - u_A(t) = (1 - F_P(t))(u_A^{FB}(t') - u_A^{FB}(t)).$$

**Lemma 6.** *Suppose (4). If  $a(t) < 1$  and  $F_P(t) < 1$  for some  $t$ , then*

- (i) *an  $\varepsilon' > 0$  such that  $a(t') < 1$  for  $t' \in [t, t + \varepsilon')$ .*
- (ii) *an  $\varepsilon'' > 0$  exists such that for  $a(t'') < 1$  for  $t'' \in (t - \varepsilon'', t]$ , provided  $t > 0$ .*

**Lemma 7.** *Suppose an interval  $0 \leq t_1 < t_2 \leq \tau_P$  exist such that  $F_A(t_1) = F_A(t_2)$ . If (5) and (6), then*

- (i)  $f_P(t) = 0$  for all  $t \in [t_1, t_2)$ .

(ii) If  $t_1 > 0$ , then there exists  $t_0 < t_1$  such that  $f_P(t) = 0$  for all  $t \in [t_0, t_2)$ .

**Lemma 8.** Suppose an interval  $[t_1, t_2)$  exists such that  $F_P(t_1) = F_P(t_2) < 1$ . In equilibrium,  $f_A(t) = 0$  for all  $t \in [t_1, t_2)$ .

**Lemma 9.** Suppose  $0 < t_1 < t_2 \leq \tau_P$  and  $F_P(t_1) = F_P(t_2) < 1$ . In equilibrium,  $F_P(t) = F_P(t_2)$  for all  $t \in [0, t_2]$ .

**Lemma 10.** Suppose  $0 < t_1 < t_2$ ,  $F_P(t_2) < 1$ , and  $F_A(t_1) = F_A(t_2)$ . In equilibrium,  $F_A(t) = F_A(t_2)$  for all  $t \in [0, t_2]$ .

### 6.3. Proof of Proposition 1

The Proposition is proved by establishing a number of claims. Each claim assumes the existence of an equilibrium with beneficial search, and its conclusion describes a property of such an equilibrium.

Let  $\bar{t}_i$ , for  $i \in \{P, A\}$  denote the maximum of the support of the principal's and agent's stopping strategies, respectively.

**Claim 1.**  $\bar{t}_A = \bar{t}_P = \tau_P$ .

**Proof of Claim 1.** *Step 1:*  $\bar{t}_P > 0$ . Suppose  $\bar{t}_P = 0$ . It follows immediately that  $F_P(t) = 1$  for all  $t \geq 0$ . That is, the principal terminates the search at time 0. Regardless of the agent's strategy,  $u_P(0) = \theta$ . Therefore, the principal does not benefit from search.

*Step 2:*  $\bar{t}_P \leq \tau_P$ . Follows immediately from Lemma 4.

*Step 3:*  $\bar{t}_A > 0$ . Suppose  $\bar{t}_A = 0$ , i.e.,  $F_A(t) = 1$  for  $t \geq 0$ . It follows immediately that  $f_A(t) = 0$  for all  $t > 0$ . From Lemma 7, we have  $f_P(t) = 0$  for all  $t \in [0, \tau_P)$ . From Lemma 4, we have  $F_P(t) = \mathcal{I}(t \geq \tau_P)$ . From Lemma 5 part (iii), the  $u_A(t) - u_A(0) = u_A^{FB}(t) - u_A^{FB}(0)$  for  $0 < t < \tau_P$ . Because  $t < \tau_P < \tau_A$ , Lemma 1 implies  $u_A^{FB}(t) - u_A^{FB}(0) > 0$ . Thus  $f_A(0) = \delta(0)$  violates 4.

*Step 4:*  $\bar{t}_A \geq \bar{t}_P$ . Suppose  $\bar{t}_A < \bar{t}_P$ . For  $t \in (\bar{t}_A, \bar{t}_P)$ , we have  $f_A(t) = 0$ , and therefore  $F_A(t) = F_A(\bar{t}_A)$ . Furthermore,  $\bar{t}_A < \bar{t}_P$  implies  $F_P(\bar{t}_A) < 1$ . Applying Lemma 10, we have  $F_A(t) = F_A(\bar{t}_A)$  for all  $t \in [0, \bar{t}_A]$ . In other words, the CDF of the agent's strategy is constant from 0 up to and including the top of the support  $\bar{t}_A$ . It follows that  $\bar{t}_A = 0$ , contradicting Step 3.

*Step 5:* For any  $\varepsilon > 0$ , there exists  $t \in (\bar{t}_P - \varepsilon, \bar{t}_P)$  such that  $a(t) = 1$ .

Suppose not. An  $\varepsilon > 0$  exists such that  $a(t) < 1$  for all  $t \in (\bar{t}_P - \varepsilon, \bar{t}_P)$ . It follows that for all  $t \in (\bar{t}_P - \varepsilon, \bar{t}_P]$ , we have  $f_A(t) > 0$  and  $\mu_1(t) = \theta$ . Therefore, for  $\bar{t}_P - \varepsilon \leq t'' < t \leq \bar{t}_P$  we have

$$u_P(t) - u_P(t'') = \int_{t''}^t \exp(-\rho s)(w_0^P(s) + w_1^P(s))\theta ds + \exp(-\rho t)W_\phi^P(t)\theta - \exp(-\rho t'')W_\phi^P(t'')\theta$$

$$\leq \exp(-\rho t'') \int_{t''}^t (w_0^P(s) + w_1^P(s)) \theta ds + \exp(-\rho t) W_\phi^P(t) \theta - \exp(-\rho t'') W_\phi^P(t'') \theta.$$

Using Observation 1, we have

$$\begin{aligned} u_P(t) - u_P(t'') &\leq \exp(-\rho t'') (W_\phi^P(t'') - W_\phi^P(t)) \theta + \exp(-\rho t) W_\phi^P(t) \theta - \exp(-\rho t'') W_\phi^P(t'') \theta \\ &= - \{ \exp(-\rho t'') - \exp(-\rho t) \} W_\phi^P(t) \theta < 0, \end{aligned}$$

where the last inequality follows from  $W_\phi^P(t) = (1 - \sigma F_A(t))(1 - G(t)) > 0$  for all  $t \geq 0$ . It follows from (6) that  $f_P(t) = 0$  for all  $t \in (\bar{t}_P - \varepsilon, \bar{t}_P]$ , contradicting the assumption that  $\bar{t}_P$  is the maximum of the support of the principal's stopping strategy.

*Step 6:*  $\bar{t}_A \leq \bar{t}_P$ . Consider any  $t' > \bar{t}_P$ . From Step 5 we have that for any  $\varepsilon$ , there exists some  $t_\varepsilon \in (\bar{t}_P - \varepsilon, \bar{t}_P)$  such that  $a(t_\varepsilon) = 1$ . It follows that

$$u_A(t') - u_A(t_\varepsilon) = \underbrace{u_A(t') - u_A(\bar{t}_P)}_A + \underbrace{u_A(\bar{t}_P) - u_A(t_\varepsilon)}_B.$$

Focus first on  $A$ . For  $t \in [\bar{t}_P, t']$  we have  $F_P(t) = 1$ , and for  $t \in (\bar{t}_P, t']$  we have  $f_P(t) = 0$ . Thus,

$$u_A(t') - u_A(\bar{t}_P) = 0.$$

Next, focus on  $B$ . From Lemma 5, we have

$$\begin{aligned} u_A(\bar{t}_P) - u_A(t_\varepsilon) &\leq (1 - F_P(t_\varepsilon)) \{ v(\bar{t}_P) - v(t_\varepsilon) \} - \exp(-\rho t_\varepsilon) W_\phi^A(t_\varepsilon) (\mu - \beta) \Rightarrow \\ \frac{u_A(\bar{t}_P) - u_A(t_\varepsilon)}{1 - F_P(t_\varepsilon)} &\leq v(\bar{t}_P) - v(t_\varepsilon) - \exp(-\rho t_\varepsilon) (1 - G(t_\varepsilon)) (\mu - \beta). \end{aligned}$$

Because  $v(\cdot)$  is continuous, for  $\varepsilon$  sufficiently small, we have  $v(\bar{t}_P) - v(t_\varepsilon) \approx 0$ . Simultaneously,  $\exp(-\rho t_\varepsilon) (1 - G(t_\varepsilon)) (\mu - \beta) > \exp(-\rho \tau_P) (1 - G(\tau_P)) (\mu - \beta) > 0$ , where the last inequality follows because  $\exp(-\rho \cdot) (1 - G(\cdot))$  is decreasing and Step 2 ( $t_\varepsilon < \bar{t}_P < \tau_P$ ). Therefore, for  $\varepsilon$  sufficiently small, there exists  $t_\varepsilon$  such that RHS is strictly negative. It follows that  $B$  is strictly negative. Recalling that  $A$  is 0, we have  $u_A(t') - u_A(t_\varepsilon) < 0$ . Therefore, a  $t_\varepsilon$  exists such that  $u_A(t') - u_A(t_\varepsilon) < 0$  for all  $t' > \bar{t}_P$ . From (4), we have  $f_A(t') = 0$  for all such  $t'$ . It follows that  $\bar{t}_A \leq \bar{t}_P$ .

*Step 7:*  $\bar{t}_P = \bar{t}_A = \tau_P$ . Combining Steps 2, 4, and 6 we have  $\bar{t}_P = \bar{t}_A \leq \tau_P$ . Suppose that the inequality is strict, i.e.,  $\bar{t}_P < \tau_P$ . Because  $\bar{t}_A = \bar{t}_P$ , for all  $t \in [\bar{t}_P, \tau_P]$  we have  $F_A(t) = 1$ . Using Lemma 2 part (iv),  $u_P(t) - u_P(\bar{t}_P) = (1 - \sigma F_A(\bar{t}_P)) (u^{FB}(t) - u^{FB}(\bar{t}_P))$  for all  $t \in [\bar{t}_P, \tau_P]$ . Note

that  $u^{FB}(t) - u^{FB}(\bar{t}_P) > 0$  for such  $t$  (see Lemma 1). Therefore,  $u_P(t) > u_P(\bar{t}_P)$ . Note further that continuity of  $u_P(\cdot)$  (see Lemma 3) implies that  $u_P(\bar{t}_P) = u_P^*$ , the principal's equilibrium payoff. Thus,  $u_P(t) > u_P^*$ , which contradicts (6).  $\square$

Let  $\underline{t}_i, i \in \{P, A\}$ , denote the bottom of the support of the principal and agent stopping strategy, respectively.

**Claim 2.**  $\underline{t}_P = \underline{t}_A = \tau_M$ , for some  $\tau_M > 0$ .

**Proof of Claim 2.** *Step 1:*  $\underline{t}_P > 0$ . From Lemma 3, in equilibrium  $u_P(\cdot)$  is continuous. It follows that  $u_P(\underline{t}_P) = u_P^*$ , the principal's equilibrium payoff. Note that  $u_P(0) = \theta$ , regardless of the agent's strategy. Therefore, if  $\underline{t}_P = 0$ , then  $u_P^* = \theta$ . In this case, the principal does not benefit from search.

*Step 2:*  $\underline{t}_A \leq \underline{t}_P$ . First, suppose  $\underline{t}_A = 0$ . From Step 1, we have  $\underline{t}_P > 0$ ; thus Step 2 follows immediately. Second, consider  $\underline{t}_A > 0$ . Note that  $\underline{t}_A \leq \bar{t}_A = \tau_P$ , where the equality is Claim 1. Next, consider any  $t' \in (0, \underline{t}_A)$ . Because  $t' < \underline{t}_A$ , we have  $F_A(t') = F_A(0) = 0$ . From Lemma 7, we have  $f_P(t) = 0$  for all  $t \in [0, t')$ . Because  $t'$  can be chosen arbitrarily close to  $\underline{t}_A$ , we have  $f_P(t) = 0$  for all  $t \in [0, \underline{t}_A)$ . Therefore,  $\underline{t}_P \geq \underline{t}_A$ .

*Step 3:*  $\underline{t}_A \geq \underline{t}_P$ . Consider any  $t' \in [0, \underline{t}_P)$ . For such  $t'$ , we have  $F_P(0) = F_P(t') = 0$ . Using Lemma 8, we have  $f_A(t) = 0$  for all  $t \in [0, t')$ . Because  $t'$  can be chosen arbitrarily close to  $\underline{t}_P$ , we have  $f_A(t) = 0$  for all  $t \in [0, \underline{t}_P)$ . Therefore,  $\underline{t}_A \geq \underline{t}_P$ .

*Step 4:*  $\underline{t}_A = \underline{t}_P = \tau_M > 0$ . Combining Steps 2 and 3 we have  $\underline{t}_A = \underline{t}_P = \tau_M$ , for some  $\tau_M \geq 0$ . Using Step 1, we have  $\tau_M > 0$ .  $\square$

**Claim 3.** *The support of  $F_A(\cdot)$  is the interval  $[\tau_M, \tau_P]$ .*

**Proof of Claim 3.** It must be shown that  $F_A(t_2) - F_A(t_1) > 0$  for any interval  $(t_1, t_2)$  where  $\tau_M < t_1 < t_2 < \tau_P$ . Suppose that there exists an interval  $(t_1, t_2)$  satisfying the required bounds such that  $F_A(t_1) = F_A(t_2)$ . Furthermore, because  $\tau_M < t_2 < \tau_P$ , the max and min of the support, we have  $0 < F_A(t_2) < 1$ . From Lemma 10, we have  $F_A(t) = F_A(t_2)$  for all  $t \in [0, t_2)$ . Therefore,  $F_A(0) > 0$ . By implication  $\tau_M = 0$ , which contradicts Step 4.  $\square$

**Claim 4.** *The support of  $F_P(\cdot)$  is the interval  $[\tau_M, \tau_P]$ .*

**Proof of Claim 4.** It must be shown that  $F_P(t_2) - F_P(t_1) > 0$  for any interval  $(t_1, t_2)$  where  $\tau_M < t_1 < t_2 < \tau_P$ . Suppose that there exists an interval  $(t_1, t_2)$  satisfying the required bounds

such that  $F_P(t_1) = F_P(t_2)$ . Furthermore, because  $\tau_M < t_2 < \tau_P$ , the max and min of the support, we have  $0 < F_P(t_2) < 1$ . From Lemma 9, we have  $F_P(t) = F_P(t_2)$  for all  $t \in [0, t_2]$ . Therefore,  $F_P(0) > 0$ . By implication  $\tau_M = 0$ , which contradicts Step 4.  $\square$

**Claim 5.** *The principal acts on type-1 news at all times, i.e.  $a(t) = 1$  for all  $t \geq 0$ .*

**Proof of Claim 5.** First, consider  $t < \tau_M$  and suppose that  $a(t) < 1$  for such  $t$ . From Lemma 6, an interval  $(t_1, t_2)$  exists such that  $t_1 \leq t < t_2$  and  $a(t') < 1$  for all  $t' \in (t_1, t_2)$ . Furthermore,  $a(t') < 1$  implies  $\mu_1(t') \leq \theta$  (consult 5). In turn,  $\mu_1(t') \leq \theta$  implies  $f_A(t') > 0$  (consult 1). Hence, for  $t' \in (t_1, t_2)$ , we have  $f_A(t') > 0$ . Recall  $t_1 \leq t < \tau_M$ . Therefore,  $F_A(\cdot)$  is strictly increasing on interval  $(t_1, \min\{\tau_M, t_2\})$ , which contradicts Claim 3. The argument for  $t > \tau_P$  is similar.

Next, consider  $t \in [\tau_M, \tau_P]$ , and suppose  $a(t) < 1$ . From Lemma 6, an interval  $(t_1, t_2) \subset [\tau_M, \tau_P]$  exists such that  $a(t) < 1$  for  $t \in (t_1, t_2)$ . Furthermore,  $a(t) < 1 \Rightarrow \mu_1(t) = \theta$  for  $t \in (t_1, t_2)$  (see 5). It follows that for  $t_1 < t'' < t' < t_2$  we have

$$\begin{aligned} u_P(t') - u_P(t'') &= \int_{t''}^{t'} \exp(-\rho s) \{w_0^P(s)\theta + w_1^P(s)\theta\} ds + \\ &\quad \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t'') W_\phi^P(t'')\theta \leq \\ \exp(-\rho t') \int_{t''}^{t'} \{w_0^P(s)\theta + w_1^P(s)\theta\} ds &+ \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t'') W_\phi^P(t'')\theta = \\ \exp(-\rho t') (W_\phi^P(t'') - W_\phi^P(t'))\theta &+ \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t'') W_\phi^P(t'')\theta = \\ &-(\exp(-\rho t'') - \exp(-\rho t')) W_\phi^P(t'')\theta < 0, \end{aligned}$$

where the last inequality uses  $t'' < t'$  and  $W_\phi^P(\cdot) > 0$ . It immediately follows for all such  $t'$ , we must have  $f_P(t') = 0$ , contradicting Claim 4.  $\square$

**Claim 6.** *The agent's strategy has no mass points.*

**Proof of Claim 6.** Suppose the agent's strategy has a mass point at  $t \in [\tau_M, \tau_P]$ . Because distribution of the arrival time of real news has no mass points, an arrival of type-1 news at  $t$  has no effect on the principal's about the payoff of the risky arm, i.e.  $\mu_1(t) = \mu$ . It follows that  $a(t) = 0$ , which contradicts Claim 5.  $\square$

**Claim 7.** *The principal's stopping strategy has no mass points for  $t \in [\tau_M, \tau_P]$ .*



**Proof of Claim 7.** Consider  $\tau_M \leq t < t' < \tau_P$ . From Lemma 5, we have

$$u_A(t') - u_A(t) \leq (1 - F_P(t))\{v(t') - v(t)\} + \exp(-\rho t')W_\phi^A(t')(\mu - \beta)a(t') - \exp(-\rho t)W_\phi^A(t)(\mu - \beta)a(t).$$

Because  $\tau_M \leq t < t' < \tau_P$ , using Claim 3 and condition (4) we have  $u_A(t) = u_A(t')$ . From Claim 5, we have  $a(t) = a(t') = 1$ . Substituting, we have

$$\begin{aligned} (1 - F_P(t))\{v(t') - v(t)\} + \exp(-\rho t')W_\phi^A(t')(\mu - \beta) - \exp(-\rho t)W_\phi^A(t)(\mu - \beta) &\geq 0 \Rightarrow \\ W_\phi^A(t') &\geq \exp(\rho(t' - t))W_\phi^A(t) - \exp(\rho t')(1 - F_P(t))\frac{v(t') - v(t)}{\mu - \beta} \\ 1 - F_P(t') &\geq \underbrace{(1 - F_P(t))\{\exp(\rho(t' - t))\frac{1 - G(t)}{1 - G(t')}\}}_A - \underbrace{\exp(\rho t')\frac{v(t') - v(t)}{\mu - \beta}}_B. \end{aligned}$$

Consider the limit of both sides as  $t' \rightarrow t^+$ . Because  $F_P(\cdot)$  is right-continuous, the limit on the left hand side exists. Next, note that as  $t' \rightarrow t^+$  the limit of term  $A$  is 1, and the limit of term  $B$  is 0, which follows from the continuity of function  $v(\cdot)$ . Therefore, the limit of the right hand side exists, and is equal to  $1 - F_P(t)$ . Combining these observations, we have  $\lim_{t' \rightarrow t^+} (1 - F_P(t')) \geq 1 - F_P(t)$ . Simplifying, we have  $\lim_{t' \rightarrow t^+} F_P(t') \leq F_P(t)$ . Because  $F_P(\cdot)$  is increasing, we also have  $\lim_{t' \rightarrow t^+} F_P(t') \geq F_P(t)$ . Therefore  $\lim_{t' \rightarrow t^+} F_P(t') = F_P(t)$ . Thus,  $F_P(\cdot)$  is continuous at  $t$ .  $\square$

## 6.4. Proof of Proposition 2 and Corollary 1

Corollary 2 is proved in the Supplemental Appendix.

**Proof of Proposition 2.** I first show that the distributions presented in the proposition are the only ones that could constitute an equilibrium with beneficial search. That is, if an equilibrium with beneficial search exists, then it must take the form presented in the statement of the proposition. I then show that these distributions, together with  $a(\cdot) = 1$ , are an equilibrium.

*Step 1.* If an equilibrium with beneficial search exists, then the distributions must be the ones from the statement Proposition 2.

*Principal Strategy.* From Proposition 1, the support of the principal's mixed strategy is  $[\tau_M, \tau_P]$  for some  $\tau_M > 0$ . It follows that the set of  $t$  such that  $f_P(t) > 0$  is dense in  $(\tau_M, \tau_P)$ . From Lemma 3,  $u_P(\cdot)$  is continuous in equilibrium. From (6)  $u_P(t) = u_P^*$  for all  $t$  such that  $f_P(t) > 0$ .

It follows that  $u_P(t) = u_P^*$  for all  $t \in [\tau_M, \tau_P]$ , where  $u_P^*$  is the principal's equilibrium payoff. Furthermore, according to Proposition 1,  $a(\cdot) = 1$ . Therefore, for all  $t \in [\tau_M, \tau_P]$  we have

$$\begin{aligned} u_P(t) - u_P(\tau_M) &= \int_{\tau_M}^t \exp(-\rho s) \{w_0^P(s)\theta + w_1^P(s)\mu_1(s)\} + \exp(-\rho t)W_\phi^P(t)\theta - \exp(-\rho\tau_M)W_\phi^P(\tau_M)\theta \\ &= \int_{\tau_M}^t \exp(-\rho s) \{((1-\mu)\theta + \mu)g(s)(1-\sigma F_A(s)) + \mu\sigma f_A(s)(1-G(s))\} ds \\ &\quad + \exp(-\rho t)(1-G(t))(1-\sigma F_A(t))\theta - \exp(-\rho\tau_M)(1-G(\tau_M))\theta = 0. \end{aligned}$$

From Proposition 1  $F_A(\cdot)$  has no mass points. It follows that  $F_A(\tau_M) = 0$  and  $F_A(\cdot)$  is absolutely continuous. By implication, for  $t \in (\tau_M, \tau_P)$  we have

$$\begin{aligned} &\exp(-\rho t)(1-G(t))(1-\sigma F_A(t))\theta - \exp(-\rho\tau_M)(1-G(\tau_M))\theta = \\ &= - \int_{\tau_M}^t \exp(-\rho s) \{g(s)(1-\sigma F_A(s)) + \sigma f_A(s)(1-G(s)) + \rho(1-G(s))(1-\sigma F_A(s))\} \theta ds. \end{aligned}$$

To see this, note that the integrand is equal to the derivative of  $\exp(-\rho t)(1-G(t))(1-\sigma F_A(t))\theta$  wherever the derivative exists, and the equality above follows from absolute continuity. Substituting the previous equality and simplifying,

$$\begin{aligned} &\int_{\tau_M}^t \exp(-\rho s) \{\mu(1-\theta)g(s)(1-\sigma F_A(s)) + (\mu-\theta)\sigma f_A(s)(1-G(s)) - \rho\theta(1-\sigma F_A(s))(1-G(s))\} ds = \\ &\int_{\tau_M}^t \exp(-\rho s) (1-\sigma F_A(s))(1-G(s)) (\theta-\mu) \left\{ \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta-\mu} - H_A(s) \right\} ds = 0, \end{aligned}$$

where  $H_A(s) \equiv \sigma f_A(s)/(1-\sigma F_A(s))$ . Because the first four terms in the integrand are strictly positive and the equation holds for all  $t \in (\tau_M, \tau_P)$ , we have

$$\int_{\tau_M}^t \left\{ \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta-\mu} - H_A(s) \right\} ds = 0.$$

Note that  $\ln(1-\sigma F_A(t))$  is absolutely continuous on  $[\tau_M, \tau_P]$  and its derivative is  $-H_A(t)$  (whenever  $F_A(\cdot)$  is differentiable). By implication

$$\ln(1-\sigma F_A(t)) - \ln(1-\sigma F_A(\tau_M)) = - \int_{\tau_M}^t \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta-\mu} ds.$$

Recall that  $F_A(\cdot)$  has no mass points, and the bottom of the support is  $\tau_M$ . It follows that

$F_A(\tau_M) = 0$ , and hence

$$F_A(t) = \frac{1}{\sigma} \left( 1 - \exp \left\{ - \int_{\tau_M}^t \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds \right\} \right).$$

*Determining  $\tau_M$ .* From Proposition 1, in an equilibrium with valuable search (i) the agent's strategy has no atoms and (ii), the the top of the support is  $\tau_P$ . It follows that

$$\frac{1}{\sigma} \left( 1 - \exp \left\{ - \int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds \right\} \right) = 1 \Leftrightarrow 1 - \exp \left\{ - \int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds \right\} = \sigma.$$

Next, note that the LHS is a strictly decreasing function of  $\tau_M$  and its value is 0 at  $\tau_M = \tau_P$ . Therefore, the previous equation is satisfied by at most one  $\tau_M$ . Furthermore, from Proposition 1, in a beneficial search equilibrium  $\tau_M > 0$ . Such a solution exists if and only if

$$\exp \left\{ - \int_0^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\mu - \theta} ds \right\} < 1 - \sigma \Leftrightarrow \sigma < \bar{\sigma}.$$

*Agent Strategy.* As a preliminary step, I show that  $u_A(\cdot)$  is continuous on  $[\tau_M, \tau_P)$ . Consider  $t'' \leq t' < \tau_P$ . From Proposition 1, we have that  $a(\cdot) = 1$ . Therefore, for such  $t'', t'$  we have

$$\begin{aligned} u_A(t') - u_A(t'') &= \int_{t''}^{t'} \exp(-\rho s) \{ (w_0^A(s) + w_S^A(s))\beta + w_1^A(s) \} ds \\ &\quad + \exp(-\rho t') W_\phi^A(t') \mu - \exp(-\rho t) W_\phi^A(t) \mu. \end{aligned}$$

Furthermore, according to Proposition 1,  $F_P(\cdot)$  has no atoms in  $[\tau_M, \tau_P)$ . Therefore,  $w_0^A(\cdot), w_1^A(\cdot), w_S^A(\cdot), W_\phi^A(\cdot)$  are all continuous. Therefore  $u_A(\cdot)$  is continuous on  $[\tau_M, \tau_P)$ .

From Proposition, the support of the agent's mixed strategy is  $[\tau_M, \tau_P]$ ; that is, the set of  $t$  such that  $f_A(t) > 0$  is dense in  $(\tau_M, \tau_P)$ . As shown above  $u_A(\cdot)$  is continuous on  $[\tau_M, \tau_P)$ . From (6)  $u_A(t) = u_A^*$  for all  $t$  such that  $f_A(t) > 0$ . It follows that  $u_A(t) = u_A^*$  for all  $t \in [\tau_M, \tau_P)$ , where  $u_A^*$  is the agent's equilibrium payoff. Furthermore, Proposition 1,  $a(t) = 1$  for all  $t$ . By implication, for  $\tau_M < t < \tau_P$  we have

$$\begin{aligned} u_A(t) - u_A(\tau_M) &= \\ &\int_t^{\tau_M} \exp(-\rho s) (w_0^A(s) + w_S^A(s))\beta + w_1^A(s) ds + \exp(-\rho t) W_\phi^A(t) \mu - \exp(-\rho \tau_M) W_\phi^A(\tau_M) \mu \\ &\int_t^{\tau_M} \exp(-\rho s) \{ (\mu + (1-\mu)\beta)g(s)(1 - F_P(s)) + f_P(s)(1 - G(s))\beta \} ds + \end{aligned}$$

$$\exp(-\rho t)(1 - G(t))(1 - F_P(t))\mu - \exp(-\rho\tau_M)(1 - G(\tau_M))\mu = 0.$$

Because  $F_P(\cdot)$  has no mass points in  $[\tau_M, \tau_P)$ , (1)  $F_P(\tau_M) = 0$ , and (2)  $F_P(\cdot)$  is absolutely continuous on this interval. By implication, for  $t \in [\tau_M, \tau_P)$  we have

$$\begin{aligned} & \exp(-\rho t)(1 - G(t))(1 - F_P(t))\mu - \exp(-\rho\tau_M)(1 - G(\tau_M))\mu = \\ & - \int_t^{\tau_M} \exp(-\rho s)\{g(s)(1 - F_P(s)) + f_P(s)(1 - G(s)) + \rho(1 - G(s))(1 - F_P(s))\}\mu ds. \end{aligned}$$

To see this, note that the integrand is equal to the derivative of  $\exp(-\rho t)(1 - G(t))(1 - F_P(t))\mu$  wherever the derivative exists, and the equality above follows from absolute continuity.

Combining and simplifying, we have

$$\begin{aligned} & \int_{\tau_M}^t \exp(-\rho s)\{\beta(1 - \mu)g(s)(1 - F_P(s)) - f_P(s)(1 - G(s))(\mu - \beta) - \rho\mu(1 - G(s))(1 - F_P(s))\}ds = \\ & \int_{\tau_M}^t \exp(-\rho s)(1 - G(s))(1 - F_P(s))(\mu - \beta)\left\{\frac{\beta(1 - \mu)H_R(s) - \rho\mu}{\mu - \beta} - H_P(s)\right\}ds = 0. \end{aligned}$$

Where  $H_P(s) \equiv f_A(s)/(1 - F_A(s))$ . Because the first four terms in the integrand are strictly positive and the equation holds for all  $t \in (\tau_M, \tau_P)$ , we have

$$\int_{\tau_M}^t \frac{\beta(1 - \mu)H_R(s) - \rho\mu}{\mu - \beta} - H_P(s)ds = 0.$$

Note that  $\ln(1 - F_P(t))$  is absolutely continuous on  $[\tau_M, \tau_P)$  and its derivative is  $-H_P(t)$  (whenever  $F_A(\cdot)$  is differentiable). By implication

$$\ln(1 - F_P(t)) - \ln(1 - F_P(\tau_M)) = - \int_{\tau_M}^t \frac{\beta(1 - \mu)H_R(s) - \rho\mu}{\mu - \beta} ds.$$

Recall that  $F_P(\cdot)$  has no mass points on  $[\tau_M, \tau_P)$ , and the bottom of the support is  $\tau_M$ . It follows that  $F_P(\tau_M) = 0$ , and hence

$$F_P(t) = 1 - \exp\left\{- \int_{\tau_M}^t \frac{\beta(1 - \mu)H_R(s) - \rho\mu}{\mu - \beta} ds\right\},$$

for  $t \in [\tau_M, \tau_P)$ . Proposition 1 requires that the support of the principal's mixed strategy is  $[\tau_M, \tau_P]$ , and a mass point on  $\tau_P$  is allowed. In particular, because  $F_P(\cdot)$  has a mass point at  $\tau_P$ , agent's payoff  $u_A(\cdot)$  has a downward jump at  $\tau_P$ . It is straightforward to verify that

$f_A(\tau_P) = 0$ , and therefore the agent's strategy does not violate (4).

*Sufficiency.* The preceding calculations establish indifference for the principal in  $[\tau_M, \tau_P]$  and for the agent in  $[\tau_M, \tau_P]$ . To establish sufficiency, it remains to show (1) the proposed distributions are valid CDFs, (2) for  $t$  at which  $f_i(t) = 0$ , each player's payoff does not exceed  $u_i^*$ , and (3) under the proposed distribution for  $F_A(\cdot)$ , we have  $a(\cdot) = 1$  in equilibrium. (1) follows from elementary calculations. (2) is immediate. (3) is proved as part of Corollary 1, below.  $\square$

**Proof of Corollary 1.** Any type-1 arrival at time  $t \notin [\tau_M, \tau_P]$  reveals  $\omega = 1$ , because the likelihood of a fake is 0. Consider times  $t \in [\tau_M, \tau_P]$ . At such a time, the principal's posterior belief is given by (1),

$$\mu_1(t) = \frac{\mu g(t)(1 - \sigma F_A(t)) + \mu \sigma f_A(t)(1 - G(t))}{\mu g(t)(1 - \sigma F_A(t)) + \sigma f_A(t)(1 - G(t))} = \frac{\mu[H_R(t) + H_A(t)]}{\mu H_R(t) + H_A(t)},$$

where  $H_A(t) \equiv \sigma f_A(t)/(1 - \sigma F_A(t))$ . Using Proposition 2, it is straightforward to show that in an equilibrium with beneficial search,

$$H_A(t) = \frac{\mu(1 - \theta)H_R(s) - \rho\theta}{\theta - \mu}.$$

Routine simplification yields the expression in the proposition. Direct substitution establishes continuity at  $\tau_P$ . That the belief is increasing follows from the assumption that  $H_R(\cdot)$  is decreasing. Finally, it must be demonstrated that an  $\varepsilon > 0$  exists such that  $\mu_1(t) \geq \theta + \varepsilon$  for  $t \in [\tau_M, \tau_P]$ . Note that  $H_R(t) \geq \rho\theta/(\mu(1 - \theta))$  for  $t \in [\tau_M, \tau_P]$ , and hence,

$$H_R(t)\mu(1 - \mu) - \rho\theta \geq \rho\theta\left(\frac{1 - \mu}{1 - \theta} - 1\right) > 0.$$

Let

$$\varepsilon \equiv (\theta - \mu) \frac{\rho\theta}{H_R(\tau_M)\mu(1 - \mu) - \rho\theta}.$$

As shown above,  $\varepsilon > 0$ . Furthermore, for  $t \in [\tau_M, \tau_P]$ , we have  $\mu_1(t) > \mu_1(\tau_M) = \theta + \varepsilon$ .  $\square$

## 6.5. Proof of Proposition 3

As argued in the text, an equilibrium in which the principal does not benefit from search exists if and only if there exists an agent faking strategy  $F_A(\cdot)$ , such that  $u_P(t) \leq \theta$  for all  $t \geq 0$ , with  $a(\cdot)$  chosen optimally (as in (5)). That is such an equilibrium exists if and only if there exists

$F_A(\cdot)$ , such that

$$u_P(t) = \int_0^t \exp(-\rho s) \{w_0^P(s)\theta + w_1^P(s)((1-a(s))\theta + a(s)\mu_1(s))\} ds + \exp(-\rho t)W_\phi^P(t)\theta \leq \theta, \quad (7)$$

for all  $t \geq 0$ , where  $a(t) = \mathcal{I}(\mu_1(t) \geq \theta)$ .

First, note that attention can be restricted to  $t \leq \tau_P$ .

*Step 1.* An equilibrium without beneficial search exists if and only if there exists a faking strategy  $F_A(\cdot)$  such that  $u_P(t) \leq \theta$  for  $t \in [0, \tau_P]$ , with  $a(\cdot)$  chosen optimally. The necessary and sufficient condition (7) must hold for all  $t \geq 0$ , and therefore must hold for  $t \in [0, \tau_P]$ . The proposed condition is therefore necessary. To see that it is sufficient, Recall Lemma 2 (ii), which establishes that if (5), then  $u(t) < u(\tau_P)$  for  $t > \tau_P$ . Therefore, if  $u(\tau_P) \leq \theta$ , then for  $t > \tau_P$  we have  $u(t) < u(\tau_P) \leq \theta$ .

*Step 2.* An equilibrium without beneficial search exists only if  $\sigma > 1 - \frac{\theta}{u_P^{FB}(\tau_P)}$ . Note that constraining the principal to select  $a(\cdot) = 1$ , weakly reduces her payoff, i.e.,

$$u_P(t) \geq \int_0^t \exp(-\rho s) \{w_0^P(s)\theta + w_1^P(s)\mu_1(s)\} ds + \exp(-\rho t)W_\phi^P(t)\theta.$$

Substituting into the right hand side, we have

$$\begin{aligned} u_P(t) &\geq \int_0^t \exp(-\rho s) \{((1-\mu)\theta + \mu)g(s)(1-\sigma F_A(s)) + \mu\sigma f_A(s)(1-G(s))\} ds \\ &\quad + \exp(-\rho t)(1-G(t))(1-\sigma F_A(t))\theta \\ &\geq (1-\sigma) \left\{ \int_0^t \exp(-\rho s) ((1-\mu)\theta + \mu)g(s) ds + \exp(-\rho t)(1-G(t))\theta \right\} = (1-\sigma)u_P^{FB}(t). \end{aligned}$$

If an equilibrium without beneficial search exists, then

$$\theta \geq u_P(t) \geq (1-\sigma)u_P^{FB}(t)$$

for all  $t \in [0, \tau_P]$ . From Lemma 1, the RHS is largest at  $\tau_P$  and  $u_P^{FB}(\tau_P) > \theta$ . Therefore, such an equilibrium exists only if  $\theta \geq (1-\sigma)u_P^{FB}(\tau_P)$ . Step 2 follows by rearranging.

*Step 3.* If an equilibrium without beneficial search exists at  $\sigma$ , then an equilibrium without beneficial search exists for all  $\sigma' > \sigma$ . Suppose that an equilibrium without beneficial search exists at  $\sigma$ , i.e, there exists some  $F_A(\cdot)$  such that  $u_P(t) \leq \theta$  for  $t \in [0, \tau_P]$ . Suppose  $\sigma' > \sigma$ . Consider

an alternative faking strategy  $\widehat{F}_A(t) = \frac{\sigma}{\sigma'} F_A(t)$  if  $t \leq \tau_P$  and  $\widehat{F}_A(t) = F_A(\tau_P)$  for  $t \geq \tau_P$ . That is, in the interval  $[0, \tau_P]$ , the alternative strategy is the original strategy scaled down by factor  $\sigma/\sigma' < 1$ , and any remaining mass is allocated to  $t = \infty$ . Note that this is a valid strategy. Next, note that for  $t \leq \tau_P$ , we have  $\sigma' \widehat{F}_A(t) = \sigma F_A(t)$  and thus,  $\sigma' \widehat{f}_A(t) = \sigma f_A(t)$ . By implication  $\mu_1(t)$ ,  $w_0^P(t)$ ,  $w_1^P(t)$ ,  $W_\phi^P(t)$  are identical for the original strategy at  $\sigma$ , and for the alternative strategy at  $\sigma'$ . Thus, the principal's payoff functions are also identical, and weakly bounded by  $\theta$ .

*Step 4. An equilibrium without beneficial search exists if  $\sigma > \bar{\sigma}$ .* Consider continuous faking strategy

$$F_A(t) = \frac{1}{\sigma} \left( 1 - \exp\left\{ - \int_0^t \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds \right\} \right),$$

supported on  $[0, \tau_X]$ . It is readily verified that the proposed CDF is increasing and that for  $\sigma > \bar{\sigma}$ , the proposed CDF attains 1 at some  $\tau_X > \tau_P$ . Furthermore, it is straightforward to show that  $u_P(t) = \theta$ , for all  $t \leq \tau_X$ . Thus, the proposed agent strategy, coupled with immediate stopping constitute an equilibrium without beneficial search.

Step 4 tells us that the set of  $\sigma$  at which an equilibrium without beneficial search exists is an interval,  $[\underline{\sigma}, 1]$ . Step 2 tells us that  $\underline{\sigma} < 1$ . Step 4 tells us that  $\underline{\sigma} \leq \bar{\sigma}$ .

## 6.6. Proofs for Section 4

**Proof of Proposition 4.** *Step 1. Commitment to naive search delivers principal a higher payoff than the unique equilibrium if  $\sigma < \bar{\sigma}$ .* If  $\sigma < \bar{\sigma}$ , then the unique equilibrium without commitment has beneficial search. Consider the payoff difference as a function of  $\sigma$ ,

$$\Delta(\sigma) = u_P^N(\sigma) - u_P^{FB}(\tau_M(\sigma)).$$

Note that for  $\sigma = 0$ , we have  $u_P^N(0) = u_P^{FB}(\tau_M(0)) = u_P^{FB}(\tau_P)$ . Thus,  $\Delta(0) = 0$ . Furthermore,

$$\frac{du_P^N(\sigma)}{d\sigma} = -\exp(-\rho\tau_P)(1 - G(\tau_P))(\theta - \mu).$$

Moreover,

$$\frac{du_P^{FB}(\tau_M(\sigma))}{d\sigma} = \frac{du_P^{FB}(\tau_M)}{d\tau_M} \frac{d\tau_M}{d\sigma}.$$

From proof of Lemma 1, we have

$$\frac{du_P^{FB}(\tau_M)}{d\tau_M} = \exp(-\rho\tau_M)(1 - G(\tau_M))\mu(1 - \theta)\left\{H_R(\tau_M) - \frac{\rho\theta}{\mu(1 - \theta)}\right\}.$$

From Proposition 2, we have the following implicit function relating  $\tau_M$  and  $\sigma$ .

$$\exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1 - \theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = 1 - \sigma.$$

Differentiating, with respect to  $\sigma$ , we have

$$\frac{d\tau_M}{d\sigma} \left\{ \frac{\mu(1 - \theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} \exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1 - \theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = -1.$$

Simplifying, we have

$$\begin{aligned} \frac{d\tau_M}{d\sigma} \left\{ \frac{\mu(1 - \theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} (1 - \sigma) &= -1 \Rightarrow \\ \frac{d\tau_M}{d\sigma} &= \left\{ \frac{\mu(1 - \theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} (1 - \sigma) = -\frac{\theta - \mu}{1 - \sigma} \frac{1}{\mu(1 - \theta)(H_R(\tau_M) - \frac{\rho\theta}{\mu(1 - \theta)})}. \end{aligned}$$

Combining these expressions,

$$\begin{aligned} \frac{du_P^{FB}(\tau_M(\sigma))}{d\sigma} &= \\ &\left( \exp(-\rho\tau_M)(1 - G(\tau_M))\mu(1 - \theta)\left\{H_R(\tau_M) - \frac{\rho\theta}{\mu(1 - \theta)}\right\} \right) \left( -\frac{\theta - \mu}{1 - \sigma} \frac{1}{\mu(1 - \theta)(H_R(\tau_M) - \frac{\rho\theta}{\mu(1 - \theta)})} \right) = \\ &-\frac{\theta - \mu}{1 - \sigma} \exp(-\rho\tau_M)(1 - G(\tau_M)). \end{aligned}$$

It follows that

$$\frac{d\Delta}{d\sigma} = -\exp(-\rho\tau_P)(1 - G(\tau_P))(\theta - \mu) + \frac{\theta - \mu}{1 - \sigma} \exp(-\rho\tau_M)(1 - G(\tau_M)).$$

Recalling that (1)  $\tau_M < \tau_P$ , (2)  $\exp(-\rho\cdot)(1 - G(\cdot))$  is a strictly decreasing function, we have

$$\frac{d\Delta}{d\sigma} > (\theta - \mu) \exp(-\rho\tau_M)(1 - G(\tau_M)) \left( \frac{1}{1 - \sigma} - 1 \right) > 0.$$

Hence,  $\Delta(0) = 0$  and  $\Delta(\cdot)$  is increasing in  $\sigma$  for  $\sigma < \bar{\sigma}$ . Therefore for all  $\sigma \in (0, \bar{\sigma}]$ , we have  $\Delta(\sigma) > 0$ .



Step 2. A  $\tilde{\sigma} \leq 1$  exists such that commitment to naive search delivers principal a higher payoff than any equilibrium without commitment if  $\bar{\sigma} < \sigma < \tilde{\sigma}$ . Without commitment, the principal's payoff in any equilibrium is  $u_P^{FB}(0) = \theta$  for  $\sigma > \bar{\sigma}$ . Thus,

$$\Delta(\sigma) = u_P^N(\sigma) - \theta.$$

Note that  $\Delta(\cdot)$  is continuous at  $\bar{\sigma}$ , because  $\tau_M(\sigma) \rightarrow 0$  as  $\sigma \rightarrow \bar{\sigma}^-$ , and  $u_P^{FB}(0) = \theta$ . It follows from Step 1 that  $\Delta(\bar{\sigma}) > \theta$ . Furthermore,  $\Delta(\sigma)$  is linear for  $\sigma > \bar{\sigma}$ . The result follows.  $\square$

**Proof of Proposition 6.** To prove the first claim of the proposition it is sufficient to show that  $U'(\theta|\theta) < 0$ , where  $U'(\cdot|\theta)$  is the derivative of  $U(\cdot|\cdot)$  with respect to  $\theta_I$ . In this case, the principal's payoff of delegating  $U(\theta_I|\theta)$  is decreasing in  $\theta_I$  at  $\theta_I = \theta$ . Thus, for an interval of  $\theta_I$  just below  $\theta$ , we have  $U(\theta_I|\theta) > U(\theta|\theta)$ .

To simplify the notation of the proof, let  $u_t^{FB}(\cdot) \equiv du_P^{FB}(t)/dt$  and  $u'_I(t|\theta_I, \theta) = \partial^2 u_P(t|\theta_I, \theta)/\partial t \partial \theta_I$ , i.e., the derivative of  $u'_P(t|\theta_I, \theta)$  with respect to  $\theta_I$ .

Step 1. Recover the expression in the text.

$$U(\theta_I|\theta) = \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} f_P(s|\theta_I) u_P(s|\theta_I, \theta) ds + (1 - F_P(\tau_P^-(\theta_I)|\theta_I)) u_P(\tau_P(\theta_I)|\theta_I, \theta)$$

Note that  $f_P(\cdot|\theta_I)$  is continuous, and  $u_P(\cdot|\theta_I, \theta)$  is continuous and differentiable on  $[\tau_M(\theta_I), \tau_P(\theta_I)]$ . Integrating by parts, we have

$$\begin{aligned} &= - \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} u_P(s|\theta_I, \theta) d[1 - F_P(s|\theta_I)] + (1 - F_P(\tau_P^-(\theta_I)|\theta_I)) u_P(\tau_P(\theta_I)|\theta_I, \theta) \\ &= (1 - F_P(\tau_M(\theta_I)|\theta_I)) u_P(\tau_M(\theta_I)|\theta_I) - (1 - F_P(\tau_P^-(\theta_I)|\theta_I)) u_P(\tau_P(\theta_I)|\theta_I, \theta) \\ &\quad + \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} (1 - F_P(s|\theta_I)) u'_P(s|\theta_I, \theta) ds + (1 - F_P(\tau_P^-(\theta_I)|\theta_I)) u_P(\tau_P(\theta_I)|\theta_I, \theta) \\ &= \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} (1 - F_P(s|\theta_I)) u'_P(s|\theta_I, \theta) ds + (1 - F_P(\tau_M(\theta_I)|\theta_I)) u_P(\tau_M(\theta_I)|\theta_I). \end{aligned}$$

From Proposition 2, we have  $(1 - F_P(\tau_M(\theta_I)|\theta_I)) = 1$  and  $u_P(\tau_M(\theta_I)|\theta_I) = u_P^{FB}(\tau_M(\theta_I))$ .

Step 2. Differentiate  $U(\theta_I|\theta)$  with respect to  $\theta_I$ .

$$\begin{aligned} U'(\theta_I|\theta) &= (1 - F_P(\tau_P^-(\theta_I)|\theta_I)) u'_P(\tau_P^-(\theta_I)|\theta_I, \theta) \tau'_P(\theta_I) - (1 - F_P(\tau_M(\theta_I)|\theta_I)) u'_P(\tau_M(\theta_I)|\theta_I, \theta) \tau'_M(\theta_I) \\ &\quad + \int_{\tau_M(\theta_I)}^{\tau_P(\theta_I)} (1 - F_P(s|\theta_I)) u'_I(s|\theta_I, \theta) - \frac{dF_P(s|\theta_I)}{d\theta_I} u'_P(s|\theta_I, \theta) ds + u_t^{FB}(\tau_M(\theta_I)) \tau'_M(\theta_I). \end{aligned}$$

Step 3. Evaluate at  $\theta_I = \theta$ .

Note that at  $\theta_I = \theta$ , we have  $u'_P(\cdot|\theta, \theta) = 0$  on  $[\tau_M, \tau_P]$ . That is, the principal's equilibrium payoff is constant inside the support, and furthermore, all equilibrium objects are continuous in  $\theta_I$ . Therefore,

$$U'(\theta|\theta) = \int_{\tau_M}^{\tau_P} (1 - F_P(s))u'_I(s|\theta, \theta)ds + u_t^{FB}(\tau_M)\tau'_M(\theta).$$

Step 4. Derive an expression for  $u'_I(s|\theta, \theta)$ .

Substituting the definitions in the text,

$$\begin{aligned} u_P(t|\theta_I, \theta) &= u_P(\tau_M(\theta_I)|\theta_I, \theta) \\ &+ \int_{\tau_M(\theta_I)}^t \exp(-\rho s) \{((1 - \mu)\theta + \mu)g(s)(1 - \sigma F_A(s|\theta_I)) + \mu\sigma f_A(s|\theta_I)(1 - G(s))\} ds \\ &+ \exp(-\rho t)(1 - G(t))(1 - \sigma F_A(t|\theta_I))\theta - \exp(-\rho\tau_M(\theta_I))(1 - G(\tau_M(\theta_I)))\theta. \end{aligned}$$

From Proposition 2  $F_A(\cdot)$  has no mass points. It follows that  $F_A(\tau_M) = 0$  and  $F_A(\cdot)$  is absolutely continuous. By implication, for  $t \in (\tau_M(\theta_I), \tau_P(\theta_I))$  we have

$$\begin{aligned} &\exp(-\rho t)(1 - G(t))(1 - \sigma F_A(t|\theta_I))\theta - \exp(-\rho\tau_M(\theta_I))(1 - G(\tau_M(\theta_I)))\theta = \\ &- \int_{\tau_M(\theta_I)}^t \exp(-\rho s) \{g(s)(1 - \sigma F_A(s|\theta_I)) + \sigma f_A(s|\theta_I)(1 - G(s)) + \rho(1 - G(s))(1 - \sigma F_A(s|\theta_I))\} \theta ds. \end{aligned}$$

To see this, note that the integrand is equal to the derivative in  $t$  of  $\exp(-\rho t)(1 - G(t))(1 - \sigma F_A(t|\theta_I))\theta$  wherever the derivative exists, and the equality above follows from absolute continuity. Substituting the previous equality and simplifying,

$$\begin{aligned} u_P(t|\theta_I, \theta) &= u_P(\tau_M(\theta_I)|\theta_I, \theta) + \\ &\int_{\tau_M(\theta_I)}^t \exp(-\rho s) \{ \mu(1 - \theta)g(s)(1 - \sigma F_A(s|\theta_I)) + (\mu - \theta)\sigma f_A(s|\theta_I)(1 - G(s)) \\ &\quad - \rho\theta(1 - \sigma F_A(s|\theta_I))(1 - G(s)) \} ds = \\ &\int_{\tau_M(\theta_I)}^t \exp(-\rho s) (1 - \sigma F_A(s|\theta_I))(1 - G(s)) (\theta - \mu) \left\{ \frac{\mu(1 - \theta)H_R(s) - \rho\theta}{\theta - \mu} - H_A(s|\theta_I) \right\} ds, \end{aligned}$$

where  $H_A(s|\theta_I) = \sigma f_A(s|\theta_I)/(1 - \sigma F_A(s|\theta_I))$ . From Proposition 2, it is readily verified that in

equilibrium,

$$H_A(s|\theta_I) = \frac{\mu(1 - \theta_I)H_R(s) - \rho\theta_I}{\theta_I - \mu}.$$

It is also helpful to denote the derivative of  $H_A(s|\theta_I)$  with respect to  $\theta_I$ . Let

$$\begin{aligned} h_A(s|\theta_I) &\equiv \frac{dH_A(s|\theta_I)}{d\theta_I} = \frac{-(\mu H_R(s) + \rho)(\theta_I - \mu) - \mu(1 - \theta_I)H_R(s) + \rho\theta_I}{(\theta_I - \mu)^2} \\ &= -\frac{\mu(1 - \mu)}{(\theta_I - \mu)^2} \left( H_R(s) - \frac{\rho}{1 - \mu} \right). \end{aligned}$$

With this notation, we have

$$\begin{aligned} u_P(t|\theta_I, \theta) &= u_P(\tau_M(\theta_I)|\theta_I, \theta) + \\ &\int_{\tau_M(\theta_I)}^t \exp(-\rho s)(1 - \sigma F_A(s|\theta_I))(1 - G(s))(\theta - \mu)\{H_A(s|\theta) - H_A(s|\theta_I)\} ds. \end{aligned}$$

Differentiating with respect to  $t$ , we have

$$u'_P(t|\theta_I, \theta) = \exp(-\rho t)(1 - \sigma F_A(t|\theta_I))(1 - G(t))(\theta - \mu)\{H_A(t|\theta) - H_A(t|\theta_I)\}.$$

Differentiating with respect to  $\theta_I$  and then evaluating at  $\theta_I = \theta$ , we have

$$u'_I(t|\theta, \theta) = -\exp(-\rho t)(1 - \sigma F_A(t))(1 - G(t))(\theta - \mu)h_A(t|\theta).$$

Note that the term involving differentiation of  $1 - \sigma F_A(t|\theta_I)$  vanishes at  $\theta_I = \theta$ , because  $H_A(t|\theta) - H_A(t|\theta) = 0$ .

*Step 5. Derive an expression for  $u_t^{FB}(\tau_M)\tau'_M(\theta)$ .*

Differentiating  $u_P^{FB}(\cdot)$  with respect to  $t$  and substituting  $\tau_M$  gives

$$\begin{aligned} u_t^{FB}(\tau_M) &= \frac{du_P^{FB}(\tau_M)}{dt} = \exp(-\rho\tau_M)(1 - G(\tau_M))\mu(1 - \theta)\left\{H_R(\tau_M) - \frac{\rho\theta}{\mu(1 - \theta)}\right\} \\ &= \exp(-\rho\tau_M)(1 - G(\tau_M))\{\mu(1 - \theta)H_R(\tau_M) - \rho\theta\} \\ &= \exp(-\rho\tau_M)(1 - G(\tau_M))(\theta - \mu)H_A(\tau_M|\theta) \end{aligned}$$

(see also the proof of Lemma 1).

To find  $\tau'_M(\theta)$  recall that, Proposition 2 gives an implicit function relating  $\tau_M$ ,  $\tau_P$ , and  $\sigma$ ,

$$\exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = 1 - \sigma \iff \exp\left\{-\int_{\tau_M}^{\tau_P} H_A(s|\theta) ds\right\} = 1 - \sigma.$$

Differentiating with respect to  $\theta$ , we have

$$\exp\left\{-\int_{\tau_M}^{\tau_P} H_A(s|\theta) ds\right\} \left( H_A(\tau_M|\theta)\tau'_M(\theta) - H_A(\tau_P|\theta)\tau'_P(\theta) - \int_{\tau_M(\theta)}^{\tau_P(\theta)} h_A(s|\theta) ds \right) = 0.$$

Note that  $H_R(\tau_P) = \rho\theta/(\mu(1-\theta))$ , which implies  $H_A(\tau_P|\theta) = 0$ . By implication

$$H_A(\tau_M|\theta)\tau'_M(\theta) - \int_{\tau_M}^{\tau_P} h_A(s|\theta) ds = 0 \Rightarrow \tau'_M(\theta) = \frac{\int_{\tau_M}^{\tau_P} h_A(s|\theta) ds}{H_A(\tau_M|\theta)}.$$

Combining these calculations, we have

$$u_t^{FB}(\tau_M)\tau'_M(\theta) = \exp(-\rho\tau_M)(1 - G(\tau_M))(\theta - \mu) \cdot \int_{\tau_M}^{\tau_P} h_A(s|\theta) ds.$$

*Step 6. Show that  $U'(\theta|\theta) < 0$ .*

From Step 3,

$$U'(\theta|\theta) = \int_{\tau_M}^{\tau_P} (1 - F_P(s))u'_I(s|\theta, \theta) ds + u_t^{FB}(\tau_M)\tau'_M(\theta).$$

Substituting the expressions derived in Steps 4 and 5, we have

$$U'(\theta|\theta) = - \int_{\tau_M}^{\tau_P} \exp(-\rho s)(1 - \sigma F_A(s))(1 - G(s))(\theta - \mu)(1 - F_P(s))h_A(s|\theta) ds + \exp(-\rho\tau_M)(1 - G(\tau_M))(\theta - \mu) \cdot \int_{\tau_M}^{\tau_P} h_A(s|\theta) ds.$$

Let  $Q(s) \equiv \exp(-\rho s)(1 - \sigma F_A(s))(1 - G(s))(1 - F_P(s))(\theta - \mu)$ . Note that  $Q(s) > 0$  for all  $s \geq 0$  and  $Q(s)$  is strictly decreasing. With this notation,

$$U'(\theta|\theta) = \int_{\tau_M}^{\tau_P} \{Q(\tau_M) - Q(s)\} h_A(s|\theta) ds.$$

Because  $Q(\cdot)$  is strictly decreasing, the term in curly braces is strictly positive. To complete

the proof, it suffices to show that  $h_A(s|\theta) < 0$  for  $s \in [\tau_M, \tau_P]$ . As shown in Step 4,

$$h_A(s|\theta) = -\frac{\mu(1-\mu)}{(\theta-\mu)^2} \left( H_R(s) - \frac{\rho}{1-\mu} \right).$$

For  $s \leq \tau_P$ , we have  $H_R(s) \geq \frac{\rho\theta}{\mu(1-\theta)}$ . Hence, for such  $s$ ,

$$h_A(s|\theta) \leq -\frac{\mu(1-\mu)}{(\theta-\mu)^2} \left( \frac{\rho\theta}{\mu(1-\theta)} - \frac{\rho}{1-\mu} \right) = -\frac{\rho\theta(1-\mu) - \rho\mu(1-\theta)}{(\theta-\mu)^2(1-\theta)} = -\frac{\rho}{(\theta-\mu)(1-\theta)} < 0.$$

This completes the proof of the first claim.

To prove the second claim, suppose that the principal has delegated to an intermediary with  $\theta_I < \theta$ . It will be established that if  $\theta_I$  sufficiently close to  $\theta$ , then the principal does not want to overrule the agent's chosen action following a news arrival. Obviously, type-0 news reveals  $\omega = 0$  and both principal and intermediary prefer the safe action in this case. Recall from Proposition 2 that in equilibrium the intermediary selects the risky action with probability 1 following type-1 news at all times  $[0, \tau_P(\theta_I)]$ . Type-1 news arrivals before  $\tau_M(\theta_I)$  are real with probability 1, and both principal and intermediary prefer the risky action for such arrivals. What remains is to show that the principal prefers the risky action following type-1 news arrivals that occur in  $[\tau_M(\theta_I), \tau_P(\theta_I)]$ . Given an arrival at such a time, the principal's posterior belief is given by (1),

$$\mu_1(s|\theta_I) = \frac{\mu g(s)(1 - \sigma F_A(s|\theta_I)) + \mu \sigma f_A(s|\theta_I)(1 - G(s))}{\mu g(s)(1 - \sigma F_A(s|\theta_I)) + \sigma f_A(s|\theta_I)(1 - G(s))} = \frac{\mu[H_R(s) + H_A(s|\theta_I)]}{\mu H_R(s) + H_A(s|\theta_I)}.$$

Thus, the principal strictly prefers the risky action following a type-1 news arrival if

$$\frac{\mu[H_R(s) + H_A(s|\theta_I)]}{\mu H_R(s) + H_A(s|\theta_I)} > \theta \iff H_A(s|\theta_I) < \frac{\mu(1-\theta)H_R(s)}{\theta-\mu}.$$

It is readily verified that in the equilibrium of Proposition 2,

$$H_A(s|\theta_I) = \frac{\mu(1-\theta_I)H_R(s) - \rho\theta_I}{\theta_I - \mu}.$$

Thus, the principal strictly prefers risky following a type-1 news arrival if

$$\frac{\mu(1-\theta_I)H_R(s) - \rho\theta_I}{\theta_I - \mu} < \frac{\mu(1-\theta)H_R(s)}{\theta - \mu}.$$

Obviously, for  $\theta_I = \theta$ , left hand side is strictly smaller. Because left hand side is continuous,

this inequality holds if  $\theta_I$  and  $\theta$  are close. □

## 7. Supplemental Appendix

### 7.1. Proof of Lemmas used in Proposition 1

**Observation 1.** Note that

$$\int_0^t w_0^P(s) + w_1^P(s) ds + W_\phi^P(t) = 1.$$

Indeed, the integral is the probability that a type-0 or type-1 arrival occurs at times less than or equal to  $t$ , while  $W_\phi^P(t)$  is the probability that an arrival of news takes longer than  $t$ . Rearranging, for  $t_1 \leq t_2$  we have

$$\int_{t_1}^{t_2} w_0^P(s) + w_1^P(s) ds = W_\phi^P(t_1) - W_\phi^P(t_2).$$

Similarly, for the agent we have,

$$\int_{t_1}^{t_2} w_0^A(s) + w_1^A(s) + w_S^A(s) ds = W_\phi^A(t_1) - W_\phi^A(t_2).$$

**Proof of Lemma 2.** *Proof of Parts (i) and (ii):* By definition,

$$\begin{aligned} u_P(t') - u_P(t) = \\ \int_t^{t'} \exp(-\rho s) (w_0^P(s)\theta + w_1^P(s)(\theta + a(s)(\mu_1(s) - \theta))) ds + \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t) W_\phi^P(t)\theta. \end{aligned}$$

*Step 1: Derive a bound on the second term of the integrand.*

$$\begin{aligned} w_1^P(s)(\theta + a(s)(\mu_1(s) - \theta)) &= w_1^P(s)\theta(1 - a(s)) + w_1^P(s)\mu_1(s)a(s) = \\ &= \underbrace{(\mu g(s)(1 - \sigma F_A(s)) + \sigma f_A(s)(1 - G(s)))}_{w_1^P(s)} \theta(1 - a(s)) + \underbrace{(\mu g(s)(1 - \sigma F_A(s)) + \mu \sigma f_A(s)(1 - G(s)))}_{w_1^P(s)\mu_1(s)} a(s). \end{aligned}$$

Combining terms, we have

$$\begin{aligned} w_1^P(s)(\theta + a(s)(\mu_1(s) - \theta)) &= \\ \mu g(s)(1 - \sigma F_A(s))(\theta(1 - a(s)) + a(s)) &+ \sigma f_A(s)(1 - G(s))(\theta(1 - a(s)) + a(s)\mu). \end{aligned}$$

Recall that  $\mu < \theta < 1$ , and hence,

$$w_1^P(s)(\theta + a(s)(\mu_1(s) - \theta)) \leq \mu g(s)(1 - \sigma F_A(s)) + \sigma f_A(s)(1 - G(s))\theta. \quad (8)$$

*Step 2: Apply the bound derived in Step 1.*

Recall that  $w_0^P(s) = (1 - \mu)g(s)(1 - \sigma F_A(s))$ . Using this and (8), we have

$$u_P(t') - u_P(t) \leq \int_t^{t'} \exp(-\rho s) \{((1 - \mu)\theta + \mu)g(s)(1 - \sigma F_A(s)) + \sigma f_A(s)(1 - G(s))\theta\} ds + \quad (9)$$

$$\exp(-\rho t')(1 - G(t'))(1 - \sigma F_A(t'))\theta - \exp(-\rho t)(1 - G(t))(1 - \sigma F_A(t))\theta.$$

*Step 3: Derive an alternative expression for the RHS of inequality (9).*

Consider the following integral,

$$\int_t^\infty \sigma f_A(s) \min\{u_P^{FB}(s) - u_P^{FB}(t), u_P^{FB}(t') - u_P^{FB}(t)\} ds =$$

$$\underbrace{\int_t^{t'} \sigma f_A(s) (u_P^{FB}(s) - u_P^{FB}(t)) ds}_A + \underbrace{(1 - \sigma F_A(t'))(u_P^{FB}(t') - u_P^{FB}(t))}_B.$$

Focus on term A. Substituting, we have

$$\int_t^{t'} \sigma f_A(s) \left[ \int_t^s \exp(-\rho x) g(x) (\mu + (1 - \mu)\theta) dx + \exp(-\rho s) (1 - G(s))\theta - \exp(-\rho t) (1 - G(t))\theta \right] ds =$$

$$\int_t^{t'} \int_t^s \sigma f_A(s) \{ \exp(-\rho x) g(x) (\mu + (1 - \mu)\theta) dx ds + \int_t^{t'} \sigma f_A(s) \exp(-\rho s) (1 - G(s))\theta ds$$

$$- \sigma (F_A(t') - F_A(t)) (\exp(-\rho t) (1 - G(t))) \}.$$

Reversing the order of integration of the double integral, we have

$$\int_t^{t'} \sigma (F_A(t') - F_A(s)) \exp(-\rho s) g(s) (\mu + (1 - \mu)\theta) + \sigma f_A(s) \exp(-\rho s) (1 - G(s))\theta ds$$

$$- \sigma (F_A(t') - F_A(t)) (\exp(-\rho t) (1 - G(t))).$$

Now consider term B. We have

$$(1 - \sigma F_A(t')) \{ u_P^{FB}(t') - u_P^{FB}(t) \} =$$

$$(1 - \sigma F_A(t')) \left\{ \int_t^{t'} \exp(-\rho s) g(s) (\mu + (1 - \mu)\theta) ds + \exp(-\rho t') (1 - G(t'))\theta - \exp(-\rho t) (1 - G(t))\theta \right\}.$$

Adding terms A and B we have,

$$\begin{aligned} & \int_t^{t'} (1 - \sigma F_A(s)) \exp(-\rho s) g(s) (\mu + (1 - \mu)\theta) + \sigma f_A(s) \exp(-\rho s) (1 - G(s))\theta ds \\ & + (1 - \sigma F_A(t')) \exp(-\rho t') (1 - G(t)) (1 - \sigma F_A(t)) \exp(-\rho t) (1 - G(t)), \end{aligned}$$

which is the RHS of inequality (9). Therefore, inequality (9) can be written

$$u_P(t') - u_P(t) \leq \int_t^\infty \sigma f_A(s) \min\{u_P^{FB}(s) - u_P^{FB}(t), u_P^{FB}(t') - u_P^{FB}(t)\} ds. \quad (10)$$

*Proof of (i).* Notice that

$$\min\{u_P^{FB}(s) - u_P^{FB}(t), u_P^{FB}(t') - u_P^{FB}(t)\} \leq u_P^{FB}(t') - u_P^{FB}(t).$$

Substituting into (10), we have

$$u_P(t') - u_P(t) \leq \int_t^\infty \sigma f_A(s) (u_P^{FB}(t') - u_P^{FB}(t)) ds = (1 - \sigma F_A(t)) (u_P^{FB}(t') - u_P^{FB}(t)). \quad (11)$$

*Proof of (ii).* Suppose  $t > \tau_P$ . Applying inequality (11), we have,

$$u_P(t) - u_P(\tau_P) \leq (1 - \sigma F_A(\tau_P)) (u_P^{FB}(t) - u_P^{FB}(\tau_P)).$$

From Lemma 1 we have that the single peak of  $u^{FB}(\cdot)$  is at  $\tau_P$ . By implication, the right hand side is strictly negative for all  $\tau_P < t < \infty$ . Furthermore, the right hand side is also strictly decreasing in  $t$ , and hence  $\lim_{t \rightarrow \infty} u_P^{FB}(t) - u_P^{FB}(\tau_P) < 0$ .

*Proof of (iii).* By definition,

$$\begin{aligned} & u_P(t') - u_P(t) = \\ & \int_t^{t'} \exp(-\rho s) (w_0^P(s)\theta + w_1^P(s)(\theta + a(s)(\mu_1(s) - \theta)) ds + \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t) W_\phi^P(t)\theta. \end{aligned}$$

Given (5), we have  $\theta + a(s)(\mu_1(s) - \theta) \geq \theta$ , and hence,

$$u_P(t') - u_P(t) \geq \theta \int_t^{t'} \exp(-\rho s) (w_0^P(s) + w_1^P(s)) ds + \exp(-\rho t') W_\phi^P(t')\theta - \exp(-\rho t) W_\phi^P(t)\theta \geq$$



$$\exp(-\rho t')\theta \int_t^{t'} (w_0^P(s) + w_1^P(s))ds + \exp(-\rho t')W_\phi^P(t')\theta - \exp(-\rho t)W_\phi^P(t)\theta.$$

Using Observation 1, we have

$$u_P(t') - u_P(t) \geq \exp(-\rho t')\theta(W_\phi^P(t) - W_\phi^P(t')) + \exp(-\rho t')W_\phi^P(t')\theta - \exp(-\rho t)W_\phi^P(t)\theta = -[\exp(-\rho t) - \exp(-\rho t')]W_\phi^P(t)\theta.$$

*Proof of (iv).* If  $F_A(t) = F_A(t')$ , then  $f_A(s) = 0$  for almost all  $s \in (t, t')$ . It follows that for almost all  $s \in (t, t')$ ,  $\mu_1(s) = 1$ ,  $a(s) = 1$  (from (5)),  $w_0^P(s) = (1 - \mu)g(s)(1 - \sigma F_A(t))$ ,  $w_1^P(s) = \mu g(s)(1 - \sigma F_A(t))$ . Furthermore,  $W_\phi^P(t) = (1 - G(t))(1 - \sigma F_A(t))$  and  $W_\phi^P(t') = (1 - G(t'))(1 - \sigma F_A(t))$ . The result follows from routine substitution.  $\square$

**Proof of Lemma 3.** Using Parts (i) and (iii) of Lemma 2, for  $t' > t$

$$-(\exp(-\rho t) - (\exp(-\rho t')))W_\phi^P(t)\theta \leq u_P(t') - u_P(t) \leq (1 - \sigma F_A(t))\{u_P^{FB}(t') - u_P^{FB}(t)\}.$$

By continuity of  $u_P^{FB}(\cdot)$  and  $\exp(\cdot)$ , we have  $\lim_{t' \rightarrow t^+} u_P(t') = u_P(t)$ . Similarly, for  $t' < t$

$$-(\exp(-\rho t') - (\exp(-\rho t)))W_\phi^P(t')\theta \leq u_P(t) - u_P(t') \leq (1 - \sigma F_A(t'))\{u_P^{FB}(t) - u_P^{FB}(t')\},$$

Note that (1)  $W_\phi^P(t') \leq W_\phi^P(t)$ , (2)  $(\exp(-\rho t') - (\exp(-\rho t))) > 0$ , and (3)  $F_A(t') \geq 0$ . It follows that

$$-(\exp(-\rho t') - (\exp(-\rho t)))W_\phi^P(t)\theta \leq u_P(t) - u_P(t') \leq u_P^{FB}(t) - u_P^{FB}(t').$$

Continuity of  $u_P^{FB}(\cdot)$  and  $\exp(\cdot)$  imply  $\lim_{t' \rightarrow t^-} u_P(t') = u_P(t)$ .  $\square$

**Proof of Lemma 4.** From Lemma 2, we have  $u_P(t) < u_P(\tau_P)$  for  $\tau_P < t \leq \infty$ . From (6), we have  $f_P(t) = 0$  for such  $t$ . By implication,  $F_P(\tau_P) = 1$ .  $\square$

**Definition:** Let

$$v(t) \equiv \int_0^t \exp(-\rho x)g(x)(\mu + (1 - \mu)\beta)dx + \exp(-\rho t)(1 - G(t))\beta.$$

**Proof of Lemma 5.** Suppose  $t < t' \leq \tau_P$ . By definition,

$$u_A(t') - u_A(t) =$$

$$\int_t^{t'} \exp(-\rho s) \{(w_0^A(s) + w_S^A(s))\beta + w_1^A(s)((1 - a(s))\beta + a(s))\} ds + \exp(-\rho t') W_\phi^A(t') (\mu a(t') + (1 - a(t'))\beta) - \exp(-\rho t) W_\phi^A(t) (\mu a(t) + (1 - a(t))\beta).$$

*Proof of (i).* First, rearrange the previous expression:

$$u_A(t') - u_A(t) = \int_t^{t'} \exp(-\rho s) \{(w_0^A(s) + w_S^A(s))\beta + w_1^A(s)((1 - a(s))\beta + a(s))\} ds + \exp(-\rho t') W_\phi^A(t') \beta - \exp(-\rho t) W_\phi^A(t) \beta + \exp(-\rho t') W_\phi^A(t') (\mu - \beta) a(t') - \exp(-\rho t) W_\phi^A(t) (\mu - \beta) a(t).$$

Because  $a(s) \leq 1$  and  $\beta < \mu < 1$ , we have

$$u_A(t') - u_A(t) \leq \int_t^{t'} \exp(-\rho s) \{(w_0^A(s) + w_S^A(s))\beta + w_1^A(s)\} ds + \exp(-\rho t') W_\phi^A(t') \beta - \exp(-\rho t) W_\phi^A(t) \beta + \exp(-\rho t') W_\phi^A(t') (\mu - \beta) a(t') - \exp(-\rho t) W_\phi^A(t) (\mu - \beta) a(t). \quad (12)$$

*Step 1: Provide a different expression the first line of the RHS of (12).*

The expression to be rewritten is

$$\int_t^{t'} \exp(-\rho s) \{(w_0^A(s) + w_S^A(s))\beta + w_1^A(s)\} ds + \exp(-\rho t') W_\phi^A(t') \beta - \exp(-\rho t) W_\phi^A(t) \beta = \int_t^{t'} \exp(-\rho s) \{g(s)(1 - F_P(s))(\mu + (1 - \mu)\beta) + \beta f_P(s)(1 - G(s))\} ds + \exp(-\rho t') (1 - F_P(t')) (1 - G(t')) \beta - \exp(-\rho t) \exp(-\rho t) (1 - F_P(t)) (1 - G(t)) \beta. \quad (13)$$

For  $t \leq \tau_P$ , we have  $H_R(t) \geq \phi_P > \rho\beta/(\mu(1 - \beta))$ . Thus, for  $t \leq \tau_P$ , function  $v(\cdot)$  is strictly increasing.

Consider the integral

$$\int_t^\infty f_P(s) \min\{v(s) - v(t), v(t') - v(t)\} ds = \underbrace{\int_t^\infty f_P(s) \{v(s) - v(t)\} ds}_A + \underbrace{(1 - F_P(t')) (v(t') - v(t))}_B.$$

which follows because  $v(\cdot)$  is increasing for  $t \leq \tau_P$ .

First, consider term A.

$$\int_t^{t'} f_P(s) \left\{ \int_t^s \exp(-\rho x) g(x) (\mu + (1 - \mu)\beta) dx + \exp(-\rho s)(1 - G(s))\beta - \exp(-\rho t)(1 - G(t))\beta \right\} ds,$$

which is equal to

$$\begin{aligned} & \int_t^{t'} \int_t^s f_P(s) \{ \exp(-\rho x) g(x) (\mu + (1 - \mu)\beta) \} dx ds + \int_t^{t'} f_P(s) \exp(-\rho s)(1 - G(s))\beta ds \\ & - (F_P(t') - F_P(t))(\exp(-\rho t)(1 - G(t))\beta). \end{aligned}$$

Reversing the order of integration in the double integral, we have

$$\begin{aligned} & \int_t^{t'} \int_x^{t'} f_P(s) \exp(-\rho x) g(x) (\mu + (1 - \mu)\beta) ds dx + \int_t^{t'} f_P(s) \exp(-\rho s)(1 - G(s))\beta ds \\ & - (F_P(t') - F_P(t))(\exp(-\rho t)(1 - G(t))\beta) = \\ & \int_t^{t'} (F_P(t') - F_P(x)) \exp(-\rho x) g(x) (\mu + (1 - \mu)\beta) dx + \int_t^{t'} f_P(s) \exp(-\rho s)(1 - G(s))\beta ds \\ & - (F_P(t') - F_P(t))(\exp(-\rho t)(1 - G(t))\beta). \end{aligned}$$

Therefore, term A is

$$\begin{aligned} & \int_t^{t'} (F_P(t') - F_P(s)) \exp(-\rho s) g(s) (\mu + (1 - \mu)\beta) + f_P(s) \exp(-\rho s)(1 - G(s))\beta ds \\ & - (F_P(t') - F_P(t))(\exp(-\rho t)(1 - G(t))\beta). \end{aligned}$$

Next, consider term B:

$$(1 - F_P(t')) \left\{ \int_t^{t'} \exp(-\rho s) g(s) (\mu + (1 - \mu)\beta) ds + \exp(-\rho t')(1 - G(t'))\beta - \exp(-\rho t)(1 - G(t))\beta \right\}.$$

Adding terms A and B, we have

$$\begin{aligned} & \int_t^{t'} (1 - F_P(s)) \exp(-\rho s) g(s) (\mu + (1 - \mu)\beta) ds + f_P(s) \exp(-\rho s)(1 - G(s))\beta ds \\ & + (1 - F_P(t'))(\exp(-\rho t')(1 - G(t'))\beta) - (1 - F_P(t))(\exp(-\rho t)(1 - G(t))\beta), \end{aligned}$$

which is exactly the expression in (13).

*Proof of (i).* Note first that by straightforward differentiation.

$$v'(t) > 0 \iff H_R(t) > \frac{\rho\beta}{\mu(1-\beta)}.$$

Note that  $\beta < \theta \Rightarrow \frac{\rho\beta}{\mu(1-\beta)} < \phi_P$ . Therefore,  $v(\cdot)$  is strictly increasing for  $t \leq \tau_P$ .

Using Step 1, we have,

$$\begin{aligned} u_A(t') - u_A(t) &\leq \\ &\int_t^\infty f_P(s) \min\{v(s) - v(t), v(t') - v(t)\} ds + \exp(-\rho t') W_\phi^A(t') (\mu - \beta) a(t') - \exp(-\rho t) W_\phi^A(t) (\mu - \beta) a(t) \leq \\ &(1 - F_P(t)) (v(t') - v(t)) + \exp(-\rho t') W_\phi^A(t') (\mu - \beta) a(t') - \exp(-\rho t) W_\phi^A(t) (\mu - \beta) a(t), \end{aligned}$$

where use has been made of the observation  $\min\{v(s) - v(t), v(t') - v(t)\} \leq v(t') - v(t)$ , completing the proof of (i).

*Proof of (ii).* Consider the expression for  $u_A(t') - u(t)$ ; noting that  $\beta < 1$  and  $a(t) \geq 0$ , we have

$$\begin{aligned} u_A(t') - u_A(t) &\geq \exp(-\rho t') \beta \int_t^{t'} \{w_0^A(s) + w_0^S(s) + w_1^A(s)\} ds + \\ &\exp(-\rho t') W_\phi^A(t') (\mu a(t') + (1 - a(t')) \beta) - \exp(-\rho t) W_\phi^A(t) (\mu a(t) + (1 - a(t)) \beta). \end{aligned}$$

Using Observation 1, we have

$$\begin{aligned} u_A(t') - u_A(t) &\geq \exp(-\rho t') (W_\phi^A(t) - W_\phi^A(t')) \beta + \\ &\exp(-\rho t') W_\phi^A(t') (\mu a(t') + (1 - a(t')) \beta) - \exp(-\rho t) W_\phi^A(t) (\mu a(t) + (1 - a(t)) \beta). \end{aligned}$$

Simplifying,

$$u_A(t') - u_A(t) \geq -(\exp(-\rho t) - \exp(-\rho t')) W_\phi^A(t) \beta + (\exp(-\rho t') W_\phi^A(t') a(t') - \exp(-\rho t) W_\phi^A(t) a(t)) (\mu - \beta)$$

*Proof of (iii).*

If  $F_P(t) = F_P(t')$ , then  $f_P(s) = 0$  almost everywhere in  $(t, t')$ . This observation, together with the other assumptions implies  $w_0^A(s) = (1 - \mu)g(s)(1 - F_P(t))$ ,  $w_1^A(s) = \mu g(s)(1 - F_P(t))$ ,  $w_0^S(s) = 0$ , for almost all  $s \in (t, t')$ . Furthermore,  $W_\phi^A(t) = (1 - G(t))(1 - F_P(t))$  and  $W_\phi^A(t') = (1 - G(t'))(1 - F_P(t))$ . The result follows from routine substitution.  $\square$

*Remark 1.* If  $f_P(s) = 0$  and  $a(s) = 1$  for  $s \in [t, t']$ , then the Right hand side of the inequality

in Lemma 5 reduces to  $(1 - F_P(t))(u_A^{FB}(t') - u_A^{FB}(t))$ . Thus, parts (i) and (iii) of Lemma 5 are mutually consistent.

**Proof of Lemma 6.** *Part (i).* Note that  $a(t) < 1 \Rightarrow f_A(t) > 0$  (see 1). It follows from (4) that  $u_A(t') - u_A(t) \leq 0$  for any  $t' > t$ . From Lemma 5, we have

$$\{\exp(-\rho t')W_\phi^A(t')a(t') - \exp(-\rho t)W_\phi^A(t)a(t)\}(\mu - \beta) - \{\exp(-\rho t) - \exp(-\rho t')\}W_\phi^A(t)\beta \leq u_A(t') - u_A(t).$$

Combining these inequalities, we have

$$\{\exp(-\rho t')W_\phi^A(t')a(t') - \exp(-\rho t)W_\phi^A(t)a(t)\}(\mu - \beta) - \{\exp(-\rho t) - \exp(-\rho t')\}W_\phi^A(t)\beta \leq 0.$$

Simplifying, we have

$$\exp(-\rho t')W_\phi^A(t')a(t') \leq \exp(-\rho t)W_\phi^A(t)a(t) + \frac{\{\exp(-\rho t) - \exp(-\rho t')\}W_\phi^A(t)\beta}{\mu - \beta}.$$

Because  $F_P(t) < 1$ , for  $\varepsilon \in (0, \bar{\varepsilon})$  we have  $F_P(t') < 1$  for all  $t' \in [t, t + \varepsilon)$ . In other words,  $W_\phi^A(t') > 0$  for such  $t'$ . Therefore, we have

$$a(t') \leq \underbrace{\frac{\exp(-\rho t)W_\phi^A(t)}{\exp(-\rho t')W_\phi^A(t')}}_A a(t) + \underbrace{\frac{\{\exp(-\rho t) - \exp(-\rho t')\}W_\phi^A(t)\beta}{(\mu - \beta)\exp(-\rho t')W_\phi^A(t')}}_B.$$

Because  $W_\phi^A(\cdot)$  is right-continuous, A is arbitrarily close to 1 provided  $\varepsilon$  sufficiently small. Simultaneously, B is arbitrarily close to 0. Because  $a(t) < 1$ , we have that the RHS is less than 1.

*Part (ii).* Note that  $a(t) < 1 \Rightarrow f_A(t) > 0$  (see 1). It follows from (4) that  $u_A(t) - u_A(t'') \geq 0$  for any  $t'' < t$ . From Lemma 5, we have

$$(1 - F_P(t''))(v(t) - v(t'')) + \exp(-\rho t)W_\phi^A(t)(\mu - \beta)a(t) - \exp(-\rho t'')W_\phi^A(t'')(\mu - \beta)a(t'') \geq u_A(t) - u_A(t'').$$

Combining these inequalities we have

$$(1 - F_P(t''))(v(t) - v(t'')) + \exp(-\rho t)W_\phi^A(t)(\mu - \beta)a(t) - \exp(-\rho t'')W_\phi^A(t'')((\mu - \beta)a(t'')) \geq 0.$$

Because  $t'' < t$ , we have  $\exp(-\rho t'')W_\phi^A(t'') \geq \exp(-\rho t)W_\phi^A(t)$ . Therefore,

$$(1 - F_P(t''))(v(t) - v(t'')) + \exp(-\rho t'')W_\phi^A(t'')\{(\mu - \beta)a(t)\} - \exp(-\rho t'')W_\phi^A(t'')((\mu - \beta)a(t'')) \geq 0.$$

Furthermore,  $F_P(t'') \leq F_P(t) < 1$ . Therefore,

$$(v(t) - v(t'')) + \exp(-\rho t'')(1 - G(t''))(\mu - \beta)(a(t) - a(t'')) \geq 0.$$

Obviously,  $\exp(-\rho t'')(1 - G(t''))(\mu - \beta) > 0$ . It follows that

$$a(t'') \leq a(t) + \frac{v(t) - v(t'')}{\exp(-\rho t'')(1 - G(t''))(\mu - \beta)}.$$

Because  $v(\cdot)$  is continuous and  $a(t) < 1$ , for  $t''$  sufficiently close to  $t$ , RHS is smaller than 1.  $\square$

**Proof of Lemma 7.** *Part (i).* Note that if  $F_A(t_1) = F_A(t_2)$ , then  $F_A(t_1) = F_A(t) = F_A(t_2)$  for all  $t \in [t_1, t_2]$ . In addition we have assumed (5). From Lemma 2 part (iv), we have  $u_P(t_2) - u_P(t) = (1 - \sigma F_A(t_1))(u_P^{FB}(t_2) - u_P^{FB}(t))$ . We have assumed  $t_2 \leq \tau_P$ . From Lemma 1,  $u_P^{FB}(t_2) > u_P^{FB}(t)$ . Therefore,  $u_P(t_2) > u_P(t)$ . From (6), we have  $f_P(t) = 0$ .

*Part (ii).* In the proof of (i) we have shown that  $u_P(t_1) < u_P(t_2)$ . From continuity of  $u_P(\cdot)$  (Lemma 3), we have  $u_P(t) < u_P(t_2)$  for all  $t \in [t_1 - \varepsilon, t_1)$ , for sufficiently small  $\varepsilon > 0$ . From (6), we have  $f_P(t) = 0$  for such  $t$ . Together with part (i) we have  $f_P(t) = 0$  for  $t \in [t_1 - \varepsilon, t_2]$ .  $\square$

**Proof of Lemma 8.** Consider  $T_P = \{t | F_P(t) = F_P(t_2)\}$ . Because  $F_P(t_2) < 1$ ,  $F_P(\cdot)$  is weakly increasing, and  $F_P(\tau_P) = 1$ , the set  $T_P$  is bounded from above by  $\tau_P$ . Therefore, it has a supremum,  $t'_2$ . In other words, there exists  $t'_2 \leq \tau_P$  and  $\varepsilon' > 0$  such that for any  $\varepsilon \in (0, \varepsilon')$ : (1)  $F_P(t) = F_P(t_2)$  for  $t \in (t'_2 - \varepsilon, t'_2)$ , and (2)  $F_P(t) > F_P(t_2)$  for  $t \in (t'_2, t'_2 + \varepsilon)$ . Note that continuity of  $u_P(\cdot)$  (Lemma 3) combined with (2) implies that  $u_P(t'_2) = u_P^*$ , the principal's equilibrium payoff, and hence  $u_P(t'_2) \geq u_P(t)$  for all  $t \geq 0$ .

*Step 1:* Show that either (I)  $a(t) = 1$  for all almost all  $t \in [t_1, t'_2)$ , or (II) some  $t \in (t_1, t'_2)$  and  $\varepsilon > 0$  exist such that  $a(t'') < 1$  for almost all  $t'' \in (t - \varepsilon, t)$  and  $a(t') = 1$  for all  $t' \in (t, t + \varepsilon)$ .

Consider  $T_A = \{t | t \in (t_1, t'_2) \cap a(t) < 1\}$  (note the exclusion of  $t_1$ ). If  $T_A$  has measure zero, then (I).

Suppose  $T_A$  has strictly positive measure. Consider  $t \in T_A$ . Using Lemma 6, we conclude that an interval  $(t - \varepsilon, t + \varepsilon) \subset T_A$  for some  $\varepsilon > 0$ . That is, the set  $T$  is open. It follows that  $T_A$  is a union of disjoint open intervals, i.e.  $T_A = \cup_{k=1}^K (q_k, p_k)$ , where  $K \leq \infty$  and  $t_1 \leq q_1 < p_1 \leq q_2 < p_2 \dots \leq q_k < p_k \dots \leq t_2$ . That is,  $q_k$  is the bottom of interval  $k$  and  $p_k$  is its top. Note that  $p_k = q_{k+1}$  is allowed; in this case, the top of open interval  $k$  coincides with the bottom of open interval  $k + 1$ . For ease of notation, if  $K = \infty$ , let  $p_K \equiv \lim_{k \rightarrow \infty} p_k$ , which exists because

$\{p_k\}$  is a bounded monotone sequence. In addition, whether  $K$  is finite or infinite,  $p_K \leq t'_2$  by construction.

Note that if a  $k \in K$  exists such that  $p_k < q_{k+1}$ , then (II). Indeed, consider time  $p_k$  and  $0 < \varepsilon < \min\{p_k - q_k, q_{k+1} - p_k\}$ . If  $t'' \in (p_k - \varepsilon, p_k)$ , then  $q_k < t'' < p_k$ , and hence  $a(t'') < 1$ . Simultaneously, if  $t' \in (p_k, p_k + \varepsilon)$ , then  $p_k < t' < q_{k+1}$ , and hence  $a(t') = 1$ .

To proceed, consider the remaining possibility that  $p_k = q_{k+1}$  for all  $k \in K$ . In this case, the upper bound of interval  $k$  always coincides with the lower bound of interval  $k + 1$ . Thus, the set  $T$  is an interval  $(q_1, p_K)$ , with a countable number of points removed, i.e.  $T = \{(q_1, p_K) - \{p_k\}_{k=1}^\infty\}$ . Therefore, for almost all  $t \in (q_1, p_K)$ , we have  $a(t) < 1$ . In turn (5) implies that  $\mu_1(t) = \theta$  for almost all such  $t$ . Thus,,

$$u_p(p_K) - u_P(q_1) = \int_{q_1}^{p_K} \exp(-\rho s) \{(w_0^P(s) + w_1^P(s) + w_S^P(s))\} \theta ds + \exp(-\rho p_K) W_\phi^P(p_K) \theta - \exp(-\rho q_1) W_\phi^P(q_1) \theta.$$

Using Observation 1, and noting that  $\exp(-\rho s) < \exp(-\rho q_1)$  for  $s \in (q_1, p_K]$ , we have

$$\begin{aligned} u_P(p_K) - u_P(q_1) &\leq \exp(-\rho q_1) (W_\phi^P(q_1) - W_\phi^P(p_K)) \theta + \exp(-\rho p_K) W_\phi^P(p_K) \theta - \exp(-\rho q_1) W_\phi^P(q_1) \theta = \\ &= -(\exp(-\rho q_1) - \exp(-\rho p_K)) W_\phi^P(p_K) \theta = \\ &= -(\exp(-\rho q_1) - \exp(-\rho p_K)) (1 - G(p_K)) (1 - F_P(t_2)) \theta < 0. \end{aligned}$$

Therefore,  $u_P(p_K) < u_P(q_1) \leq u_P(t'_2)$ . Because  $u_P(\cdot)$  is continuous (see Lemma 3), it must be the case that  $p_K < t'_2$ .

Finally, consider time  $p_K$  and  $0 < \varepsilon < \min\{p_K - q_1, p_K - t'_2\}$ . If  $t'' \in (p_K - \varepsilon, p_K)$ , then  $t'' \in (q_1, p_K)$ . Recall that  $a(t'') < 1$  for almost all such  $t''$ . Furthermore, if  $t' \in (p_K, p_K + \varepsilon)$ , then  $t' \in (p_K, t'_2)$ , and hence,  $a(t') = 1$ . This completes the proof of Step 1.

*Step 2: Show that  $a(t) = 1$  for all almost all  $t \in [t_1, t'_2]$ .*

Step 2 is proved by establishing that Case (II) in the statement of Step 1 leads to a contradiction. Therefore, suppose that for some  $t \in (t_1, t'_2)$  and  $\varepsilon > 0$ , exist such that  $a(t'') < 1$  for almost all  $t'' \in (t - \varepsilon, t)$ , and  $a(t') = 1$  for all  $t' \in (t, t + \varepsilon)$ .

Consider  $t - \varepsilon < t_L < t < t_M < t_H < t + \varepsilon \leq t'_2$ . By construction,  $a(t) = 1$  for all  $t \in [t_M, t_H]$ . In addition,  $F_P(t_M) = F_P(t_H) = F_P(t_2) < 1$ . From Lemma 5 part (iii), we have

$$\frac{u_A(t_H) - u_A(t_M)}{1 - F_P(t_2)} = u_A^{FB}(t_H) - u_A^{FB}(t_M).$$

Using Lemma 5 part (ii), we have

$$u_A(t_M) - u_A(t_L) \geq \{\exp(-\rho t_M)W_\phi^A(t_M) - \exp(-\rho t_L)W_\phi^A(t_L)a(t_L)\}(\mu - \beta) - \{\exp(-\rho t_L) - \exp(-\rho t_M)\}W_\phi^A(t_L)\beta.$$

Note that  $a(t_L) \leq 1$  and  $W_\phi^A(t_L) \leq 1$ . It follows that

$$u_A(t_M) - u_A(t_L) \geq \{\exp(-\rho t_M)W_\phi^A(t_M) - \exp(-\rho t_L)W_\phi^A(t_L)\}(\mu - \beta) - \{\exp(-\rho t_L) - \exp(-\rho t_M)\}\beta.$$

Next, note that  $F_P(t_M) = F_P(t_L) = F_P(t_2) < 1$ . Substituting and simplifying, we have

$$\begin{aligned} \frac{u_A(t_M) - u_A(t_L)}{1 - F_P(t_2)} &\geq \\ &- \{\exp(-\rho t_L)(1 - G(t_L)) - \exp(-\rho t_M)(1 - G(t_M))\}(\mu - \beta) - \{\exp(-\rho t_L) - \exp(-\rho t_M)\}\beta \\ &= -(\gamma(t_L) - \gamma(t_M)), \end{aligned}$$

where  $\gamma(t) \equiv \exp(-\rho t)\{(1 - G(t))(\mu - \beta) + \beta\}$ .

Next, note that

$$\frac{u_A^{FB}(t_H) - u_A^{FB}(t_L)}{1 - F_P(t_2)} = \frac{u_A^{FB}(t_H) - u_A^{FB}(t_M)}{1 - F_P(t_2)} + \frac{u_A^{FB}(t_M) - u_A^{FB}(t_L)}{1 - F_P(t_2)}.$$

Using the previous bounds, it follows that

$$\frac{u_A(t_H) - u_A(t_L)}{1 - F_P(t_2)} \geq u_A^{FB}(t_H) - u_A^{FB}(t_M) - (\gamma(t_L) - \gamma(t_M)).$$

Choose  $t_M$  just above  $t$ . By continuity of  $\gamma(\cdot)$ , for  $t_L$  sufficiently close to  $t$ , the difference  $\gamma(t_L) - \gamma(t_M) \approx 0$ . At the same time Lemma 1 guarantees  $u_A^{FB}(t_H) - u_A^{FB}(t_M) > 0$ . Therefore for some  $\varepsilon'' > 0$ , we have that  $t'' \in (t - \varepsilon'', t)$  implies  $u_A(t_H) - u_A(t'') > 0$ . From (4), we have  $f_A(t'') = 0$ . Hence,  $\mu_1(t'') = 1$ , which in turn implies  $a(t'') = 1$ . This contradicts the initial assumption that  $a(t'') < 1$  for almost all  $t'' \in (t - \varepsilon, t)$ .

*Step 3: Show that  $f_A(t) = 0$  for all  $t \in [t_1, t'_2]$ .*

From Step 2,  $a(t) = 1$  for almost all  $t \in [t_1, t'_2]$ . Consider  $t_1 \leq t < t'_2$ . For any such  $t$ , it is



possible to find a  $t' \in (t, t'_2)$  for which  $a(t') = 1$ . Note that

$$\begin{aligned} \frac{u_A(t') - u_A(t)}{1 - F_P(t_2)} &= \\ \int_t^{t'} \exp(-\rho s)g(s)(\mu + (1 - \mu)\beta)ds + \exp(-\rho t')(1 - G(t'))\mu - \exp(-\rho t)(1 - G(t))(\beta + a(t)(\mu - \beta)) &\geq \\ \int_t^{t'} \exp(-\rho s)g(s)(\mu + (1 - \mu)\beta)ds + \exp(-\rho t')(1 - G(t'))\mu - \exp(-\rho t)(1 - G(t))\mu &= \\ u_A^{FB}(t') - u_A^{FB}(t) &> 0, \end{aligned}$$

where the last equality follows from  $t' < t'_2 \leq \tau_P < \tau_A$ . It follows that for any  $t \in [t_1, t'_2)$ , there exists some  $t'$  such that  $u_A(t') > u_A(t)$ . From (4), we have  $f_A(t) = 0$ . Obviously  $t'_2 \geq t_2$ , and  $f_A(t) = 0$  holds for  $t \in [t_1, t_2)$ .  $\square$

**Proof of Lemma 9.** Consider  $T_P = \{t | F_P(t) = F_P(t_2)\}$ . Because  $T_P$  is bounded from below by 0 and  $F_P(\cdot)$  is right-continuous,  $T_P$  has a minimum,  $\underline{t}_1 \in [0, t_1]$ . To prove the result, it will be shown that  $\underline{t}_1 = 0$ .

Suppose that  $\underline{t}_1 > 0$ . Consider interval  $[\underline{t}_1, t'_2)$  where  $t_1 < t'_2 < t_2$ . Applying Lemma 8, we have  $f_A(t) = 0$  for all  $t \in [\underline{t}_1, t'_2)$ . Next, consider interval  $[\underline{t}_1, t''_2]$  where  $t_1 < t''_2 < t'_2$ . Because  $f_A(t) = 0$  for  $t \in [t_1, t''_2]$ , we have  $F_A(t_1) = F_A(t''_2)$ . Furthermore,  $t''_2 < t_2 < \tau_P$ , where the last inequality follows from  $F_P(t_2) < 1$  and Lemma 4. Using Lemma 7 part (ii), there exists  $t_0 < \underline{t}_1$  such that  $f_P(t) = 0$  for all  $t \in [t_0, t''_2]$ . Consider interval  $[t_0, t'''_2]$  where  $t_1 < t'''_2 < t''_2$ . Because  $f_P(t) = 0$  for all such  $t$ , we have  $F_P(t_0) = F_P(t'''_2)$ . Furthermore,  $t'''_2 \in [t_1, t_2]$ , and therefore  $F_P(t'''_2) = F_P(t_2)$ . Hence,  $F_P(t_0) = F_P(t_2)$ . Because  $t_0 < \underline{t}_1$ , we have that  $\underline{t}_1$  is not the minimum of  $T_P$ .  $\square$

**Proof of Lemma 10.** Consider  $T_A = \{t | F_A(t) = F_A(t_2)\}$ . Because  $T_A$  is bounded from below by 0 and  $F_A(\cdot)$  is right-continuous,  $T_A$  has a minimum,  $\underline{t}_1 \in [0, t_1]$ . To prove the result, it will be shown that  $\underline{t}_1 = 0$ .

Suppose that  $\underline{t}_1 > 0$ . Consider interval  $[\underline{t}_1, t'_2)$  where  $t_1 < t'_2 < t_2$ . Applying Lemma 7 part (ii), we have  $f_P(t) = 0$  for all  $t \in [t_0, t'_2)$ , for some  $t_0 < \underline{t}_1$ . By implication,  $f_P(t) = 0$  for all  $t \in [t_0, t''_2]$ , where  $t_1 < t''_2 < t'_2$ . Hence,  $F_P(t_0) = F_P(t''_2)$ . Furthermore, because  $t''_2 < t_2$ , we have  $F_P(t''_2) \leq F_P(t_2) < 1$ . Applying Lemma 8, we have  $f_A(t) = 0$  for all  $t \in [t_0, t''_2)$ . By implication,  $f_A(t) = 0$  for all  $t \in [t_0, t'''_2]$ , where  $t_1 < t'''_2 < t''_2$ . Therefore,  $F_A(t_0) = F_A(t'''_2) = F_A(t_2)$ , where the last equality follows from  $t'''_2 \in (t_1, t_2)$ . Because  $t_0 < \underline{t}_1$  and  $F_A(t_0) = F_A(t_2)$ , we have that  $\underline{t}_1$  is not the minimum of  $T_A$ .  $\square$

## 7.2. Proof of Corollary 2

**Proof of Corollary 2.** *Part (i).* Obviously,  $\sigma$  has no effect on  $\tau_P$ . To see that it reduces  $\tau_M$ , recall the following implicit function relating  $\tau_M$  and  $\sigma$  derived in Proposition 2.

$$\exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = 1 - \sigma.$$

Differentiating with respect to  $\sigma$ ,

$$\frac{d\tau_M}{d\sigma} \left\{ \frac{\mu(1-\theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} \exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = -1.$$

Simplifying,

$$\begin{aligned} \frac{d\tau_M}{d\sigma} \left\{ \frac{\mu(1-\theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} (1 - \sigma) &= -1 \Rightarrow \\ \frac{d\tau_M}{d\sigma} &= \left\{ \frac{\mu(1-\theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \right\} (1 - \sigma) = -\frac{\theta - \mu}{1 - \sigma} \frac{1}{\mu(1-\theta)(H_R(\tau_M) - \frac{\rho\theta}{\mu(1-\theta)})} < 0, \end{aligned}$$

where the last inequality follows because  $H_R(s) > \rho\theta/(\mu(1-\theta))$  for  $s < \tau_P$ , and  $\tau_M < \tau_P$ .

Next, note that

$$\begin{aligned} \frac{\partial F_A(t)}{\partial \tau_M} \frac{d\tau_M}{d\sigma} &= -\frac{1}{\sigma} (1 - \sigma F_A(t)) \frac{\mu(1-\theta)H_R(\tau_M) - \rho\theta}{\theta - \mu} \left( -\frac{\theta - \mu}{1 - \sigma} \frac{1}{\mu(1-\theta)(H_R(\tau_M) - \frac{\rho\theta}{\mu(1-\theta)})} \right) \\ &= \frac{1 - \sigma F_A(t)}{\sigma(1 - \sigma)} \geq \frac{1}{\sigma}. \end{aligned}$$

Meanwhile, we have

$$\frac{\partial F_A(t)}{\partial \sigma} = -\frac{1}{\sigma^2} F_A(t) \geq -\frac{1}{\sigma^2}.$$

Hence,  $dF_A(t)/d\sigma \geq 1/\sigma - 1/\sigma^2 > 0$  for  $\sigma \in (0, 1)$ . Therefore, an increase in  $\sigma$  increases  $F_A(\cdot)$ , resulting in a FOSD shift toward smaller  $t$ .

Note that  $F_P(t)$  depends on  $\sigma$  only via  $\tau_M$  and

$$\frac{dF_P(t)}{d\tau_M} = -(1 - F_P(t)) \frac{\beta(1-\mu)H_R(\tau_M) - \rho\mu}{\mu - \beta} < 0,$$

where the last inequality follows because  $H_R(\cdot)$  is decreasing, and  $\tau_M < \tau_A$ . Because an increase in  $\sigma$  reduces  $\tau_M$ , it increases  $F_P(\cdot)$ , resulting in a FOSD shift toward smaller  $t$ .

Part (ii). First note that

$$H_R(\tau_P) = \frac{\rho\theta}{\mu(1-\theta)} \Rightarrow \tau'_P(\theta) = \frac{\rho}{H'_R(\tau_P)\mu(1-\theta)^2} < 0,$$

where the inequality follows from  $H_R(\cdot) < 0$ .

Next, recall that Proposition 2 gives an implicit function relating  $\tau_M$ ,  $\tau_P$ , and  $\sigma$ ,

$$\exp\left\{-\int_{\tau_M}^{\tau_P} \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu} ds\right\} = 1 - \sigma.$$

Let

$$H_A(s) \equiv \frac{\mu(1-\theta)H_R(s) - \rho\theta}{\theta - \mu}$$

$$h_A(s) \equiv \frac{dH_A(s)}{d\theta} = \frac{-(\mu H_R(s) + \rho)(\theta - \mu) - \mu(1-\theta)H_R(s) + \rho\theta}{(\theta - \mu)^2} = -\frac{\mu(1-\mu)}{(\theta - \mu)^2} \left(H_R(s) - \frac{\rho}{1-\mu}\right).$$

Differentiating with respect to  $\theta$ , we have

$$\exp\left\{-\int_{\tau_M}^{\tau_P} H_A(s) ds\right\} \left(H_A(\tau_M)\tau'_M(\theta) - H_A(\tau_P)\tau'_P(\theta) - \int_{\tau_M(\theta)}^{\tau_P(\theta)} h_A(s) ds\right) = 0.$$

Note that  $H_R(\tau_P) = \rho\theta/(\mu(1-\theta))$ , which implies  $H_A(\tau_P) = 0$ . By implication,

$$H_A(\tau_M)\tau'_M(\theta) - \int_{\tau_M}^{\tau_P} h_A(s) ds = 0 \Rightarrow \tau'_M(\theta) = \frac{\int_{\tau_M}^{\tau_P} h_A(s) ds}{H_A(\tau_M)}.$$

Next, note that for  $s \in [\tau_M, \tau_P]$ ,

$$H_R(s) - \frac{\rho}{1-\mu} \geq \frac{\rho\theta}{\mu(1-\theta)} - \frac{\rho}{1-\mu} = \frac{\rho(\theta - \mu)}{\mu(1-\mu)(1-\theta)} > 0.$$

It follows that  $h_A(s) < 0$  for such  $s$ . Simultaneously,  $H_R(s) > \rho\theta/(\mu(1-\theta))$  and hence,  $H_A(s) > 0$ . Combining these observations,  $\tau'_M(\theta) < 0$ .

Next, note that

$$F_A(t) = \frac{1}{\sigma} \left(1 - \exp\left(-\int_{\tau_M}^t H_A(s) ds\right)\right).$$

It follows that

$$\frac{dF_A(t)}{d\theta} = \frac{1}{\sigma} \exp\left(-\int_{\tau_M}^t H_A(s)ds\right) \left(\int_{\tau_M}^t h_A(s)ds - H_A(\tau_M)\tau'_M(\theta)\right).$$

Substituting for  $\tau'_M(\theta)$ , we have

$$\begin{aligned} \frac{dF_A(t)}{d\theta} &= \frac{1}{\sigma} \exp\left(-\int_{\tau_M}^t H_A(s)ds\right) \left(\int_{\tau_M}^t h_A(s)ds - \int_{\tau_M}^{\tau_P} h_A(s)ds\right) \\ &= \frac{1}{\sigma} \exp\left(-\int_{\tau_M}^t H_A(s)ds\right) \left(-\int_t^{\tau_P} h_A(s)ds\right). \end{aligned}$$

Recalling that  $h_A(\cdot) < 0$  for such  $s$ , we have  $dF_A(t)/d\theta > 0$ , and hence an increase in  $\theta$  results in a FOSD shift toward smaller  $t$ .

Finally, note that  $F_P(\cdot)$  depends on  $\theta$  only through  $\tau_M$ , and for  $t \in [\tau_M, \tau_P)$

$$\frac{dF_P(t)}{d\tau_M} = (1 - F_P(t)) \frac{\beta(1 - \mu)H_R(\tau_M) - \rho\mu}{\beta - \mu} > 0,$$

where the last inequality follows because  $\tau_M < \tau_A$ . Thus, for all  $t \in [\tau_M, \tau_P)$  the distribution shifts up, and the upper bound  $\tau_P$  where the distribution has an atom, also shifts left, resulting in a FOSD shift toward smaller  $t$ .  $\square$