The Dissertation Committee for Yiman Sun
certifies that this is the approved version of the following dissertation:

# Essays on Experimentation in Agency Models

Committee:

Venkataraman Bhaskar, Supervisor

Caroline D. Thomas, Co-Supervisor

Thomas E. Wiseman

John W. Hatfield

# Essays on Experimentation in Agency Models

by

**Yiman Sun**

**DISSERTATION**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2019

Dedicated to my wife Zhenzhen, my parents, and my grandparents.

# Acknowledgments

I am deeply grateful to my advisors, V. Bhaskar and Caroline Thomas, for their invaluable guidance and support. I would like to thank John Hatfield and Thomas Wiseman for their insightful comments and suggestions on this dissertation. My gratitude also goes to other members of the theory group at UT Austin Svetlana Boyarchenko, William Fuchs, Vasiliki Skreta, Dale Stahl, Maxwell Stinchcombe, and Rodrigo Velez. I thank Cowles Foundation at Yale for their hospitality, where this dissertation was partially written, as well as their theory group members Joyee Deb, Laura Doval, Mira Frick, John Geanakoplos, Johannes Hörner, Erik Madsen, Dov Samet, Larry Samuelson, Takuo Sugaya, and Jidong Zhou. I also wish to express my gratitude to Jason Abrevaya, Daniel Ackerberg, Sandra Black, Andrew Glover, Leigh Linden, Stephen Ryan, Robert Town, Haiqing Xu, and Gordan Žitković for supporting me in different stages of my research, as well as my career development.

I would like to thank my supportive friends and schoolmates Hassan Afrouzi, Ian Ball, Yi Chen, Sangwoo Choi, Jessie Coe, Gleb Domnenko, Jessica Fears, Wayne Gao, Xinchen Gu, Xinying Hao, Ziwei He, Sev Hou, Vasudha Jain, Shaofei Jiang, Colin Koutney, Mike Lan, Minghao Li, Xiangliang Li, Xue Li, Fan Liu, Shuo Liu, Xiang Liu, Yalin Liu, Julian Ludwig, Anthony Maxam, Daria Pus, Andres Ruiz, Peter Toth, Xinyang Wang, Neil White, Mark Whitmyer, Lezhen Wu, Jiangang Zeng, Yingda Zhai, Chi Zhang, Ge Zhang, Rui Zhang, Xinyu Zhang, and Ross Zhu.

Last, I would like to express my special thanks to my wife Zhenzhen, my parents, and my grandparents for their love, understanding, support, and everything.

**Essays on Experimentation in Agency Models**

Publication No. _______________

Yiman Sun, Ph.D.
The University of Texas at Austin, 2019

Supervisors: Venkataraman Bhaskar
Caroline D. Thomas

This dissertation consists of three chapters in microeconomic theory with a focus on dynamic games and learning. It has applications in political economy, contracts, and industrial organization.

In the first chapter, I study censorship in a dynamic game between an informed agent and an uninformed evaluator. Two types of public news are informative about the agents ability – a conclusive good news process and a bad news process. However, the agent can censor bad news, at some cost, and will censor it if and only if this secures her a significant increase in tenure. Thus, the evaluator faces a bandit problem with an endogenous news process. When bad news is conclusive, the agent always censors when the public belief is sufficiently high, but below a threshold, she either stops censoring or only censors with some probability, depending on the information structure. The possibility of censorship hurts the evaluator and the good agent, and it may also hurt the bad agent. However, when bad news is inconclusive, I show that

the good agent censors bad news more aggressively than the bad agent does. This improves the quality of information, and may benefit all players – the evaluator, the bad agent, and the good agent.

The second chapter examines the nature of contracts that optimally reward innovations in a risky environment, when the innovator is privately informed about the quality of her innovation and must engage an agent to develop it. I model the innovator as a principal who has private but imperfect information about the quality of her project: the project might be worth exploring or not, but even a project of high quality may fail. I characterize the best equilibrium for the high type principal, which is either a separating equilibrium or a pooling one. Due to the interaction between the signaling incentives of the principal and dynamic moral hazard of the agent, the best equilibrium induces inefficiently early termination of the high quality project. The high type principal is forced to share the surplus – with the agent in the separating equilibrium, or the low type principal in the pooling equilibrium. A mediator, who offers a menu of contracts and keeps the agent uncertain about which contract will be implemented, can increase the payoff of the high type principal to approximate her full information surplus.

In the third chapter, I study how competition between platforms affects the process of social learning. Especially, how product differentiation affects that process. Che and Hörner (2018) show that a monopolistic platform may want to over-recommend consumers in the early phase to gather and learn information for the sake of future consumers. I show that when platforms

do not differentiate their products, duopoly competition dramatically reduces the early experimentation, and the Full Transparency policy is the unique equilibrium strategy for both platforms. When platforms differentiate their products, I show that the equilibrium strategy is in between the Full Transparency policy and the optimal policy in the monopolistic case, and depends on how differentiated the products are.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# A Dynamic Model of Censorship

## 1.1 Introduction

Individuals in positions of power, be they political leaders or the managers of firms, often suppress or censor bad news in order to improve their standing and prevent any threats to their authority. Such censorship is widely regarded to be undesirable. Nonetheless, we can imagine situations where the suppression of bad news may lead to better outcomes. For example, a political leader may be embarking on a radical reform that has the potential to be transformative. Being radical, the reform is also subject to teething troubles, and if the public were to become aware of all the difficulties, it might prematurely lose faith in the leader and replace her.

The decades of the 1960's - 1980's witnessed rapid industrialization and exceptionally high growth rates in many Asian developing countries and regions, including Singapore, South Korea, and Taiwan. This has been called an economic "miracle". Controversially, all these countries and regions were under authoritarian rule at the time. One hypothesis from a review by Sirowy and Inkeles (1990) links economic growth with authoritarianism: "the superior ability of an authoritarian regime to govern that facilitates economic growth is expressed indirectly by the social and political stability it fosters" (p. 130). This "political stability" may be sustained by different means, one of which is censorship.[1] As Rodan (2004) noticed, "[a]lmost by definition, authoritarian regimes involve censorship" (p. 1). On the other hand, we have many examples of leaders who persist with foolhardy projects, hiding all negative evidence. For instance, during China's Great Leap Forward (1958-1962), the central government's failure to access up-to-date local information due to local officials' concealment was partially responsible for the ensuing famine.[2] These considerations suggest that it is important to formally examine the implications of censorship in a dynamic context.

More specifically, this paper studies the interplay between learning and censorship in a dynamic environment. We consider a relationship between an agent who seeks to remain in office, and an evaluator who learns the competence of the agent from public news about the quality of the agent's project. The evaluator wants to retain the competent agent and dismiss the incompetent one, thereby terminating her project. The project of the agent gives rise to two news processes – good news and bad news. Good news is publicly observed, and confirms that the project is a good one. However, bad news can also arise, and is more likely when the project is a bad one. The agent can always suppress bad news when it materializes, but this is costly. Thus, the information that the evaluator receives is endogenously determined by the

[1] Guriev and Treisman (2018) study how repression, co-option, propaganda, and censorship are used for an authoritarian regime to survive.

[2] See Li and Yang (2005).

agent's censorship policy. We assume that the agent knows her own competence level, and consequently, competent agents and incompetent ones may well censor differently.

There are several applications of our model. The agent can be the manager of a division of a large firm, while the evaluator is the firm's overall manager or CEO. Alternatively, the agent might be an entrepreneur, with the evaluator being a venture capitalist who is funding the project. Finally, the agent might be a political leader, with the evaluator standing for the population. In these contexts, bad news can take several forms – mechanical breakdowns, reports of malpractice or customer complaints. Bad news can be suppressed – accounts can be "cooked", log files can be faked, and unhappy consumers could be mollified with refunds or gifts. None of these measures are costless; they take time and money, and psychological costs may be associated with dishonest behavior.[3] Similarly, politicians can arrest reporters, bribe witnesses, or shut down Internet forums, but this is also costly.

Basic economic intuition suggests that concealing information necessarily hurts the evaluator. Moreover, the possibility of censorship makes the evaluator more suspicious about the agent's performance. He does not know whether the reason that no bad news arrives is because the agent is competent, or because the agent is censoring. Thus the possibility of censorship also hurts a competent agent who has no way to prove that she was not hiding anything.

[3] See Rosenbaum et al. (2014) for a review on honesty experiments.

It can only benefit an incompetent agent, since censorship helps her survive bad news. If the above intuition is correct, then the policy implication would be to reduce censorship by making it as hard and costly as possible.

However, this paper shows that the above intuition is only partially correct, and it crucially depends on the details of the information structure. Specifically, it depends on whether negative evidence is *conclusive*, i.e. it can only arise when the agent is incompetent. Indeed, when bad news can also arise when the agent is competent, we show that censorship can potentially increase the welfare of all parties, including the evaluator.

We now turn to the details of our model and its basic insights. All news is modeled as exponential/Poisson news. Good news can only arise for a competent agent, and is therefore conclusive. We consider two qualitatively different information structures. We first assume that bad news is also conclusive, i.e. it can only arise for an incompetent agent. We then consider the more general case of inconclusive bad news. The evolution of the evaluator's posterior belief about the competence of the agent in the absence of news depends on the agent's censorship policy, and on the (exogenous) parameters of the news processes. In our analysis, it will be useful to distinguish information structures according to whether the absence of news is "good" or "bad" news, i.e. whether the evaluator's posterior belief drifts up or down. We say that *good news arrives faster* when good news arrives faster than the incompetent agent's bad news. If good news arrives faster, then, regardless of the censoring decisions of the agent, the evaluator's belief is updated downwards in the absence of news. We say that *good news arrives slower*, when it arrives slower than the incompetent agent's bad news. In this case, the direction in which the evaluator's belief moves in the absence of news depends upon the censorship policy of the agent.

The simplest case is when bad news is conclusive, i.e. it only arises for the incompetent agent, and when good news arrives faster. Thus the public belief about the agent being competent will drift down in the absence of news regardless of the censorship policy, but censorship will accelerate the downward process. Our first observation is that the belief threshold at which the evaluator fires the agent is *independent* of the (incompetent) agent's censorship strategy, due to the fact that the option value of continuation depends only upon the possibility that good news arises. This observation allows us to use the logic of backwards induction to pin down behavior in any equilibrium. Since censorship is costly, the agent will incur the cost if and only if this secures her a sufficient increase in tenure. Thus, there is a unique belief threshold, at which the agent switches from *Full-Censorship* to *No-Censorship*. Obviously, the evaluator and the competent agent are worse off with censorship, which exactly confirms our initial intuition. However, the incompetent agent may also be worse off with censorship, since she has a shorter tenure, conditional on no bad news occurring, than when censorship is not possible. This happens whenever the Full-Censorship period is short. The benefit of censoring for the incompetent agent is that she survives bad news in the Full-Censorship period. When it is short, it means that the benefit of censoring is small, thus it will be overcome by the cost of censoring, i.e. the acceleration of the downward drifting of public belief.

Consider now the case where good news arrives faster, but where bad news is inconclusive, i.e. it can arise for both types of agent. Our main insight arises from the fact that the competent agent has greater incentives to censor bad news than the incompetent one. This is due to the fact that the competent agent knows that *good news may arise and secure her permanency in tenure*, which is not possible for the incompetent agent. Furthermore, the competent agent is also less likely to get further bad news. Therefore, the competent agent has a higher continuation value on the job, and the belief threshold at which she stops censoring, $p^G$, is lower than the threshold at which the incompetent agent stops censoring, $p^{B\dagger}$. Consequently, in the interval $[p^G, p^{B\dagger})$, bad news becomes conclusive *endogenously*; bad news can only come from the incompetent agent since the competent one always censors. By improving the quality of information, censorship can increase the payoff of the evaluator.[4]

Finally, let us consider the case where good news arrives slower. When bad news is conclusive, we cannot have a pure strategy equilibrium where the incompetent agent stops censoring at some threshold – if this were the case, the evaluator's belief would drift upwards, which would then make censoring bad news attractive. There is a critical threshold such that when this belief threshold is reached, both parties randomize. The incompetent agent ran-

[4]Indeed, for some parameter values it can also improve the payoffs of both types of agent, and therefore benefit all players.

domizes between censoring and not, while the evaluator randomizes between firing the agent and retaining her.[5] Above this belief threshold, the incompetent agent censors for sure. It is worth noting that the evaluator dismisses the agent at a higher threshold belief than he does in the absence of censorship, since the existence of censorship reduces the availability of future bad news, and since the option value of the evaluator also depends upon the arrival rate of bad news. This only happens when good news arrives slower, because the dismissal threshold depends critically on the arrival of bad news only if the belief updating process in the absence of news is upward, but not downward. We call the increase in the dismissal threshold belief the *discouragement effect* since censorship discourages the agent's incentive for learning. When bad news is inconclusive, we show that the very similar equilibrium still exists when the censoring cost is low. The evaluator and the incompetent agent use the same equilibrium strategy as in the conclusive bad news case, but the competent agent censors with probability 1 even when the belief is at the dismissal threshold. Basically, in this equilibrium, all bad news from both types of agent will be censored, except some bad news at the dismissal threshold. This hurts the evaluator, since her information quality is worsened. However, when the censoring cost is intermediate, there exists an equilibrium in which only the competent agent finds it optimal to censor when the belief is sufficiently high, but the incompetent agent never censors since the cost is

[5] More precisely, in our continuous time model, the evaluator randomizes over stopping times.

too high for her. Again, this separation improves the quality of information and benefits the evaluator. We also find an *encouragement effect*; that is, the evaluator's dismissal threshold belief decreases since his incentive for learning is encouraged.

Our model has implications for how we should interpret censorship, and institutional measures against it. Let us interpret the cost of censorship as reflecting institutional strictures against it. In some circumstances, i.e. when bad news only arises for the incompetent agent, we find that censorship is unambiguously bad, and thus the cost of censorship should be as high as possible. However, when bad news also arises for the competent agent, the evaluator prefers neither a very strong institution that prevents censorship, nor a very weak institution that allows for too much censorship. Mild censorship may be better than the two extreme cases.

The rest of the paper is organized as follows. Section 2.2 discusses the related literature. Section 2.3 introduces the model. Section 1.4 and Section 1.5 characterize equilibria and discuss the welfare effect for the conclusive bad news case and the inconclusive bad news case, respectively. Section 2.7 concludes. The Appendix provides all proofs.

## 1.2 Related Literature

This paper studies censorship in a dynamic environment of learning, where multiple signals arrives gradually. It mainly relates to four strands of literature.

First, it closely relates to a small literature on political censorship. Shadmehr and Bernhardt (2015) study a game between a ruler and a representative citizen, in which the ruler can censor a bad media report at a cost in order to mitigate the likelihood of revolution. Their paper, as well as this paper, assumes the role of the media is passive. Thus the focus is solely on the relation between the ruler and the citizen. They find that at the ex ante stage, before the ruler knows her type, she can increase her expected payoff by committing to censoring slightly less than she does in equilibrium where such a commitment power does not exist. One of our results also demonstrates that not only the agent before she knows her type, but also the agent who knows her type is bad can benefit from committing to no censorship.

Besley and Prat (2006) study the media capture problem. In their model, a media outlet maximizes his profits either from his audience who are interested in the informative news, or from a bad government who bribes him to keep silent about the bad news. Their paper focuses on the role of media outlets, and explicitly models the censoring cost as a direct or indirect transfer to a media outlet that is willing to forgo its readership and take a bribe. Gehlbach and Sonin (2014) and Eraslan and Ozerturk (2017) also examine the role of media outlets. The former takes a Bayesian persuasion approach as in Kamenica and Gentzkow (2011) such that a government can commit to an editorial policy for news release, while the latter studies a media outlets reputation concerns that arise due to an information gatekeeping policy.

Egorov et al. (2009) consider the role of media outlets in monitoring bureaucrats.[6] In their model, a dictator needs a bureaucrat to implement her policy, but faces a moral hazard problem. To incentivize the bureaucrat to work, the dictator could rely only on the media report. However, a free media that solves the monitoring problem also exposes the dictator's incompetence to the public. Thus, if the dictator's income mainly depends on the bureaucrat's performance, she would rather risk being overthrown but allow a free media to monitor the bureaucrat. Lorentzen (2014) studies a regime change game in which a regime trades off the risk being overthrown and an informative media capable of monitoring the lower-level officials.

Edmond (2013) and Redlicki (2017) study other types of information manipulation in politics. Specifically, they study a global game played by citizens to attack a regime, in which the private signal of the citizens can be manipulated at a cost by the regime, either through shifting the mean of the signal in Edmond (2013), or through increasing the noise of the signal in Redlicki (2017). Guriev and Treisman (2018) study other means – propaganda, censorship, co-optation, and repression – that a dictator can employ to survive.

Although almost all papers on censorship consider one-shot interaction in a political setting, a notable exception is Smirnov and Starkov (2018).[7]

[6]Egorov et al. (2009) also extend their one-shot censorship model into a dynamic setting by assuming the dictator faces a stationary environment where censoring public bad news can effectively keep the ruler in power, and obtain a similar result as in their static model. In contrast, our model analyzes the (non-stationary) dynamics of the public confidence about the competence of the agent.

[7]Hauser (2017) also studies costly information manipulation in a dynamic environment. Different from the censorship literature that focuses on the ex post information suppression, he explores the case where ex ante effort can be exerted to slow down the arrival of bad

They study censorship in product reviews. The main difference with our setup is that they assume censorship is costless, and some of the myopic consumers are naive, i.e. they do not understand that reviews can be censored. In contrast, we examine censorship with a strategic and forward-looking evaluator, and study the interdependence of censorship and learning, where costs play a crucial role.

Second, another related literature is on disclosure of verifiable information, beginning with Grossman (1981) and Milgrom (1981). A subset of the literature following Dye (1985) and Jung and Kwon (1988) studies the scenario where a receiver is uncertain about whether a sender possesses a piece of evidence. This is also the case in our setup since the evaluator's lack of news could happen either due to censorship or because no news has arrived. While the early literature exclusively focuses on static models, later papers extend them into different dynamic settings where multiple signals may arise in multiple periods, e.g. Shin (2003), Grubb (2011), Acharya et al. (2011), Guttman et al. (2014).

Although the literature on disclosure games is large, very little attention has been paid to the case in which concealing information is costly. Three recent papers explore different implications of such a cost, all of which are very different from those that we focus on. Dye (2017) studies a static model of voluntary disclosure, in which a seller who has withheld information may

news.

be caught by a fact finder after the sale of an asset. In such an event, she has to make a damages payment to the buyer, the amount of which equals the product of the buyer's overpayment and a "damages multiplier". An important implication of this particular cost structure is that the seller would withhold more information as the punishment of the withholding goes up. Daughety and Reinganum (2018) study the problem of suppression of exculpatory evidence in prosecutions. In their model, a prosecutor wants to convict a defendant but also incurs a moral cost if she convicts an innocent defendant. The prosecutor also receives a penalty if she is caught for suppressing evidence by a reviewing judge. They extend their model to incorporate the teamwork of two prosecutors and show that this results in the concentration of authority regarding suppressing evidence. Kartik et al. (2017) provide a result on Bayesian updating, and apply it in a multi-sender disclosure game. They show that competition leads to more disclosure in the presence of a cost of concealment.

Third, a recent literature on dynamic information design is also related, e.g. Ely (2017), Renault et al. (2017), Che and Hörner (2018). Their papers, as well as ours, study information manipulation in a dynamic environment. Their models rely on the principal's commitment power to design a flexible information disclosure policy to induce an agent to choose a desirable action for the principal, and also assume manipulating information is costless. In contrast, we focus on one particular kind of information manipulation – censorship. In addition, we do not assume the commitment power, and assume censorship is costly.

Finally, this paper relates to two-armed Poisson bandit models in continuous time, e.g. Presman (1991), Keller et al. (2005), Keller and Rady (2010, 2015). The evaluator here faces a two-armed bandit problem, in which the information generated by the risky arm is endogenous and partially controlled by the agent. We will borrow results from this literature to solve for the benchmark case in the absence of censorship.

## 1.3 The model

Two risk neutral players, an agent (she) and an evaluator (he), play a game in continuous time $t \in [0, \infty)$. Their discount rates are $\rho_0 > 0$ and $\rho_1 > 0$, respectively.

At time $t = 0$, nature chooses the type of the agent $\theta$ and the type of the project $\alpha$, both from $\Theta = \{G, B\}$. We assume that $\theta = \alpha$, so that nature's choices are perfectly correlated. We assume that the agent observes nature's choice, while the evaluator does not. Let $p_0 \in (0, 1)$ denote the probability that nature chooses a type $G$ project/agent.

The agent enjoys a flow payoff $w > 0$ that is independent with her type while she stays in her job, and has no payoffs after she is dismissed by the evaluator. Only the type $G$ agent can succeed in her project. A success is publicly observable, and it arrives according to a Poisson process $\boldsymbol{S} = \{S_t\}_{t \geq 0}$ with an arrival rate $\gamma > 0$. The first success reveals that the type of the agent is $G$.

A success yields a lump-sum payoff $k > 0$ to the evaluator. The evaluator can choose a time $t \geq 0$ to irreversibly dismiss the agent. After the dismissal, the game ends, and the evaluator receives his outside option, the present discounted value of which is normalized to $m$.[8] We assume $h := \gamma k > m > 0$. Thus, the evaluator prefers the type $G$ agent to the outside option, and prefers the outside option to the type $B$ agent.

The project also potentially produces adverse news events, such as breakdowns. The arrival rate of such news depends upon the type of the project. If the project's type is $G$, it generates a piece of news at each jumping time of a Poisson process $\boldsymbol{N}^G = \{N_t^G\}_{t\geq 0}$ with an arrival rate $\beta^G \geq 0$. Conditional on the project's type being $G$, the success process $\boldsymbol{S}$ and the news process $\boldsymbol{N}^G$ are independent. If the project's type is $B$, it generates a piece of news at each jumping time of a Poisson process $\boldsymbol{N}^B = \{N_t^B\}_{t\geq 0}$ with an arrival rate $\beta^B$, where $\beta^B > \beta^G$. We call a piece of such news *bad news* since it happens more often to the type $B$ agent, though it has no payoff consequence. We also call a success *good news*.

The agent can observe the bad news process. When a piece of bad news arrives, she can incur a lump-sum cost $c > 0$ to censor it. The evaluator can observe bad news if and only if the agent does not censor it, but he cannot distinguish whether a piece of bad news comes from $\boldsymbol{N}^G$ or $\boldsymbol{N}^B$, unless $\beta^G = 0$.

[8]It can also be interpreted as the flow payoff from the outside option, since they are equal after normalization (i.e. $\int_0^\infty \rho_1 e^{-\rho_1 t} m \, \mathrm{d}t = m$).

Let $X_t^\theta \in \{1, 0\}$ be the type $\theta \in \Theta$ agent's censoring decision at time $t \geq 0$ when a piece of bad news has arrived at time $t$; $X_t^\theta = 1$ denotes censoring. A piece of bad news is *revealed* to the evaluator if it is not censored by the agent.

**Histories and Strategies** – Some care must be taken while defining the agent's and the evaluator's decision nodes in the game. Before the game ends,[9] the agent needs to make her censoring decisions only at dates when a piece of bad news arrives. On the other hand, the evaluator can dismiss the agent at any time.

At time $t$, a private history of the type $\theta \in \Theta$ agent is denoted by $h_t^\theta$. It consists of a finite sequence of news realizations before and including date $t$, and her censoring decision for those bad news realizations before but not including date $t$. Let $\bar{h}_t^\theta$ be a typical private history for her at $t$ when a piece of bad news has just arrived at $t$. Her strategy specifies a censoring probability $x_t^\theta \in [0, 1]$ at time $t$ for each history $\bar{h}_t^\theta$; this strategy $\boldsymbol{x}^\theta = \{x_t^\theta\}_{t \geq 0}$ is progressively measurable with respect to the filtration induced by those histories. Additional requirements on the strategies will be imposed later to ensure that public beliefs are well-defined.

The evaluator only observes the censored news processes (i.e. the public history). At time $t$, a public history $h_t$ is a finite sequence of news realizations that have not been censored before and including date $t$. Let $\mathscr{F} = \{\mathscr{F}_t\}_{t \geq 0}$ be the filtration induced by those public histories. The evaluator's pure strategy

[9]All strategies are defined conditional on the game has not yet ended.

at time $t$ for a public history $h_t$ is a $\mathscr{F}$-stopping time $T^t$,[10] at which time he dismisses the agent.

Pure strategies are not sufficient to study equilibria. Thus, we introduce mixed strategies, which are cumulative distribution functions over stopping times. To be specific, a mixed strategy for the evaluator at time $t$ for a public history $h_t$ is a $\mathscr{F}$-adapted process $\boldsymbol{r}^t = \{r^t_\nu\}_{\nu \geq t}$, such that pathwise,

a) $r^t_\nu$ is non-decreasing and right continuous in $\nu \geq t$, and takes values in $[0, 1]$;

b) for any time $t^{'} \geq t$, and $\nu \geq t^{'}$, $r^{t^{'}}_\nu$ is related to $r^t_\nu$ as follows:

$$r^t_\nu = r^t_{t^{'}\text{-}} + (1 - r^t_{t^{'}\text{-}}) r^{t^{'}}_\nu, \quad \text{if} \quad r^t_{t^{'}\text{-}} < 1,$$

where $r^t_{t^{'}\text{-}} = \lim_{s \uparrow t^{'}} r^t_s$ for $s > t$, and $r^t_{t-} = 0$.

a) requires $r^t$ to be a cumulative distribution function over the stopping times $T^t$. b) requires $r^t$ to be time consistent,[11] but imposes no restriction for $r^{t^{'}}_\nu$ if $r^t_{t^{'}\text{-}} = 1$. We also call $\frac{\mathrm{d}r^t_\nu / \mathrm{d}\nu}{1 - r^t_\nu}$ the (instantaneous) hazard rate whenever it exists.

**Beliefs** – The public belief at time $t$ about the agent's type being $G$ is $p_t = \mathbb{P}[\theta = G | \mathscr{F}_t]$, where the probability measure is induced by the public conjectured strategy $\tilde{\boldsymbol{x}}^\theta = \{\tilde{x}^\theta_t\}_{t \geq 0}$ of the type $\theta$ agent. In equilibrium, we

[10]The superscript $t$ means the strategy is defined for a history at time $t$. This also applies to the strategy $\boldsymbol{r}^t$ which will be defined later.

[11]See Laraki et al. (2005) and Riedel and Steg (2017) for details.

must have $\tilde{x}_t^\theta = \mathbb{E}[x_t^\theta | h_t]$ for any $t$ and and $h_t$. When a success arrives, the public belief jumps to 1. If at time $t$ a piece of bad news is revealed to the evaluator, the public belief jumps from $p_{t\text{-}}$ to

$$J(p_{t\text{-}}, \tilde{x}_t^G, \tilde{x}_t^B) := \frac{p_{t\text{-}}\beta^G(1-\tilde{x}_t^G)}{p_{t\text{-}}\beta^G(1-\tilde{x}_t^G) + (1-p_{t\text{-}})\beta^B(1-\tilde{x}_t^B)},$$

when the denominator is positive. Observe that in the absence of censorship (i.e. $\tilde{x}_t^G = \tilde{x}_t^B = 0$), a piece of bad news makes the public belief jump from $p_{t\text{-}}$ down to $j(p_{t\text{-}}) := J(p_{t\text{-}}, 0, 0) < p_{t\text{-}}$.

If no news is revealed to the evaluator over the time interval $[t, t+\mathrm{d}t)$, the public belief $p_t$ evolves continuously according to the following ordinary differential equation (ODE),

$$\dot{p}_t = g(p_t, \tilde{x}_t^G, \tilde{x}_t^B) := -p_t(1-p_t)[\gamma + (1-\tilde{x}_t^G)\beta^G - (1-\tilde{x}_t^B)\beta^B]. \quad \text{(ODE)}$$

To make sure the above (ODE) with initial condition $p_0 \in (0,1)$ admits a well-defined solution, we require $\tilde{\boldsymbol{x}}^\theta$ to satisfy the following requirement. For $\theta \in \Theta$, we say $\tilde{\boldsymbol{x}}^\theta$ is *admissible* if for any public history $h_{t'}$ and $t' > 0$, $\{\tilde{x}_t^\theta\}_{t \le t'}$ is piecewise continuous in $t$ with finite cutoffs, and it is right continuous with left limits at the cutoffs. Given any admissible $\tilde{\boldsymbol{x}}^G$ and $\tilde{\boldsymbol{x}}^B$, the ODE with initial condition $p_0 \in (0,1)$ admits a unique solution (or trajectory) $\boldsymbol{p} = \{p_t\}_{t\ge 0}$ which is piecewise continuously differentiable.[12]

[12]Given piecewise continuous $\tilde{\boldsymbol{x}}^\theta$, it is easy to see that $g(p, \tilde{x}_t^G, \tilde{x}_t^B)$ and $g_p(p, \tilde{x}_t^G, \tilde{x}_t^B)$ are piecewise continuous in $t$ for a fixed $p$, and $g(p, \tilde{x}_t^G, \tilde{x}_t^B)$ is continuously differentiable in $p$ for a fixed $t$. Thus, the standard local existence and uniqueness results for ODE ensure a unique solution in the interval where both $\tilde{x}_t^G$ and $\tilde{x}_t^B$ are continuous. Also, a unique global solution can be obtained through concatenation.

**Payoffs** – For $\bar{h}_t^\theta$, given $\boldsymbol{r}^t$ and $\boldsymbol{x}_t^\theta = \{x_\nu^\theta\}_{\nu \geq t}$, the type $\theta$ agent's expected payoff at time $t$ is

$$v_\theta^{\boldsymbol{r},\boldsymbol{x}^\theta}(\bar{h}_t^\theta) = \mathbb{E}\left[ -\rho_0 \; c \; X_t^\theta + \int_t^{T^t} \rho_0 \; e^{-\rho_0(\nu - t)}(w \, \mathrm{d}\nu - c \; X_\nu^\theta \, \mathrm{d}N_\nu^\theta) \middle| \; \bar{h}_t^\theta \right],$$

where the expectation is taking over $T^t$, $\{X_\nu^\theta\}_{\nu \geq t}$, and $\{N_\nu^\theta\}_{\nu \geq t}$.

For $h_t$, given $\boldsymbol{r}^t$ and $\tilde{\boldsymbol{x}}_t^\theta = \{\tilde{x}_\nu^\theta\}_{\nu \geq t}$, the evaluator's expected payoff at time $t$ is

$$u^{\boldsymbol{r},\tilde{\boldsymbol{x}}}(h_t) = \mathbb{E}\left[ \int_t^{T^t} \rho_1 \; e^{-\rho_1(\nu - t)} \; \mathbb{1}_{\theta = G} \, k \, \mathrm{d}S_\nu + e^{-\rho_1(T^t - t)} \; m \middle| \; h_t \right],$$

where the superscript $\tilde{\boldsymbol{x}} := \{\tilde{\boldsymbol{x}}^G, \tilde{\boldsymbol{x}}^B\}$, and the expectation is taking over $T^t$, $\{S_\nu\}_{\nu \geq t}$, $\{\tilde{X}_\nu^\theta\}_{\nu \geq t}$, and $\{N_\nu^\theta\}_{\nu \geq t}$.

Let $V_\theta^{\boldsymbol{r}}(\bar{h}_t^\theta) = \sup_{\{x_\nu^\theta\}_{\nu \geq t}} v_\theta^{\boldsymbol{r},\boldsymbol{x}^\theta}(\bar{h}_t^\theta)$ be the value function of the type $\theta$ agent, and $U^{\tilde{\boldsymbol{x}}}(h_t) = \sup_{\{r_\nu^t\}_{\nu \geq t}} u^{\boldsymbol{r},\tilde{\boldsymbol{x}}}(h_t)$ be the value function of the evaluator, given the strategies of other players.

**Equilibria** – We study Perfect Bayesian Equilibria (PBE), which consist of strategies $\boldsymbol{x}^\theta = \{x_t^\theta\}_{t \geq 0}$ and $\boldsymbol{r} = \{r_\nu^t\}_{t \geq 0, \nu \geq t}$, a public belief $\boldsymbol{p} = \{p_t\}_{t \geq 0}$, and the conjectured strategy $\tilde{\boldsymbol{x}}^\theta = \{\tilde{x}_t^\theta\}_{t \geq 0}$ such that

1. For each $\bar{h}_t^\theta$, $\boldsymbol{x}_t^\theta$ is optimal for the type $\theta$ agent, given $\boldsymbol{p}$ and $\boldsymbol{r}$;
2. For each $h_t$, $\boldsymbol{r}^t$ is is optimal for the evaluator, given $\boldsymbol{p}$ and $\tilde{\boldsymbol{x}}^\theta$;
3. $\boldsymbol{p}$ is updated according to $\tilde{\boldsymbol{x}}^\theta$ and Bayes rule whenever possible, and $\tilde{\boldsymbol{x}}^\theta = \mathbb{E}[\boldsymbol{x}^\theta | \mathscr{F}]$.

It is well known in the two-armed bandit literature that the optimal allocation rule is a cutoff rule with respect to the public belief. In our setup, cutoff strategies are of particular interest. We say the evaluator's strategy is a pure cutoff strategy with a cutoff public belief $\hat{p} \in (0,1)$ if it is a pure strategy $T^t = \inf\{\nu \geq t : p_\nu \leq \hat{p}\}$ for any public history $h_t$. This means that the evaluator dismisses the agent whenever the public belief is below or equal to the cutoff $\hat{p}$. The evaluator's strategy is a mixed cutoff strategy with a cutoff $\hat{p} \in (0,1)$ if for any public history $h_t$ and $\nu \geq t$,

$$r_\nu^t = \begin{cases} 1 & \exists s \in [t,\nu], p_s < \hat{p}, \\ 0 & \forall s \in [t,\nu], p_s > \hat{p}. \end{cases}$$

This means that the evaluator never dismisses the agent when the public belief is above the cutoff $\hat{p}$, but dismisses the agent immediately when the belief is below $\hat{p}$.

Similarly, we say the type $\theta$ agent's strategy is a pure (resp. mixed) cutoff strategy with a cutoff $\hat{p} \in (0,1)$ if for any history $\bar{h}_t^\theta$,

$$x_t^\theta = \begin{cases} 1 & \text{if } p_t \in (\hat{p}, 1), \\ 0 & \text{if } p_t \leq \hat{p} \text{ (resp. if } p_t < \hat{p}). \end{cases}$$

This means that the agent censors bad news when the public belief is above the cutoff $\hat{p}$ (but below 1), and she stops censoring when the belief is below $\hat{p}$. The difference between the pure and the mixed strategies is that the mixed strategy allows mixing at the cutoff belief.[13] We call a PBE a *cutoff equilibrium* if the evaluator uses a cutoff strategy.

[13] A mixed cutoff strategy first needs to satisfy the definition of a strategy; thus, the mixing is also a contingent plan.

## 1.4 Conclusive bad news

We first study the baseline model when the type $G$ agent does not have any bad news (i.e. $\beta^G = 0$). Thus, the type $G$ agent is a passive player who does not have any action. A piece of revealed bad news is conclusive evidence to the evaluator that the agent's type is $B$. In this case, we will show that censorship not only always makes the evaluator and the type $G$ agent worse off, but also sometimes makes the type $B$ agent worse off.

Now a revealed signal (good or bad news) resolves all uncertainty about the type of the agent. Hence, after seeing a piece of good news, the evaluator never dismisses the agent. After seeing a piece of revealed bad news, the evaluator dismisses the agent immediately. It follows that the type $B$ agent will not censor any bad news after a piece of bad news has already been revealed to the evaluator, even if the evaluator has deviated and has not yet dismissed her. Therefore, what remains to be determined is simply (1) the evaluator's strategy under the public history at time $t$ (denoted by $\varnothing_t$) where no signal has been revealed to the evaluator, and (2) the type $B$ agent's strategy under her private history at time $t$ (denoted by $\bar{\varnothing}_t^B$) where no signal has been revealed to the evaluator as yet and a piece of bad news has just arrived. We say the evaluator experiments when he does not dismiss the agent under the public history $\varnothing_t$.

### 1.4.1 No censorship when the censoring cost is high

A prohibitively high cost will prevent censorship. This section determines the threshold cost above which censorship does not exist in equilibrium.

Observe that it is strictly dominant for the evaluator to dismiss the agent when he sees bad news. Given this restriction, the best possible strategy of the evaluator for the type $B$ agent is one where she is only dismissed when bad news arises. We use $\boldsymbol{r}_\infty$ to denote this strategy. Thus, the dismissal time is the first time that the type $B$ agent does not censor a piece of bad news when it arrives. Given this strategy, the type $B$ agent faces a stationary problem: consider the type $B$ agent's problem at the history $\bar{\varnothing}^B_t$. The agent can choose between censoring or not. Her value function satisfies the following Bellman equation,

$$V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B_t) = \max\left\{\mathbb{E}[-\rho_0\ c + \int_0^{\tau^B} \rho_0\ e^{-\rho_0 \nu} w\,\mathrm{d}\nu + e^{-\rho_0 \tau^B} V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B_{\tau^B})], 0\right\},$$

where $\tau^B$ is the arrival time of the next piece of bad news from the type $B$ agent. She will be dismissed and obtain 0 payoff if she does not censor this piece of bad news. Otherwise, by censoring it, she incurs a cost $\rho_0 c$, but can stay in her job until the next piece of bad news arrives at $\tau^B$, and enjoy her continuation value after $\tau^B$. Clearly, $V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B_t)$ does not depend on $t$. Thus we use $V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B) := V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B_t)$.

It is easy to check that

$$V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}^B) = \max\left\{w - (\rho_0 + \beta^B)c, 0\right\}.$$

Hence, the type $B$ agent has a strict incentive not to censor any bad news if and only if $c > \underline{c}$, where $\underline{c} = \frac{w}{\rho_0 + \beta^B}$. Since any equilibrium strategy of the evaluator gives a weakly lower continuation payoff to the type $B$ agent than $\boldsymbol{r}_\infty$, she will never censor any bad news when $c > \underline{c}$. Consequently, the evaluator faces a decision problem in the absence of censorship. His optimal policy is a cutoff strategy in which the cutoff belief is determined by the evaluator's trade-off between exploitation and exploration, as demonstrated in the bandit literature. The following proposition summarizes the result.

*Proposition* 1.1. Assume $\beta^G = 0$ and $c > \underline{c}$. There is a unique PBE with the following features:

- Assume that $\gamma > \beta^B$. The type $B$ agent never censors any bad news, and the evaluator dismisses the agent whenever the public belief is below $p_{fast} = \frac{\rho_1 m}{\rho_1 h + \gamma(h-m)}$.
- Assume that $\gamma < \beta^B$. The type $B$ agent never censors any bad news, and the evaluator dismisses the agent whenever the public belief is below $p_{slow} = \frac{\rho_1 m}{\rho_1 h + \beta^B(h-m)}$.

Notice that we have made a distinction about what kind of news arrives faster. We say *good news arrives faster* if it arrives faster than bad news from the type $B$ agent (i.e. $\gamma > \beta^B$), and *good news arrives slower* if it arrives slower than bad news from the type $B$ agent (i.e. $\gamma < \beta^B$). This distinction is often made in Poisson bandit models where news arrives according to exogenous

Poisson processes. It determines the direction of the evolution of the public belief in the absence of news. In a Poisson bandit problem, if good news arrives faster, then the public belief drifts up in the absence of news, and the cutoff belief in the evaluator's optimal strategy is $p_{fast}$. This is similar to the breakthrough case as in Keller et al. (2005). If good news arrives slower, then the public belief drifts down in the absence of news, and the cutoff belief in the evaluator's optimal strategy is $p_{slow}$. This is similar to the breakdown case as in Keller and Rady (2015). We will see later that this distinction is also useful for our analysis where the news process is endogenously determined by the equilibrium censorship strategy.

This equilibrium also demonstrates what would happen if censorship were not possible, i.e. the *no censorship benchmark* (NCB). To analyze the welfare effect of censorship later in the paper, we often compare a player's equilibrium payoff with his or her expected payoff in the NCB.

For the rest of the baseline model (conclusive bad news), we will assume that the censoring cost is not prohibitively high: $0 < c < \underline{c}$. We again distinguish two cases, according to the direction of drift of the posterior belief in the absence of news and in the absence of censorship.

### 1.4.2 Faster good news

#### 1.4.2.1 Equilibrium

First we consider the case where good news arrives faster (i.e. $\gamma > \beta^B$). For such a project, (ODE) indicates that the public belief drifts down in the absence of news, regardless of the agent's censorship strategy. The greater the intensity of censorship, the faster the public belief declines. However, the evaluator's optimal strategy (as a function of the public belief) does not depend on how much information has been censored, as illustrated in Figure 1.1.

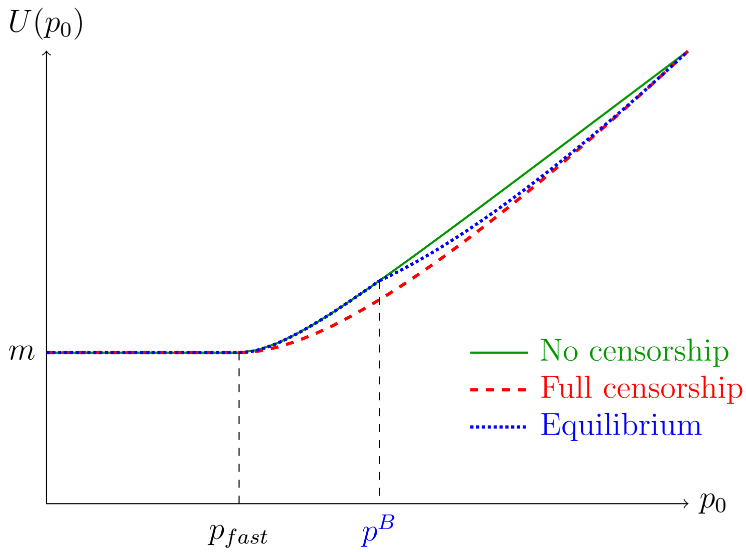


Figure 1.1: The evaluator's value functions (conclusive bad news & faster good news)

Figure 1.1 shows the evaluator's value functions under different censorship policies. The horizontal axis is the prior belief. Consider two extreme censorship policies – no censorship (green solid curve) where the agent never censors bad news, and full censorship (red dashed curve) where the agent always censors bad news. The no censorship policy gives the evaluator the highest payoff for any prior belief, while the full censorship policy gives him the lowest payoff. The evaluator's best response as a function of the public belief, however, is actually the same for those two extreme censorship policies, i.e. he dismisses the agent if and only if the public belief is below $p_{fast}$. This is because the evaluator learns from good news at the margin in both cases – he will dismiss the agent if no news is revealed or if a piece of bad news is revealed, so only good news can change his action. Hence, the optimal cutoff belief $p_{fast}$ does not depend on the arrival of bad news, and therefore is the same for the two extreme cases. From the evaluator's perspective, the informativeness of any censorship strategy of the agent lies between those two extreme strategies, in term of the Blackwell order, at each instant. Thus, for any admissible strategy of the agent, the evaluator's value must be in between his values when he faces those two extreme censorship policies. Hence, he will use the same cutoff strategy. The blue dotted curve in Figure 1.1 illustrates the evaluator's value function when the type $B$ agent uses a pure cutoff strategy with the cutoff $p^B$. We will show later in Proposition 1.2 that this strategy is also the evaluator's equilibrium strategy. We should keep in mind that although the evaluator uses the same cutoff strategy, the greater the agent censors, the faster the same cutoff belief is reached in the absence of news.

Given the evaluator's strategy, we can pin down the type $B$ agent's strategy and the unique PBE.

*Proposition* 1.2. Assume $\beta^G = 0$, $c < \underline{c}$ and $\gamma > \beta^B$. There exists a unique PBE. The equilibrium features a pair of pure cutoff strategies.

1. The evaluator dismisses the agent whenever the public belief is below $p_{fast}$;

2. The type $B$ agent censors bad news whenever the public belief is above $p^B$, where $p^B > p_{fast}$.

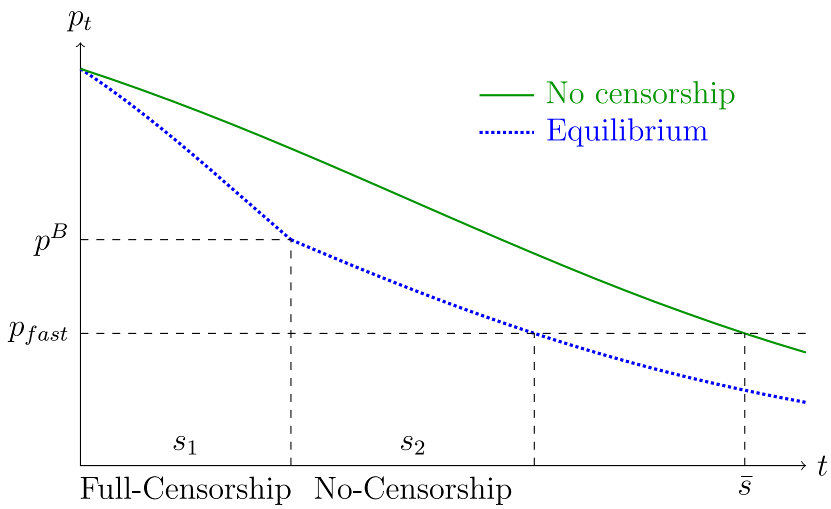


Figure 1.2: The belief evolution in the absence of news (conclusive bad news & faster good news)

Figure 1.2 shows the evolution of the public belief in the absence of news in the equilibrium (blue dotted curve). The evaluator will dismiss the agent whenever the public belief is below $p_{fast}$, which will happen eventually without the arrival of good news. Hence, the type $B$ agent knows that she will be dismissed in the end. As the public belief approaches the threshold $p_{fast}$, she knows that she has little time left before she is fired. Her benefit of censoring – the extra time she is able to stay on the job – becomes relatively small, compared with the cost of censoring. Thus, the type $B$ agent will give up censorship before the public belief reaches $p_{fast}$. In addition, before giving up censorship, the type $B$ agent would censor all bad news with probability one. This is because the type $B$ agent's continuation value after censoring is increasing in the public belief which is a decreasing function of time. Whenever it is weakly optimal to censor bad news at time $t$ before the public belief drifts down to $p_{fast}$, it is strictly optimal to censor all bad news before time $t$. Hence, the censorship policy has two phases. At the beginning when the public belief is high (i.e. higher than $p^B$), the agent censors all bad news, we call it the Full-Censorship period; the length of this period is denoted by $s_1$ in Figure 1.2. After the Full-Censorship period when the public belief is low (i.e. lower than $p^B$), the agent stops censoring. We call this the No-Censorship period; the length of this period is stochastic since it is ended by the arrival of bad news, and the maximal length is denoted by $s_2$ in Figure 1.2. If the agent could commit to a policy of never censoring, then her maximal duration on the job would be $\bar{s}$ in Figure 1.2.

The following indifference condition determines the threshold at which the agent stops censoring:

$$\mathbb{E}[-\rho_0 \ c + \int_0^{s_2 \wedge \tau^B} \rho_0 \ e^{-\rho_0 \nu} w \, \mathrm{d}\nu] = 0,$$

where $s_2$ the maximal length of the No-Censorship period, and $\tau^B$ is the arrival time of the next piece of bad news from the type $B$ agent. Thus, the expected continuation value in the No-Censorship period must be equal to the censoring cost. Clearly, the larger the censoring cost, the sooner it becomes not profitable to censor bad news, and the longer the No-Censorship period is. At last, $s_2$ determines the cutoff belief $p^B$, at which the agent switches from the Full-Censorship policy to the No-Censorship policy.

#### 1.4.2.2 Welfare: Equilibrium versus NCB

We now consider the welfare consequence of censorship. Specifically, we compare two scenarios: players' equilibrium payoffs, and their expected payoffs in the absence of censorship (i.e. their payoffs in the NCB). Since the equilibrium strategy of the type $B$ agent is to not censor any bad news when $p_t \leq p^B$, nothing would be different between these two scenarios. Thus, let us only consider $p_0 > p^B$.

The evaluator is worse off in equilibrium, compared with the NCB.[14] Figure 1.1 illustrates the evaluator's welfare loss due to censorship. This is because censorship garbles information. No news in equilibrium means either nothing really happened, or bad news has arrived but has been censored. Hence, the public belief drifts down faster when no news is revealed, compared with the NCB. The type $G$ agent is also worse off in equilibrium, since, in order to stay in her job, she now has to have a success within a shorter period of time before the public belief drifts down below the evaluator's cutoff belief.

It is worth noticing that the type $B$ agent may also be worse off in equilibrium. Since the type $B$ agent can survive bad news when she censors in the Full-Censorship period, censorship does provide her a higher expected flow payoff. This is her benefit from censorship. However, the cost is that the possibility of censorship also drives the public belief down faster in the absence of news, compared with the NCB. Another way to look at this is the

[14]We always compare the equilibrium payoff with the payoff in the NCB, unless stated otherwise. Hence, we will not repeat it every time.

martingale property of the public belief. Since bad news would be censored, the fact that the public belief cannot jump to 0 means that the drifting down process must be accelerated. We can show that whenever the actual censoring period (i.e. the Full-Censorship period) is short, the benefit from censorship is dominated by the cost. Hence, we have the following result.

*Proposition* 1.3. Assume $\beta^G = 0$, $c < \underline{c}$, $\gamma > \beta^B$ and $p_0 > p^B$. There exists a $s_1^* > 0$ such that the type $B$ agent is worse off in equilibrium than she is in the NCB if and only if $s_1 < s_1^*$.

The Full-Censorship period is an equilibrium object, but we can express this result in terms of primitive variables. The most straightforward comparative statics is for the prior belief. When the prior belief is high, the Full-Censorship period is long, thus the type $B$ agent is better off in equilibrium. When the prior belief is low but higher than $p^B$, it means the Full-Censorship period is short, thus the type $B$ agent is worse off in equilibrium.[15]

Another interesting comparative statics is for the censoring cost. The censoring cost stands for how hard it is to hide evidence from the evaluator. It also represents the strength of anti-censorship institutions. The cost directly determines when the agent is willing to give up censorship (i.e. the threshold belief $p^B$) and the length of the No-Censorship period. Thus, it also indi-

[15]Similarly, if the value of the outside option $m$ is low, then the type $B$ agent is better off in equilibrium. If $m$ is high but not too high such that the Full-Censorship period still exists, then the type $B$ agent is worse off in equilibrium.

rectly determines how long the agent censors bad news (i.e. the length of the Full-Censorship period). For a given prior belief, if the censoring cost is very high (but still below $\underline{c}$), it means the No-Censorship period is very long such that the Full-Censorship period disappears. Then, obviously, the equilibrium payoff is the same as the payoff in the NCB. If the censoring cost is very low, it means the No-Censorship period is very short, which in turn means the Full-Censorship is long. If it is long enough, then the type $B$ agent is better off in equilibrium. However, when the censoring cost is intermediate, then the Full-Censorship exists but is short, and the type $B$ agent is worse off in equilibrium. The result is summarized below.

*Corollary* 1.1. Assume $\beta^G = 0$, $c < \underline{c}$, $\gamma > \beta^B$ and $p_0 > p_{fast}$. There exists a $0 \leq \mathfrak{c}_1 < \mathfrak{c}_3 < \bar{c}$, the type $B$ agent is better off (resp. worse off) in the equilibrium than she is in the NCB when $c \in (0, \mathfrak{c}_1)$ (resp. when $c \in (\mathfrak{c}_1, \mathfrak{c}_3)$). In addition, $\mathfrak{c}_1 > 0$ if and only if $\bar{s} > s^*$ for some $s^* > 0$.

Another way to understand the censoring cost is to examine the dependence of the equilibrium payoff with respect to the cost, which is shown in the following result.

*Corollary* 1.2. Assume $\beta^G = 0$, $c < \underline{c}$, $\gamma > \beta^B$ and $p_0 > p_{fast}$. There exists a $\mathfrak{c}_2 \in [\mathfrak{c}_1, \mathfrak{c}_3)$, the equilibrium payoff of the type $B$ agent is decreasing (resp. increasing) in $c$ when $c \in [0, \mathfrak{c}_2]$ (resp. when $c \in [\mathfrak{c}_2, \mathfrak{c}_3]$). In addition, there exists a $s^{**} \in (0, s^*)$, such that $\mathfrak{c}_2 > \mathfrak{c}_1 \geq 0$ (resp. $\mathfrak{c}_2 = \mathfrak{c}_1 = 0$) when $\bar{s} > s^{**}$

(resp. when $\bar{s} \leq s^{**}$).

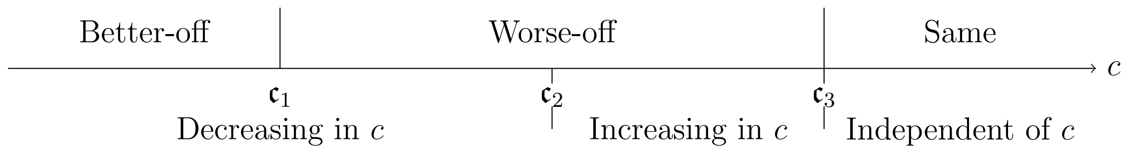


Figure 1.3: Comparative statics with respect to the censoring cost

The censoring cost servers as a "commitment" device for the agent. An increase in the censoring cost decreases the type $B$ agent's expected flow payoff in the Full-Censorship period since censoring is more costly, but it increases her payoff in the No-Censorship period. Moreover, an increase in the cost also increases the length of the No-Censorship period, but decreases the length of the Full-Censorship period.

As we illustrated in Corollary 1.1, when the censoring cost is very small (i.e. $c \in (0, \mathfrak{c}_1)$), the Full-Censorship period is relatively long, and the type $B$ agent is better off in equilibrium. In this case, the negative effect of an increasing censoring cost would dominate the positive effect, since the expected flow payoff in the Full-Censorship period is more important. In this case, the type $B$ agent's equilibrium payoff is decreasing in the censoring cost. This means that the type $B$ agent who is better off with censorship must prefer a weaker institution (i.e. a smaller censoring cost).

The inverse statement is not true. When the censoring cost is high such that the Full-Censorship exists but is short (i.e. $c \in (\mathfrak{c}_1, \mathfrak{c}_3)$), the type $B$ agent is worse off in equilibrium. However, it is not necessary for her to prefer a stronger institution. In general, her equilibrium payoff is not monotonic in the cost. Only when the cost is high enough (i.e. $c \in (\mathfrak{c}_2, \mathfrak{c}_3)$), the commitment (positive) effect of an increasing censoring cost dominates the negative effect. In this case, her equilibrium payoff is increasing in the censoring cost, and a stronger institution would be preferred by the type $B$ agent.

### 1.4.3 Slower good news

#### 1.4.3.1 Equilibrium

Now we turn to the other case where good news arrives lower (i.e. $\gamma < \beta^B$). In the NCB, since the evaluator would expect to see bad news more often, the public belief will drift up when no news arrives. Hence, "no news is good news", and the evaluator only dismisses the agent when a piece of bad news arrives, if he starts with a prior belief that is not too low ($p_0 > p_{slow}$).

However, the possibility of censorship changes everything. Foremost, the evaluator's optimal dismissal policy does depend on the agent's censorship policy, which is different from the case where good news arrives faster. In the NCB, the evaluator's optimal policy is to dismiss the agent whenever the public belief is below $p_{slow}$. However, if the agent censors all bad news, then the evaluator's optimal policy is to dismiss the agent when the public belief is below $p_{fast}$, as shown in the last section. Hence, censorship discourages the evaluator's incentive for exploring the competence of the agent. In addition, we also know from the above two extreme cases that the evaluator would never dismiss the agent when the public belief is higher than $p_{fast}$, and he would dismiss the agent immediately when the public belief is below $p_{slow}$, since again the evaluator's value function for any censorship policy is in between his values given by the two extreme censorship policies. This is illustrated in Figure 1.4.

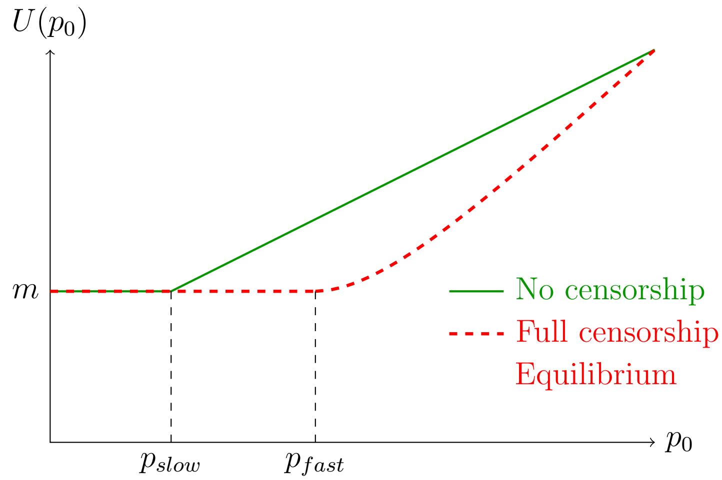


Figure 1.4: The evaluator's value functions (conclusive bad news & slower good news)

When the public belief is higher than $p_{fast}$, perhaps ironically, the fact that the type $B$ agent can keep her job for sure also implies that she must censor all bad news in equilibrium when the public belief is above $p_{fast}$. First, when no news is revealed to the evaluator, the public belief must eventually drift down below $p_{fast}$, otherwise the type $B$ agent can stay in her job forever by censoring all bad news, which in turn drives the public belief down – a contradiction. Second, the only way that the public belief drifting down below $p_{fast}$ is that the agent censors bad news before the public is below $p_{fast}$. In addition, if the agent censors bad news at a public belief that is above $p_{fast}$, she must also find it optimal to censor bad news when the public belief is even higher, since a longer stay in her job gives her a higher incentive for censorship. Those together imply that the type $B$ agent will censor all bad news when the public belief is above $p_{fast}$, and the only direction that the public belief drifts is downward, when it is above $p_{fast}$. Essentially, the evaluator faces the worst information, and would use a cutoff $p_{fast}$ to dismiss the agent.

However, a pure strategy that dismisses the agent when the cutoff belief hits $p_{fast}$ cannot be supported in equilibrium. If the evaluator uses such a pure strategy, he needs to dismiss the agent deterministically if no news arrives for some time, since the type $B$ agent censors all bad news when the public belief is above $p_{fast}$. However, this means that the type $B$ agent would like to give up censorship even when the public belief is above $p_{fast}$, if the evaluator would dismiss her deterministically in a short time. Such a contradiction implies that we need a mixed strategy for the evaluator. This also means that the type $B$ agent must also mix at the cutoff belief $p_{fast}$, since censoring too much would induce the evaluator to use the pure strategy, while censoring too little would induce the public belief to drift up above $p_{fast}$. The former contradicts with the fact that the evaluator must uses a mixed strategy, and the latter contradicts with the fact that belief cannot drift up above $p_{fast}$.

Hence, the unique cutoff equilibrium is summarized in the following result.

*Proposition* 1.4. Assume $\beta^G = 0$, $c < \underline{c}$ and $\gamma < \beta^B$. There exists a unique cutoff equilibrium, which features a pair of mixed cutoff strategies with a common cutoff $p_{fast}$.

1. The evaluator dismisses (resp. does not dismiss) the agent when the public belief is below (resp. above) $p_{fast}$. He uses a constant hazard rate $z^* := \frac{w}{c} - \rho_0 - \beta^B$ to dismiss the agent at the cutoff belief $p_{fast}$.

2. The type $B$ agent censors (resp. does not censor) bad news when the public belief is above (resp. below) $p_{fast}$. She censors bad news with a constant probability $x^{B*} := 1 - \frac{\gamma}{\beta^B}$ at the cutoff belief $p_{fast}$.

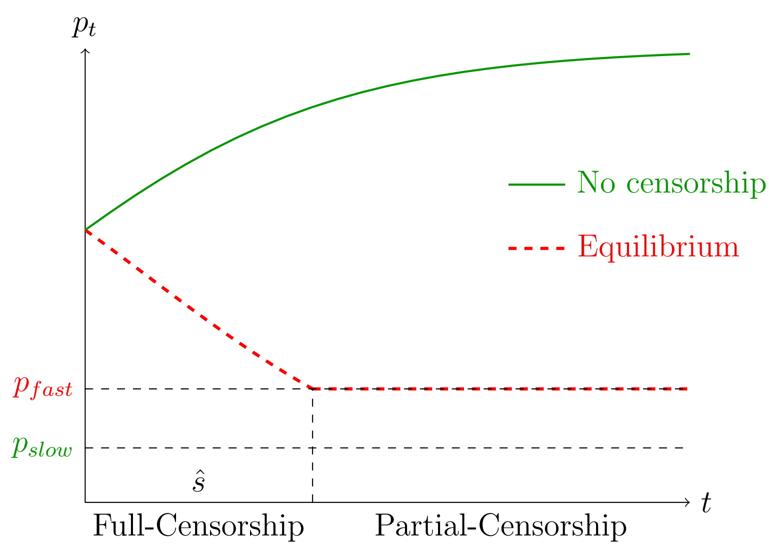


Figure 1.5: The belief evolution in the absence of news (conclusive bad news & slower good news)

Figure 1.5 shows the evolution of the public belief in the absence of news in the equilibrium (red dashed curve). The censorship policy also has two phases. Since the type $B$ agent censors all bad news when the public belief is above $p_{fast}$, we still have the Full-Censorship period at the beginning with length $\hat{s}$. When the public belief reaches $p_{fast}$, the type $B$ agent censors bad news with a constant probability such that the arrival rate of good news equals to the arrival rate of the non-censored bad news (i.e. $\gamma = (1-x^{B*})\beta^B$). Hence, the public belief would stay constant at $p_{fast}$, if no news arrives, which makes the problem becomes stationary. Moreover, the evaluator's constant hazard rate is also chosen to make the type $B$ agent indifferent between censoring and not. We call this period as the Partial-Censorship period.

This is the unique cutoff equilibrium. If the evaluator uses a cutoff strategy with a different cutoff belief $\hat{p}$, then it must be below $p_{fast}$, because the evaluator never dismisses the agent when the public belief is above $p_{fast}$. However, the exact argument which shows the type $B$ agent censors all bad news when the public belief is above $p_{fast}$ in equilibrium implies that now she must censor all bad news when the public belief is above $\hat{p}$. This means even less information is left for the evaluator. Hence, he does not want to use a cutoff strategy with a lower cutoff, since his incentive for learning is discouraged. Therefore, we obtain the unique cutoff equilibrium.

#### 1.4.3.2 Welfare: Equilibrium versus NCB

We still compare players' payoffs in the cutoff equilibrium with their payoffs in the NCB. Clearly, the evaluator is worse off in equilibrium, as illustrated in Figure 1.4. Actually, this is the worst payoff he can obtain from any censorship policy. Equivalently, the evaluator cannot use any information from bad news; he can only rely on good news. That is why his incentive for learning is discouraged so that he increases the cutoff belief from $p_{slow}$ to $p_{fast}$ in his cutoff strategy. This is a new effect that does not appear when the good news arrives faster; we call it the *discouragement effect*. Due to this discouragement effect, when the prior belief is in between $p_{slow}$ and $p_{fast}$, neither type of agent would have a chance to start her job in equilibrium, but she could at least have some positive time in her job in the NCB. Thus, when $p_0 \in (p_{slow}, p_{fast})$, both types of agent are worse off in equilibrium, compared with the NCB.

When the prior belief is above $p_{fast}$, the type $G$ agent would stay in her job forever in the NCB, while she could be dismissed with a strictly positive probability in the equilibrium. Hence, she is worse off in equilibrium, compared with the NCB. For the type $B$ agent, She has a higher payoff in the Full-Censorship period since she survives bad news, but a lower payoff in the Partial-Censorship period. Since her total payoff is the average payoff in those two periods, she is also worse off in equilibrium whenever the Full-Censorship period is short. We summarize the result below.

*Proposition* 1.5. Assume $\beta^G = 0$, $c < \underline{c}$ and $\gamma < \beta^B$.

1. When $p_0 \in (p_{slow}, p_{fast}]$, the type $B$ agent's is worse off in equilibrium than she is in the NCB;

2. When $p_0 > p_{fast}$, there exists a $\hat{s}^* > 0$ such that the type $B$ agent's is worse off in equilibrium than she is in the NCB if and only if $\hat{s} < \hat{s}^*$.

The Full-Censorship period, again, is an equilibrium object, but we can easily express the condition in terms of primitive variables. For example, when the prior belief is high, it means the Full-Censorship period is long, thus the type $B$ agent is better off in equilibrium. When it is low but still higher than $p_{fast}$, it means the Full-Censorship period exists but is short, thus the type $B$ agent is worse off in equilibrium.

For the censoring cost, different from the faster good news case, it does not determine the length of the Full-Censorship period, since the type $B$ agent switches her strategy to partial censorship at a particular belief $p_{fast}$ that does not depend on the cost. Hence, whether the type $B$ agent is better off in equilibrium does not depend on the cost. However, an increase in the censoring cost decreases the type $B$ agent's payoff in the Full-Censorship period since censoring is more costly, but increases her payoff in the Partial-Censorship period. Her total equilibrium payoff is the average payoff in those two periods. Therefore, if she is better off in equilibrium, it means the Full-Censorship period must be long and important. Hence, an increase in the cost decreases her equilibrium payoff. The inverse statement is also true. If the type $B$ agent is worse off in equilibrium, it means the Full-Censorship period must be short and not important. In this case, an increase in the cost increases her equilibrium payoff. Therefore, the type $B$ agent's payoff in equilibrium is monotonic in the cost, thus she prefers either a very strong institution or a very weak institution, depending on how long the Full-Censorship period is. The result is summarized below.

*Corollary* 1.3. Assume $\beta^G = 0$, $c < \underline{c}$ and $\gamma < \beta^B$.

- Whether the type $B$ agent is better off in the equilibrium than she is in the NCB does not depend on the censoring cost $c$;
- When $p_0 > p_{fast}$, the type $B$ agent's expected payoff in equilibrium is monotonic in the censoring cost $c$. Moreover, it is decreasing (resp. increasing) in the cost, if and only if, she is better off (resp. worse off) in equilibrium than in the NCB.

## 1.5 Inconclusive bad news

In this section, we consider the case where the good agent also has some bad news, but less frequently than the bad agent has (i.e. $0 < \beta^G < \beta^B$). Hence, in the NCB,[16] a piece of bad news is an informative but imperfect signal; the posterior public belief after a piece of bad news jumps down, but is above 0. Hence, after a piece of bad news, the evaluator knows that the agent is more likely to be a bad type, but he is not certain. The optimal policy for the evaluator's decision problem in the NCB is still a cutoff strategy, in which

[16] Though the NCB in the inconclusive news case is different from the NCB in the conclusive news case, we define them in the same way – the NCB means censorship is not possible, for example, due to a prohibitively high censoring cost.

the cutoff belief depends on the information structure. However, since bad news is a noisy signal, the evaluator may make both type I and type II errors based on bad news.

Inconclusive bad news both enriches and complicates the problem. We will only focus on new insights derived from inconclusive bad news, but will not fully characterize all equilibria, nor conduct case-by-case discussion. In particular, we will show that it is possible that censorship may benefit the evaluator, as well as both types of the agent.

### 1.5.1 Faster good news

#### 1.5.1.1 Equilibrium

We begin with the faster good news case; that is, good news arrives faster than the type $B$ agent's bad news (i.e. $\gamma > \beta^B$). In the NCB, the public belief will drift down if no news arrives.[17]

Now, both types of agent need to choose a censorship strategy. The basic trade-off does not change – the agent censors bad news if and only if her equilibrium continuation value is higher than the censoring cost. The cost is the same for both types of agent, however, the continuation value is different. The good agent has a higher continuation value than the bad agent due to the following two reasons. First, since the good agent has a chance to succeed in

[17] In the NCB, the direction of the belief drifting process is determined by the sign of $\gamma + \beta^G - \beta^B$. However, as we will show later, the corresponding equilibrium term is still determined by the sign of $\gamma - \beta^B$, since the type $G$ agent censors bad news more aggressively than the type $B$ agent does.

her project, she will stay in her job longer in expectation than the bad agent. Second, the same censorship strategy is less costly for the good agent than the bad agent, since bad news happens more often to the bad agent. Therefore, it is the good agent who has a higher incentive to censor bad news. This observation significantly changes the welfare effect of censorship, which will be discussed after we introduce the following equilibrium.

*Proposition* 1.6. Assume $c < \underline{c}$ and $\gamma > \beta^B$. There exists a $\bar{\beta} \in (0, \beta^B)$, such that for any $\beta^G \in (0, \bar{\beta}]$, there exists a PBE in which every player uses a pure cutoff strategy.

1. The evaluator dismisses the agent whenever the public belief is below $p_{fast}$;
2. The type $G$ agent censors bad news whenever the public belief is above $p^G$ (but not equal to 1), where $p^G > p_{fast}$;
3. The type $B$ agent censors bad news whenever the public belief is above $p^{B\dagger}$, where $p^{B\dagger} > p^G$.

Here we maintain the assumption that the censoring cost is low (i.e. $c < \underline{c}$). In addition, we assume that the good agent's bad news does not happen too often (i.e. $\beta^G \in (0, \bar{\beta}]$). Figure 1.6 depicts the evolution of the public belief in the absence of news in equilibrium (blue dotted curve) and in the NCB (green solid curve). The first observation is that the evaluator's

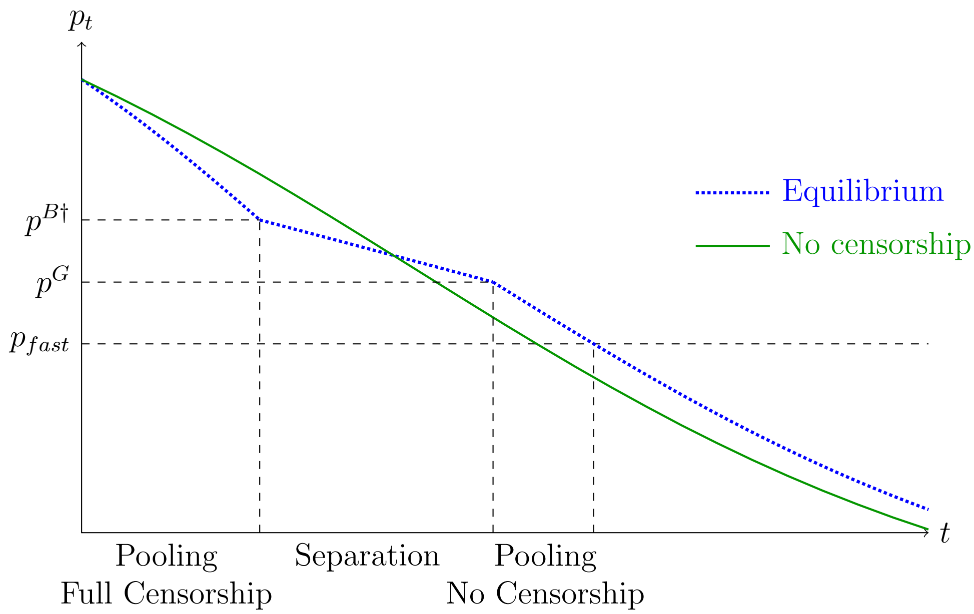


Figure 1.6: The belief evolution in the absence of news (inconclusive bad news & faster good news)

strategy is still a cutoff strategy with the cutoff belief $p_{fast}$ both in the NCB and in equilibrium, due to the same reason in the conclusive bad news case, i.e. the cutoff belief does not depend on the arrival rate of bad news, thus does not depend on the censorship policy. Second, there are three different phases in equilibrium. Here, $p^G$ (resp. $p^{B\dagger}$) is the cutoff belief below which the type $G$ (resp. $B$) agent stops censoring. When the public belief is high (i.e. $p_t > p^{B\dagger}$), both types of agent censor bad news for sure – they are pooling on full censorship. Hence, there is no bad news in equilibrium. This implies that the belief drifts down faster than it does in the NCB (i.e. $\gamma > \gamma + \beta^G - \beta^B$). The off-path belief after a piece of bad news in this phase still needs to be specified. We set the public belief jump down to 0 after a piece of bad news in this phase. When the public belief is intermediate (i.e. $p_t \in (p^G, p^{B\dagger})$), only the type $G$ agent censors for sure, but the type $B$ agent stops censoring; we call it the Separation period. As we explained, since the type $G$ agent has a higher incentive to censor bad news, she censors more aggressively than the type $B$ agent in the Separation period. In equilibrium, bad news may arise in this phase. Since the separation between the two types of agent, a piece of bad news is endogenously conclusive. When it arrives, the public belief jumps down to 0 (i.e. $J(p_t, 1, 0) = 0$). The noisy inconclusive bad news becomes to a perfect conclusive signal because of the separation of censorship strategies. It also implies that the belief drifts down slower than it does in the NCB (i.e. $\gamma - \beta^B < \gamma + \beta^G - \beta^B$). This is due to the martingale property of public belief – the jumping process is amplified, thus the drifting process must be mitigated. When the public belief is low (i.e. $p_t < p^G$), both types of agent stops censoring – they are pooling on no censorship. In this phase, bad news is never censored by any type of agent. A piece of bad news makes the belief jump down from $p_t$ to $j(p_t)$. We require $\beta^G \leq \bar{\beta}$, which is equivalent to $j(p^G) \leq p_{fast}$. Hence, in the last phase, a piece of bad news makes the public belief jump down below $p_{fast}$, which results in the dismissal of the agent.

#### 1.5.1.2 Welfare: Equilibrium versus NCB

We now compare all players' expected payoffs from the above equilibrium and their expected payoffs in the NCB. We find that, under some conditions, every player has a strictly higher payoff from the above equilibrium than he or she has in the NCB.

*Proposition* 1.7. Assume $c < \underline{c}$ and $\gamma > \beta^B$. There exists a $\underline{\beta} \in (0, \bar{\beta})$, when $\beta^G \in (0, \underline{\beta}]$ and $p_0 \in (p^G, p^{B\dagger}]$, the evaluator and both types of agent have strictly higher payoffs in the equilibrium of Proposition 1.6, compared with their payoffs in the NCB.

The conditions we need for the above result are first the game starts from the Separation period (i.e. $p_0 \in (p^G, p^{B\dagger}]$), and second the good agent's bad news does not happen too often (i.e. $\beta^G \in (0, \underline{\beta}]$).[18] In the Separation period, due to the differential censorship strategies, a piece of bad news becomes a conclusive signal that the agent is type $B$. This improves the quality of the evaluator's information. In this period, he would never dismiss a type $G$ agent, and the agent he might dismiss must be a type $B$ agent. The endogenously conclusive signal helps the evaluator to avoid making type I and type II errors. This is why his payoff is strictly higher than his payoff in the NCB, if the game starts from the Separation period. Figure 1.7 shows the evaluator's value function in equilibrium, and the horizontal axis is the prior belief. The value functions in equilibrium and in the NCB may cross when the prior belief is high enough, but not in this numerical example. When the prior belief is high ($p_0 > p^{B\dagger}$), the evaluator expects to have low quality information first,

[18]Those are sufficient, but not necessary conditions to obtain the above result. By continuation, the above result still holds when the prior belief is slightly higher than $p^{B\dagger}$, but we restrict to those conditions to make a clear explanation about the underling economic forces.

i.e. no bad news can arise, then high quality information, i.e. conclusive bad news can arise. Whether he is better off in equilibrium compared with the NCB depends on whether the benefit from the Separation period dominates the loss from the period when both types of agent pool on full censorship.

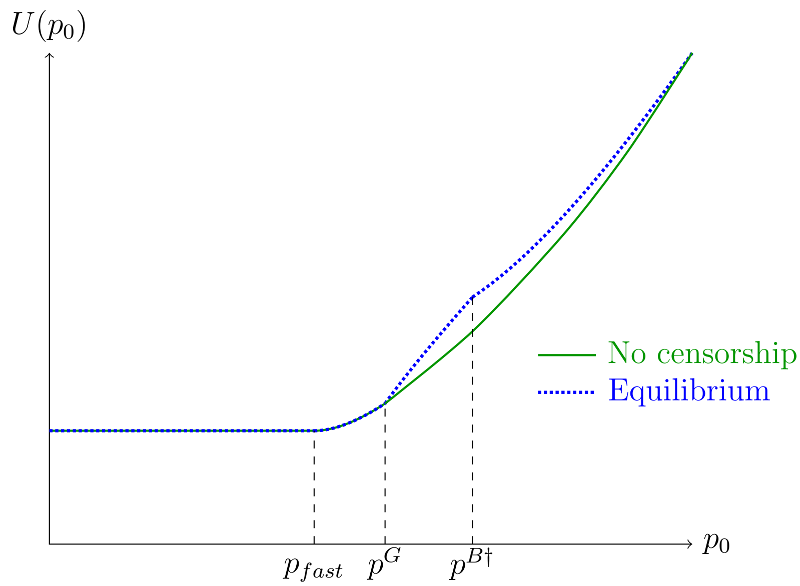


Figure 1.7: The evaluator's value functions (inconclusive bad news & faster good news)

Another implication from the Separation period is that no news becomes less severe, which is why both types of agent are better off in equilibrium. Though the public belief still drifts down in the absence of news, it drifts down slower than it does in the NCB. It means the evaluator would give more time to the agent in her job if no news arises, compared with the NCB. For the type $B$ agent, when the game starts from the Separation period, she does not censor bad news both in the equilibrium and in the NCB. We also require $\beta^G \leq \underline{\beta}$, which is equivalent to $j(p^{B\dagger}) \leq p_{fast}$. Hence, a piece of bad news makes public belief jump down below $p_{fast}$, which results in the dismissal of the type $B$ agent. Thus, the type $B$ agent has the same payoff both in the equilibrium and in the NCB when a piece of bad news arrives. However, if the type $B$ agent is lucky that bad news does not arise, she can stay longer in her job in the equilibrium, compared with the NCB. Thus, the type $B$ agent is strictly better off in equilibrium. For the type $G$ agent, she is also strictly better off in equilibrium. On the one hand side, similar to the conclusive bad news case, censoring bad news in the Separation period gives her a higher expected flow payoff in equilibrium. On the other hand side, she is given more time in her job if no news arrives in equilibrium, thus she is also more likely to succeed. Notice that censorship has opposite effect in the conclusive and the inconclusive bad news cases. In the conclusive news case, no news makes the belief drift down faster, while in the inconclusive news case, no news makes the belief drift down slower. Thus, the type $G$ agent has a strictly higher payoff in the equilibrium than she has in the NCB.

When the censoring cost increases, the type $B$ agent first gives up censorship, since she has a lower value in her job than the type $G$ agent. If the censoring cost is intermediate (i.e. $c \in (\underline{c}, \bar{c})$, where $\bar{c} := \frac{\gamma+\rho_0}{\beta^G+\gamma+\rho_0}\frac{w}{\rho_0}$) and the type $G$ agent's bad news does not happen too often (i.e. $\beta^G \leq \bar{\beta}$), we can show that there exists an equilibrium in which the evaluator and the type $G$ agent use the same strategy as in Proposition 1.6, and the type $B$ agent never censors. Notice that $\underline{c}$ is the cost threshold for the type $B$ agent that determines whether she is willing to censor all bad news in exchange for staying in her job forever, while $\bar{c}$ is the cost threshold for the type $G$ agent that determines whether she is willing to censor all bad news before she succeeds in exchange for staying in her job forever. Hence, the pooling on full censorship period disappears. We only have the Separation period and the pooling on no censorship period in equilibrium. Thus, we still have a similar welfare effect as in Proposition 1.7 under similar conditions.

### 1.5.2 Slower good news

#### 1.5.2.1 Equilibrium

In this section, we consider the slower good news case, i.e. good news arrives slower than the type $B$ agent's bad news ($\gamma < \beta^B$). "Slower good news" here does not necessary mean "no news is good news" in the NCB. The type $G$ agent now has two possible signals – good news with intensity $\gamma$ and bad news with intensity $\beta^G$, while the type $B$ agent has only bad news with intensity $\beta^B$. Only when the signal intensity from the type $G$ agent is smaller than the signal intensity from the type $B$ agent (i.e. $\gamma + \beta^G < \beta^B$), "no news is good news", i.e. the public belief will drift up if no news arrives. Hence, when the good news arrives very slow (i.e. $\gamma < \beta^B - \beta^G$), "no news is good news" in the NCB. However, if good news arrives just mildly slow (i.e. $\beta^B - \beta^G < \gamma < \beta^B$), the public belief will drift down if no news arrives in the NCB. In both cases, the evaluator's optimal policy in the NCB is a cutoff strategy; he dismisses the agent whenever the public belief is below the cutoff

$p^* \in (p_{slow}, p_{fast}]$. The cutoff belief $p^*$ is strictly below $p_{fast}$ when $\gamma < \beta^B - \beta^G$, and it is equal to $p_{fast}$ when $\gamma \in (\beta^B - \beta^G, \beta^B)$.

Similar to the faster good news case, the type $G$ agent has a higher incentive to censor bad news than the type $B$ agent does. If we maintain the assumption that the censoring cost is low, $c < \underline{c}$, then the cutoff equilibrium in the conclusive bad news case still exists when bad news is actually inconclusive. That is, there exists an equilibrium in which the evaluator and the type $B$ agent use the same strategies as in Proposition 1.4, and the type $G$ agent censors bad news if and only if the public belief is above or equal to $p_{fast}$. First, it is easy to see that, given the strategies of both types of agent, the evaluator faces the exact same problem as in the conclusive bad news case when the public belief is above or equal to $p_{fast}$, since the type $G$ agent censors all bad news. In addition, it is not hard to show that the evaluator wants to dismiss the agent whenever the public belief is below $p_{fast}$. Second, given the evaluator's strategy, the type $B$ agent faces the exact same problem as in the conclusive bad news case, hence she also uses the same strategy. At last, since the type $G$ agent has a higher incentive to censor bad news than the type $B$ agent does, the former censors bad news with probability one whenever the latter censors with positive probability. Therefore, the equilibrium in the conclusive bad news case still exists.

When the censoring cost increases, the type $B$ agent first gives up censorship. When the cost is intermediate (i.e. $c \in (\underline{c}, \bar{c})$), there exists an equilibrium in which the type $B$ agent never censors, but the type $G$ agent censors bad news whenever the public belief is above $p_{slow}$. In addition, the equilibrium strategy of the evaluator is to dismiss the agent whenever the public belief is below $p_{slow}$. It is worth emphasizing that $p_{slow}$ is the cutoff belief in the evaluator's optimal policy in the NCB when bad news is conclusive. In this equilibrium, separation between the two types of agent makes a piece of bad news endogenously conclusive. Hence, the evaluator's equilibrium strategy is the cutoff strategy with the cutoff belief $p_{slow}$. In this equilibrium, since good news arrives slower than the type $B$ agent's bad news, "no news is good news"; the public belief drifts up in the absence of news. Since the censoring cost is intermediate, it is too costly for the type $B$ agent to censor any news, but it is worth for the type $G$ agent censoring bad news in exchange for staying in the job.

The following proposition summarizes the result.

*Proposition* 1.8. Assume $\gamma < \beta^B$ and $\beta^G \in (0, \beta^B)$.

1. (Low cost) When $c < \underline{c}$, there exists an equilibrium as follows:
   - The evaluator and the type $B$ agent use the same strategies as in Proposition 1.4;
   - The type $G$ agent censors bad news whenever the public belief is above and equal to $p_{fast}$, but below 1.
2. (Intermediate cost) When $c \in (\underline{c}, \bar{c})$, there exists an equilibrium as follows:

- The evaluator dismisses the agent whenever the public belief is below or equals to $p_{slow}$;
- The type $G$ agent censors bad news whenever the public belief is above $p_{slow}$, but below 1, and the type $B$ agent never censors.

#### 1.5.2.2 Welfare: Equilibrium versus NCB

We still want to compare equilibrium payoffs with the payoffs in the NCB. However, the comparison would be tedious if we try to cover every possible scenario, especially now the NCB has two very different cases, depending on whether the good news arrives very slow or mildly slow. Hence, we only give a big picture to see how the change of the censoring cost may change the welfare comparison from the evaluator's point of view.

*Proposition* 1.9. Assume $\gamma < \beta^B$ and $\beta^G \in (0, \beta^B)$.

1. (Low cost) Assume $c < \underline{c}$. When $p_0 \in (p^*, 1)$, the evaluator has a strictly lower payoff in the equilibrium from Proposition 1.8 than he has in the NCB.
2. (Intermediate cost) Assume $c \in (\underline{c}, \bar{c})$. When $p_0 \in (p_{slow}, 1)$, the evaluator has a strictly higher payoff in the equilibrium from Proposition 1.8 than he has in the NCB.

Figure 1.8 illustrates the evaluator's value functions in the NCB and in the equilibrium from Proposition 1.8. When the censoring cost is low, since both types of agent censors very aggressively, the evaluator essentially cannot rely on bad news to learn the type of the agent. This is the worst information he could have, thus he is always worse off in the equilibrium, compared with the NCB. In addition, the evaluator's incentive for learning is discouraged when the good news arrives very slow. That is, when $\gamma < \beta^B - \beta^G$, the cutoff belief in the evaluator's optimal policy in the NCB increases from $p^*$ to the cutoff belief in his equilibrium strategy $p_{fast}$. We have seen this discouragement effect from the conclusive bad news case. When the censoring cost is intermediate, the differential censorship policy between two types of agent dramatically improves the quality of information – a piece of bad news is endogenously conclusive. Hence, the evaluator is better off in the equilibrium, compared with the NCB. Moreover, the improvement in the quality of information encourages the evaluator to explore the type of the agent even when the public belief is in between $p_{slow}$ and $p^*$. We call it the *encouragement effect*; the cutoff belief in the evaluator's optimal policy in the NCB decreases from $p^*$ to the cutoff belief in his equilibrium strategy $p_{slow}$.

Both types of agent are also directly affected by the discouragement and the encouragement effects of the evaluator. When the censoring cost is intermediate, because of the encouragement effect, both types of agent are better off in the equilibrium when the prior belief is in between $p_{slow}$ and $p^*$. In the equilibrium, the type $G$ agent will stay in her job forever, and the type $B$ agent will stay in her job for some positive time. While in the NCB, both types of agent will not have a chance to start. When the censoring cost is

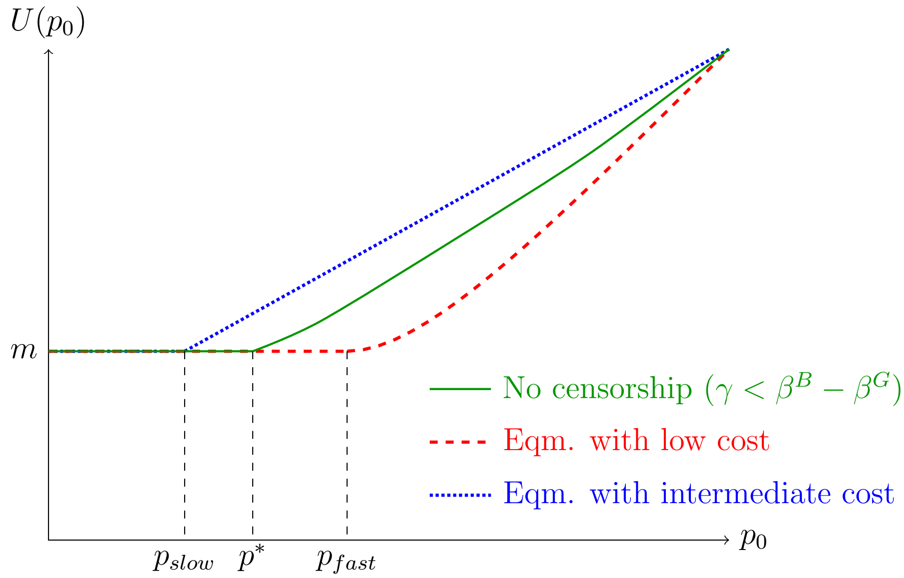


Figure 1.8: The evaluator's value functions (inconclusive bad news & slower good news)

low and the good news arrives very low ($\gamma < \beta^B - \beta^G$), the effect is opposite; both types of agent are worse off in the equilibrium when the prior belief is in between $p^*$ and $p_{fast}$, due to the discouragement effect.

## 1.6 Conclusion

This paper studies costly censorship in a dynamic environment. An evaluator tries to learn the agent's competence from a good news process and a bad news process. However, an informed agent who knows her competence level can conceal bad news at some cost.

When bad news is conclusive, only the incompetent agent censors. The evaluator suffers from censorship since otherwise valuable information has been suppressed. This also makes her interpret no news more severely, which hurts the competent agent. It may also hurt the incompetent agent, even though censorship helps her survive bad news. The incompetent agent is worse off with censorship when the period that she finds it optimal to censor is short. When good news arrives slower than the bad news from the incompetent agent, a discouragement effect occurs – the evaluator increases the threshold public belief of dismissal.

When bad news is inconclusive, both competent and incompetent agents censor bad news. However, it is the competent agent who has a higher incentive to censor since she has a higher value in the job. When the two types of agent separate in their censorship strategies (i.e. only the competent agent censors), the quality of information for the evaluator is improved – the inconclusive noisy bad news becomes conclusive and perfect. Not only the evaluator benefits from that, but also both types of agent, since no news now is less severe and the agent may stay in her job longer. When good news arrives very slow, an encouragement effect may occur – the evaluator decreases the threshold public belief of dismissal, when the censoring cost is not too high or too low such that only the competent agent finds it optimal to censor.

# Chapter 2

# Contracts that Reward Innovation: Delegated Experimentation with an Informed Principal

## 2.1 Introduction

How should we design contracts in order to reward innovation, when poor ideas can masquerade as worthwhile ones? Consider an innovator who has a new idea, and has private but imperfect information about the quality of the idea. Specifically, the idea may be worth exploring, but even high quality ideas may result in failure, just as ideas that are of low quality ex ante may sometimes be successful. The innovator needs to engage an agent to explore the idea, and thus the agent's moral hazard must also be confronted. Since the agent becomes more pessimistic over time about the probability of success in the absence of a breakthrough, moral hazard is dynamic. Moreover, the innovator also needs to convince the agent that the idea's quality is high and it is worth exploring. Our question is: what are the contracts that provide maximal rewards for high quality innovations? If high quality innovations are not rewarded properly, then innovators would not invest in better ideas in the first place. The question is particularly important in the knowledge economy, where economic growth is driven by innovations. Furthermore, all the ingredients listed are important considerations in knowledge industries:

innovators have private information, but even the best ones are not omniscient, and sometimes come up with unworkable ideas; the innovators themselves may not be the best ones to undertake project development, and may have to delegate the task to specialized bodies; only these specialized bodies know how intensively they are working on the innovators project.

We analyze a model where a privately informed innovator engages an agent to work on a project. A senior professor may need to hire a junior research assistant to conduct some lab experiments for a research project; a tech startup founder may need to hire a professional team to validate the needs of the new product. In drug development, scientists in biotech companies may have strong insights into the fundamental mechanism of diseases. Other entities, like Contract Research Organizations (CRO) and Contract Manufacturing Organizations (CMO) may be specialized in other research support services.[1] When a biotech company hires a CRO or a CMO to undertake some experiments for a drug development project, the biotech company needs to not only incentivize them to exert efforts towards exploring the viability of the project, but also convince them of the quality of the project. The problem is particularly severe when high quality projects are relatively scarce, and cannot be

[1]According to Wikipedia, the services provided by a CRO may include biopharmaceutical development, biologic assay development, commercialization, preclinical research, clinical research, clinical trials management, and pharmacovigilance. The services provided by a CMO may include pre-formulation, formulation development, stability studies, method development, pre-clinical and Phase I clinical trial materials, late-stage clinical trial materials, formal stability, scale-up, registration batches and commercial production. See https://en.wikipedia.org/wiki/Contract_research_organization and https://en.wikipedia.org/wiki/Contract_manufacturing_organization.

distinguished from low quality projects. This is typically the case in research industries.[2]

This paper has two parts. In the first part, we combine a signaling game with an exponential bandit model to study the dynamic agency problem. An innovator or principal (she) is endowed with a project. Neither she nor the hired agent (he) knows whether the project is viable, i.e. whether it will generate any profits. However, the principal is privately aware of the project's quality, i.e. how long the project, if it is viable, will take in expectation to generate profits.[3] Only the high quality project is worth exploring. The principal can commit to a long-term contract that specifies how payments depend on outcomes to incentivize the agent to exert private effort. The contract also serves as a signal of the quality of her project, or her type. When the agent works on the project, he also gradually learns about both the viability and the quality of the project. Hence, the principal has two tasks at the same time: signal her type and provide incentives to the agent to overcome the dynamic moral hazard problem. To isolate the conflict between signaling and providing incentives, we assume both players are risk neutral and have unlimited liability.

Since our focus is on rewarding innovation, we characterize the best

[2]Stevens and Burley (1997) estimate that there is only one commercial success in 3,000 raw ideas of innovation across most industries. For drug development, Klees and Joines (1996) report that only one compound is approved for marketing among every 5,000 to 10,000 compounds that enter preclinical testing.

[3]To be more specific, the quality of a project is modeled as the success probability in one period when the project is viable and the agent exerts effort.

equilibrium contract for the high type principal. The high type principal can never obtain her full information surplus, i.e. the payoff she would get if she was known to be a high type. Depending on the proportion of high quality projects in the population, the best equilibrium for her is either a separating equilibrium or a pooling equilibrium.

When high quality projects are scarce, a separating equilibrium gives the best outcome to the high type principal. Without limited liability, the high type principal could sell her project to the agent, leaving him to be the residual claimant that solves the dynamic moral hazard problem. Indeed, selling the project would be the optimal contract if the quality of the project were publicly known. That is, the agent makes an upfront payment that equals the expected surplus of the project, in exchange for the entire profit when the project succeeds. When only the principal knows her type, though, such a contract does not do a good job separating the two types. Both types value the upfront payment the same, and the low type principal has incentives to exaggerate the quality of her project to sell it for a good price. To separate from the low type principal and convince the agent of the project's quality, the high type principal would like to pay some positive base wage to the agent that does not depend on the outcome, as well as to share only a part of the profits as bonus payments with him when the project succeeds. The base wage serves as a signal for the high type principal. The bonus payments are increasing over time to incentivize an increasingly pessimistic agent to continue working until the principal's desired termination date. Thus, the high type principal shares a portion of the total surplus with the agent. Moreover, since the high type principal cannot ask for compensation for the dynamic moral hazard costs, she would like to terminate the project inefficiently early. Thus, both the inefficient termination and sharing the surplus with the agent reduce the payoff of the high type principal.

On the other hand, when high quality projects are not rare, it is better for the high type principal to pool with the low type to avoid signaling costs. In that case, the best equilibrium for the high type principal is a pooling equilibrium, where both types charge the agent a positive sign-up fee, then share the profits with him once the project succeeds. Thus, instead of leaving rents to the agent, the high type principal "shares" rents with the low type principal. Moreover, the equilibrium contract still features inefficiently early termination, compared with a project whose quality is ex ante unknown.

The above results show that the high quality project is operated inefficiently, and the high type principal cannot obtain her full information surplus. This provides a role for a mediator to facilitate the agency process. Online platforms, such as Science Exchange[4], Scientist[5] and Upwork[6], are such mediators. Science Exchange and Scientist help scientists outsource their research to other scientific institutions around the world. Upwork is a global freelancing platform that connects independent professionals.

[4]https://www.scienceexchange.com.
[5]https://www.scientist.com.
[6]https://www.upwork.com.

In the second part of the paper, we analyze the mechanism design problem faced by a mediator who seeks to maximize the reward for superior innovations, i.e. who seeks to maximize the payoff of the principal with a high quality project. The designer offers a menu consisting of two contracts, one for each type of principal. This menu must satisfy incentive compatibility for the principal, i.e. it must induce truthful revelation by each type (it must also satisfy individual rationality). The agent observes the menu, and infers that each of the contracts will be implemented with a probability corresponding to his prior, and decides whether to accept or reject the menu. However, the agent is not told which element is being implemented until it is essential for him to know. In other words, the agent is confronted with an opaque contract, and remains uncertain about his exact rewards for some time. This relaxes the agent's individual rationality and incentive compatibility constraints; they only need to hold on average. In addition, the opaqueness of the contingent transfers allows the two types of principal to bet on a success, which provides an additional device for the high type principal to separate herself. In this way, the mediator designs a mechanism that improves the payoff of the high type principal. Moreover, the inefficiency costs are minimal, since the mediator could recommend the low type project running for only one period. In the optimal mechanism for the high type principal with pure recommendations, the high type principal obtains approximately all the surplus from her innovation when the period length is small. Thus the contract with the mediator allows the innovator of a superior project to appropriate almost her entire contribution to social surplus. Furthermore, an innovator who comes up with an inferior project is left with no surplus, thus the menu simultaneously minimizes the rewards for wasteful innovations.

The rest of the paper is organized as follows. Section 2.2 discusses the related literature. Section 2.3 introduces the model. Section 2.4 discusses the efficient solution and other benchmarks. Section 2.5 characterizes the best equilibrium for the high type principal in the signaling game, and discusses its properties. Section 2.6 considers the third-party mechanism design problem. Section 2.7 concludes. The Appendix provides proofs, and studies the implementation problem of the full information benchmark.

## 2.2 Related Literature

This paper examines the incentives for experimentation when the principal is informed. It mainly relates to three strands of literature.

First, it relates to an increasing literature on incentives for experimentation. These papers, as well as the current one, build on a two-armed bandit model of learning, as in Keller et al. (2005), and focus on how to provide incentives for agents to experiment through contingent contracts. Most papers, such as Bergemann and Hege (1998, 2005), and Hörner and Samuelson (2013), consider how to incentivize one party to experiment, who is subject to the moral hazard problem. They typically consider a repeated interaction between the principal and the agent, assume limited liability, and demonstrate an inefficiency result due to the agency costs. Guo (2016) studies a dynamic relationship in which a principal delegates experimentation to a biased agent who has private information about the prior belief that the state is good. Thus, the incentive problem comes from the hidden information of an informed but biased agent. Halac et al. (2016) is a closely related paper. They examine an agency problem subjected to both moral hazard and adverse selection in the context of experimentation. They assume that the principal can commit to a long-term contract, and no limited liability. The major difference is that, instead of considering the private information on the side of the agent, our paper examines the case when the private information is on the side of the principal. In the screening problem of Halac et al. (2016), the high type agent has an incentive to pretend to be the low type. In our signaling problem, the low type principal has an incentive to pretend to be the high type. Therefore, Halac et al. (2016) show that the optimal contract has no distortion for the high type agent, but requires the low type agent to terminate the project inefficiently early.[7] By contrast, we show that it is the high type project that is terminated inefficiently early in the best equilibrium for the high type principal. Moreover, in that equilibrium, either there is no distortion for the low type project, or it is distorted towards over-experimentation, depending on the prior belief about the low type principal. Thus the economic forces underlying the analyses are very different in the two papers. Furthermore, when we allow for a mediator, we show that we can achieve approximate efficiency, a result

[7] It is a typical result in screening problems that distortions occur at the bottom, but not at the top.

that has no counter-part in Halac et al. (2016).

Second, our paper relates to a relatively small literature on the informed principal problem with moral hazard. Myerson (1983) first considers the informed principal problem from an axiomatic point of view. Maskin and Tirole (1990, 1992) develop a noncooperative game framework to analyze the informed principal problem with no moral hazard. According to their categorization, our model is the "common value" informed principal problem, since the agent cares directly about the type of principal. Beaudry (1994) considers the informed principal problem with moral hazard. He characterizes a separating equilibrium where the principal leaves rents to the agent. More recent papers, such as Silvers (2012), Wagner et al. (2015), Bedard (2016), and Karle et al. (2016) also consider the informed principal problem in the presence of moral hazard. In all these papers, the moral hazard problem is static. Few papers examine this problem in a dynamic setting.[8] One exception is Kaya (2010). She studies a similar informed principal problem with moral hazard when the principal and the agent interact repeatedly. She considers a situation when the principal and the agent start with symmetric uncertainty about the productivity information. In each period, the principal has a choice to acquire that information without costs, but would rather delay to acquiring it in order to save the costs for signaling. However, in our paper, the agent learns both the type of the principal and the viability of the project during

[8]Fryer and Holden (2012) consider a two period informed principal problem with moral hazard. However, they make a behavioral assumption that the agent only learns from a noisy signal, but not from the contract proposed by the principal.

experimentation, which makes the incentive problem very different.[9]

Last, our paper is also related to the information design problem with moral hazard. Jehiel (2015) examines whether a principal with private signals prefers to commit to a non-transparent information disclosure policy to overcome the agent's moral hazard. He finds that full transparency is generically suboptimal under some mild conditions. Our result in the mechanism design part echoes his result - keeping the agent in the dark improves the payoff of the high type principal. However, we focus on the signaling problem of the principal after her private information is realized, while Jehiel (2015) assumes the principal can commit to an information disclosure policy before the realization of the private information. Ely and Szydlowski (2016) study a model where a principal is privately informed about the duration of required effort for completing a project. The principal's objective is to induce the agent to work as much as possible, thus they find that the optimal information disclosure policy is "moving the goalposts": at the outset, the principal tries to make the agent optimistic that the task is easy in order to induce him to start working, but persuades him that the task is hard when the difficult goal is within reach. In their setting, there is no signaling consideration, and the principal's problem is to keep the agent from quitting.

[9]Kaya (2010) uses "money burning", i.e. a pure production-irrelevant cost, as a signaling device for the principal after acquiring information. This money burning signaling device, together with the assumption that the agent has limited liability, are essential to her results. See Kaya (2010) Footnote 15.

## 2.3 The Model

Time is discrete and the horizon is infinite, with a small but strictly positive period length $\Delta > 0$. The per period common discount factor is $\delta = e^{-\rho\Delta}$, where $\rho \geq 0$. To get rid of integer problems, we will look at the case as $\Delta \to 0$ for some results.[10]

There are two risk neutral players, a principal $P$, and an agent $A$. The principal (she) hires an agent (he) to complete a project with uncertain viability. The agent, after accepting the principal's offer, could choose whether to exert efforts in every period before the game ends. The agent has flow costs $c\Delta$, where $c > 0$, when he exerts efforts in any period. The efforts are not observable or verifiable by the principal.

The principal is privately informed about the *quality* of the project, $\theta \in \Theta = \{H, L\}$. $H$ represents a high quality project, and $L$ represents a low quality project. The quality is persistent, and is chosen by nature at the beginning. We can also regard the quality as the *type* of the project/principal. The agent doesn't know the principal's type, but has a prior $\beta_0 \in (0, 1)$ on the $H$ type, which is common knowledge.

A project of either type may be in one of the two *states*: a good state

[10] Those results approximate the circumstance when $\Delta$ is small. We do not use a continuous time model to avoid technical difficulties, since stochastic integration with respect to general transfer scheme may not be well defined. We also use the following notational conventions. $N$ and $n$ are used for time periods when we talk about the model with a fixed $\Delta > 0$; $T$ and $t$ are used for time when we examine the model in the limit $\Delta \to 0$. For a variable $x_n$ that has time subscript $n$, let $x_t$ be the limit of $x_n$ as $\Delta \to 0$ and $n\Delta \to t$ (when it exists).

$G$, or a bad state $B$. The state is also persistent and chosen by nature at the beginning. However, neither party knows the state of the project. They have a common prior $q_0 \in (0, 1)$ on the $G$ state, and the distribution of states is independent of the distribution of types.

The following figures summarize the information structure when nature moves. $P$ knows the types, but not the states. $A$ knows neither of them.

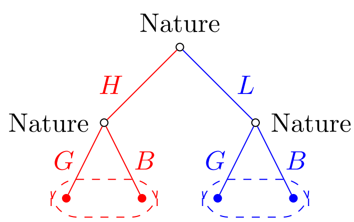


Figure 2.1: Principal's Information

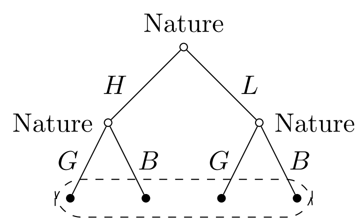


Figure 2.2: Agent's Information

In the $B$ state, a project, independent of its type and the agent's efforts, never generates any profits. In the $G$ state, when the agent exerts efforts in one period, a $\theta \in \Theta$ type project can generate a lump-sum profit $h > 0$ to the principal in that period with probability $\lambda^\theta \Delta$, and $\lambda^H > \lambda^L > 0$. It is a *success* of a project in some period when the lump-sum profit is generated, otherwise we call it a *failure*. A success of the project can be publicly observed and verified[11], and will end the game. Hence, the difference between the $H$

[11]The observability and verifiability of a success is a harmless assumption in the model. If a success is only observed by the agent, then the agent could choose to misreport a failure to the principal when a success actually arrives. However, the agent has no incentive to hide a success from the principal in any equilibrium contract in this paper. When a contract gives exact incentives for the agent to work if a success is contractible, it also gives enough incentives for the agent to truthfully report the success if it is not contractible. See Section 2.5.2.

and the $L$ type projects is the arrival rate of success conditional on the project is in the $G$ state and the agent exerts efforts. In addition, we will maintain the following assumption in this paper:

*Assumption.* $q_0\lambda^H h > c \geq q_0\lambda^L h$.

Thus, given the cost and the benefit of the project, and the prior belief about the state, only the $H$ type project is worth experimenting on. However, if it was commonly known that the project's type was $L$, then it would exit the market.

The timing of the game is the following.

At the beginning of the game, the principal learns her type, $H$ or $L$, then proposes a take-it-or-leave-it long-term contract to the agent. The principal can fully commit to her contract. A contract will specify the following:

- A termination date $N \in \mathbb{N}_0$ of the project conditional on no success.[12] This means the principal would shut down the project after $N$ failures.
- A lump-sum payment, $W \in \mathbb{R}$ from the principal to agent at time 0. It can be positive or negative.
- A contingent payment plan during experimentation, two vectors $\boldsymbol{b} \in \mathbb{R}^N$ and $\boldsymbol{p} \in \mathbb{R}^N$, from the principal to the agent. $\boldsymbol{b}$ is for bonuses, it

[12]The principal, whatever type she is, never wants to experiment forever. Therefore, without loss, we do not need to consider the option to run the project forever.

determines the payment from the principal to the agent when the project succeeds in any period; $\boldsymbol{p}$ is for penalties, it determines the payment from the principal to the agent when the project fails in any period. In other words, if the project succeeds in the $n$-th period, where $1 \leq n \leq N$, then the payments to the agent are $p_k$ in the $k$-th period for $1 \leq k < n$, and $b_n$ in the $n$-th period. Bonuses and penalties can be either positive or negative.

Hence, a *contract* is a quadruple $C = \{N, W, \boldsymbol{b}, \boldsymbol{p}\}$. Note that all above terms depend on $\Delta$, however, we omit that to save notation. A contract is a *bonus contract* if $\boldsymbol{p} = \mathbf{0}$, and a *penalty contract* if $\boldsymbol{b} = \mathbf{0}$. If $N = 0$, then the payment in the contract only consists of $W$. A contract is a *null contract* if $N = 0$ and $W = 0$.

After the principal proposes a contract, the agent chooses whether to reject or accept it. If the agent rejects it, then both parties obtain their reservation payoffs 0 and the game ends. If the agent accepts it, then the contract is implemented and the agent will decide whether to exert efforts in each period, until the project succeeds or the termination date is reached.

Since the informed player (the principal) moves first, this is a signaling game. We will consider Perfect Bayesian Equilibria (PBE).

A contract allows any transfers between the principal and the agent. However, the set of the contracts is unnecessarily large. We can restrict attention to a smaller set of contracts, i.e. the set of bonus contracts or the set of penalty contracts, without loss of generality.

Let $\beta$ be the probability of the $H$ type, $C$ be some arbitrary contract with terminating date $N$, and $\boldsymbol{a} \in A^N = \{0,1\}^N$ be agent's action plan, where 0 means shirking and 1 means working. Given the contract, the belief and the action plan, then

- $\Pi^\theta(C, \boldsymbol{a})$ denote the expected discounted payoff for type $\theta$ principal;
- $W_A^\beta(C, \boldsymbol{a})$ denote the expected discounted payoff for the agent.

We define *payoff-equivalent* contracts as follows:

*Definition.* Two contracts $C$ and $\hat{C}$, with the same termination date $N$, are payoff-equivalent, if for any $\boldsymbol{a} \in A^N$, any $\theta \in \Theta$ and any $\beta \in [0,1]$, $\Pi^\theta(C, \boldsymbol{a}) = \Pi^\theta(\hat{C}, \boldsymbol{a})$ and $W_A^\beta(C, \boldsymbol{a}) = W_A^\beta(\hat{C}, \boldsymbol{a})$.

The transfer structure is rich enough to allow many different contracts that are payoff-equivalent. If two contracts induce the same discounted transfers on any realized event, then they are payoff-equivalent, since both parties are risk neutral and share the same time preference. Furthermore, for any realized event, i.e. whether and when the project succeeds, bonuses (or penalties) together with the lump sum payment $W$ could deliver any discounted transfers that a general contract could. Formally, *Lemma* 2.1. For any contract, there exist both an equivalent bonus contract and an equivalent penalty contract.

The idea and the proof are the same as in the Proposition 1 of Halac et al. (2016). Our setup differs from theirs, since we have a signaling game rather than a screening problem. The agent may form different beliefs on payoff-equivalent contracts. However, the principal and the agent would feel indifferent to payoff-equivalent contracts as long as the agent forms the same beliefs, since they would induce the same action plan of the agent and deliver the same expected payoffs. Hence, we shall have no reason to assume that the agent would regard different payoff-equivalent contracts as different signals. Formally, we assume

*Assumption.* For any payoff-equivalent contracts, the agent forms the same belief about the principal's type.

Hence, without loss of generality, we can focus only on bonus contracts or penalty contracts. In this paper, we will restrict attention to bonus contracts, or *contracts* for simplicity. Thus, a contract is a triple $C = \{N, W, \boldsymbol{b}\}$.

## 2.4 Benchmarks

### 2.4.1 Efficient Solution

We first consider the efficient solution without the agency problem. Since we have a quasi-linear environment, consider a social planner who seeks to maximize total surplus, given that he knows both the private type of the principal and the hidden actions of the agent, but not the state of the project.

Clearly, the social planner would solve an optimal stopping problem. The optimal strategy is to stop the project as the posterior belief about the state of the project being $G$ dropping to some cutoff belief. This strategy is also equivalent to specifying how long to experiment.[13]

Let $V^\theta(N)$ be the expected discounted value of the $\theta \in \Theta$ project when the social planner experiments $N$ times. Then it is given by:

$$V^\theta(N) = \sum_{n=1}^{N} \delta^n f^\theta_{n-1}(q_0)(q^\theta_n \lambda^\theta h - c)\Delta, \tag{2.1}$$

where $f^\theta_m(q) = q(1-\lambda^\theta\Delta)^m + 1 - q$ is the probability that a $\theta$ project, with prior $q$ being in the $G$ state, fails $m$ times. $(q^\theta_n \lambda^\theta h - c)\Delta$ is the expected payoff for the $n$-th experiment conditional on a success not having arrived yet. Here, $q^\theta_n$ is the posterior belief about the state being $G$ for a $\theta$ project before the start of the $n$-th experiment,[14]

$$q^\theta_n = \frac{q_0(1-\lambda^\theta\Delta)^{n-1}}{q_0(1-\lambda^\theta\Delta)^{n-1} + 1 - q_0}.$$

Thus, the optimal policy for the social planner, with a $\theta$ project, is to conduct the project as long as

$$q^\theta_n \lambda^\theta h \geq c.$$

[13] More precisely, a strategy specifies how long to experiment without a success. Since a success ends the game, we omit repeating "without a success".

[14] After a project fails $n-1$ previous experiments.

Given the assumption $q_0\lambda^L h \leq c$, it is easy to see that the social planner would like to abort the $L$ type project immediately; the $L$ type project is a lemon. Thus, the *efficient termination date for the L type project* is 0. On the other hand, given the assumption $q_0\lambda^H h > c$, the *efficient termination date for the H type project*, denoted by $N_*^H$, is strictly positive. To get rid of the integer problem of $N_*^H$, it will be helpful to examine the limit of the experimenting time,

$$T_*^H := \lim_{\Delta \to 0} N_*^H \Delta = \frac{\ln l_0 - \ln l^H}{\lambda^H},$$

where $l_0 = \frac{q_0}{1-q_0}$ is the likelihood ratio of the prior belief that the state of the project is $G$, and $l^H = \frac{c}{\lambda^H h - c}$ is the likelihood ratio of the efficient cutoff posterior belief that the state of the $H$ project is $G$. We also denote $V_0^\theta(T) := \lim_{\Delta \to 0, N\Delta \to T} V^\theta(N)$.

The efficient solution is obtained assuming away both the private information of the principal and the hidden action of the agent. In the next section, we consider the benchmarks where the incentive problem is one-sided.

### 2.4.2 One-sided Incentive Problem

Our model has both private information on the side of the principal, and hidden actions on the side of the agent. Both are crucial, because the incentive problem would be trivial without either of them.

**No Private Information** - Let us first consider the case when the private type of the project is public information, while the agent's actions are still private. Although the principal still needs to incentivize the agent to experiment on the commonly unknown state of the project, the incentive problem is trivial. The principal can just sell the project to the agent, since the agent has no financial constraint. After the agent becomes the residual claimant of the project, he would like to implement the efficient solution. Thus, the $H$ type can extract all the surplus by selling her project with a price at its expected value, and the $L$ type would exit the market.

**Observable Efforts** - Another simple case is when the agent's efforts are publicly observable and verifiable. Thus, the principal can contract directly on the efforts. Even though the principal has private information on the type of the project, the incentive problem is also trivial. There exists a separating equilibrium in which both projects are conducted efficiently and the $H$ type principal extracts all the surplus of her project. The $H$ type principal proposes an "honest" contract that pays the agent each period until her efficient termination date for his costly efforts if and only if he exerts efforts, while the $L$ type exits the market. The agent would always like to accept such "honest" contract and exert efforts, and believe contracts other than the "honest" contract are offered by the $L$ type.

Therefore, in the above two cases, both projects are implemented efficiently, and all the surplus goes to the principal. We call the result as the *full information benchmark*, or *FIB*. As it will be shown later, in the presence of both private information on the side of the principal and hidden actions on the side of the agent, the principal needs to signal her type and incentivize the agent to work at the same time. Then, the FIB can never be achieved.

## 2.5 Equilibrium Characterization

In this section, we first show that there is no equilibrium that achieves the FIB. Multiple equilibria exist, since a PBE does not restrict the off-equilibrium beliefs. Because a $L$ type project is a lemon, we examine how much a $H$ type project can be rewarded in the equilibrium. Hence, we study the best equilibrium for the $H$ type. We characterize the best equilibrium for the $H$ type principal, which is either a separating equilibrium or a pooling equilibrium, depending on the prior belief about the $H$ type principal.

### 2.5.1 The Impossibility of Achieving the FIB

If an equilibrium implements the FIB, then it is a separating equilibrium, in which the $L$ type principal aborts the project, and the $H$ type principal incentivizes the agent to work until the efficient termination date and extracts all the surplus. More specifically, in such an equilibrium, the $H$ type contract should leave no rents to the agent, and incentivize him to work until period $N_*^H$, when he believes the principal is the $H$ type. In addition, the $L$ type principal cannot get strictly positive payoff from mimicking the $H$ type.

We first characterize all contracts that satisfy the agent's binding individual rationality (IR) constraint, and the agent's incentive compatible (IC) constraints until period $N_*^H$, when the agent believes the principal's type is $H$. Then we find the worst one among that set of contracts for the $L$ type, and show that the $L$ type can still obtain strictly positive payoff from the worst contract. Thus, any contract that implements the FIB must violate the IC constraint for the $L$ type. Hence, the FIB cannot be achieved in any equilibrium.

When the agent believes that the principal's type is $\theta \in \Theta$, a contract $C = \{N, W, \boldsymbol{b}\}$ satisfies the IC constraints for the agent to work from period 1 to period $N$, when for all $1 \leq n \leq N$,

$$\sum_{s=n}^{N} \delta^{s-n} f_{s-n}^{\theta}(q_n^{\theta})(q_s^{\theta}\lambda^{\theta}b_s - c)\Delta \geq \sum_{s=n+1}^{N} \delta^{s-n} f_{s-n-1}^{\theta}(q_n^{\theta})(q_{s-1}^{\theta}\lambda^{\theta}b_s - c)\Delta. \tag{$IC_A^{\theta}(N)$}$$

$IC_A^{\theta}(N)$ contains $N$ inequality constraints. The left hand side (LHS) is the agent's expected discounted payoff when he experiments from period $n$ until the termination date $N$, after failing $n-1$ previous experiments. It has the same structure as the social planner's expected discounted value in expression (2.1), except the agent's value of a success in period $s$ is $b_s$, not $h$. The right hand side (RHS) is the agent's expected discounted payoff when he shirks in period $n$ but experiments from then on until the termination date $N$, after failing $n-1$ previous experiments. Thus, $IC_A^{\theta}(N)$ prevents a one time profitable deviation (shirking) of the agent in all histories when he never shirks before. If the agent deviates and shirks in some periods before, then it is still optimal for him to work thereafter, since he is more optimistic than he would be had he worked.

The moral hazard problem is dynamic. In the $n$-th period, the expected payoff of working within the current period is $(q_n^{\theta}\lambda^{\theta}b_n - c)\Delta$, and shirking gives 0 current payoff. The continuation values of working and shirking are also different, due to two effects. First, the *learning effect*: the (unexpected) shirking makes the agent more optimistic about the project than the principal. After (unexpected) shirking, the principal's belief about the state being $G$ is $q_{n+1}^{\theta} = \frac{q_n^{\theta}(1-\lambda^{\theta}\Delta)}{1-q_n^{\theta}\lambda^{\theta}\Delta} < q_n^{\theta}$, while the agent's belief is still $q_n^{\theta}$. Second, the *end-of-game effect*: the game has only $f_1^{\theta}(q_n^{\theta}) = 1 - q_n^{\theta}\lambda^{\theta}\Delta$ probability to continue if the agent works, while it continues for sure if the agent shirks. Those two effects would increase the continuation value of shirking compared to working. Therefore, the bonus must compensate both the current period costs, and the loss of the continuation value for the agent.

When the agent believes that the principal's type is $\theta \in \Theta$, a contract $C = \{N, W, \boldsymbol{b}\}$ satisfies the IR constraint for the agent, when

$$W_A^{\theta} = W + \sum_{n=1}^{N} \delta^n f_{n-1}^{\theta}(q_0)(q_n^{\theta}\lambda^{\theta}b_n - c)\Delta \geq 0. \qquad IR_A^{\theta}(N)$$

Here, the agent receives both an upfront payment $W$ before the experimentation, and a bonus for success during the experimentation.

Let $S(N)$ be the set of contracts that satisfy both $IC_A^H(N)$ and a binding $IR_A^H(N)$. Hence, the set of contracts, which implement the FIB for the $H$ type when the agent believes her type is $H$, is $S(N_*^H)$.

The worst contract $C^{wt} = \{N_*^H, W^{wt}, \boldsymbol{b^{wt}}\}$ for the $L$ type in the above set of contracts solves the following Program I:

$$\min_{\boldsymbol{b},W} \Pi^L = -W + \sum_{n=1}^{N_*^H} \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n)\Delta$$
$$\text{s.t. } IC_A^H(N_*^H) \text{ and a binding } IR_A^H(N_*^H).$$

Note that the expected payoff of the $L$ type is determined by the payment transfered to the agent before experimentation $-W$, and the kept share of profits once the project succeeds. Here, $q_n^L \lambda^L (h - b_n)\Delta$ is expected payoff of the $L$ type principal for the $n$-th experiment conditional on a success not having arrived yet.

*Lemma* 2.2. The worst contract for the $L$ type in the set $S(N_*^H)$ of contracts is the contract in which the agent's IC constraints bind for every period $n \in \{1, 2, \ldots, N_*^H\}$.

We now provide intuition for why all the IC constraints must bind. The set of contracts that implement the FIB for the $H$ type is large, because the principal has the discretion to give more high powered incentives than required, so that the agent's IC constraints are slack, and then extract the surplus so conferred via a larger sign-up fee, i.e. a lower value of $W$. However, such a contract is more attractive for the $L$ type. We now explain why.

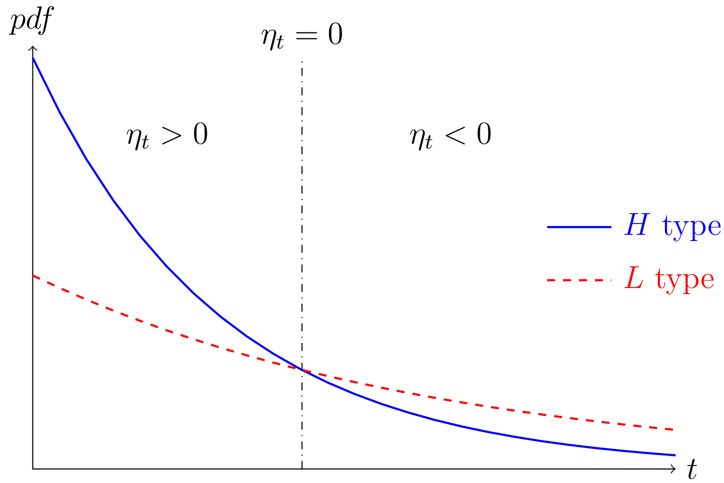


Figure 2.3: The probability of success: the $H$ type and the $L$ type

Figure 2.3 shows the probability of success over time for the two types. Let $\eta_n := f^H_{n-1}(q_0)q^H_n\lambda^H - f^L_{n-1}(q_0)q^L_n\lambda^L$, then $\eta_n\Delta$ denote the difference of probabilities of success for the $n$-th experiment between the $H$ and $L$ types. Though the $H$ type is more likely to succeed than the $L$ type conditional on the state being $G$, the posterior beliefs about the state being $G$ conditional on no success decrease faster for the $H$ type than for the $L$ type, since the $H$ type also learns faster. Therefore, the two curves cross only once. In the earlier periods, the $H$ type is more likely to succeed ($\eta_n > 0$), while in the later periods, the $L$ type is more likely to succeed ($\eta_n < 0$).

Suppose that the agent believes that the principal's type is $H$, but her type is actually $L$. Thus, the agent overvalues the bonus payments when the probability of success is greater for the $H$ type than that for the $L$ type, and is willing to pay a higher sign-up fee. Consequently, to minimize the incentive of the $L$ type to mimic the $H$ type, bonus payments should be minimized in any period where the probability of success is greater for the $H$ type than that for the $L$ type.

Consider now any period where the probability of success is greater for the $L$ type than that for the $H$ type. By the preceding argument, it would seem that the bonus payment in such a period should be increased, since they reduce the incentives of the $L$ type to mimic the $H$ type. However, this is not true. Consider the period right after the crossing point. Raising the bonus in that period makes the agent's continuation value of shirking, in all previous periods, higher. To incentivize the agent to work, the principal needs to raise bonus payments in all previous periods proportionally. It actually makes the $L$ type better off, because the distribution of success for the $L$ type first-order stochastic dominates that for the $H$ type, i.e. $\phi_n := \sum_{s=1}^{n} \eta_s \Delta > 0$ for any $n \geq 1$. To minimize the $L$ type's incentive to mimic the $H$ type, the bonus payment should also be minimized in the period right after the crossing point. By induction, we can show the bonus payments in any period should be minimized in the worst contract for the $L$ type. Thus, all IC constraints should bind in the worst contract for the $L$ type. This result is important, and it will hold whenever the $H$ type has a signaling concern.[15]

Even in the worst contract for the $L$ type, the bonus scheme is an increasing sharing plan that leaves the $L$ type with positive expected payoff

[15]See Lemma B.1 in the Appendix.

during experimentation. Moreover, the agent is also willing to pay some positive sign-up fee, if he believes the principal's type is $H$. Clearly, the $L$ type can obtain strictly positive payoff from that worst contract, and has incentive to mimic the $H$ type. Hence, we conclude:

*Proposition* 2.1. There is no equilibrium that implements the FIB.

### 2.5.2 Separating Equilibrium

The FIB cannot be implemented in any equilibrium, since any contract that achieves the FIB for the $H$ type must violate the IC constraint for the $L$ type. Thus, how can the $H$ type principal separate herself from the $L$ type?

The simplest way is to increase the upfront payment $W$. Since the $L$ type has a lower probability of success than the $H$ type, an upfront payment $W$ that makes the $H$ type profitable could be unprofitable for the $L$ type. Consider the contract $C^{wt}$ we found. Let $\Pi^L(C^{wt})$ be the payoff of the $L$ type from $C^{wt}$, when the agent believes her type is $H$. Now, we introduce a new contract, $C^1 = \{N_*^H, W^1, \boldsymbol{b^{wt}}\}$, with the same termination date and bonus payments, but a higher upfront payment, where $W^1 = W^{wt} + \Pi^L(C^{wt})$. Hence, the $L$ type would obtain 0 payoff from $C^1$, even when the agent believes her type is $H$. However, the $H$ type can obtain strictly positive payoff from $C^1$, when the agent believes her type is $H$. This can be seen from the comparison between contracts $C^1$ and $C^0 = \{N_*^H, 0, \boldsymbol{h}\}$, where $\boldsymbol{h} \in \mathbb{R}^{N_*^H}$ is the vector whose entries are all $h$. Clearly, $C^0$ gives the project to the agent for free, and leaves the principal 0 payoff. In addition, the argument from Lemma 2.2 asserts that the $H$ type prefers $C^1$ to $C^0$, since the former has tight IC constraints for the agent, while the latter has slack IC constraints.

Therefore, $C^1$ can separate the $H$ type from the $L$ type. Furthermore, the $H$ and $L$ type proposing $C^1$ and the null contract respectively, constitute part of an equilibrium. In other words, for any other contract, we can find some belief that prevents a profitable deviation by either type. In fact, we have the following result:

*Lemma* 2.3. For any contract, if the agent believes the principal's type is $L$, then the principal, whatever type she actually is, cannot obtain a strictly positive payoff.

Hence, a pessimistic belief that assigns probability 1 to type $L$ for any off-equilibrium path contract can support the above equilibrium.

The contract $C^1$ simply uses the upfront payment to separate the two types. The $H$ type could design a more sophisticated contract to separate herself, and improve her payoff. Now, we consider the *best separating equilibrium for the $H$ type*, or *BSEH*, i.e. the equilibrium that gives the $H$ type the highest payoff among all separating equilibria.

Given the structure of a contract, the $H$ type principal could use three potential devices to separate herself. First, she could increase the upfront payment $W$ as we have seen. Second, she could reduce the experimentation by terminating it earlier. Since the $L$ type learns slower than the $H$ type, the value of an additional experiment decreases more slowly for the former than for the latter. Hence, an experiment could be more valuable to the $L$ type than to the $H$ type after some time. Thus, shutting down the project earlier makes mimicry less attractive. Third, she could delay the project, by having the agent not work in some periods before terminating the project permanently. Those devices are all costly to the $H$ type; the problem is how to use them in the least costly way.

Though it is not clear at first whether delaying the project temporarily permits separation, we can show that this is not the case. Delaying the project essentially increases the discount factor between two consecutive experiments.[16] Each discount factor affects discounted payoffs linearly, if all terms in the payoff functions are independent with the discount factor. To make the agent follow the induced action plan, bonus payments have to be correlated with future discount factors. However, allowing for the possibility to delay the project does not change the fact that the least costly separation and the least costly pooling make the agent's IC constraints bind. It, in turn, makes the current bonus payment linearly correlate with any future discount factor. Thus, discounted payoffs are linear in each discount factor between two consecutive experiments. Therefore, it is optimal to either never delay the project, or delay the project forever (i.e. terminating the project), from the

[16] Given the time preference $\delta$, the discount factor between the $(n-1)$-th and $n$-th experiments is $\delta_n = \delta^k$, where $k \in \mathbb{N} \cup \{\infty\}$ is the number of delaying periods. $k = 1$ means no delay, and $k = \infty$ means terminating the project.

$H$ type's point of view. Hence, the $H$ type project will never be temporarily delayed in the best equilibrium for the $H$ type. From now on, we will not consider delaying the project temporarily in the main body of the paper. More details can be found in the Appendix.

In the BSEH, the $L$ type must obtain a payoff of 0; let her propose the null contract in the equilibrium. We also assume that the agent assigns probability 1 to the belief about the principal being the $L$ type for any off-equilibrium path contract. Thus, Lemma 2.3 ensures that both types of principal would not deviate to those contracts. We shall not worry about the $H$ type's IR constraint and her incentives to choose the $L$ type's equilibrium contract in the BSEH. The last constraint for the principal we need to consider, is the $L$ type's incentives to deviate to the $H$ type's equilibrium contract. Given the $H$ type's contract $C = \{N, W, \boldsymbol{b}\}$, the $L$ type's IC constraint is

$$\Pi^L = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n)\Delta \leq 0. \qquad IC_L^H(N)$$

The agent's IR and IC constraints are the same as before. Hence, the $H$'s equilibrium contract, $C^{sep} = \{N^{sep}, W^{sep}, \boldsymbol{b^{sep}}\}$, in the BSEH will solve the following Program II:

$$\max_{N,W,\boldsymbol{b}} \Pi^H = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^H(q_0) q_n^H \lambda^H (h - b_n)\Delta$$
$$\text{s.t. } IC_A^H(N)\text{ , } IR_A^H(N)\text{ , and } IC_L^H(N)$$

To solve this program, we drop the constraint $IR_A^H(N)$, and verify it later. In the relaxed program, $IC_L^H(N)$ must bind, which determines the time zero transfer $W$, otherwise we can decrease $W$ to obtain a higher payoff for the $H$ type. For any given $N$, the agent's IC constraints in $IC_A^H(N)$ also bind for every period $1 \leq n \leq N$ due to the same reason as in Lemma 2.2, which determines the bonus scheme $\boldsymbol{b}$. The last thing is to determine the termination date $N$. We take the limit $\Delta \to 0$ to avoid the integer problem, and obtain the following result:

*Proposition* 2.2. In the limit $\Delta \to 0$, there is a unique equilibrium contract for the $H$ type $C^{sep} = \{T^{sep}, W^{sep}, \boldsymbol{b^{sep}}\}$ in the BSEH, and it has the following features:

- Under experimentation: $0 < T^{sep} < \frac{1}{2}T_*^H$, where $T^{sep}$ satisfies

$$\lambda^H e^{-\lambda^H T^{sep}} - \lambda^L e^{-\lambda^L T^{sep}} = (\lambda^H - \lambda^L)e^{-\lambda^H T_*^H} e^{(\lambda^H - \lambda^L)T^{sep}};$$

- A positive base wage: $W^{sep} > 0$;
- A (weakly) increasing bonus plan: $b_t^{sep}$ is weakly increasing in $t$, and $0 < b_t^{sep} < h$ for $0 \leq t \leq T^{sep}$.

**Rent Sharing** - In such an equilibrium, the $H$ type principal shares some surplus with the agent. She obtains less than $1 - \lambda^L/\lambda^H$ of the total surplus $V_0^H(T^{sep})$, and the agent obtains more than $\lambda^L/\lambda^H$ of it. The surplus is shared through both a positive base wage $W^{sep}$ before experimentation, and an increasing bonus scheme in the period when the success arrives. Even though the principal has all the bargaining power and the agent has no financial constraint, she leaves some rents to the agent, to signal that her type is $H$. This phenomenon is first shown in Beaudry (1994). The $L$ type is left with no rents, and exits the market in this equilibrium.

**Role of Learning** - An increasing bonus scheme means the principal would give a larger reward to the agent if the success comes later. This is due to the decline of the posterior beliefs about the viability of the project. A failed experiment drives down the agent's belief about the viability of the project, and greater incentives are needed in later periods.

Inefficiently early termination is the second type of signaling cost. Since the initial payment $W$ must equal the $L$ type's expected share of the profits if the $L$ type proposes the $H$ type's contract, the $H$ type principal's payoff can be seen as the difference between her expected share of the profits and that of the $L$ type. Extending the experiment has two marginal effects.

Consider only the last period payoffs, the marginal net benefit for extending the experiment is the difference between the $H$ type's expected share of her profits and that of the $L$ type in that period:

$$q_0 e^{-\rho T}(\lambda^H e^{-\lambda^H T} - \lambda^L e^{-\lambda^L T})(h - b_T), \tag{2.2}$$

where $\lambda^\theta e^{-\lambda^\theta T}$ is the probability of success in the last period $T$ for the $\theta \in \Theta$ type project in the $G$ state, and $h - b_T$ is the share retained by the principal.

Consider all other periods payoffs, due to the dynamic moral hazard problem, the marginal net cost for extending the experiment is the difference between the increased accumulated bonuses that are given up by the $H$ type and that of the $L$ type:

$$q_0 e^{-\rho T}(e^{-\lambda^L T} - e^{-\lambda^H T})(\lambda^H b_T - c), \tag{2.3}$$

where $e^{-\lambda^L T} - e^{-\lambda^H T}$ is the probability difference of failures until the last period $T$ between the $H$ and $L$ type in the $G$ state, and $\lambda^H b_T - c$ represents the increased bonuses.[17]

Equalizing the above two effects gives the equilibrium terminating date $T^{sep}$. The inefficiency comes from the combination of the private information and actions on both sides. Moreover, it is aggravated because of the dynamic moral hazard problem due to *learning*. If the state of the project is known, then extending the experiment will not increase the agent's share of the profits in previous periods; the principal does not need to give more incentives to the agent than the current experimenting cost. Hence, the effect represented by expression (2.3) is zero, when there is no learning towards the state of the project. Then, the inefficiency is determined only by expression (2.2), and it would diminish, i.e. the terminating date goes to be infinity, as $\lambda^L$ goes to 0. However, in our model, the state of the project is unknown; thus, prolonging experimentation increases the bonus payments in all periods. Even when $\lambda^L$ goes to 0, expression (2.3) is strictly positive. In that case when $\lambda^L = 0$, the

[17]The marginal increase of a bonus at time $t$ by extending the experiment is $\lambda^H b_T - c$ with proper discounting, i.e. $\frac{\mathrm{d}b_t}{\mathrm{d}T} = e^{-\rho(T-t)}(\lambda^H b_T - c)$.

equilibrium termination date, $T^{sep}$, is determined by

$$q_0 e^{-\rho T} e^{-\lambda^H T} \lambda^H (h - b_T) = q_0 e^{-\rho T} (1 - e^{-\lambda^H T})(\lambda^H b_T - c),$$

where $\lambda^H(h - b_T)$ and $(\lambda^H b_T - c)$ are the expected payoffs for the principal and the agent in the $G$ state, respectively. Therefore, for $T = T^{sep}$

$$e^{-\lambda^H T} = \frac{\lambda^H b_T - c}{\lambda^H h - c} = \frac{l^H}{l_T} = e^{-\lambda^H (T_*^H - T)},$$

where $l^H = \frac{c}{\lambda^H h - c}$ and $l_T = \frac{c}{\lambda^H b_T - c}$ are the likelihood ratio of the posterior beliefs that the state of the $H$ project is $G$ at the efficient termination time $T_*^H$ and the termination time $T$, respectively. Hence, we have $T^{sep} = \frac{1}{2} T_*^H$ when $\lambda^L = 0$.

We can show that the equilibrium termination date $T^{sep}$ is decreasing in $\lambda^L$, therefore the $H$ type project is only operated less than half of the efficient time.

**Limited Liability** - Consider a different setting where the principal is publicly known to be the $H$ type, but the agent is protected by limited liability, i.e. $W$ and $\boldsymbol{b}$ must be positive. Then, it is easy to show that the optimal contract in that case is the $H$ type's equilibrium contract in the BSEH of our original model when $\lambda^L = 0$, in which $W^{sep} = 0$. Hence, our model provides an alternative explanation for the limited liability. Economists observe that the agent is sometimes protected by limited liability. The principal cannot charge anything from the agent, or sell the project to him. A standard argument is that the agent may have some financial constraints. However, in our model, the principal is willing not to charge anything from the agent, not due to any limited liability protection, but to signal her type. Without financial constraints, we still may observe the "limited liability" phenomenon in equilibrium.

**Unobservable Successes** - We assume a success is publicly observable and verifiable so far. However, it is sometimes very costly for the principal to verify a success. If those costs are prohibitively high, then a contract can only depend on the agent's report on successes, but not successes per se. Hence, the principal may face an additional incentive problem: the agent may delay reporting a success to obtain a higher discounted payoff if the increase of bonuses over time offsets the costs of discounting.

However, this is not the case for the equilibrium contract. The agent has no incentive to hide a success from the principal, as long as he is given the exact incentives to experiment when a success can be costlessly verified, as in our original model. In other words, given the binding IC constraints for the agent to experiment, the agent would like to report a success truthfully when it arrives.

Formally, Lemma B.2 in the Appendix shows that binding IC constraints for the agent implies, for any $n$ before the termination date,

$$b_n \geq \delta b_{n+1}.$$

When $\delta < 1$, the above inequality is strict. Since all the remaining results in this paper feature no temporary suspension and binging IC constraints for the agent, the agent has no incentive to delay reporting a success even when it cannot be observed by the principal.

**Equilibrium Refinement** - We have shown two different separating equilibria, there are many others. Let us consider a commonly used equilibrium refinement, i.e. the intuitive criterion from Cho and Kreps (1987). We have the following result:

*Proposition* 2.3. The BSEH survives the intuitive criterion, and all equilibria which give a lower payoff to the $H$ type fail the intuitive criterion.

Note that all other separating equilibria gives less payoff to the $H$ type, thus fail the intuitive criterion. In addition, according to Lemma 2.5 in Section 2.5.4, all other equilibria that give 0 payoff to the $L$ type also give less equilibrium payoff to the $H$ type than the BSEH does, thus fail the intuitive criterion. However, there may be some pooling equilibrium gives more payoff for $H$ type. We examine them in the next section.

### 2.5.3 Pooling Equilibrium

A pooling equilibrium is an equilibrium in which both types of principal propose the same equilibrium contract, and the agent holds the prior belief $\beta_0$ about the $H$ type after the equilibrium contract is proposed.

Let us first consider the efficient solution when the type of the project is unknown at the ex ante stage. We call it the *mixed* project. When the prior belief is $\beta_0$, the value of the mixed project with the termination date $N$ is

$$V(N) := \beta_0 V^H(N) + (1-\beta_0)V^L(N).$$

Let $N_*$ be the *efficient termination date for the mixed project*, which maximizes $V(N)$, and let $T_*$ be the limit of $N_*\Delta$ when $\Delta \to 0$. When the prior $\beta_0$ is too pessimistic, i.e. $q_0\lambda_0 h \leq c$, where $\lambda_0 := \beta_0\lambda^H + (1-\beta_0)\lambda^L$, the efficient termination date is 0. Thus, any pooling equilibrium is trivial, and features no experiments or transfers. Otherwise, when the prior $\beta_0$ is not too small, i.e. $q_0\lambda_0 h > c$, the efficient termination date for the mixed project is in between 0 and the efficient termination date for the $H$ type. We will only consider the latter case.

One simple pooling equilibrium is that both types of principal sell the mixed project to the agent, and extract all expected surplus of the mixed project. Lemma 2.3 ensures no one would like to deviate to any off-equilibrium path contract, where the agent believes the deviator is the $L$ type. However, the $H$ type could achieve a higher payoff in some other pooling equilibrium. Again, we consider the *best pooling equilibrium for the $H$ type*, or *BPEH*.

When both types of principal pool together, the agent would update his beliefs on both the type and the state of the project during the experimentation. Furthermore, even the type and the state of the project are independent at the outset, the posterior beliefs after some failures will be correlated. Let $q_n$ be the posterior belief about the state of the project being $G$, after $n-1$ failures, and $\beta_n$ be the posterior belief about the type of the project being

$H$ conditional on the state being $G$, after $n-1$ failures. The posterior belief about the type of the project being $H$ conditional on the state being $B$ is the prior $\beta_0$, no matter how many failures the project has. Hence,

$$q_n = \frac{q_0[\beta_0(1-\lambda^H\Delta)^{n-1} + (1-\beta_0)(1-\lambda^L\Delta)^{n-1}]}{q_0[\beta_0(1-\lambda^H\Delta)^{n-1} + (1-\beta_0)(1-\lambda^L\Delta)^{n-1}] + 1 - q_0},$$

and

$$\beta_n = \frac{\beta_0(1-\lambda^H\Delta)^{n-1}}{\beta_0(1-\lambda^H\Delta)^{n-1} + (1-\beta_0)(1-\lambda^L\Delta)^{n-1}}.$$

Let $\lambda_n := \beta_n\lambda^H + (1-\beta_n)\lambda^L$ be the expected arrival rate of success conditional on the state being $G$, after $n-1$ failures. Both $q_n$ are $\beta_n$ are strictly decreasing over time towards 0, and $\lambda_n$ is strictly decreasing over time towards $\lambda^L$. That is, both the posterior belief about the $G$ state and the posterior belief about the $H$ type conditional on the state being $G$ go down after a failure.

Now the agent's IC constraints are more complex due to the learning process. However, those constrains can be easily transformed from the constraints in the separating equilibrium by using correct beliefs. Given $C = \{N, W, \boldsymbol{b}\}$, the agent's IC constraints become, for $1 \leq n \leq N$

$$\sum_{s=n}^{N} \delta^{s-n} f_{s-n}(q_n, \beta_n)(q_s\lambda_s b_s - c)\Delta \geq \sum_{s=n+1}^{N} \delta^{s-n} f_{s-n-1}(q_n, \beta_n)(q_{s-1}\lambda_{s-1} b_s - c)\Delta, \tag{$IC_A(N)$}$$

where $f_m(q, \beta) = q[\beta(1-\lambda^H\Delta)^m + (1-\beta)(1-\lambda^L\Delta)^m] + 1 - q$ is the probability of failing $m$ times for a project starting with probability $q$ being in state $G$, and probability $\beta$ being type $H$ conditional on being in state $G$.

The agent's IR constraint becomes

$$W_A = W + \sum_{n=1}^{N} \delta^n f_{n-1}(q_0, \beta_0)(q_n \lambda_n b_n - c) \geq 0. \qquad IR_A(N)$$

In the BPEH, the principal's incentive is trivially satisfied on the equilibrium path. By Lemma 2.3, no type has an incentive to deviate to any other contract once we assume that the agent assigns probability 1 to the belief about the principal being the $L$ type for those contracts. We shall not worry about the IR constraint for the $H$ type. The IR constraint for the $L$ type is

$$\Pi^L = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n)\Delta \geq 0. \qquad IR_L(N)$$

Then, the equilibrium contract, $C^{pl} = \{N^{pl}, W^{pl}, \boldsymbol{b^{pl}}\}$, in the BPEH will solve the following Program III:

$$\max_{N,W,\boldsymbol{b}} \Pi^H = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^H(q_0) q_n^H \lambda^H (h - b_n)\Delta$$
$$\text{s.t. } IC_A(N),\ IR_A(N), \text{ and } IR_L(N)$$

$IR_L(N)$ is slack in the above program, so we drop it and verify it after solving the relaxed program. $IR_A(N)$ must bind in the relaxed program, otherwise we can decrease $W$ to obtain more payoff for both types. Thus, it determines the time zero transfer $W$. Given any $N$, we can show that the agent's IC constraints in $IC_A(N)$ must bind for all $1 \leq n \leq N$, which determines the bonus scheme $\boldsymbol{b}$.[18] Then, we can solve the program and find the equilibrium

[18]The reason for binding IC constraints of the agent is similar to Lemma 2.2 and Proposition 2.2. Notice that the agent obtaining no rents implies that the payoff of the $H$ type principal equals the sum of the project's value and $(1 - \beta_0)$ portion of the payoff difference between the $H$ and $L$ type. Hence, given a fixed amount of experimentation, the pooling contract in the BPEH must maximize the payoff difference.

termination time.

*Proposition* 2.4. Assume that $q_0\lambda_0 h > c$. In the limit $\Delta \to 0$, there is a unique equilibrium contract for the $H$ type $C^{pl} = \{T^{pl}, W^{pl}, \boldsymbol{b^{pl}}\}$ in the BPEH, and it has the following features:

- Under experimentation: $0 < T^{pl} < T_*$;
- Sign-up fee: $W^{pl} < 0$;
- A (weakly) increasing bonus plan: $b_t^{pl}$ is weakly increasing in $t$, and $0 < b_t^{pl} < h$ for $0 \le t \le T^{pl}$.

Moreover, the equilibrium payoff for the $H$ type in the BPEH is strictly increasing in the prior belief about the $H$ type $\beta_0$, and converges to $V_0^H(T_*^H)$ as $\beta_0 \to 1$.

The above contract shares some common features with the $H$ type's equilibrium contract in the BSEH. It gives the agent exact enough incentives to work, i.e. the IC constraints for the agent always bind. When there is no discounting, i.e. $\rho = 0$, the bonus payments are constant over time. When $\rho > 0$, they are strictly increasing. Furthermore, as in the BSEH, the increase of bonus payments cannot offset the cost for discounting. Hence, even when a success is not publicly observed, the agent has no incentive to delay reporting a success.

**Pooling Costs** - Unlike the case in the BSEH, where the $H$ type gives a positive base wage to the agent, in the BPEH, both types charge a sign-up fee $-W^{pl} > 0$ from the agent before experimentation, leaving no rents to the agent. Since the agent has no financial constraint, he is willing to pay the sign-up fee to exchange for some share of profits, i.e. bonuses, once the project succeeds. The amount that the $H$ type principal extracts from the agent cannot recover her FIB surplus due to the pooling costs. When the $L$ type pools with the $H$ type, the latter has to compensate more costs to the agent, but only partially extracts the promised bonuses back from the agent. Both the additional compensation costs and the non-recoverable bonuses are proportional to the population of the $L$ type, i.e. $1 - \beta_0$. Thus, those pooling costs are diminishing as the prior belief about the $H$ type goes to 1. On the other hand, the $L$ type is left with strictly positive rents by pooling with the $H$ type.

**Learning Viability and Quality** - When both types pool together, the agent needs to learn both the viability and the quality of the project. A failure means the agent would be more pessimistic about both the viability and the (conditional) quality[19], thus learning creates dynamic moral hazard costs. However, learning the viability and the quality has different consequences. If what needs to learn is the viability ($0 < q_0 < 1$), it is really costly to extend the experiment, since the posterior beliefs of generating success eventually

[19] The conditional quality means the project's quality conditional on it being viable. Note that the conditional quality is what matters for the agent's incentives, and the unconditional quality could be increasing with one more failure after some point.

goes to 0. However, if what needs to learn is solely the quality ($q_0 = 1$, and $0 < \beta_0 < 1$), the principal can always incentivize the agent to work with bounded bonuses, i.e. $b_t \leq \frac{c}{\lambda^L}$ for any $t > 0$. Formally, let $T$ be the termination date, then the limit of the marginal effect of extending experiment for the last period bonus is

$$\lim_{T\to\infty} \frac{\mathrm{d}b_T}{\mathrm{d}T} = \begin{cases} \infty & \text{if } q \in (0,1), \\ 0 & \text{if } q = 1. \end{cases}$$

**Under Experimentation**[20] - It is straightforward that the equilibrium termination time is less than the efficient termination time for the $H$ type. In addition, it is also less than the efficient termination time for the mixed project. We now provide some intuition.

Since the agent has no rents in the equilibrium, the value generated from the mixed project is shared by two types of principal. Thus, the equilibrium payoff of the $H$ type is equal to the sum of (1) the generated value of the mixed project, and (2) $(1 - \beta_0)$ portion of the equilibrium payoff difference between the $H$ and $L$ type.

We now consider the marginal effect of extending the mixed project at its efficient termination time $T_*$, and show that it is negative. Thus, it is better to terminate the project earlier.

Consider the first part of the $H$ type's payoff, i.e. the value generated from the mixed project. By the definition of the efficient termination time

[20]The contract in the BPEH features under experimentation only for the $H$ type principal, while the same contract requires over experimentation for the $L$ type principal.

for the mixed project, extending the project has zero marginal effect on the generated value. For the second part of the $H$ type's payoff, extending the project at $T_*$ does not generate any marginal benefit for the principal, since the required bonus (for incentivizing the agent to work) has to be equal to the lump-sum profit of success. However, due to the dynamic moral hazard problem and the fact that the $H$ type is more likely to succeed before the termination date than the $L$ type, the costs for sharing more bonuses to the agent in all periods incurred by the $H$ type is larger than that incurred by the $L$ type. Thus, extending the project does generate some additional costs. Therefore, the marginal effect of extending the project at $T_*$ is negative; the $H$ type principal would like to terminate the project earlier than the efficient termination time for the mixed project.

**Equilibrium Refinement** - Though the BPEH may give a higher payoff to the $H$ type than the BSEH does, we can show that it fails the intuitive criterion:

*Proposition* 2.5. The BPEH fails the intuitive criterion.

This is because there exists a contract that enlarges the payoff difference between the two types of principal without scarifying efficiency, when the agent believes the deviator is the $H$ type. Enlarging the payoff difference between the two types leaves room to increase the $H$ type's payoff and decrease the $L$ type's payoff at the same time. Keeping the efficiency makes the agent still participate voluntarily when the $H$ type asks for more share of surplus. Hence, such a contract could make the BPEH fail the intuitive criterion.

### 2.5.4 Best Equilibrium for the High Type

We have examined both the BSEH and the BPEH. In the BSEH, the equilibrium payoff for the $H$ type is independent of the prior belief $\beta_0$, and is less than her FIB surplus. However, in the BPEH, the equilibrium payoff for the $H$ type is strictly increasing in $\beta_0$ when her payoff is strictly positive, i.e. when $q_0\lambda_0 h > c$, and converges to her FIB surplus as $\beta_0$ goes to 1. Obviously, we can find a cutoff prior belief about the $H$ type, $\beta_c \in (0, 1)$, such that when $\beta_0 < \beta_c$, the BSEH gives the $H$ type a higher payoff than the BPEH does, and when $\beta_0 > \beta_c$, the BPEH gives the $H$ type a higher payoff than the BSEH does.

For the purpose to find the best equilibrium for the $H$ type, we still need to examine all other equilibria, like partial separating equilibrium which allows the principal to randomize over different contracts.

To use the results we have already established, we categorize all equilibria into two classes: (1) the set of equilibria that gives the $L$ type strictly positive payoff, and (2) the set of equilibria that gives the $L$ type 0 payoff.

We will show that (1) among all equilibria that gives the $L$ type strictly positive payoff, the BPEH gives the $H$ type the highest payoff, and (2) among all equilibria that gives the $L$ type 0 payoff, the BSEH gives the $H$ type the highest payoff. Thus, we can conclude that the best equilibrium for the $H$

type is either BSEH or BPEH, depending on $\beta_0$.

First, consider the set of equilibria that gives the $L$ type strictly positive payoff. This means that no equilibrium path contract can fully reveal the $L$ type. In other words, the $L$ never chooses any contract by herself alone in the equilibrium, she always pools with the $H$ type. Moreover, there must be at least one contract proposed by the $L$ type in the equilibrium, such that the agent's belief about the principal being the $H$ type does not exceed the prior $\beta_0$. Together with Proposition 2.4, we can show that:

*Lemma* 2.4. For any equilibrium such that the $L$ type obtains a strictly positive payoff, the $H$ type's payoff cannot exceed her payoff in the BPEH.

Now, consider the set of equilibria that gives the $L$ type 0 payoff. Fix any such equilibrium $E$, and any equilibrium contract $C$ proposed by the $H$ type in $E$. Suppose the agent forms belief $\beta \in [0, 1]$ about $H$ type after $C$ is proposed. Then $C$ must satisfy the agent's IC and IR constraints, and the $L$ type cannot get strictly positive payoff by proposing the same contract. We solve the best contract for the $H$ type subjected to the above constraints, and show that the best outcome the $H$ type can obtain is increasing in $\beta$. In addition, we come back to our BSEH result when $\beta = 1$. Hence, we show the following result:

*Lemma* 2.5. For any equilibrium such that the $L$ type obtains 0 payoff, the $H$ type's payoff cannot exceed her payoff in the BSEH.

With the above two lemmas, we can conclude:

*Proposition* 2.6. There exits a cutoff belief $\beta_c \in (0, 1)$,

- when $\beta_0 < \beta_c$, the BSEH gives the $H$ type the highest payoff among all equilibria;
- when $\beta_0 > \beta_c$, the BPEH gives the $H$ type the highest payoff among all equilibria.

Thus, in the best equilibrium for the $H$ type, the termination time of the $H$ type project is inefficiently early, and she has to share rents with either the agent (in the BSEH), or the $L$ type (in the BPEH). Since the equilibrium contract for the $H$ type in the BSEH also maximizes the payoff difference between the $H$ and $L$ type when the agent assigns probability 1 to the belief about the principal being the $H$ type, together with Lemma B.3 in the Appendix, we can conclude that the BSEH is the equilibrium that maximizes the payoff difference between the two types of principal. Hence, if we consider a pre-game where the principal can invest in the quality of a project before the signaling game, namely exerting efforts to increase the probability $\beta_0$ of having a $H$ type project, her investment decision, i.e. the prior $\beta_0$, is endogenously determined by the payoff difference between the two types in the equilibrium of the signaling game. Thus, the BSEH permits the highest prior, as long as the marginal investment cost is increasing in $\beta_0$.

## 2.6 Introducing a Mediator

The above results have shown the conflict between signaling and providing incentives to the agent. We now consider the implication of allowing for a mediator who designs a mechanism (i.e. a menu of contracts).

As in the rest of the paper, we assume that the goal of the mediator is to maximize the payoff of the $H$ type principal. This is a natural benchmark that facilitates the comparison across the signaling game and the mechanism design problem. The mediator needs to attract the $H$ type project, which generates the social surplus. Thus, it must provide her with a higher payoff than she can obtain in the signaling game without the mediator. Furthermore, if there are multiple profit-oriented mediators who compete in a Bertrand fashion, they would maximize the $H$ type's payoff in the equilibrium. We also consider the FIB as the objective of the mediator in the Appendix.

The basic role of the mediator is to communicate with both the principal and the agent, and disclose information at a proper time. Thus, we can achieve immediate separation of the two types of the principal, but this information can be concealed from the agent, at least temporarily. In this way, the agent's IR and IC constraints only need to hold in expectation, though different principals offer different contracts.

Science Exchange, an online platform that connects scientific researchers with experimental service providers, is an example of a mediator. The researcher corresponds to the principal in our model, while the service provider is the agent. The platform has its own system to verify the qualification of service providers, so one shall not worry much about their unobservable abilities. The major problem for service providers and for the platform is that the quality of ideas brought in by researchers is, by their nature, hard to evaluate. Furthermore, moral hazard on the part of service providers is likely to be important, given the uncertainties associated with the research process. The role played by the platform is not only to reduce search and transaction costs, but also to design contracts, protect intellectual properties and confidentiality, and facilitate communications.[21] Hence, researchers' confidential information can be concealed from service providers, and disclosed to them only when necessary.

Let us formally examine the mediator's mechanism design problem. A mechanism, $\mathcal{M} = \{\mathcal{C}^H, \mathcal{C}^L\}$, is an extensive form game containing a menu of two contracts. A contract, $\mathcal{C}^\theta = \{N^\theta, W^\theta, \boldsymbol{b}^{\boldsymbol{\theta}}\}$, is a triple as before, where $\theta \in \Theta$, $N^\theta \in \mathbb{N}_0$, $W^\theta \in \mathbb{R}$, and $\boldsymbol{b}^{\boldsymbol{\theta}} \in \mathbb{R}^{N^\theta}$.

Given a mechanism, $\mathcal{M} = \{\mathcal{C}^H, \mathcal{C}^L\}$, the agent and the principal simultaneously make their participation decision, and the principal reports her type to the mediator. If both parties accept the mechanism and the principal reports $\theta \in \Theta$, then the mediator implements the contract $\mathcal{C}^\theta$ as follows:

- In any period $n \leq N^\theta$ (before success), the mediator will recommend that the agent work. If the project succeeds in that period, the principal

[21] See https://www.scienceexchange.com/trust and other sites under the same domain for details.

will transfer $b_n^\theta$ to the agent, and the game ends; otherwise, there is no payment.

- In the period $n = N^\theta + 1$ (before success), the mediator will recommend that the agent not work, and the game ends.

- The principal will transfer $W^\theta$, which is measured in the time-zero discounted value, to the agent when the game ends.

The agent's action set is the same as before. He can choose whether to follow the mediator's recommendations (i.e. whether to exert effort), which is unobservable to both the principal and the mediator.

Before we go any further, we should explain the space of mechanisms that we study. They are direct mechanisms with two restrictions. First, recommendations for delaying a project temporarily are not considered, but this is without loss for the purpose of finding the *optimal mechanism for the $H$ type principal*. Second, random recommendations are not allowed. It has some loss. However, when the period length shrinks, we show that the $H$ type principal can obtain a payoff that converges to her FIB surplus via pure recommendations. In the Appendix, we will introduce random recommendations, through which the $H$ type's payoff could be further improved for any fixed $\Delta > 0$.

Now we consider constraints in this new environment. We say a mechanism $\mathcal{M} = \{\mathcal{C}^H, \mathcal{C}^L\}$ is *feasible*, if both types of principal and the agent are willing to participate, both types of principal are willing to report truthfully, and the agent is willing to follow recommendations. In other words, a feasible mechanism satisfies the following IR and IC constraints.

The type $\theta \in \Theta$ principal's IR constraint is

$$\Pi^\theta = -W^\theta + \sum_{n=1}^{N^\theta} \delta^n f_{n-1}^\theta(q_0) q_n^\theta \lambda^\theta (h - b_n^\theta)\Delta \geq 0. \qquad IR_\theta(\mathcal{M})$$

The agent's IR constraint is

$$W_A = \sum_{\theta \in \Theta} \beta_0^\theta [-W^\theta + \sum_{n=1}^{N^\theta} \delta^n f_{n-1}^\theta(q_0)(q_n^\theta \lambda^\theta b_n^\theta - c)\Delta] \geq 0, \qquad IR_A(\mathcal{M})$$

where $\beta_0^H = \beta_0$ and $\beta_0^L = 1 - \beta_0$.

The type $\theta \in \Theta$ principal's IC constraint is

$$-W^\theta + \sum_{n=1}^{N^\theta} \delta^n f_{n-1}^\theta(q_0) q_n^\theta \lambda^\theta (h - b_n^\theta)\Delta \geq -W^{\theta'} + \sum_{n=1}^{N^{\theta'}} \delta^n f_{n-1}^\theta(q_0) q_n^\theta \lambda^\theta (h - b_n^{\theta'})\Delta, \qquad IC_\theta^{\theta'}(\mathcal{M})$$

where $\theta' \in \Theta \backslash \{\theta\}$.

The above constraints are straightforward extensions from the signaling game. The only difference is that the transfers now depend on the reported type.

The agent's IC constraints are more involved, since signaling can take place at any time. Suppose the $H$ type experiments longer than the $L$ type, i.e. $N^H \geq N^L$.[22] Since recommendations for both types are the same until $N^L$, the agent's beliefs about the type of the principal and about the state of

[22] When $N^H < N^L$, we can define the IC constraints for the agent in the same way, though it is irrelevant for the optimal mechanism for the $H$ type principal.

the project evolve in the same way as in the pooling equilibrium. However, if $N^H > N^L$ and if the agent is recommended to work in the period $N^L + 1$, his posterior belief about the type of principal being $H$ will jump to 1, but his posterior belief about the state of the project being $G$ will jump down, since a failure of the $H$ type is more informative about the state being $B$ than a failure of the $L$ type. Specifically, the agent's IC constraints are

$$\begin{cases} \nu_n \chi_n^H + (1-\nu_n)\chi_n^L \geq 0, & \text{for } 1 \leq n \leq N^L \ , \\ \chi_n^H \geq 0, & \text{for } N^L + 1 \leq n \leq N^H \text{ if } N^H > N^L, \end{cases} \qquad IC_A(\mathcal{M})$$

where $\nu_n = q_n\beta_n + (1-q_n)\beta_0$ is the posterior belief about the principal's type being $H$ after $n-1$ failures, and $\chi_n^\theta$ is the difference of the expected payoffs for the agent between always following recommendations and shirking in period $n$, but following all remaining recommendations thereafter, conditional on the principal's type being $\theta$. Hence, $\chi_n^\theta = \sum_{s=n}^{N^\theta} \delta^{s-n} f_{s-n}^\theta(q_n^\theta)(q_s^\theta \lambda^\theta b_s^\theta - c)\Delta - \sum_{s=n+1}^{N^\theta} \delta^{s-n} f_{s-n-1}^\theta(q_n^\theta)(q_{s-1}^\theta \lambda^\theta b_s^\theta - c)\Delta$.

*Remark.* The equilibrium contracts in the signaling game remain feasible. Moreover, Proposition 2.1 implies that there is no feasible mechanism that implements the FIB.

We now show how a mediator can achieve a strictly higher payoff for the $H$ type, compared with both the BPEH and the BSEH, by relaxing only the IR constraint of the agent.

Consider feasible mechanisms in which the agent is recommended to not work at all if the principal reports $L$, i.e. $N^L = 0$. Hence, the agent would learn the type of the principal before he starts to work, but after he accepts the offer. The reason that the mediator can achieve a strictly higher payoff for the $H$ type compared with the BPEH is as follows. First, by recommending not working for the $L$ type, the total social surplus increases. Second, it is also cheaper to incentivize the agent to work without the $L$ type pooling during the experimentation. Comparing with the BSEH, the mediator can also achieve a strictly higher payoff for the $H$ type. Recall that the $H$ type leaves some rents to the agent in the BSEH. The mediator can lower the lump-sum transfers in both contracts by the same amount, while keeping the agent's IR constraint satisfied in expectation before he learns the principal's type. This will not change the incentive of the $L$ type, but increase the payoffs for both types. Thus, separation is less costly when it occurs right after the offer acceptance.

The mediator can further improve the $H$ type's payoff by also relaxing the IC constraints of the agent. This requires that the mediator also recommends that the agent work on the $L$ type project for some time. Thus, separation takes place in the period when recommendations differ for different types of principal.

In our quasi-linear environment, the total surplus generated in any feasible mechanism is shared by both types of principal and the agent. Let $\Pi^\theta$ be the payoff obtained by the $\theta \in \Theta$ type of principal, and $W_A$ be the payoff obtained by the agent. Then, we have

$$\beta_0 V^H(N^H) + (1-\beta_0)V^L(N^L) = \beta_0 \Pi^H + (1-\beta_0)\Pi^L + W_A.$$

On the one hand, feasibility requires that both types of principal and the agent must obtain non-negative payoffs. On the other hand, the value generated by the $L$ type, $V^L(N^L)$, is strictly decreasing in $N^L$ for $N^L \geq 1$, and the value generated by the $H$ type, $V^H(N^H)$, is maximized at $N^H = N^H_*$. Thus, for any $N^L \geq 1$, we have

$$\begin{aligned}\Pi^H =& V^H(N^H) + \frac{1-\beta_0}{\beta_0}(V^L(N^L) - \Pi^L) - \frac{1}{\beta_0}W_A \\ \leq& V^H(N^H_*) + \frac{1-\beta_0}{\beta_0}V^L(1).\end{aligned}$$

This means that in any feasible mechanism with $N^L \geq 1$, the payoff of $H$ type principal has an upper bound, which is strictly positive for a small $\Delta$.

We now show that there exists a feasible mechanism with $N^L \geq 1$ that achieves the upper bound. Clearly, such a mechanism must (1) satisfy feasibility, (2) leave no rents to both the agent and the $L$ type principal, and (3) implement the $H$ type project efficiently ($N^H = N^H_*$) and the $L$ type project almost efficiently ($N^L = 1$). Thus, the last condition determines the efficiency of both projects, and so the time when recommendations differ for different reports of the principal. Together with the second condition, they determine the lump-sum transfers:

$$W^L = \delta q_0 \lambda^L (h - b^L_1)\Delta, \tag{2.4}$$

$$W^H = -[V^H(N^H_*) + \frac{1-\beta_0}{\beta_0}V^L(1)] + \sum_{n=1}^{N^H_*} \delta^n f^H_{n-1}(q_0) q^H_n \lambda^H (h - b^H_n)\Delta. \tag{2.5}$$

Now we show how to use bonus payments to incentivize both the agent (to work) and the $L$ type principal (to report truthfully) at the same time.

The mediator will reveal the principal's report – which is her type if she reports truthfully – in period 2 by recommending differently. If the mediator recommends working, then the agent will learn that the principal's type is $H$, and his IC constraints from period 2 to $N_*^H$ are the same as in the separating equilibrium. However, in the first period, the agent's belief about the principal being type $H$ is the prior $\beta_0$, and his IC constraint is simplified to

$$q_0(\beta_0\lambda^H b_1^H + (1-\beta_0)\lambda^L b_1^L) - c \geq \beta_0 q_0 \lambda^H \sum_{s=2}^{N_*^H} \delta^{s-1}(1-\lambda^H\Delta)^{s-2}(\lambda^H b_s^H - c)\Delta. \tag{2.6}$$

The agent requires enough bonus payment in expectation to work in period 1. Thus, given the bonus payment from the $H$ type, the inequality (2.6) determines a lower bound for the bonus payment from the $L$ type $b_1^L$.

The $L$ type's IC constraint is $-W^H + \sum_{n=1}^{N_*^H} \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n^H)\Delta \leq 0$, i.e.

$$V^H(N_*^H) + \frac{1-\beta_0}{\beta_0} V^L(1) - \sum_{n=1}^{N_*^H} \delta^n \eta_n (h - b_n^H)\Delta \leq 0. \tag{2.7}$$

Since the $H$ type is more likely to succeed in period 1 than the $L$ type (i.e. $\eta_1 > 0$), the inequality (2.7) determines an upper bound for the bonus payment from the $H$ type $b_1^H$.

Thus, distinct bonus payments for the first experiment can be used to incentivize both the $L$ type principal and the agent. We can make $b_1^H$ small to prevent the $L$ type from mimicking the $H$ type, but keep $b_1^L$ large such that the expected bonus payment is high enough for the agent to work. This virtually resolves the conflicts between signaling and providing incentives to the agent.

When the period length is small, the inefficiency is negligible, and the $H$ type principal would obtain approximately her FIB payoff. We have the following result:

*Proposition* 2.7. There exists a $\bar{\Delta} > 0$, such that for any $\Delta \in (0, \bar{\Delta})$, there exists an optimal mechanism for the $H$ type, in which the $H$ type obtains a payoff of $V^H(N_*^H) + \frac{1-\beta_0}{\beta_0} V^L(1)$. As $\Delta \to 0$, the $H$ type's payoff converges to her FIB surplus.

The optimal mechanism features a trial period, i.e. the first period. During the trial period, the agent is always recommended to work, regardless of the principal's report. Hence, the agent does not know which project he is working on. The principal's private information about her type will be revealed to the agent only after the trial period, i.e. the agent will be told to stop working for the $L$ project, but to keep working for the $H$ project until the efficient termination time. The existence of the trial period relaxes the agent's incentive constraint for working. In addition, the payment structure during the trial period is designed to elicit the private information of the principal. Both types of principal believe that the chance to succeed during the trial period for the $H$ type is larger than that for the $L$ type. Hence, the mediator can use a mixture of base payments and bonus payments to explore such a differential belief. The contract designed for the $H$ type contains a high base payment but a low bonus payment for the success during the trial period, while the contract designed for the $L$ type contains a low base payment but a high bonus payment. Thus, the two contracts are two bets on the event of the trial period success, which induces the principal to truthfully report her type due to the differential belief about the event. Therefore, the average bonus payment during the trial period – a high bonus from the $L$ type and a low bonus from the $H$ type – provides an incentive for the agent to work, while the structure of the payments provides an incentive for the principal to reveal her type. From the efficiency point of view, this mechanism also induces efficient experimentation for the $H$ type project. However, running the $L$ type project for the trial period is inefficient. But this trial period is short. It lasts only one period, and it is negligible when the period length is small. Hence, the $H$ type principal can obtain a payoff close to her FIB surplus for a small period length, while the $L$ type principal and the agent are left with no rents.[23]

We also introduce random recommendations and study the implementation problem of the FIB in the Appendix. We show that, for any fixed $\Delta > 0$, the FIB can be virtually implemented by recommending that the agent experiment once with an arbitrarily small probability $\mu \in (0, 1)$ for the $L$ type. Thus, introducing a mediator with a menu of contracts helps the $H$ type keep virtually all the surplus of her innovation, while the $L$ type who has an inferior innovation is left with nothing. Hence, this mechanism provides proper incentives for innovators to invest in better ideas ex ante.[24]

[23]The agent would gain some rents from working for the $H$ type, and lose some rents from working for the $L$ type, but break even on average.

[24]This mechanism is also approximately ex ante optimal for the principal.

The reader may ask, can we achieve approximate efficiency without a mediator? Suppose we consider a game where the principal can propose an arbitrary menu of contracts, with the provision that the exact contract that is to be implemented will depend upon a private message sent by the principal. Furthermore, the menu provides a disclosure date, when it will be revealed to the agent which contract has been chosen. Is there a PBE where both types of principal propose the above mechanism? This is an appropriate generalization, to our context, of the question that has been examined in the three-stage mechanism proposing game of Maskin and Tirole (1992). The key problem here is ensuring that deviations, by either type of principal to a different menu of contracts, are unprofitable. In our context, the set of possible deviating contracts is extremely large and complex, and we have been unable to show that every such deviation is unprofitable. More specifically, we need to show that for any other menu, there exists a belief of the agent about the type of the principal that prevents a profitable deviation by either type. Our Lemma 2.3 shows that a pessimistic belief that assigns probability 1 about the principal being the $L$ type can prevent a profitable deviation in the original signaling game, but such a simple belief does not work in the menu proposing game.[25] An explicit construction of beliefs that prevents profitable deviations appears to be intractable. In this context, the paper by Wagner et al. (2015) is relevant. They examine moral hazard in a static setting and show that

[25] There exists a menu that the $H$ type would find profitable offering the agent if he believed that she was the $L$ type.

when the FIB is feasible, it is an equilibrium outcome. However, they do not study the case where the FIB is not feasible. The key feature of the present model is that the FIB is not feasible, even though we can approximate it.[26] Consequently, their results and methods cannot be applied in our context.

## 2.7 Conclusion

We analyze a model where an informed principal engages an agent to explore the viability of her project. The principal has private information about the quality of her project. Thus, in addition to providing incentives to the agent to experiment on the project, the high type principal has to convince the agent of her project's quality. We examine the best outcome for the high type principal in equilibrium.

When there is a large prior probability that the project is of low quality, the best equilibrium for the high type principal is a separating equilibrium. The high type principal separates himself from the low type by leaving rents to the agent, and also by terminating the project inefficiently early. When the prior probability of a low quality project is small, the best equilibrium for the high type is a pooling equilibrium. We find that the best equilibrium, from the point of view of rewarding innovations, is either a separating equilibrium or a pooling equilibrium.

In neither equilibrium of the signaling game is the innovator with a

[26] Recall that the full information benchmark requires immediate termination of the $L$ types project and a zero payoff for this type. This is not feasible.

superior project able to capture her contribution to social surplus. This leads us to study the role of a mediator. The mediator offers a menu of two contracts, and does not disclose the type of the principal to the agent unless it is necessary to do so. We find that an opaque contract is able to approximately implement the optimal outcome, in terms of both inducing efficient experimentation, and ensuring high rewards for the high type principal. It is necessary to induce experimentation by the low type only for a single period, and as the period length becomes small, this is approximately efficient.

# Chapter 3

# Competition in Social Learning

## 3.1 Introduction

Recently, there are a growing number of online platforms who help consumers make their decisions to choose experience products. An experience product is a product with uncertain quality or value for consumers ex ante, thus consumers make purchase decisions based on their beliefs about the quality of the product. The uncertainty of quality is a common feature for digital products, like movies, TV shows, e-books, songs, video games, softwares, apps, etc. Take movies and TV shows as an example, platforms like *Netflix*, *Amazon Prime Instant Video*(hereinafter to be referred as *Amazon*), *Hulu*, *IMDb* and *Rotten Tomatoes* can collect moviegoers' experience or rating about a movie, and give movie recommendation to other potential consumers in their platforms. Apparently, those platforms differ in terms of whether they also provide the digital products that they recommend. *Netflix*, *Amazon* and *Hulu* provide both recommendations and products, while *IMDb* and *Rotten Tomatoes* only provide recommendations.[1] It should be noticed that the model in this paper is more suitable for the former platforms who provide

[1] They provide links to other platforms like *Amazon* who has that product.

both recommendations and products.

As platforms mediate the information about the quality of the product from previous consumers to future potential consumers, there are two different issues to consider in social learning. One is consumers' rating problem, which involves how to incentivize consumers to truthfully report their experience about the product and recover the quality information from the feedback. Another is platforms' recommendation problem, which involves how to incentivize or recommend consumers to experiment the product and learn quality information from the experimentation. While both issues are of interest theoretically, I focus on the latter issue in this paper and hold the former one as simple as possible.

Generally speaking, there are two aspects related to the value of the recommendation to potential consumers. One is the knowledge of a platform about the quality of the product. On one hand, a platform may have her own research group to study the quality of the product, on the other hand, the platform gets feedback about the quality from previous consumers. The stronger the research group is and the more previous consumers she has, the more knowledge about the quality of the product she may attain. Another aspect is how truthfully a platform recommends the product. Even for the platform who cares about the consumers' benefits, she may "spam" them (give them over recommendation to conduct experimentation when her belief about the quality of the product is not good enough to cover the cost), because inducing more consumption could get more feedback about the quality and thus increase the knowledge about the quality, which helps the platform benefit future potential consumers. Therefore, there's a trade-off between "exploration" and "exploitation".

Che and Hörner (2018) explores the optimal design of such recommendation policy when there is only one benevolent platform whose objective is to maximize the discounted expected surplus of consumers. They show that in their baseline model the optimal design of such policy when no conclusive news arrived is a "bang-bang" policy, under which the platform over-recommends the product in the early phase to conduct early experimentation but not too much compared to the first-best policy, because the platform needs to incentivize consumers to follow her recommendation. Therefore, experimentation under this policy occurs faster than under the full transparency policy but slower than under the first-best policy.

However, one monopolistic platform seems unrealistic. As in the movie example I mentioned, there are more than two platforms who compete for potential consumers. Like *Netflix* and *Amazon*, they compete in providing valuable recommendation to consumers, as well as other services. When competition exists, the platforms cannot afford to "spam" their consumers too much, otherwise they would lose their consumers to other platforms. However, without over recommendation in the early phase, their knowledge about the quality of the product would be less than those in the monopolistic scenario since they conduct less experimentation in the early phase.

The current paper seeks to explore the competition effect in social learning, i.e. how the introducing of competition limits the scope of experimentation and the social learning process. Especially, I study how the differentiation of the products affect the social learning process. Currently, there're two very popular political TV series, *House of Cards* produced by *Netflix* and *Alpha House* produced by *Amazon*. Notice that the former is a drama, and the latter is a comedy. Why do the two platforms choose different genres of the TV series and how does this affect the social learning process?

To my knowledge, this is the first paper that studies the competition issue between information intermediaries in social learning. I focus on a duopoly competition, and use the baseline model in Che and Hörner (2018) as my "workhorse", thus I can compare my results to theirs to examine the competition effect. In my current model, I consider "private recommendation", i.e. platforms give personalized recommendation to their consumers through consumers' personal accounts which could be seen only by the account owners themselves. Another form of recommendation is "public recommendation", i.e. platforms give general and uniform recommendation to all consumers which could be seen by all. Both private recommendation and public recommendation are widely used in all platforms I mentioned above. The current version of paper studies only private recommendation, and I will study public recommendation as an extension of the current model in future research. It should be noticed that private recommendation is more basic to study when competition exists, because it isolates the competition effect from the information signaling effect between platforms. When platforms publicly recommend their products, we must consider not only the competition effect, but also how public recommendation serves as a signal of a platform's knowledge about the product's quality to other platforms and consumers. I start with a simpler information environment when platforms use private recommendation before going to a more complex information environment when platforms use public recommendation.

An important benchmark here is the Full Transparency policy, that is the policy under which the platform commits to always recommending truthfully. I show that when platforms do not differentiate their products, the Full Transparency policy for both platforms is the unique equilibrium strategy, and platforms behave myopically and maximize consumers' immediate payoffs given the knowledge they have. When platforms differentiate their products, then the equilibrium strategy is in between the Full Transparency policy and the optimal policy in the monopolistic scenario.[2] The equilibrium outcome thus depends on how differentiated the products are, which is denoted by $\chi$ in my model. When $\chi = 1$, the two products can be seen as in two separated markets, and platforms act as if they were monopolistic. When $\chi = 0$, the two products are homogeneous ex ante, platforms have no room to experiment and have to behave myopically and choose the Full Transparency policy. When $\chi \in (0, 1)$, platforms conduct less experimentation compared to the monopolistic case before the posterior belief about the quality of products hits the same threshold as in the monopolistic case. Therefore, there would be less (at

[2]That is the second-best policy in Che and Hörner (2018).

each time) but longer experimentation in this case.

This paper relates to five strands of literature. The first one is strategic experimentation (Bolton and Harris (1999); Keller et al. (2005); Keller and Rady (2010, 2015)). Actually, the model in Che and Hörner (2018) is built on Keller et al. (2005), since spaming and collecting information from consumers is just experimentation. The second literature is information herding or observational learning (Banerjee (1992); Bikhchandani et al. (1992); Smith and Sørensen (2000)). In observational learning, consumers learn from the actions taken by previous consumers. Here instead, the actions taken by previous consumers cannot be observed by current consumers, they learn only from the platforms' recommendation. The third literature is belief manipulation (Aumann and Maschler (1995); Kamenica and Gentzkow (2011)). The platforms mix their recommendation in the case that they actually know the quality of the product and in the case that they have no conclusive news about the quality of the product together to manipulate Bayesian consumers' beliefs to conduct social desirable experimentation. The fourth literature is learning in R&D competition (Acemoglu et al. (2011); Moscarini and Squintani (2010); Halac et al. (2014)), in which different firms compete directly in their knowledge or the probability of innovation. In this paper, different platforms compete for consumers by giving valuable recommendation, which is determined by both the knowledge of the quality of the products and how truthfully they recommend. The last literature is competition between search engines (Eliaz and Spiegler (2011); Argenton and Prüfer (2012); Taylor (2013); De Corniere

(2013); Hagiu and Jullien (2014)). Search engines compete in providing valuable search results, and that literature usually focuses on how search engines trade off consumer traffic with sponsored advertising. In addition, it usually assumes the quality of search results is known and controlled by search engines. However, though platforms in my model also compete in providing valuable recommendation, they need to learn quality information by experimentation.

## 3.2 Baseline model

### 3.2.1 Model setup

There are two platforms, which are characterized and indexed by their different products $i \in \{a, b\}$ in each of the two platforms. In reality, platforms like *Netflix* and *Amazon* provide many variety of movies and TV shows. Some of the products are provided in both platforms, and some of them are exclusive in only one platform. To provide exclusive content is a common way to attract consumers to platforms. There are two reasons for the exclusivity of products. First, platforms make their own original content, like *House of Cards* produced by *Netflix* and *Alpha House* produced by *Amazon*. Second, platforms have exclusive right to distribute some content (in some regions) that are produced by other producers, like *ABC*, *BBC*, *HBO*, etc.

The two different products in the model could be in the same genre or different genres. In the baseline model, I assume that platforms have different products but they are in the same genre. In the extension, I assume that platforms have different products and they are in different genres. Hence, to choose different genres is a way to differentiate the products.

At each time $t$, there are two different types of consumers $j \in \{c, d\}$, each of whom has measure one. A consumer who is type $j$ is uniformly located at $n \in [0, 1]$. Consumers are short-lived, i.e. they participate the market at time $t$ and then leave the market forever. Thus each time-$t$-lived consumer has only one unit of demand of the products at time $t$, which may result from the time constraint. Long-lived consumers could be studied as a extension of the current model. Consumers are different in their tastes of genres. $c$ could be understood as "comedy", and $d$ could be understood as "drama". Thus type $c$ consumers prefer comedy to drama, while type $d$ consumers prefer drama to comedy. Notice that in the example I mentioned, both *House of Cards* and *Alpha House* are politically-themed TV series. However, the former is a drama, and the latter is a comedy. We may ask what role does product differentiation play in the competition of social learning.

Time is continuous with infinite horizon.

#### 3.2.1.1 Consumers and platforms

At time $t$, the utility of a type $j$ consumer who is located at $n$, and consumes product $i$ is

$$V_i - k_{ij} + \omega_{ijnt},$$

and if the consumer does not consume anything, I normalize the utility to be 0. Therefore, the value to consume some product has three components. First, $V_i \in \{1, 0\}$ is the common value of product $i$ to all consumers, which is irrelevant to tastes and unknown to all at the beginning but the prior of which is common knowledge. This uncertain component is the essential part in the social learning. Thus there are four states of the world concerning the common value of the two products, $\{11, 10, 01, 00\}$, where the first number means the value of product $a$ and the second means that of product $b$. State is persistent and is chosen by nature on the outset of the game according to the probability distribution: $\{\pi_{11}, \pi_{10}, \pi_{01}, \pi_{00}\}$. Let $p_{a0} = \pi_{11} + \pi_{10}$ and $p_{b0} = \pi_{11} + \pi_{01}$ be the prior probabilities of each product's value being 1, the subscript 0 for $p_{i0}$ means the beginning time 0.[3] I'm only interested in the symmetric case, thus I assume that $p_{i0} \equiv p_0 > 0$ for both platforms. Second, $k_{ij} \in \{k_1, k_2\}$ is the matching (negative) value or cost for product $i$ and consumer $j$, which is known and observable to all.[4] Moreover, the value of $k_{ij}$ depends only on the genre of the product and the taste of the consumer. Therefore, the products could be horizontally differentiated. If product $i$ is a comedy, then $0 < k_{ic} = k_1 < k_{id} = k_2 < 1$ because type $c$ consumers prefer comedy to drama,

[3]In the current model, the correlation of quality between the two products does not affect any result. However, when platforms use public recommendation, the correlation could play an important role.

[4]Essentially, this is to assume that the types of consumers are observable to platforms and the consumers themselves. This assumption could be justified by the following reasons. (See `https://help.netflix.com/en/node/9898` for example.) First, as I mentioned above, platforms use private recommendation through consumers' accounts. They ask their consumers to sign up by using their email addresses, *Facebook*, *Google* accounts, etc. Platforms could analyze their social media accounts to acquire tastes information. Second, they usually encourage their consumers to fill out their "taste profiles" and claim that they give their consumers personalized services including personalized recommendation. Third, they can collect consumers' histories of streaming and rating to obtain their preference information. Note that the assumption of observability of consumers' types does not change any results when platforms do not differentiate themselves as in the baseline model, but that assumption matters when platforms do differentiate themselves as in the extension.

while type $d$ consumers prefer drama to comedy. To keep it symmetric,[5] if product $i$ is a drama, then $0 < k_{id} = k_1 < k_{ic} = k_2 < 1$ because type $d$ consumers prefer drama to comedy, while type $c$ consumers prefer comedy to drama. As I pointed out earlier, in the baseline model, I assume that both platforms choose the same genre, say comedy, thus $k_{ic} = k_1 < k_{id} = k_2$ for both $i \in \{a, b\}$. In the extension, I assume that they choose different genres. Last, $\omega_{ijnt}$ is the idiosyncratic value for the consumer $jn$ to consume product $i$ at time $t$. This value is unknown to all ex ante, and has zero mean. Therefore, $\omega_{ijnt}$ does not affect the consumer's decisions. Moreover, $\omega_{ijnt}$ is distributed independently but not identically, and its variance is sufficiently large.[6] Hence, the law of large number can not be applied here. Therefore, even consumers report their utility to platforms truthfully, they cannot recover information of $V_i$.

I assume that $0 < p_0 < k_1 < k_2 < 1$, thus based on the prior, both types of consumers do not want to consume the products ex ante, and there's no consumption without social learning. If some product's value is 1, then both types of consumers can obtain (expected) positive surplus by consuming it. Therefore, social learning is valuable, without which the market would fail.

Consumers do not observe the decisions and experiences by previous

[5]This could be modeled as a linear city model. Consumers are located at the two extreme points in a linear city. Two platforms can choose one of the two extreme points to position their products. The cost of consuming one product for some consumer is linear in their distance.

[6]The more "subjective" the product's quality is, the larger variance it has. This is the case for some digital products, like movies.

consumers, but platforms mediate some information, i.e. private recommendation, through them. Consumers need to choose one platform to enter or sign up, and have an option to consume the product or not based on whether being recommended. Notice that I assume for simplicity that because of time constraint, each consumer can choose to consume at most one unit of product in the platform he enters. In addition, one consumer must enter one and only one platform.[7]

Platforms cannot instantly learn $V_i$ by directly collecting consumers' report on their utilities, how do they learn about the quality information? Motivated by the strategic experimentation literature, Keller et al. (2005) for example, I assume that consumers who consume product $i$ may together receive a signal that indicates the quality information.[8] If there are $\psi_{it}$ consumers who consume product $i$ at time $t$ and $V_i = 1$, then they will get a good signal or

[7]Even though I do not explicitly model the cost and other benefits of signing up a platform, the assumption of entering only one platform could be justified by the explicit and implicit cost and other benefits for signing up a platform. The explicit cost is the monthly fee for membership, which is very similar across platforms in reality. The implicit cost includes the time to manage your accounts, receive advertisement in your emails and keep your personal information safe and secure. Other benefits includes all kinds of other services provided by platforms. Therefore, signing up one platform could be strictly better than no participation. When platforms provide some common products and some different products, consumers may don't want to participate in both platforms because the expected marginal benefit may be smaller than the marginal cost. The reason why the expected marginal benefit is relatively small can be seen from my current model. If I allow consumers to enter additional platforms, consumers could benefit only when the platform who gives them larger expected payoffs does not recommend her product and the platform who gives them smaller expected payoffs does recommend her product, because one consumer has at most one unit of demand. This could happen with relatively small chance in any symmetric equilibria, especially when the quality of the two products are independent or positively correlated.

[8]The signal arrival is independent across any two exclusive groups of consumers.

news with probability $\lambda^g \psi_{it} dt$ in the time interval $[t, t + dt)$. If there are $\psi_{it}$ consumers who consume product $i$ at time $t$ and $V_i = 0$, then they will get a bad signal or news with probability $\lambda^b \psi_{it} dt$ in the time interval $[t, t + dt)$. $\lambda^g$ and $\lambda^b$ are the corresponding arrival rates of news, and $\psi_{it}$ is the intensity of experimentation for product $i$ at time $t$.[9] The signals are independent between products. Given product $i$'s quality, the process of arrival of news for product $i$ is independent with the quality of product $-i$. In addition, given both products' quality, the processes of arriving of news for platforms are independent with each other. The good news indicates that the product's quality is 1 for sure, and the bad news indicates the product's quality is 0 for sure. However, if no news is received by now, the product's quality could be either 0 or 1. Because consumers live only one period and have no cost to report their signals, so I assume they will give truthful report on their signals. I consider the case that $\lambda^g > \lambda^b > 0$ in the current model, thus a good news comes faster than a bad news, and "no news means bad news".

Usually, platforms do not solely learn the quality information from experimenting their consumers. They have their own research groups, who conduct independent experimentation for the quality information. The measure of the research group in each platform is $\rho > 0$. Therefore, if there are $\psi_{it}$ consumers who consume product $i$ at time $t$ and $V_i = 1$, then they together with the research group will get and report a good news with probability

[9]The underlying processes of news arrivals could be modeled as Poisson processes with corresponding arrival rates in different situations

$\lambda^g(\rho + \psi_{it})dt$ in the time interval $[t, t + dt)$. If there are $\psi_{it}$ consumers who consume product $i$ at time $t$ and $V_i = 0$, then they together with the research group will get and report a bad news with probability $\lambda^b(\rho+\psi_{it})dt$ in the time interval $[t, t + dt)$.[10]

Following Che and Hörner (2018), in the baseline model, I consider platforms that commit to maximizing the expected welfare of their own consumers.[11] Platforms need to choose whether to recommend their consumers to consume at each time $t$.

#### 3.2.1.2 Timing, information structure and strategy

At time 0, each platform commits to a private recommendation policy, i.e. the percentage of consumers in her platform she would recommend to consume at each time. Experimentation is valuable in the early phase of product release, thus platforms have incentives to over recommend their consumers to experiment the products if consumers would like to follow their recommendations. Therefore, a platform's recommendation is incredible without com-

[10]Because consumers and the research group conduct independent experimentation, if there are $\alpha_{it}$ consumers who consume product $i$ at time $t$ and $V_i = 1$, then they together with the research group will get and report a good news with probability $\lambda^g\psi_{it}dt + (1 - \lambda^g\psi_{it}dt)\lambda^g\rho dt = \lambda^g(\rho\psi_{it})dt + o(dt) \approx \lambda^g(\rho + \psi_{it})dt$. Similarly, we can get the above formula for the bad news.

[11]This could result from the fact that platforms face different price elasticities for consumers and advertisers. Therefore, when the price elasticity for consumers is sufficiently large, and the price elasticity for advertisers is relatively small, then platforms may want to commit to maximizing the expected welfare of their own consumers but charge fees from advertisers. In addition, as mentioned in Che and Hörner (2018), when platforms provide recommendation for a large collection of products of varying vintages, "recommendation on a number of products with different vintages means that the social welfare gain from optimal experimentation will be spread evenly across users arriving in different times".

mitment power, and consumers would not like to follow her recommendation in the early phase of product release, even it is possible to get a conclusive good news in the early phase. To get rid of this problem, Che and Hörner (2018) assume that the benevolent platform has commitment power to her strategy, which may result from some reputation concerns, thus the platform can benefit by tying herself not to spam or over recommend too much, which makes her recommendation credible. I also assume the two platforms have commitment power to their strategies in my model. However, unlike Che and Hörner (2018), besides consumers, there are two platforms involving strategic interaction. What kind of strategies they can commit themselves to is consequential here. Let's first see what kind of information they know. At each time $t$, the platform $i$ knows (1) the signals she received until time $t$, whose arrival process is defined in the previous sections, (2) the platform $i$'s previous actions before time $t$, and both platforms' committed policies at time 0, and (3) her consumers' action of entering and consuming before time $t$. Generally, a private recommendation policy could depend on all these information, but commitment power makes cooperation or collusion between platforms trivial if I allow them to commit to somewhat general strategies. Technically, commitment power makes sequentially irrational strategy valid. Thus a platform can punish her opponent who deviates from the collusive action by committing to a strategy which is not sequentially rational without commitment power. This seems not realistic when regulation authorities exist. In fact, how much commitment power one platform has is an empirical question. Here I try to keep the commitment power as small as possible[12] by assuming that each platform can commit to a strategy which only depends on her own signals, as well as the types of consumers and time.

Specifically, at the beginning, platform $i$ commits to the following private recommendation policy: $(\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt})$:

- At time $t$, she recommends the product to a fraction $\gamma_{ijt} \in [0, 1]$ of type $j$ consumers if she learns the product to be good (or good news has arrived by time $t$), a fraction $\beta_{ijt} \in [0, 1]$ of type $j$ consumers if she learns it to be bad (or bad news has arrived by time $t$), and she recommends or spams to a fraction $\alpha_{ijt} \in [0, 1]$ of type $j$ consumers if no news has arrived by $t$.

As for consumers, they do not observe the decisions and experiences by previous consumers. They only know their own types, the calendar time $t$ at which they arrive in the system, and the committed policies by the two platforms. They choose and only choose one platform to enter according to their beliefs. If consumers feel indifferent between the two platforms, I assume they will choose to enter either one with 50% chance. The entering action for type $j$ consumers at time $t$ is $Y_{jt} \in \{(1, 0), (0, 1), (1/2, 1/2)\} \subset \Delta\{a, b\}$.[13] Let $\eta_{ijt} = \mathbb{1}(Y_{jt} = i) + 1/2 \cdot \mathbb{1}(Y_{jt} = (1/2, 1/2)$ be the measure of type $j$ consumers

[12]Essentially, I assume that a platform cannot commit to some strategy that explicitly or implicitly depends on her opponent's actions, which is forbidden by regulation authorities.

[13]Sometimes, with a little abuse of notation, I denote $Y_{jt} = (1, 0)$ by $Y_{jt} = a$, and $Y_{jt} = (0, 1)$ by $Y_{jt} = b$. Moreover, I denote $Y_{jt}$ means that type $j$ consumers at time $t$ take the same strategy, this is not restrictive since a time-$t$-lived consumer's expected utility does not depend on any other time-$t$-lived consumer's strategy.

entering platform $i$ at time $t$, and $\eta_{it} = \sum_{j \in \{c,d\}} \eta_{ijt}$ be the measure of both types consumers entering platform $i$ at time $t$. They observe whether the platform recommends or not to consume the product, and choose whether to consume it. If consumers feel indifferent between consuming or not, they will consume. Therefore the consuming strategy for type $j$ consumers at time $t$ when entering platform $i$ is a function $X_{ijt} : \{R, NR\} \to \{C, NC\}$, where $R$ means the platform recommends to consume, $NR$ means the platform does not recommend to consume, $C$ means consumers choose to consume, and $NC$ means consumers choose to not consume.

### 3.2.2 Belief

Though one platform can form beliefs on all four states, however, what she really cares about is just the value of her own product and the consumers' beliefs on the value of both products. Because they don't know their opponents' signals, so they only infer the quality of their opponents' products by the prior correlation between these two products . For consumers, they don't observe any signal, they choose platforms by the knowledge about the states they have[14] and how truthfully they recommend their products $(\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt})$. Therefore, what I need is only the marginal beliefs on the value of each product separately.

Let $\psi_{it} = \sum_{j \in \{c,d\}} \eta_{ijt}[\alpha_{ijt}\mathbb{1}(X_{ijt}(R) = C) + (1 - \alpha_{ijt})\mathbb{1}(X_{ijt}(NR) = C)]$ be the measure of normal consumers who consume the product $i$ at time $t$, and

[14]That is, $g_{it}$, $b_{it}$ and $p_{it}$, which will be defined later.

$\mu_{it} = \rho + \psi_{it}$ be the measure of both research group and normal consumers who experience the product $i$ at time $t$, then conditional on no news for platform $i$, the posterior quality of product $i$, $p_{it}$, follows

$$p_{it} + \mathrm{d}p_{it} = \frac{p_{it}(1 - \lambda^g \mu_{it} \mathrm{d}t)}{p_{it}(1 - \lambda^g \mu_{it} \mathrm{d}t) + (1 - p_{it})(1 - \lambda^b \mu_{it} \mathrm{d}t)}.$$

Rearranging and simplifying, the posterior must follow the law of motion:

$$\dot{p}_{it} = -(\lambda^g - \lambda^b)\mu_{it} p_{it}(1 - p_{it}).$$

Consumers' beliefs about platform $i$ on it receiving a good news, $g_{it}$, and on it receiving a bad news, $b_{it}$, are[15]

$$g_{it} + \mathrm{d}g_{it} = g_{it} + (1 - g_{it} - b_{it}) p_{it} \lambda^g \mu_{it} \mathrm{d}t,$$

$$b_{it} + \mathrm{d}b_{it} = b_{it} + (1 - g_{it} - b_{it})(1 - p_{it}) \lambda^b \mu_{it} \mathrm{d}t,$$

and $g_{i0} = b_{i0} = 0$. Rearranging and simplifying,

$$\dot{g}_{it} = (1 - g_{it} - b_{it}) \lambda^g \mu_{it} p_{it},$$

$$\dot{b}_{it} = (1 - g_{it} - b_{it}) \lambda^b \mu_{it} (1 - p_{it}).$$

They form martingale property,

$$p_0 = g_{it} \cdot 1 + b_{it} \cdot 0 + (1 - g_{it} - b_{it}) p_{it}.$$

Note that only no-news-arrival policy $\alpha_{ijt}$ affects $g_{it}$, $b_{it}$ and $p_{it}$ through $\mu_{it}$. We can interpret no-news-arrival recommendation as experimentation,

[15] $g_{it}$ and $b_{it}$ are also consumers' beliefs that the platform $i$' posterior belief on $V_i$ to be 1 and 0, respectively.

thus how many consumers experience/consume the product when no conclusive news has arrived determines the beliefs evolution, and the platform's knowledge about the state. Moreover, since $\lambda^g - \lambda^b > 0$, i.e. a good news arrives fast than a bad news, the posterior $p_{it}$ strictly decreases in time. Nonetheless, $g_{it}$ and $b_{it}$ would be increasing in time; consumers believe that platforms would have a increasing probability of receiving a conclusive news.

### 3.2.3 Equilibrium

Note that conditional on entering platform $i$ at time $t$, the probabilities of type $j$ consumers being recommended and not being recommended are, respectively,

$$Pr_{ij}(R \mid t) = g_{it}\gamma_{ijt} + b_{it}\beta_{ijt} + (1 - g_{it} - b_{it})\alpha_{ijt},$$

and

$$Pr_{ij}(NR \mid t) = g_{it}(1 - \gamma_{ijt}) + b_{it}(1 - \beta_{ijt}) + (1 - g_{it} - b_{it})(1 - \alpha_{ijt}).$$

In addition, conditional on entering platform $i$ at time $t$ and being recommended, the conditional expected benefit of consuming that product for type $j$ consumers is

$$E(V_i \mid R, j, t) = \frac{g_{it}\gamma_{ijt} + (1 - g_{it} - b_{it})\alpha_{ijt}p_{it}}{Pr_{ij}(R \mid t)}.$$

Similarly, conditional on entering platform $i$ and not being recommended, the conditional expected benefit of consuming that product for type $j$ consumers is

$$E(V_i \mid NR, j, t) = \frac{g_{it}(1-\gamma_{ijt}) + (1-g_{it}-b_{it})(1-\alpha_{ijt})p_{it}}{Pr_{ij}(NR \mid t)}.$$

Thus the expected value of entering platform $i$ at time $t$ for type $j$ consumers is

$$\begin{aligned} m_{ijt} &= Pr_{ij}(R \mid t) \cdot \max\{0, E(V_i \mid R, j, t) - k_{ij}\} + Pr_{ij}(NR \mid t) \cdot \max\{0, E(V_i \mid NR, j, t) - k_{ij}\} \\ &= \max\{0, g_{it}\gamma_{ijt} + (1-g_{it}-b_{it})\alpha_{ijt}p_{it} - k_{ij}(g_{it}\gamma_{ijt} + b_{it}\beta_{ijt} + (1-g_{it}-b_{it})\alpha_{ijt})\} \\ &+ \max\{0, g_{it}(1-\gamma_{ijt}) + (1-g_{it}-b_{it})(1-\alpha_{ijt})p_{it} - k_{ij}(g_{it}(1-\gamma_{ijt}) \\ &+ b_{it}(1-\beta_{ijt}) + (1-g_{it}-b_{it})(1-\alpha_{ijt}))\}. \end{aligned}$$

A Pure Perfect Bayesian Equilibrium with Commitment Power is a strategy profile $\{\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}, Y_{jt}, X_{ijt}(R), X_{ijt}(NR)\}_{i\in\{a,b\}, j\in\{c,d\}, t\geq 0}$, and a belief system $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\}, t\geq 0}$ such that

1. Given $\{\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}, Y_{jt}\}_{i\in\{a,b\}, j\in\{c,d\}, t\geq 0}$, and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\}, t\geq 0}$, if $Y_{jt} = i$ or $Y_{jt} = (1/2, 1/2)$,
   - $X_{ijt}(R) = C$ iff $E(V_i \mid R, j, t) \geq k_{ij}$;
   - $X_{it}(NR) = C$ iff $E(V_i \mid NR, j, t) \geq k_{ij}$.
2. Given $\{\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}, X_{ijt}(R), X_{ijt}(NR)\}_{i\in\{a,b\}, j\in\{c,d\}, t\geq 0}$, and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\}, t\geq 0}$,
   - $Y_{jt} = i$ iff $m_{ijt} > m_{-ijt}$;
   - $Y_{jt} = (1/2, 1/2)$ iff $m_{ijt} = m_{-ijt}$.

3. Given $\{Y_{jt}, X_{ijt}(R), X_{ijt}(NR)\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\},t\geq 0}$, and the opponent's strategy $\{\gamma_{-ijt}, \beta_{-ijt}, \alpha_{-ijt}\}_{j\in\{c,d\},t\geq 0}$, the platform $i$ chooses policy $\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}_{j\in\{c,d\},t\geq 0}$ to maximize $\mathcal{W}_i(\gamma_i, \beta_i, \alpha_i)$ defined by

$$
\begin{aligned}
\int_{t\geq 0} e^{-rt} \sum_{j\in\{c,d\}} \eta_{ijt} \cdot \{ & g_{it}[\gamma_{ijt}\mathbb{1}(X_{ijt}(R)=C) + (1-\gamma_{ijt})\mathbb{1}(X_{ijt}(NR)=C)](1-k_{ij}) \\
& + b_{it}[\beta_{ijt}\mathbb{1}(X_{ijt}(R)=C) + (1-\beta_{ijt})\mathbb{1}(X_{ijt}(NR)=C)](-k_{ij}) \\
& + (1-g_{it}-b_{it})[\alpha_{ijt}\mathbb{1}(X_{ijt}(R)=C) + (1-\alpha_{ijt})\mathbb{1}(X_{ijt}(NR)=C)](p_{it}-k_{ij})\}\, \mathrm{d}t.
\end{aligned}
$$

4. Given $\{\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}, Y_{jt}, X_{ijt}(R), X_{ijt}(NR)\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$,

   - Beliefs evolve by Bayesian updating, which are defined in the last section.

## 3.3 Characterization

### 3.3.1 Meaningful Recommendation

Because "recommendation" could mean either recommendation to consume or recommendation to not consume. W.L.O.G., I assume the former one, or $R$ is weakly more valuable for consumers than $NR$, mathematically,

*Assumption* 3.1. For both platforms, they choose policy $\{\gamma_{ijt}, \beta_{ijt}, \alpha_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$, such that $E(V_i \mid R, j, t) \geq E(V_i \mid NR, j, t)$, i.e.

$$
\frac{g_{it}\gamma_{ijt} + (1-g_{it}-b_{it})\alpha_{ijt}p_{it}}{g_{it}\gamma_{ijt} + b_{it}\beta_{ijt} + (1-g_{it}-b_{it})\alpha_{ijt}} \geq \frac{g_{it}(1-\gamma_{ijt}) + (1-g_{it}-b_{it})(1-\alpha_{ijt})p_{it}}{g_{it}(1-\gamma_{ijt}) + b_{it}(1-\beta_{ijt}) + (1-g_{it}-b_{it})(1-\alpha_{ijt})}.
$$

Consider equilibrium condition 1, one may wonder whether $R$ and $NR$ can make a difference, i.e. is the recommendation trivial in some cases? Or is that possible they are both weakly larger than $k_{ij}$? If so, then it's easy to show that

$g_{it}\gamma_{ijt} + (1 - g_{it} - b_{it})\alpha_{ijt}p_{it} + g_{it}(1 - \gamma_{ijt}) + (1 - g_{it} - b_{it})(1 - \alpha_{ijt})p_{it} \geq$

$k_{ij}[g_{it}\gamma_{ijt} + b_{it}\beta_{ijt} + (1 - g_{it} - b_{it})\alpha_{ijt} + g_{it}(1 - \gamma_{ijt}) + +b_{it}(1 - \beta_{ijt}) + (1 - g_{it} - b_{it})(1 - \alpha_{ijt})]$,

i.e.

$$g_{it} + (1 - g_{it} - b_{it})p_{it} \geq k_{ij},$$

or,

$$p_0 \geq k_{ij},$$

which contradicts with $0 < p_0 < k_1 < k_2 < 1$. The immediate implication is that we must always have

$$\frac{g_{it}(1 - \gamma_{ijt}) + (1 - g_{it} - b_{it})(1 - \alpha_{ijt})p_{it}}{g_{it}(1 - \gamma_{ijt}) + b_{it}(1 - \beta_{ijt}) + (1 - g_{it} - b_{it})(1 - \alpha_{ijt})} < k_{ij},$$

or if entering to platform $i$, we must have $X_{ijt}(NR) = NC$. The following lemma summarizes this:

*Lemma* 3.1. When $p_0 < k_{ij}$, under Assumption 3.1, if type $j$ consumers enter platform $i$ at time $t$, then $X_{ijt}(NR) = NC$.

Because the prior quality is too pessimistic for both types of consumers, when platforms do not recommend the products, no one would like to consume them.

### 3.3.2 Simplification of belief

This section follows Che and Hörner (2018) directly. Let $l_{it} = \frac{p_{it}}{1-p_{it}}$ be the likelihood ratio of posterior beliefs, then $p_{it} = \frac{l_{it}}{1+l_{it}}$, $l_{it}$ and $p_{it}$ are one-to-one mapping, thus the evolution of $p_{it}$ becomes

$$\dot{l}_{it} = -(\lambda^g - \lambda^b)\mu_{it} l_{it} = -\Delta\lambda^g \mu_{it} l_{it},$$

where $\Delta = \frac{\lambda^g - \lambda^b}{\lambda^g} \in (0, 1)$.

Note that beliefs form the martingale

$$p_0 = g_{it} \cdot 1 + b_{it} \cdot 0 + (1 - g_{it} - b_{it})p_{it}.$$

Thus we can simplify the beliefs in the following way:

*Lemma* 3.2.

$$\dot{l}_{it} = -\Delta\lambda^g \mu_{it} l_{it}.$$

In addition,

$$g_{it} = p_0(1 - (\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}),$$
$$b_{it} = (1 - p_0)(1 - (\frac{l_{it}}{l_0})^{\frac{1}{\Delta}-1}).$$

*Proof.* See Appendix C.1. □

So the platform's belief $p_{it}$ or equivalently $l_{it}$ serves as a "sufficient statistic" for the posteriors that the consumers attach to the arrival of news $g_{it}$ and $b_{it}$.

### 3.3.3 Arrival of conclusive news

When platform gets a good news or a bad news, what policy should she choose? Because the objective of the platform is to maximize discounted expected surplus for her own consumers, their interests are aligned.

Assume platform $i$'s equilibrium policy for type $j$ consumers when getting a good news at time $t$ is some $\gamma^*_{ijt} < 1$ . If the policy in the equilibrium is implementable, i.e. some type $j$ consumers enter platform $i$ at time $t$ and they would like to follow the recommendation, then setting $\gamma_{ijt} = 1$ would strictly dominate $\gamma^*_{ijt} < 1$, when holding other strategies unchanged. Because on one hand, setting $\gamma_{ijt}$ larger would relax platform $i$'s constraints, thus it would also be implementable, on the other hand, setting $\gamma_{ijt}$ larger would increase platform $i$'s current benefit and keep the beliefs evolution and future benefit unchanged. If the policy in the equilibrium is not implementable, then $\gamma_{ijt} = 1$ is indifferent to $\gamma^*_{ijt} < 1$.

Similarly, setting $\beta_{ijt} = 0$ is also a weakly dominant strategy for both platforms. To summarize,

*Lemma* 3.3. $\gamma_{ijt} = 1, \beta_{ijt} = 0$ are weakly dominant strategy for both platforms.

Now, we can check whether we impose any restriction in Assumption 3.1. It turns out there is no extra restriction at all. See Appendix C.2 for details.

### 3.3.4 Competitive advantage

Given Lemma 3.1, 3.3, we can simplify $m_{ijt}$ to

$$m_{ijt} = \max\{0, g_{it}(1 - k_{ij}) + (1 - g_{it} - b_{it})\alpha_{ijt}(p_{it} - k_{ij})\}.$$

Let $z_{ijt} = g_{it}(1 - k_{ij}) + (1 - g_{it} - b_{it})\alpha_{ijt}(p_{it} - k_{ij})$, then $m_{ijt} = \max\{0, z_{ijt}\}$. Here I interpret $z_{ijt}$ as the competitive advantage of platform $i$ for type $j$

consumers at time $t$, because equilibrium condition 1 can be summarized as $X_{ijt}(R) = C$ iff $z_{ijt} \geq 0$, given Lemma 3.1, 3.3. In addition, given $X_{ijt}(R) = C$ for both platforms, what these two platforms compete for type $j$ consumers at time $t$ is just $z_{ijt}$. Moreover, $z_{ijt}$ is the platform $i$'s instant payoff for each type $j$ consumer who enters her platform and follows her recommendation at time $t$.

Furthermore, given the martingale of beliefs and the evolution of $g_{it}$ from Lemma 3.2, I can show that

$$z_{ijt} = v_{ij}[1 - (\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}(1 - \alpha_{ijt}(1 - \frac{\kappa_{ij}}{l_{it}}))],$$

where $\kappa_{ij} = \frac{k_{ij}}{1-k_{ij}}$ is the normalized cost and $v_i = \frac{p_0}{1+\kappa_{ij}} = p_0(1 - k_{ij})$ is the initial condition.

To further analyze competition, I define the set of platform $i$'s implementable policy for type $j$ consumers at time $t$ with respect to the strategy profile $\{\{\alpha_{ij\tau}\}_{0\leq\tau<t}, \{\alpha_{-ij\tau}\}_{0\leq\tau\leq t}\}_{j\in\{c,d\}}$ as

$$\mathcal{F}_{ijt} = \{\alpha_{ijt} \in [0,1] : z_{ijt} \geq 0 \text{ and } z_{ijt} \geq z_{-ijt}\}.$$

First, let's take a look at how $z_{ijt}$ depends on $\alpha_{ijt}$ and $l_{it}$. For platform $i$, $\alpha_{ijt}$ is the instant policy for type $j$ consumers at time $t$, and $l_{it}$ is a summary of the knowledge about the quality of the product $i$ until time $t$. These two items, as well as $p_0$ and $k_{ij}$, determine the competitive advantage of platform $i$ for type $j$ consumers at time $t$.

Since $\Delta \in (0,1)$, we have

$$\frac{\partial z_{ijt}}{\partial \alpha_{ijt}} = v_{ij}(\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}(1 - \frac{\kappa_{ij}}{l_{it}}) \propto l_{it} - \kappa_{ij} \propto p_{it} - k_{ij},$$

and

$$\frac{\partial z_{ijt}}{\partial l_{it}} = -\frac{v_{ij}}{\Delta}(\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}\frac{1}{l_{it}^2}[l_{it}(1-\alpha_{ijt}) + \alpha_{ijt}\kappa_{ij}(1-\Delta)] \propto -\Delta.$$

Because $\Delta > 0$, the posterior $p_{it}$ strictly decreases in time, thus $p_{it} \leq p_0 < k_{ij}$, which gives following lemma,

*Lemma* 3.4. At time $t$, platform $i$' competitive advantage for type $j$ consumers $z_{ijt}$ is strictly decreasing in the instant policy $\alpha_{ijt}$. In addition, $z_{ijt}$ is strictly decreasing in $l_{it}$.

Therefore, the smaller $l_{it}$ (compared to $l_0$) it is, the more knowledge about the quality of the product $i$ the platform obtains by time $t$. One question should be pursued. If the implementable policy exists, is it better than the unimplementable policies? Let the Full Transparency policy be $\alpha_{ijt} = \mathbb{1}(p_{it} \geq k_{ij})$.[16]

From Lemma 3.4, given $l_{it}$, as well as $p_0$ and $k_{ij}$, the Full Transparency policy at time $t$ gives platform $i$ the largest competitive advantage $z_{ijt}$ for type $j$ consumers. Therefore, if platform $i$ has implementable policy for type

[16]When $\Delta > 0$ and $p_0 < k_{ij}$, the Full Transparency policy is just $\alpha_{ijt} = 0$. I write as $\alpha_{ijt} = \mathbb{1}(p_{it} \geq k_{ij})$, because the following lemma also holds for the Full Transparency policy in a more general situation (when $\Delta \leq 0$).

$j$ consumers at time $t$, then the Full Transparency policy can also be implemented. In addition, Full Transparency policy is strictly better than any unimplementable policy. The reason is the following.

Note first that the objective function of platform $i$ simplifies to

$$\mathcal{W}_i(\alpha_i) \equiv \int_{t \geq 0} e^{-rt} \sum_{j \in \{c,d\}} z_{ijt} \eta_{ijt} \mathbb{1}(X_{ijt}(R) = C) \, \mathrm{d}t.$$

Since $\Delta > 0$, $p_{it} \leq p_0 < k_{ij}$, the Full Transparency policy prescribes $\alpha_{ijt} = 0$ for both types of consumers at any time $t$. Any $\alpha_{ijt} \notin \mathcal{F}_{ijt}$ gives 0 payoff at time $t$ to platform $i$ since $\eta_{ijt}\mathbb{1}(X_{ijt}(R) = C) = 0$. However, when $\alpha_{ijt} = 0 \in \mathcal{F}_{ijt}$, $\eta_{ijt}\mathbb{1}(X_{ijt}(R) = C) \in \{1/2, 1\}$. It's easy to see that when $t > 0$, given $\alpha_{ijt} = 0$, $z_{ijt} > 0$. Thus when $t > 0$, the Full Transparency gives strict positive payoff at time $t$ to platform $i$. Moreover, neither unimplementable policy nor $\alpha_{ijt} = 0$ could affect the beliefs evolution and thus future constraints.

Therefore, we have the following lemma:

*Lemma* 3.5. If platform $i$ has implementable policy for type $j$ consumers at time $t$, i.e. $\mathcal{F}_{ijt} \neq \varnothing$, then $\alpha_{ijt} = \mathbb{1}(p_{it} \geq k_{ij}) \in \mathcal{F}_{ijt}$. In addition $\alpha_{ijt} = \mathbb{1}(p_{it} \geq k_{ij})$ is strictly better than any unimplementable policy at time $t$ for type $j$ consumers in platform $i$.[17]

[17]It should be noted that since we have continuous time setting, the lemma holds "almost everywhere". I omit "almost everywhere" thereafter.

### 3.3.5 Equilibrium

Therefore, given the above lemmas, we can simplify the equilibrium in the following way.

A Pure Perfect Bayesian Equilibrium with Commitment Power is a strategy profile $\{\{\gamma_{ijt} = 1, \beta_{ijt} = 0, \alpha_{ijt}\}, Y_{jt}, X_{ijt}(R), X_{ijt}(NR) = NC\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$, and a belief system $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\},t\geq 0}$ such that

1. Given $\{\{\gamma_{ijt} = 1, \beta_{ijt} = 0, \alpha_{ijt}\}, Y_{jt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$, and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\},t\geq 0}$, if $Y_{jt} = i$ or $Y_{jt} = (1/2, 1/2)$,

    - $X_{ijt}(R) = C$ iff $z_{ijt} \geq 0$.

2. Given $\{\{\gamma_{ijt} = 1, \beta_{ijt} = 0, \alpha_{ijt}\}, X_{ijt}(R), X_{ijt}(NR) = NC\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$, and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\},t\geq 0}$,

    - $Y_{jt} = i$ iff $\max\{0, z_{ijt}\} > \max\{0, z_{-ijt}\}$;
    - $Y_{jt} = (1/2, 1/2)$ iff $\max\{0, z_{ijt}\} = \max\{0, z_{-ijt}\}$.

3. Given $\{Y_{jt}, X_{ijt}(R), X_{ijt}(NR) = NC\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ and $\{p_{it}, g_{it}, b_{it}\}_{i\in\{a,b\},t\geq 0}$, and the opponent's strategy $\{\gamma_{-ijt} = 1, \beta_{-ijt} = 0, \alpha_{-ijt}\}_{j\in\{c,d\},t\geq 0}$,

    - platform $i$ chooses policy $\{\gamma_{ijt} = 1, \beta_{ijt} = 0, \alpha_{ijt}\}_{j\in\{c,d\},t\geq 0}$ to maximize

$$\mathcal{W}_i(\alpha_i) \equiv \int_{t\geq 0} e^{-rt} \sum_{j\in\{c,d\}} z_{ijt}\eta_{ijt}\mathbb{1}(X_{ijt}(R) = C)\,\mathrm{d}t.$$

4. Given $\{\{\gamma_{ijt} = 1, \beta_{ijt} = 0, \alpha_{ijt}\}, Y_{jt}, X_{ijt}(R), X_{ijt}(NR) = NC\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$,

- Beliefs evolution are defined in Lemma 3.2.
- $g_{i0} = b_{i0} = 0$, and $\mu_{it} = \rho + \sum_{j \in \{c,d\}} \alpha_{ijt} \eta_{ijt} \mathbb{1}(X_{ijt}(R) = C)$.

Generally speaking, given some strategy of an opponent, it's not easy to find a platform's best response because it's very hard to solve an optimal control problem in a general setting. However, when platforms do not differentiate their products, i.e. they choose the same genre of products, the situation is relatively simple. We have the following proposition.

*Proposition* 3.1. When platforms do not differentiate their products, there exists a unique equilibrium, where both platforms prescribe the Full Transparency policy for both types of consumers.

*Proof.* See Appendix C.3. □

As we can see, when platforms do not differentiate their products, severe competition makes the social learning process very different from that of the monopolistic case in Che and Hörner (2018). No platform recommends to any type of consumers to consume the product when the posterior belief about the quality of the product is less than the cost for that type of consumers. In this sense, platforms truthfully or transparently give recommendations, even though they prefer to spam some consumers to conduct experimentation in the early phase of the product release for the sake of all consumers in their platforms. They do not spam to any type of consumers, i.e. they do not recommend when the posterior about the quality of the product is less than the cost of that type of consumers, because otherwise their opponents could spam slightly less and grab all that type of consumers' shares. By doing so, the one who spams less could benefit at that time since she still cares about that type of consumers at the instant time, and in addition gains the competitive advantage in the future, because even the percentage of that type of consumers who are experimented is slightly less than before, she has twice of that type of consumers entering her platform now, which means the total consumers who are experimented would be more than before. Thus, given her opponent's policy, to spam less than her opponent is good for both current and future payoffs. Therefore, platforms behave myopically and maximize the short-lived consumers' immediate payoffs, and the only possible equilibrium strategy is the Full Transparency policy. Note that the relationship between $k_1$ and $k_2$ does not matter here, and heterogeneous tastes of consumers have no effect on the equilibrium, because platforms do not treat different types of consumers differently when they do not differentiate their products. Thus, when $k_1 = k_2$, we have the same result.

## 3.4 Products differentiation

In this section, I assume platforms differentiate their products. In the motivating example I mentioned, as two popular concurrent political TV series, *House of Cards* produced by *Netflix* and *Alpha House* produced by *Amazon* are in different genres. The former is a drama, and the latter is a comedy. The horizontal differentiation of products is well understood to avoid severe competition in many different economic environment. Here I provide a model to show how horizontal differentiation affect the process of social learning.

Without loss of generality, I assume platform $a$ chooses a comedy, and platform $b$ chooses a drama. Therefore, type $c$ consumers are the target consumers for platform $a$, and type $d$ consumers are the target consumers for platform $b$. Mathematically, $p_0 < k_{ac} = k_{bd} = k_1 < k_{ad} = k_{bd} = k_2 < 1$. In addition, I assume type $c$ consumers would like to choose platform $a$ if they feel the two platforms are indifferent, and type $d$ consumers would like to choose platform $b$ if they feel the two platforms are indifferent. Moreover, the lemmas in the above sections hold anyway here.[18]

Because of the symmetry of the model, I consider a candidate symmetric equilibrium. Note that when two platforms choose the same policy, then their posterior beliefs $p_{it}$ or $l_{it}$ are the same at any $t$. In addition, when $l_{it}$ are the same, there always exists an implementable policy for one platform's target type of consumers at that time, because the competitive advantage $z_{ijt}$ is strictly decreasing in the cost $k_{ij}$. Therefore, the two platforms would get their corresponding target consumers in any symmetric equilibrium according to Lemma 3.5. Now consider what kind of policy for the non-target types of consumers the two platforms would like to prescribe in this candidate symmetric equilibrium? Note that when $l_{it}$ are the same, there exists no implementable

[18]Note that the definition of implementable policy changes a little bit, because when some types of consumers are indifferent between two platforms, they would like to choose the platform who targets to them.

policy for the non-target types of consumers. If one platform, say platform $a$, chooses a policy that is different from the Full Transparency policy for the non-target types of consumers at some time $t$, i.e. $\alpha_{adt} > 0$, then this would give her opponent $b$ a larger set of implementable policy for $b$'s target type of consumers $d$ at that time, since $z_{adt}$ is strictly decreasing in $\alpha_{adt}$. Moreover, if platform $b$ would like to spam more to type $d$ consumers at that time, which is implementable, this in turn would give $b$ a smaller posterior $l_{bt+\mathrm{dt}}$ than $l_{at+\mathrm{dt}}$ in the future, which would give platform $a$ a smaller set of implementable policy for type $c$ consumers, since $z_{bct+\mathrm{dt}}$ is strictly decreasing in $l_{bt+\mathrm{dt}}$. Therefore, if both platforms just attract their own target consumers at any time, it's weakly better to set the policy for the non-target consumers to be the Full Transparency policy ($\alpha_{ijt} = 0$ for $ij \in \{ad, bc\}$).

Now we consider a single platform's decision problem, say platform $a$, when she only attracts her own target type of consumers $c$,[19] i.e. $\eta_{act} = 1$ and $\eta_{adt} = 0$ for any $t$, and her opponent $b$ chooses the Full Transparency policy for type $c$ consumers, i.e. $\alpha_{bct} = 0$ for any $t$, and has $l_{bt} = l_{at}$ for any $t$. I first solve the optimal policy for platform $a$ in this situation. The problem can be summarized as the following.[20]

[19]Thus, platform $a$'s policy for type $d$ consumers, i,e. $\alpha_{adt}$, is irrelevant.

[20]When platform $b$ chooses $\alpha_{bct} = 0$, and $l_{bt} = l_{at}$, $z_{bct} = v_{bc}[1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]$ is always positive. Therefore, $z_{act} \geq z_{bct}$ implies $z_{act} \geq 0$, thus $\mathbb{1}(X_{act}(R) = C)$ for any $t$.

$$\sup_{\alpha_{ac}} \mathcal{W}_a(\alpha_{ac}) = \int_{t\geq 0} e^{-rt} z_{act}\,\mathrm{d}t$$
$$\text{s.t.}\quad z_{act} \geq v_{bc}[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}], \tag{Problem A}$$
$$\dot{l}_{at} = -\Delta\lambda^g(\rho+\alpha_{act})l_{at}, \text{ and } \alpha_{act} \in [0,1].$$

Let's consider the constraint $z_{act} \geq z_{bct} = v_{bc}[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]$ first. When platform $b$ chooses $\alpha_{bct} = 0$ at some time $t$, it would give the smallest set of implementable policy for $a$'s target type of consumers $c$ at that time. The constraint can be rewritten as

$$(1-k_1)[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}(1-\alpha_{act}(1-\frac{\kappa_1}{l_{at}}))] \geq (1-k_2)[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}],$$

where I denote $\kappa_n = \frac{k_n}{1-k_n}$ for $n \in \{1,2\}$. Simplify this, we can get

$$\alpha_{act} \leq \hat{\alpha}(l_{at}) \equiv \frac{k_2-k_1}{1-k_1}\frac{(\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}}-1}{\kappa_1 - l_{at}}l_{at}.$$

The right hand side of the inequality has two parts. The first part, call it $\chi \equiv \frac{k_2-k_1}{1-k_1}$, is constant over time, and the second part, $\frac{(\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}}-1}{\kappa_1-l_{at}}l_{at}$, is evolving over time. I can show that the second part is strictly increasing over time from 0 at time 0, and approaches infinity when time goes to infinity. Details can be found in Appendix C.4. Moreover, $z_{act} \geq 0$ iff $\alpha_{act} \leq \frac{(\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}}-1}{\kappa_1-l_{at}}l_{at}$, therefore the second part is just the constraint without considering platform $a$'s opponent. Note that $0 < p_0 < k_1 < k_2 < 1$, thus $\chi = \frac{k_2-k_1}{1-k_1} \in (0,1)$. When the difference between the cost or (negative) matching value of the target type of consumers $k_1$ and that of non-target type of consumers $k_2$ shrinks to 0, then the first part also shrinks to 0. When the cost of the non-target type of consumers $k_2$ increases to 1, then the first part also increases to 1. This is to say, when consumers have homogeneous tastes, then there is no room for spaming or experimentation; when consumers' heterogeneous tastes are too different, then two platforms act as if the markets are totally separated, because the only constraint is $z_{act} \geq 0$ and they don't need to consider their opponents' competitive advantages. Therefore, I interpret $\chi = \frac{k_2-k_1}{1-k_1}$ as a measure of how differentiated the two products are.

Note that $\alpha_{act} \in [0, 1]$, define

$$\bar{\alpha}(l_{at}) \equiv \min\{1, \hat{\alpha}(l_{at})\},$$

thus the constraint can be summarized as $\alpha_{act} \in [0, \bar{\alpha}(l_{at})]$, and platform $a$ can spam strictly more as time passes before $\hat{\alpha}(l_{at})$ reaches 1. Therefore, Problem A can be summarized as

$$\begin{aligned} \sup_{\alpha_{ac}} \mathcal{W}_a(\alpha_{ac}) &= \int_{t\geq 0} e^{-rt} z_{act} \,\mathrm{d}t \\ \text{s.t.} \quad \dot{l}_{at} &= -\Delta\lambda^g(\rho + \alpha_{act}) l_{at} \\ \alpha_{act} &\in [0, \bar{\alpha}(l_{at})]. \end{aligned} \tag{Problem A'}$$

This is very similar to the monopolistic platform's problem in Che and Hörner (2018) except $\bar{\alpha}(l_{at})$ is defined in a different way. The following lemma summarizes the optimal solution for this problem.

*Lemma* 3.6. The solution for Problem A exists and is unique almost everywhere. In terms of the posterior likelihood ratio, when $l_{at} < l^*$, then $\alpha^*_{act} = 0$; when $l_{at} > l^*$, then $\alpha^*_{act} = \bar{\alpha}(l_{at})$, where

$$l^* = \frac{r + \lambda^g \rho}{r + \lambda^g(\rho + 1)} \kappa_1.$$

Or in terms of time,

1. When $l_0 \leq l^*$, then $\alpha^*_{act} = 0$ for any $t \geq 0$.

2. When $l_0 > l^*$ then $\alpha^*_{act} = \bar{\alpha}(l_{at})$ for $t < t^*$, and $\alpha^*_{act} = 0$ for $t > t^*$, where $t^*$ and $l_{at}$ are determined by

$$\begin{cases} t^* = \inf\{t : l_{at} \leq l^*\}, \\ \dot{l}_{at} = -\Delta\lambda^g(\rho + \bar{\alpha}(l_{at}))l_{at}. \end{cases}$$

*Proof.* See Appendix C.5. □

Therefore, in Problem A, platform $a$ would like to experiment as much as she can when the posterior belief is higher than some threshold, and to stop the experimentation when the posterior belief hits the threshold. This "bang-bang" policy is very similar to the monopolistic platform's optimal policy in Che and Hörner (2018), except for the maximum experimentation that can be implementable.[21] However, note that in Problem A, I assume that platform $b$ has $l_{bt} = l_{at}$, which is true only when platform $b$ uses the exact same strategy.

Hence, I get a candidate symmetric equilibrium strategy profile $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$:

- For the non-target type of consumers, both platforms prescribes the Full Transparency policy, i.e. $\alpha^*_{adt} = \alpha^*_{bct} = 0$ for any $t$;

- For the target type of consumers, both platforms prescribe the optimal policy for Problem A, i.e.

[21] In Che and Hörner (2018), $\chi = 1$.

1. When $l_0 \leq l^*$, then $\alpha^*_{act} = \alpha^*_{bdt} = 0$ for any $t$.

2. When $l_0 > l^*$ then $\alpha^*_{act} = \alpha^*_{bdt} = \bar{\alpha}(l_t)$ for $t < t^*$, and $\alpha^*_{act} = \alpha^*_{bdt} = 0$ for $t > t^*$, where $t^*$ and $l_t$ are determined by

$$\begin{cases} t^* = \inf\{t : l_t \leq l^*\}, \\ \dot{l}_t = -\Delta\lambda^g(\rho + \bar{\alpha}(l_t))l_t. \end{cases}$$

Now, I need to show under what conditions, this candidate symmetric equilibrium is truly an equilibrium. From Problem A, we know that when both platforms get their target types of consumers, they would like to prescribe the candidate equilibrium strategy if they have no incentives to earn the extra non-target types of consumers. To earn the extra non-target types of consumers means to spam or experiment more than their opponents to obtain more quality information, this is not feasible when $t < t^*$, because both platforms have already experimented as much as possible. Thus, if one platform wants to earn the extra non-target types of consumers, she must also experiment when $t > t^*$. The trade-off is to sacrifice the target type of consumers to conduct experimentation at the current time, and to attract the non-target type of consumers in the future.

Therefore, I consider the following problem. Suppose $l_0 > l^*$, platform $b$ chooses the strategy $\alpha_{bct} = \alpha_{bdt} = 0$ for $t \geq t^*$. In addition, both platforms' posterior likelihood ratio are $l^*$ when $t = t^*$. I consider platform $a$'s problem from $t^* > 0$, that is

$$
\begin{aligned}
\sup_{\alpha_{ac},\alpha_{ad}} \int_{t\geq t^*} & e^{-r(t-t^*)}\{z_{act} + z_{adt}\eta_{adt}\}\, \mathrm{d}t \\
\text{s.t.} \quad & z_{act} \geq v_{bc}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}], \\
\eta_{adt} &= \mathbb{1}(z_{adt} \geq v_{bd}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]), \\
& l_{at^*} = l_{bt^*} = l^*, \\
& \dot{l}_{bt} = -\Delta\lambda^g \rho l_{bt}, \\
\dot{l}_{at} &= -\Delta\lambda^g(\rho + \alpha_{act} + \alpha_{adt}\eta_{adt})l_{at}, \\
& \alpha_{act}, \alpha_{adt} \in [0,1].
\end{aligned}
\tag{Problem B}
$$

Notice that when $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ is an equilibrium of the original game, we must have that $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution forProblem B. In addition, when Problem B has $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ as an optimal solution, then $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ is an equilibrium of the original game. This is because if one platform would not like to experiment in Problem B, she would not like to experiment when her opponent has an information advantage, either. Thus she would suffer if she did not experiment as much as she could before time $t^*$. Therefore we have a necessary and sufficient condition for $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ to be an equilibrium.

*Lemma* 3.7. When $l_0 > l^*$, $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ constitute an equilibrium iff $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B.

*Proof.* See Appendix C.6. □

*Remark* 3.1. When $l_0 \leq l^*$, then clearly $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ constitute an equilibrium iff $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, in which we replace $t^*$ with 0 and $l^*$ with $l_0$.

Generally, I'm unable to fully solve for Problem B.[22] Furthermore, when $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, it does not mean $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ is the unique equilibrium, especially there may exist some worse equilibrium in terms of platforms' payoffs. Even though platforms have no incentives to attract the non-target types of consumers after $t^*$, platforms may want to spam or experiment their target types of consumers to learn more quality information and to keep or not lose their target types of consumers, if their opponents still experiment after $t^*$. Therefore, it may be possible that both platforms still experimenting after $t^*$ constitutes an equilibrium in some situations.

Nonetheless, to attract the non-target type of consumers is not easy. When $t \to 0$ or $t \to \infty$, there is always an implementable policy for the target type of consumers, which means it's infeasible to attract the non-target type of consumers in these extreme cases, because any advantage of quality information $l_{it}$ is dominated by the advantage of the matching values $k_{ij}$.

In addition, it's obvious to see that when platform $a$ can never attract any type $d$ consumers after time $t^*$, i.e. when it's infeasible to attract any non-target type of consumers for one platform after time $t^*$, then $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B according to Lemma 3.6, thus

[22] It's a non-standard optimal control problem.

$\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ constitute an equilibrium.

Furthermore, when it's infeasible to attract any non-target type of consumers after time $t^*$, this equilibrium gives both platforms the worst equilibrium payoffs. First, the equilibrium strategy for the target type of consumers is always implementable before time $t^*$, which is irrelevant to her opponent's strategy. Second, when it's also impossible to attract any non-target type of consumers after time $t^*$, then the equilibrium strategy for the target type of consumers after time $t^*$ is also always implementable. Therefore, both platforms can secure this payoff by prescribing $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$, and other equilibria, if any, cannot give worse payoffs for both platforms.

Formally, consider the following dynamic system starting from $t^*$

$$\begin{aligned}
l_{at^*} &= l_{bt^*} = l^*, \\
\dot{l}_{bt} &= -\Delta\lambda^g \rho l_{bt}, \\
\dot{l}_{at} &= -\Delta\lambda^g(\rho + \min\{\alpha_{act}, 1\})l_{at}, \\
v_{ac}[1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}(1 - \alpha_{act}(1 - \frac{\kappa_1}{l_{at}}))] &= v_{bc}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}].
\end{aligned} \tag{System A}$$

Thus, the infeasibility of attracting any non-target type of consumers means there exists no $t \geq t^*$ such that $v_{ad}[1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}] \geq v_{bd}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]$ in System A. We have the following lemma.

*Lemma* 3.8. When $l_0 > l^*$, if there exists no $t \geq t^*$ such that $v_{ad}[1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}] \geq v_{bd}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]$ in System A, then $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ constitute an equilibrium, which gives both platforms the worst equilibrium payoffs.

*Remark* 3.2. When $l_0 \leq l^*$, we need to replace $t^*$ with 0 and $l^*$ with $l_0$ in System A, then Lemma 3.8 still holds.

Notice that whether it's feasible to attract any non-target type of consumers after time $t^*$ depends on whether the social learning process relies too much on the experimentation of the target type of consumers, or equivalently how large the research group $\rho$ is, and how differentiated the two products are ($\chi$). If the research group is too weak, i.e. $\rho \approx 0$, then when one platform stops experimenting her target type of consumers, it almost means that she stops learning any quality information. Thus her opponent is relatively easy to earn some information advantage by continuing the experimentation, and if the information advantage is sufficiently large compared to the cost disadvantage ($\chi$), her opponent can attract extra consumers. On the other hand, if the research group is relatively strong, i.e. $\rho$ is sufficiently large, then one platform cannot earn enough information advantage by continuing the experimentation to overcome the cost disadvantage ($\chi$), and thus could never attract any non-target type of consumers.

Now, I provide a sufficient condition, under which it's infeasible to attract the non-target type of consumers.

*Lemma* 3.9. When $l_0 > l^*$, if $\chi > \frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}[\frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}]^{\rho}$, then there exists no $t \geq t^*$ such that $v_{ad}[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}] \geq v_{bd}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]$ in System A.

*Proof.* See Appendix C.7. □

*Remark* 3.3. When $l_0 \leq l^*$, we need to replace $t^*$ with 0 and $l^*$ with $l_0$ in System A, then Lemma 3.9 still holds, and the condition becomes to $\chi > \frac{1}{1+\rho}$.

Clearly, $\frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}[\frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}]^\rho$ is increasing in $(\frac{l^*}{l_0})^{\frac{1}{\Delta}}$. We can interpret $(\frac{l^*}{l_0})^{\frac{1}{\Delta}}$ as the quality information at time $t^*$. The lower $(\frac{l^*}{l_0})^{\frac{1}{\Delta}}$, the more quality information one platform has, and the harder her opponent can attract the non-target type of consumers thereafter. Also note that $\frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}[\frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}]^\rho$ is bounded below $\frac{1}{1+\rho}$, thus we immediately have the following proposition.

*Proposition* 3.2. For any $\chi \in (0, 1)$, there exists some $\underline{\rho} > 0$, such that for any $\rho > \underline{\rho}$, $\{\alpha^*_{ijt}\}_{i\in\{a,b\},j\in\{c,d\},t\geq 0}$ constitute an equilibrium, and this equilibrium gives both platforms the worst equilibrium payoffs.

First, note that the condition is not too strict in reality, since research groups $\rho$ are usually large. This may result from the fact that their experimentation costs are relatively lower than those of consumers since they specialize in research and experimentation.

Second, note that $\chi = \frac{k_2-k_1}{1-k_1}$ is a measure of how differentiated the two products are, and it serves two roles. On one hand, the more differentiated the two products (the larger $\chi$), the harder she can lose her target type of consumers after time $t^*$. On the other hand, the more differentiated the two products (the larger $\chi$), the more she can experiment before time $t^*$. When $\chi = 1$, this equilibrium can be seen as a generalization of Che and Hörner (2018). Products are so different that they can be seen as if they were in two separated markets, thus platforms act as if they were monopolistic. In addition, when $\chi = 0$, this equilibrium prescribes the Full Transparency policy for both platform, and we go back to my baseline model. Products are homogeneous ex ante, thus platforms have no room to experiment and have to behave myopically and maximize consumers' immediate payoffs by prescribing the Full Transparency policy.

Third, when $\chi \in (0, 1)$, even though the threshold of posterior belief, at which platforms stop experimentation, is the same as in the monopolistic case, platforms conduct less (proportional to $\chi$) experimentation compared to the monopolistic case. Thus it would take a longer time until the posterior belief hits the threshold and platforms stop the experimentation. In addition, when $\chi$ increases, i.e. the products become more differentiated, platforms conduct more experimentation and it takes a shorter time until platforms stop the experimentation.

Clearly, when platforms differentiate their products, they are strictly better-off in this worst-payoff equilibrium for two reasons. First, they attract their target type of consumers respectively, which is more efficient than splitting half of the target type and half of the non-target type of consumers. Second, they can conduct experimentation in the early phase of product release, which improves the overall welfare in one platform (but sacrifices the early consumers). Now, consider a positioning game prior to my original game: two platforms choose what kind of genre (comedy or drama) of products they would like to produce or have exclusive right to distribute, and the costs of any genre of products are the same. Obviously, they will choose to differentiate their products in a pure strategy equilibrium.

## 3.5 Conclusion

Internet makes information more accessible to people, and recommendation platforms play an increasingly important role nowadays. Che and Hörner (2018) study how to optimally design a recommendation policy for a monopolistic platform. However, monopoly is not a realistic assumption. Motivated by the competition between *Netflix* and *Amazon*, this paper examines a duopoly competition between two platforms who provide both products and recommendations.

In my baseline model, I consider the case when platforms do not differentiate their products. I show that there exists a unique equilibrium, where both platforms prescribe the Full Transparency policy. They do not experiment or spam their consumers any more, because otherwise their opponents would spam slightly less and grab all the market shares. Therefore, competition, even with one more platform, would destroy any experimentation, which is worthy for the overall welfare in one platform.

I also extend my model to the case when platforms differentiate their products. Product differentiation in this model means platforms choose different genres of products. For example, as two popular political TV series, *House of Cards* produced by *Netflix* and *Alpha House* produced by *Amazon* are in different genres. The former is a drama, and the latter is a comedy. I show that for any degree of differentiation $\chi \in (0, 1)$, when the research group is large enough compared to $\chi$, there always exists an equilibrium, where platforms conduct experimentation as much as they can before the posterior beliefs hit some threshold, and thereafter they stop experimentation. The threshold is the same as in the monopolistic case, however, the experimentation before the posterior beliefs hit the threshold is limited by the degree of differentiation $\chi$. Therefore, there would be less experimentation and it would take a longer time until platforms stop the experimentation compared to the monopolistic case. When $\chi = 1$, the two products can be seen as in two separated markets, and platforms act as if they were monopolistic. Thus Che and Hörner (2018) can be seen as a special case of my model.

# Appendices

# Appendix A

# Appendix for Chapter 1

## A.1 Proofs

*Proof of Proposition 1.1.* (1) First, we show that when $c > \underline{c}$, the type $B$ agent never censors bad news under any history $\bar{\varnothing}_t^B$ in any PBE.

In any PBE, a piece of bad news results in an immediate dismissal of the agent. Fixed any equilibrium and the equilibrium strategies $\{\hat{\boldsymbol{r}}, \hat{\boldsymbol{x}}^B\}$, let $V_B^{\hat{\boldsymbol{r}}}(\bar{\varnothing}_t^B) \geq 0$ be the equilibrium payoff of the type $B$ agent under the history $\bar{\varnothing}_t^B$. Suppose $\hat{x}_t^B > 0$ for a history $\bar{\varnothing}_t^B$, then under alternative strategy of the evaluator $\boldsymbol{r}_\infty$, the type $B$ agent obtains $v_B^{\boldsymbol{r}_\infty, \hat{\boldsymbol{x}}^B}(\bar{\varnothing}_t^B)$ by censoring that piece of arrived bad news. On the other hand, under the evaluator's strategy $\boldsymbol{r}_\infty$, if the agent censors that piece of arrived bad news and resumes the optimal strategy (i.e. never censors again), she can obtain

$$-\rho_0 c + \mathbb{E}[\int_0^{\tau^B} \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu + e^{-\rho_0 \tau^B} V_B^{\boldsymbol{r}_\infty}(\bar{\varnothing}_{\tau^B}^B)] = \frac{\rho_0}{\rho_0 + \beta^B}[w - (\rho_0 + \beta)c],$$

where $\tau^B$ is the arrival time of the next piece of bad news from the type $B$ agent. Thus, $v_B^{\boldsymbol{r}_\infty, \hat{\boldsymbol{x}}^B}(\bar{\varnothing}_t^B) \leq \frac{\rho_0}{\rho_0 + \beta^B}[w - (\rho_0 + \beta)c] < 0$, since $c > \underline{c}$.

In addition, $V_B^{\hat{\boldsymbol{r}}}(\bar{\varnothing}_t^B) \leq v_B^{\boldsymbol{r}_\infty, \hat{\boldsymbol{x}}^B}(\bar{\varnothing}_t^B) < 0$ since any strategy of the evaluator is weakly worse for the agent than $\boldsymbol{r}_\infty$ as the latter subscribes that the evaluator never dismisses the agent without bad news. This contradicts $V_B^{\hat{\boldsymbol{r}}}(\bar{\varnothing}_t^B) \geq 0$. Thus, $\hat{x}_t^B = 0$ for any history $\bar{\varnothing}_t^B$.

(2) Second, since the type $B$ agent never censors bad news, the evaluator faces a standard exponential bandit model. The literature on bandit models provides the optimal strategy of the evaluator. For completeness, we sketch the proof here.

Let $\mathbf{0}$ represent the strategy of the type $B$ agent that never censors bad news, i.e. $x_t^B = 0$ for any $t$. Then the evaluator chooses $\boldsymbol{r}_t$ to maximize

$$\begin{aligned} u^{\boldsymbol{r},\mathbf{0}}(\varnothing_t) &= \mathbb{E}\left[\int_t^T \rho_1 e^{-\rho_1(\nu-t)}\, \mathbb{1}_{\theta=G}\, k\, \mathrm{d}N_\nu^G + e^{-\rho_1(T-t)}m \,\middle|\, \varnothing_t\right] \\ &= \mathbb{E}\left[\int_t^T \rho_1 e^{-\rho_1(\nu-t)} p_\nu h\, \mathrm{d}\nu + e^{-\rho_1(T-t)}m \,\middle|\, \varnothing_t\right], \end{aligned}$$

where the second equation comes from the Law of Iterated Expectations.

Since the payoff of the evaluator depends on history only through the public belief $p_t$, we rewrite the above $u^{\boldsymbol{r},\mathbf{0}}(p_t) := u^{\boldsymbol{r},\mathbf{0}}(\varnothing_t)$, and

$$U^{\mathbf{0}}(p_t) = \sup_{\boldsymbol{r}_t} u^{\boldsymbol{r},\mathbf{0}}(p_t).$$

When $\gamma > \beta^B$, the following ODE solves $U^{\mathbf{0}}(p)$ when it is optimal to experiment,

$$\rho_1 U^{\mathbf{0}}(p) = \rho_1 ph + p\gamma(h - U^{\mathbf{0}}(p)) + (1-p)\beta^B(m - U^{\mathbf{0}}(p)) - p(1-p)(\gamma - \beta^B)U^{\mathbf{0}\prime}(p).$$

When the evaluator is indifferent between experimenting and dismissal at a public belief $\hat{p}$, then $U^{\mathbf{0}}(\hat{p}) = m$ (value matching) and $U^{\mathbf{0}\prime}(\hat{p}) = 0$ (smooth pasting) since the continuation region and stopping region communicate. Hence,

$$\rho_1 m = \rho_1 \hat{p} h + \hat{p}\gamma(h - m),$$

i.e. $\hat{p} = p_{fast} = \frac{\rho_1 m}{\rho_1 h + \gamma(h-m)}$. In addition, we can show that when $p > p_{fast}$,

$$U^{\mathbf{0}}(p) = m + p(h-m)(1 - e^{-(\rho_1+\gamma)\bar{s}}) - (1-p)m\frac{\rho_1}{\rho_1 + \beta^B}(1 - e^{-(\rho_1+\beta^B)\bar{s}}).$$

Here, $\bar{s} = \frac{\ln[\frac{p}{1-p}\frac{1-p_{fast}}{p_{fast}}]}{\gamma - \beta^B}$ is the duration when the evaluator experiments without a conclusive signal.

When $\gamma < \beta^B$, the same ODE solves $U^{\mathbf{0}}(p)$ when it is optimal to experiment. However, the smooth pasting condition does not hold, since the continuation region and stopping region do not communicate. A more direct approach is to compute the optimal experimenting time without a conclusive signal. If the experimenting time is $s$, then the expected payoff of the evaluator is

$$m + p(h-m)(1 - e^{-(\rho_1+\gamma)s}) - (1-p)m\frac{\rho_1}{\rho_1 + \beta^B}(1 - e^{-(\rho_1+\beta^B)s}).$$

It is easy to see (by taking first order derivative with respect to $s$) that it is either optimal to experiment forever or to stop immediately. Therefore, value matching gives $\hat{p} = p_{slow} = \frac{\rho_1 m}{\rho_1 h + \beta^B(h-m)}$ which solves

$$m + \hat{p}(h-m) - (1-\hat{p})m\frac{\rho_1}{\rho_1 + \beta^B} = m.$$

In addition, when $p > p_{slow}$,

$$U^{\mathbf{0}}(p) = m + p(h-m) - (1-p)m\frac{\rho_1}{\rho_1 + \beta^B}.$$

□

*Proof of Proposition 1.2.* (1) First, for any admissible strategy of the type $B$ agent, the evaluator's best response, if it exists, is a pure cutoff strategy with the cutoff $p_{fast}$.

Let $\mathbf{0}$ and $\mathbf{1}$ represent the strategy of the type $B$ agent that never censors bad news ($x_t^B = 0$ for any $t$) and always censors bad news ($x_t^B = 1$ for any $t$). Standard analysis in bandit literature shows that in both cases the optimal response of the evaluator is to use a cutoff strategy with the cutoff $p_{fast} = \frac{\rho_1 m}{\rho_1 h + \gamma(h-m)}$. We have seen the case without censorship. To see the case with full censorship, let us write $U^{\mathbf{1}}(p_0)$ as the evaluator's value function. When it is optimal to experiment, it solves the following ODE,

$$\rho_1 U^{\mathbf{1}}(p) = \rho_1 ph + p\gamma(h - U^{\mathbf{1}}(p)) - p(1-p)\gamma U^{\mathbf{1}\prime}(p).$$

Value matching and smooth pasting give $\hat{p} = p_{fast}$ which solves

$$\rho_1 m = \rho_1 ph + p\gamma(h-m).$$

In addition, when $p > p_{fast}$

$$U^{\mathbf{1}}(p) = m + p(h-m)(1 - e^{-(\rho_1+\gamma)\hat{s}}) - (1-p)m(1 - e^{-\rho_1 \hat{s}}).$$

Here, $\hat{s} = \frac{\ln[\frac{p}{1-p}\frac{1-p_{fast}}{p_{fast}}]}{\gamma}$ is the duration when the evaluator experiments without a conclusive signal.

There two cases represent the maximal and minimal information that the evaluator can get. The minimal information is a garbling of the information that generated from any other admissible strategy of the type $B$ agent, which in turn is a garbling of the maximal information.

Consider the case where the type $B$ agent uses a strategy $\boldsymbol{x}^B$, and the public conjecture is $\tilde{\boldsymbol{x}}^B$. Let $\boldsymbol{r}$ be the best response of the evaluator with respect to $\tilde{\boldsymbol{x}}^B$. In addition, let $\boldsymbol{r}_0$ and $\boldsymbol{r}_1$ be the best responses of the evaluator with respect to $\mathbf{0}$ and $\mathbf{1}$, respectively. Then, we have

$$U^{\mathbf{0}}(p) = u^{\boldsymbol{r}_0,\mathbf{0}}(p) \geq u^{\boldsymbol{r},\mathbf{0}}(p) \geq U^{\tilde{\boldsymbol{x}}^B}(p) = u^{\boldsymbol{r},\tilde{\boldsymbol{x}}^B}(p) \geq u^{\boldsymbol{r}_1,\tilde{\boldsymbol{x}}^B}(p) \geq u^{\boldsymbol{r}_1,\mathbf{1}}(p) = U^{\mathbf{1}}(p).$$

This is because by using the same strategy, the evaluator has weakly higher value under the maximal information than under any other information induced by the strategy of the type $B$ agent, and he has weakly lower value under the minimal information than under any other information.

In addition, when $p > p_{fast}$, the best response is to keep experimenting since it is also the best response under the minimal information. Similarly, the best response is to dismiss the agent when $p < p_{fast}$.

(2) Second, given the evaluator uses a cutoff strategy with the cutoff $p_{fast}$ in any PBE, the type $B$ agent would be eventually dismissed. Let $s_1 = \sup\{t \geq 0 : \bar{h}_t^B \in \bar{\varnothing}_t^B, x_t^B > 0\}$ in a PBE. Clearly, without any revealed signal until $s_1$, we must have $p_{s_1} > p_{fast}$ in any PBE, otherwise censoring a piece of bad news only incurs a cost of $c$, but gives 0 continuation value to the agent. Hence, when a piece of bad news arrives at time $t > s_1$, $x_t^B = 0$.

In addition, when bad news has not been revealed to the evaluator yet at time $t < s_1$, we must have $x_t^B = 1$. Suppose not, i.e. there exists a $t < s_1$ and $x_t^B < 1$. Since $s_1 = \sup\{t \geq 0 : \bar{h}_t^B \in \bar{\varnothing}_t^B, x_t^B > 0\}$, there must a $t' \in (t, s_1]$ such that $x_t' > 0$. Obviously, the agent's equilibrium payoff cannot be negative, $V_B^{\boldsymbol{r}}(\bar{\varnothing}_{t'}^B) \geq 0$. Equivalently,

$$\mathbb{E}\left[\int_{t'}^{T} \rho_0 e^{-\rho_0(\nu - t')}(w\,\mathrm{d}\nu - cX_\nu^B\,\mathrm{d}N_\nu^B)\middle|\, \bar{\varnothing}_{t'}^B\right] \geq \rho_0 c,$$

where $T$ is induced by the equilibrium strategy of the evaluator $\boldsymbol{r}$. Consider an alternative strategy $\boldsymbol{x}^{B'}$ such that censor all bad news for time $\nu \in [t, t']$ and resume the original strategy, then

$$\begin{aligned} v_B^{\boldsymbol{r},\boldsymbol{x}^{B'}}(\bar{\varnothing}_t^B) &= -\rho_0 c + \int_t^{t'} \rho_0 e^{-\rho_0(\nu - t)}(w \,\mathrm{d}\nu - \beta^B c \,\mathrm{d}\nu) \\ &\quad + e^{-\rho_0(t'-t)}\, \mathbb{E}\left[ \int_{t'}^{T} \rho_0 e^{-\rho_0(\nu - t')}(w \,\mathrm{d}\nu - c X_\nu^B \,\mathrm{d}N_\nu^B) \middle| \bar{\varnothing}_{t'}^B \right] \\ &\geq -\rho_0 c + (w - \beta^B c)(1 - e^{-\rho_0(t'-t)}) + e^{-\rho_0(t'-t)} \rho_0 c \\ &= [w - (\rho_0 + \beta^B) c](1 - e^{-\rho_0(t'-t)}) > 0. \end{aligned}$$

Hence, it is strictly optimal to censor that piece of bad news, which contradicts with $x_t^B < 1$.

At time $s_1$, note that

$$V_B^{\boldsymbol{r}}(\bar{\varnothing}_{s_1}^B) = \max\left\{ \mathbb{E}[-\rho_0 c + \int_0^{s_2 \wedge \tau^B} \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu], 0 \right\},$$

where $s_2 = \frac{\ln[\frac{p_{s_1}}{1-p_{s_1}} \frac{1-p_{fast}}{p_{fast}}]}{\gamma - \beta^B}$ is duration when the public belief drifts from $p_{s_1}$ to $p_{fast}$ in the absence of news, and $\tau^B$ is the arrival time of the next piece of bad news. Clearly, the first term is continuous and increasing in $s_2$. Since for any $s < s_2$, $\mathbb{E}[-\rho_0 c + \int_0^{s \wedge \tau^B} \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu] \leq 0$, we must have

$$\mathbb{E}[-\rho_0 c + \int_0^{s_2 \wedge \tau^B} \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu] \leq 0.$$

Suppose that the above expression is strictly less than 0; let is be $-\epsilon < 0$. Then consider the value at time $t < s_1$. Since it is optimal to censor any news before $s_1$ and censors no news after $s_1$, thus

$$\begin{aligned} V_B^{\boldsymbol{r}}(\bar{\varnothing}_t^B) &= -\rho_0 c + \int_0^{s_1-t} \rho_0 e^{-\rho_0\nu}(w-\beta^B c)\,\mathrm{d}\nu + e^{-\rho_0(s_1-t)}\int_0^{s_2\wedge\tau^B} \rho_0 e^{-\rho_0\nu} w\,\mathrm{d}\nu \\ &= -\rho_0 c + (w-\beta^B c)(1-e^{-\rho_0(s_1-t)}) + e^{-\rho_0(s_1-t)}(\rho_0 c - \epsilon) \\ &= [w-(\rho_0+\beta^B)c] - e^{-\rho_0(s_1-t)}[w-(\rho_0+\beta^B)c+\epsilon]. \end{aligned}$$

Clearly, when $t$ is sufficiently close to $s_1$, the above expression is less than 0, which is a contradiction. Hence, we must have

$$\mathbb{E}[-\rho_0 c + \int_0^{s_2\wedge\tau^B} \rho_0 e^{-\rho_0\nu} w\,\mathrm{d}\nu] = 0.$$

This pins down

$$s_2 = \frac{\ln\frac{w}{w-(\rho_0+\beta^B)c}}{\rho_0+\beta^B},$$

which in turn determines

$$p^B = p_{s_1} = \frac{p_{fast}}{p_{fast}+(1-p_{fast})e^{-(\gamma-\beta^B)s_2}}.$$

Hence, if $p_0 > p^B$, the type $B$ ruler would censor all bad news for $s_1$ time period,

$$s_1 = \frac{\ln[\frac{p_0}{1-p_0}\frac{1-p^B}{p^B}]}{\gamma} = (1-\frac{\beta^B}{\gamma})(\bar{s}-s_2).$$

If $p_0 \le p^B$, the type $B$ ruler would not censor any bad news.

Hence, the type $B$ ruler's best response is also a cutoff strategy with the cutoff $p^B$. It is easy to verify that the best response of the evaluator to such a cutoff strategy of the type $B$ ruler exists, and it is precisely the cutoff strategy with the cutoff $p_{fast}$.

Since the strategy of the type $B$ ruler is also unique in any PBE, this is the unique PBE. $\square$

*Proof of Proposition 1.3.* When $p_0 > p^B$, we can compute the expected payoff of the type $B$ ruler in the equilibrium,

$$(w-\beta^B c)(1-e^{-\rho_0 s_1})+e^{-\rho_0 s_1}\frac{\rho_0 w}{\rho_0+\beta^B}(1-e^{-(\rho_0+\beta^B)s_2}) = (w-\beta^B c)(1-e^{-\rho_0 s_1})+e^{-\rho_0 s_1}\rho_0 c.$$

If the censorship is not possible, then her expected payoff would be

$$\frac{\rho_0 w}{\rho_0+\beta^B}(1-e^{-(\rho_0+\beta^B)\bar{s}}).$$

Hence, she would be better off in the equilibrium with censorship if and only if

$$(w-\beta^B c)(1-e^{-\rho_0 s_1})+e^{-\rho_0 s_1}\rho_0 c > \frac{\rho_0 w}{\rho_0+\beta^B}(1-e^{-(\rho_0+\beta^B)\bar{s}}).$$

Since we have

$$w-\beta^B c > \frac{\rho_0 w}{\rho_0+\beta^B} > \frac{\rho_0 w}{\rho_0+\beta^B}(1-e^{-(\rho_0+\beta^B)\bar{s}}) > \rho_0 c,$$

thus she would be better off in the equilibrium with censorship if and only if

$$e^{-\rho_0 s_1} < \frac{\beta^B+\rho_0 e^{-\frac{\gamma(\rho_0+\beta^B)}{\gamma-\beta^B}s_1}}{\rho_0+\beta^B}.$$

Let

$$f(s) = e^{-\rho_0 s} - \frac{\beta^B+\rho_0 e^{-\frac{\gamma(\rho_0+\beta^B)}{\gamma-\beta^B}s}}{\rho_0+\beta^B}.$$

Note that $f(s)$ is first increasing in $s$, then decreasing in $s$. Moreover, $f(0)=0$, and $\lim_{s\to\infty} f(s) = -\frac{\beta^B}{\rho_0+\beta^B} < 0$. Hence, there exists a $s_1^* = s_1^*(\rho_0,\gamma,\beta^B) > 0$, such that the type $B$ ruler would be worse off in the equilibrium if and only if $s_1 < s_1^*$. $\square$

*Proof of Corollary 1.1.* Let $s^* := \frac{\gamma}{\gamma-\beta^B} s_1^*$. Since $s_1 = \frac{\gamma-\beta^B}{\gamma}(\bar{s}-s_2) > 0$, $s_1 < s_1^*$ is equivalent to $\bar{s} < s_2 + s^*$.

Note that

$$p^B = \frac{p_{fast}}{p_{fast} + (1-p_{fast})e^{-(\gamma-\beta^B)s_2}}$$

is increasing in $s_2$. In addition, $s_2$ is increasing in $c$, and $\lim_{c\to 0} p^B = p_{fast}$ and $\lim_{c\to\bar{c}} p^B = 1$. Hence, $p_0 = p^B$ defines $\mathfrak{c}_3 \in (0, \underline{c})$, such that $p_0 \leq p^B$ when $c \in [\mathfrak{c}_3, \underline{c})$. Thus, when $c \in [\mathfrak{c}_3, \underline{c})$, the equilibrium payoff and the payoff in the NCB would be the same for the type $B$ agent.

The type $B$ agent would be worse off in the equilibrium if and only if $s_2 > \bar{s} - s^*$. If $\bar{s} \leq s^*$, then $s_2 > \bar{s} - s^*$ holds for all $c \in (0, \mathfrak{c}_3]$, and the type $B$ agent would be worse off in the equilibrium. Suppose $\bar{s} > s^*$. Since $s_2$ is increasing in $c$, and $\lim_{c\to 0} s_2 = 0$ and $\lim_{c\to\mathfrak{c}_3} s_2 = \bar{s}$, $s_2 = \bar{s} - s^*$ defines $\mathfrak{c}_1 \in (0, \mathfrak{c}_3)$ such that $s_2 > \bar{s} - s^*$ when $c \in (\mathfrak{c}_1, \mathfrak{c}_3]$, and $s_2 < \bar{s} - s^*$ when $c \in (0, \mathfrak{c}_1)$. We complete the proof by redefining $\mathfrak{c}_1$ as $\mathfrak{c}_1 \mathbb{1}_{\bar{s}>s^*}$. □

*Proof of Corollary 1.2.* From Corollary 1.1, when $c \in [\mathfrak{c}_3, \bar{c})$, the equilibrium payoff and the NCB payoff would be the same for the type $B$ ruler, and they do not depend on the cost $c$.

Suppose $c < \mathfrak{c}_3$, then $p_0 > p^B$. The expected payoff of the type $B$ ruler in the equilibrium is

$$(w - \beta^B c)(1 - e^{-\rho_0 s_1}) + e^{-\rho_0 s_1}\rho_0 c.$$

Note that $s_1 = \frac{\gamma-\beta^B}{\gamma}(\bar{s}-s_2)$ and $\frac{\partial s_2}{\partial c} = \frac{1}{w-(\rho_0+\beta^B)c}$. Thus,

$$\frac{\partial (w-\beta^B c)(1-e^{-\rho_0 s_1}) + e^{-\rho_0 s_1}\rho_0 c}{\partial c} = -\beta^B + \beta^B e^{-\rho_0 s_1}\frac{\rho_0+\gamma}{\gamma}.$$

Hence, the type $B$ ruler's equilibrium payoff is increasing in $c$ if and only if

$$e^{-\rho_0 s_1} > \frac{\gamma}{\rho_0+\gamma}.$$

Let $s_1^{**} = s_1^{**}(\rho_0,\gamma) > 0$ be the solution of $e^{-\rho_0 s} = \frac{\gamma}{\rho_0+\gamma}$. Hence, the type $B$ ruler's equilibrium payoff is increasing (resp. decreasing) in $c$ if and only if $s_1 < s_1^{**}$ (resp. $s_1 > s_1^{**}$).

Let $s^{**} = \frac{\gamma}{\gamma-\beta^B}s_1^{**} > 0$. Hence, $s_1 < s_1^{**}$ is equivalent to $s_2 > \bar{s} - s^{**}$. Note that $s_2$ is increasing in $c$, and $\lim_{c\to 0} s_2 = 0$ and $\lim_{c\to \mathfrak{c}_3} s_2 = \bar{s}$.

If $\bar{s} \leq s^{**}$, then $s_2 > \bar{s} - s^{**}$ holds for all $c \in (0, \mathfrak{c}_3]$, and the type $B$ ruler's equilibrium payoff is increasing in $c \in (0,\mathfrak{c}_3]$. Suppose $\bar{s} > s^{**}$, then $s_2 = \bar{s} - s^{**}$ defines $\mathfrak{c}_2 \in (0,\mathfrak{c}_3)$ such that $s_2 > \bar{s} - s^{**}$ when $c \in (\mathfrak{c}_2, \mathfrak{c}_3]$, and $s_2 < \bar{s} - s^{**}$ when $c \in (0,\mathfrak{c}_2)$. Last, we redefine $\mathfrak{c}_2$ as $\mathfrak{c}_2\,\mathbb{1}_{\bar{s}>s^{**}}$.

Finally, we show that $s^{**} \in (0, s^*)$, or equivalently $s_1^{**} \in (0, s_1^*)$. Let $\nu_1 = s_1^{**}$ be the solution of

$$e^{-\rho_0 \nu_1} = \frac{\gamma}{\rho_0+\gamma},$$

and $\nu_2$ be the solution of

$$\frac{\beta^B + \rho_0 e^{-\frac{\gamma(\rho_0+\beta^B)}{\gamma-\beta^B}\nu_2}}{\rho_0+\beta^B} = \frac{\gamma}{\rho_0+\gamma}.$$

Hence, $s_1^{**} < s_1^*$ is equivalent to $\nu_2 < \nu_1$. Note that $\nu_1 = \frac{\ln[\frac{\rho_0+\gamma}{\gamma}]}{\rho_0}$, and $\nu_2 = \ln[\frac{\rho_0+\gamma}{\gamma-\beta^B}]\frac{\gamma-\beta^B}{\gamma(\rho_0+\beta^B)}$. In addition,

$$\frac{\partial \nu_2}{\partial \beta^B} = \frac{1}{\gamma}\frac{1}{\rho_0+\beta^B}(1-\ln[\frac{\rho_0+\gamma}{\gamma-\beta^B}]\frac{\rho_0+\gamma}{\rho_0+\beta^B})$$

Let $\eta = \frac{\gamma-\beta^B}{\rho_0+\gamma} \in (0,1)$, then $\frac{\partial \nu_2}{\partial \beta^B} < 0$ if and only if

$$1-\eta+\ln\eta < 0,$$

which is always true. Hence, $\nu_2$ is decreasing in $\beta^B \in (0,\gamma)$. Also,

$$\lim_{\beta^B \to 0} \nu_2 = \nu_1.$$

Hence, for $\beta^B \in (0,\gamma)$, $\nu_2 < \nu_1$. Thus, $s_1^{**} < s_1^*$ and $s^{**} < s^*$. In addition, by the definitions of $\mathfrak{c}_1$ and $\mathfrak{c}_2$, $\mathfrak{c}_1 = \mathfrak{c}_2 = 0$ when $\bar{s} \le s^{**}$, $\mathfrak{c}_1 = 0 < \mathfrak{c}_2$ when $\bar{s} < s^{**}$ when $\bar{s} \in (s^{**}, s^*]$, and $0 < \mathfrak{c}_1 < \mathfrak{c}_2$ when $\bar{s} > s^*$.

□

*Proof of Proposition 1.4.* (1) First, let us verify that the pair of mixed cutoff strategies constitutes an equilibrium. Given the type $B$ agent's strategy, call it $\boldsymbol{x}_M$,

$$x_t^B = \begin{cases} 1 & \text{if } p_t > p_{fast}, \\ 1-\frac{\gamma}{\beta^B} & \text{if } p_t = p_{fast}, \\ 0 & \text{if } p_t < p_{fast}, \end{cases}$$

we now solve the evaluator's best response. Consider three cases: $p_0 = p_{fast}$, $p_0 > p_{fast}$ and $p_0 < p_{fast}$.

When $p_0 = p_{fast}$, the public belief does not change in the absence of news. The evaluator faces a stationary problem, and his value function solves

$$U^{\boldsymbol{x}_M}(p_{fast}) = \max\left\{p_{fast}h + \frac{\gamma}{\rho_1}[p_{fast}h + (1-p_{fast})m - U^M(p_{fast})], m\right\}.$$

It is easy to check that

$$U^{\boldsymbol{x}_M}(p_{fast}) = \max\left\{p_{fast}h + (1-p_{fast})\frac{\gamma}{\rho_1+\gamma}m, m\right\}.$$

Note that the first term is exactly $m$, thus the evaluator is indifferent between experimenting or not.

When $p_0 > p_{fast}$, the public belief drifts down until $p_{fast}$ in the absence of news. The strategy of the evaluator is just how long to experiment in the absence of any signal; we use $\boldsymbol{r}_s$ to denote this strategy, $T_s$ to denote the stopping time, and $s \in [0, \hat{s}]$ to be the length of experimentation. Then his payoff would be

$$u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0) = \mathbb{E}\left[\int_0^{T_s} \rho_1 e^{-\rho_1\nu} p_\nu h \,\mathrm{d}\nu + e^{-\rho_1 T_s} m\right].$$

We can show that

$$u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0) = p_0[\int_0^s \gamma e^{-\gamma\nu}\rho_1 e^{-\rho_1\nu}(k + \frac{h}{\rho_1})\,\mathrm{d}\nu + e^{-\rho_1 s}e^{-\gamma s}m] + (1-p_0)e^{-\rho_1 s}m.$$

Hence,

$$\begin{aligned}
\frac{\partial u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0)}{\partial s} &= p_0 e^{-(\rho_1+\gamma)s}(\rho_1+\gamma)(h-m) - (1-p_0)e^{-\rho_1 s}\rho_1 m \\
&= (1-p_0)e^{-\rho_1 s}[\frac{p_0}{1-p_0}e^{-\gamma s}(\rho_1+\gamma)(h-m) - \rho_1 m] \\
&= (1-p_0)e^{-\rho_1 s}[\frac{p_s}{1-p_s}(\rho_1+\gamma)(h-m) - \rho_1 m].
\end{aligned}$$

Since $p_s \in [p_{fast}, p_0]$, and $\frac{p_{fast}}{1-p_{fast}} = \frac{\rho_1 m}{(\rho_1+\gamma)(h-m)}$, we have $\frac{\partial u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0)}{\partial s} > 0$ for $s \in [0, \hat{s})$. Hence, it is optimal for the evaluator to experiment $\hat{s}$ time until the public belief drifts down to $p_{fast}$.

When $p_0 < p_{fast}$, the public belief drifts up until $p_{fast}$ in the absence of news. Let it be $s \in [0, \bar{s}]$ be the length that the evaluator experiments without a conclusive signal. Then his payoff would be

$$
\begin{aligned}
u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0) = p_0[&\int_0^s \gamma e^{-\gamma\nu}\rho_1 e^{-\rho_1\nu}(k+\frac{h}{\rho_1})\,\mathrm{d}\nu + e^{-\rho_1 s}e^{-\gamma s}m] \\
&+ (1-p_0)[\int_0^s \beta^B e^{-\beta^B\nu}e^{-\rho_1\nu}\,\mathrm{d}\nu + e^{-\rho_1 s}e^{-\beta^B s}]m.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{\partial u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0)}{\partial s} &= p_0 e^{-(\rho_1+\gamma)s}(\rho_1+\gamma)(h-m) - (1-p_0)e^{-(\rho_1+\beta^B)s}\rho_1 m \\
&= (1-p_0)e^{-(\rho_1+\beta^B)s}[\frac{p_0}{1-p_0}e^{-(\gamma-\beta^B)s}(\rho_1+\gamma)(h-m) - \rho_1 m] \\
&= (1-p_0)e^{-(\rho_1+\beta^B)s}[\frac{p_s}{1-p_s}(\rho_1+\gamma)(h-m) - \rho_1 m].
\end{aligned}
$$

Since $p_s \in [p_0, p_{fast}]$, and $\frac{p_{fast}}{1-p_{fast}} = \frac{\rho_1 m}{(\rho_1+\gamma)(h-m)}$, we have $\frac{\partial u^{\boldsymbol{r}_s,\boldsymbol{x}_M}(p_0)}{\partial s} < 0$ for $s \in [0, \hat{s})$. Hence, it is optimal for the evaluator to dismiss the agent immediately when $p_0 < p_{fast}$.

Hence, the best response of the evaluator to the strategy $\boldsymbol{x}_M$ is a cutoff strategy with a cutoff at $p_{fast}$.

Now, let us solve the best response of the type $B$ agent given the evaluator's mixed cutoff strategy with a constant hazard rate $z^* = \frac{w}{c} - \rho_0 - \beta^B$ at the cutoff $p_{fast}$.

Let this strategy be $\boldsymbol{r}_c$. Clearly, since the evaluator would overthrow the agent when $p_t < p_{fast}$, thus the agent would never censor any bad news when $p_t < p_{fast}$. When $p_t = p_{fast}$, the evaluator faces a stationary problem since the evaluator would dismiss the agent at a constant rate. Hence,

$$V_B^{\boldsymbol{r}_c}(\bar{\varnothing}_t^B) = \max\left\{-\rho_0 c + \mathbb{E}[\int_0^{\lambda\wedge\tau^B} \rho_0 e^{-\rho_0\nu} w \,\mathrm{d}\nu + e^{-\rho_0\tau^B} V_B^{\lambda}(\varnothing_{\tau^B}^b)\, \mathbb{1}_{\{\tau^B<\lambda\}}], 0\right\},$$

where $\lambda$ is the arrival time of dismissal induced by $z^* = \frac{w}{c} - \rho_0 - \beta^B$.

Since $\mathbb{E}[\int_0^{\lambda\wedge\tau^B} \rho_0 e^{-\rho_0\nu} w \,\mathrm{d}\nu] = \rho_0 c$, clearly the type $B$ agent would be indifferent between censoring or nor. This also means that the continuation value of the type $B$ agent is $\rho_0 c$ when no bad news arrives at $p_t = p_{fast}$.

When $p_t > p_{fast}$, the evaluator believes that the evaluator would censor all bad news until the public belief reaches $p_{fast}$. Hence, the type $B$ agent can stay in power for some positive time $s = \frac{\ln[\frac{p_t}{1-p_t}\frac{1-p_{fast}}{p_{fast}}]}{\gamma} > 0$. Censoring all bad news until then gives her a payoff

$$\begin{aligned} v_B^{\boldsymbol{r}_c,\boldsymbol{x}_M}(\bar{\varnothing}_t^B) &= -\rho_0 c + \int_0^s \rho_0 e^{-\rho_0\nu}(w - \beta^B c)\,\mathrm{d}\nu + e^{-\rho_0 s}\rho_0 c \\ &= [w - (\rho_0 + \beta^B)c](1 - e^{-\rho_0 s}) > 0. \end{aligned}$$

Hence it is optimal to censor all bad news when $p_t > p_{fast}$. Thus, the strategy $\boldsymbol{x}_M$ is the best response to the evaluator's strategy, and the PBE is verified.

(2) Now we show the above PBE is the unique cutoff equilibrium. First, we show that for any admissible strategy of the agent, the evaluator's best response, if it exists, is to experiment when $p_t > p_{fast}$ and to dismiss the agent when $p_t < p_{slow}$. This helps us restrict our attention to the evaluator's cutoff

strategy $\hat{p} \leq p_{fast}$. Second, we show some necessary conditions for the agent's equilibrium strategy in a cutoff equilibrium. At last, we combine the above results to show no other cutoff equilibrium exists.

(2a) As in the proof of Proposition 1.2, the maximal information (no censorship) and the minimal information (full censorship) provide two bounds for the value function of the evaluator for any admissible strategy of the agent.

From the Proposition 1.1 and Proposition 1.2, we know that in both extreme cases, the evaluator would experiment when $p_t > p_{fast}$, and would dismiss the agent when $p_t < p_{slow}$. Hence, for any admissible strategy, the evaluator would do the same.

(2b) Fix a cutoff equilibrium where the evaluator uses a cutoff strategy with the cutoff belief $\hat{p}$, then we show that the equilibrium strategy of the type $B$ agent is to censor all bad news whenever $p_t > \hat{p}$, and when $p_t = \hat{p}$, her censoring probability must satisfy $x_t^B \geq x^{B*} := 1 - \frac{\gamma}{\beta}$.

Suppose the equilibrium strategy of the type $B$ agent is $x_t^B < 1$ at time $t$ when $p_t > \hat{p}$.

First, note that the public belief in the absence of news will eventually drift down to $\hat{p}$ in finite time $s = \inf\{\nu > t : p_\nu = \hat{p}\} > 0$, otherwise the type $B$ ruler would censor all bad news since the cost $c$ is low (i.e. $c < \underline{c}$), which is a contradiction. Hence, there must be some $t' \in (t, s)$ such that $\dot{p}_{t'} < 0$; otherwise the pubic belief cannot drift down to $\hat{p}$. Thus, $x_{t'}^B > 0$, and $V_B^{\boldsymbol{r}}(\bar{\varnothing}_{t'}^B) \geq 0$. The exactly same argument in Proposition 1.2 shows that an alternative strategy $\boldsymbol{x}^{B'}$ such that censor all bad news for time $\nu \in [t, t']$ and resume the original strategy gives $v^{\boldsymbol{r},\boldsymbol{x}^{B'}}(\bar{\varnothing}_t^B) > 0$. Hence, it is optimal for the type $B$ agent to censor bad news at time $t$, i.e. $x_t^B = 1$. Contradiction.

Suppose the equilibrium strategy of the type $B$ agent is $x_t^B < x^{B*}$ at time $t$ when $p_t = \hat{p}$.

Clearly, when $x_t^B < 1 - \frac{\gamma}{\beta}$ and $p_t = \hat{p}$ in equilibrium, then $\dot{p}_t > 0$. Hence, for continuation games, $\dot{p}_\nu > 0$ for $\nu$ in some half-neighborhood of $t$, say $[t, t')$, since $x_t^B$ is continuous from the right. It is clear that for any $\nu \in (t, t')$, $p_\nu > p^g$. Hence, by the above result, $x_\nu^B = 1$ for $\nu \in (t, t')$, which implies that $\dot{p}_\nu < 0$ for $\nu \in (t, t')$. Contradiction.

(2c) Suppose there is another cutoff equilibrium, in which the evaluator uses a cutoff strategy with a cutoff belief $\hat{p}$.

According to (2a), $\hat{p} \leq p_{fast}$. (2b) shows that in the equilibrium, the type $B$ agent would censor all bad news when $p_t > \hat{p}$, and her censoring probability when $p_t = \hat{p}$ cannot be less than $x^{B*}$. Moreover, her censoring probability when $p_t = \hat{p}$ cannot be more than $x^{B*}$ either. Suppose not, i.e. suppose $p_t = \hat{p}$ at some time $t$, and the type $B$ agent's censoring probability is larger than $x^{B*}$. Then, the public belief would be drifting down below $\hat{p}$ after $t$. Since the evaluator uses a cutoff strategy, he must remove the agent at time $t$. However, the type $B$ agent would not censor bad news at all at time $t$. This is a contradiction. Hence, in this cutoff equilibrium, the censoring probability when $p_t = \hat{p}$ must be equal to $x^{B*} \in (0, 1)$, and the public belief would stay at $\hat{p}$ thereafter. This means the evaluator's strategy at this belief must make the type $B$ agent indifferent between censoring and not censoring bad news. This implies that the evaluator would use a constant hazard rate $z^*$ to remove the agent at this belief. When $\hat{p} = p_{fast}$, this is the cutoff equilibrium we verified. When $\hat{p} < p_{fast}$, we can show that the best response of the evaluator is actually a cutoff strategy with the cutoff belief $p_{fast}$, since more censorship weakly increases the cutoff belief in the evaluator's cutoff strategy. This contradiction completes the proof.

□

*Proof of Proposition 1.5.* The first result is obvious.

When $p_0 \geq p_{fast}$, the type $B$ agent's expected payoff in the equilibrium is

$$(w - \beta^B c)(1 - e^{-\rho_0 \hat{s}}) + e^{-\rho_0 \hat{s}} \rho_0 c.$$

Her payoff in the NCB is

$$\frac{\rho_0 w}{\rho_0 + \beta^B}.$$

Hence, she would be worse off in the equilibrium with censorship if and only if

$$(w - \beta^B c)(1 - e^{-\rho_0 \hat{s}}) + e^{-\rho_0 \hat{s}} \rho_0 c < \frac{\rho_0 w}{\rho_0 + \beta^B}.$$

Since we have

$$w - \beta^B c > \frac{\rho_0 w}{\rho_0 + \beta^B} > \rho_0 c,$$

thus she would be worse off in the equilibrium with censorship if and only if

$$e^{-\rho_0 \hat{s}} > \frac{\beta^B}{\rho_0 + \beta^B}.$$

Denote $\hat{s}^* := \frac{1}{\rho_0} \ln[\frac{\rho_0+\beta^B}{\beta^B}]$, we get the result.

□

*Proof of Corollary 1.3.* Clearly, since $\hat{s}$ is not a function of the censoring cost $c$, whether the type $B$ agent is better off in the equilibrium than she is in the NCB does not depend on the cost.

when $p_0 \geq p_{fast}$, the type $B$ agent's expected payoff in the equilibrium is

$$(w - \beta^B c)(1 - e^{-\rho_0 \hat{s}}) + e^{-\rho_0 \hat{s}} \rho_0 c.$$

Its first derivative with respect to $c$ is

$$-\beta^B + (\rho_0 + \beta^B) e^{-\rho_0 \hat{s}}.$$

It is clear that the type $B$ agent's expected payoff in the equilibrium is decreasing/increasing in the censoring cost $c$ if and only if

$$e^{-\rho_0 \hat{s}} \lessgtr \frac{\beta^B}{\rho_0 + \beta^B}.$$

From the proof of Proposition 1.5, clearly, the conditions coincide.

□

*Proof of Proposition 1.6.* (1) First, given the strategies of all players, the pubic belief will drift and jump differently in different phases. When $p_t > p^B$, no bad news is expected in equilibrium, and we assume the public belief will jump to 0 after off-path bad news. In addition, the drifting process in the absence of news follows

$$\mathrm{d}p_t = -p_t(1-p_t)\gamma \,\mathrm{d}t.$$

When $p_t \in (p^G, p^B)$, the public belief will jump to $J(p_t, 1, 0) = 0$ after a piece of bad news, and the drifting process in the absence of news follows

$$\mathrm{d}p_t = -p_t(1-p_t)(\gamma - \beta^B)\,\mathrm{d}t.$$

When $p_t < p^G$, the public belief will jump to $j(p_t) > 0$ after a piece of bad news, and the drifting process in the absence of news follows

$$\mathrm{d}p_t = -p_t(1-p_t)(\gamma + \beta^G - \beta^B)\,\mathrm{d}t.$$

In addition, we will see later that $j(p^G) \leq p_{fast}$ is equivalent to $\beta^G \leq \bar{\beta}$, hence $j(p_t) \leq j(p^G) \leq p_{fast}$ when $p_t < p^G$.

Hence, it means that the public belief $p_t$ will jump below $p_{fast}$ after a piece of bad news, no matter how high $p_t < 1$ is.

(2) Now, suppose that the evaluator uses a cutoff strategy with the cutoff $p_{fast}$, let us solve the best response of each type of the agent.

Since the type $B$ agent faces the exactly same problem as in the conclusive bad news case. Her optimal strategy is the cutoff strategy with the cutoff belief $p^B$.

The logic in Proposition 1.2 also applies to the type $G$ agent. She will eventually stop censoring when the public belief approaches $p_{fast}$. Let $\tau^G$ and

$\chi$ be the arrival times of a piece of bad news and a piece of good news (i.e. success) from the type $G$ agent, respectively. Then her payoff after stopping censoring, given the public belief $p > p_{fast}$, is

$$\mathbb{E}[\int_0^T \rho_0 e^{-\rho_0 \nu} w d\nu] = \mathbb{E}[w(1 - e^{-\rho_0 T})],$$

where

$$T = \begin{cases} \infty, & \chi < \tau^G \wedge s, \\ \tau^G \wedge s, & \chi > \tau^G \wedge s, \end{cases}$$

and $s$ is the time when the public belief drifts from $p$ to $p_{fast}$. We can show that

$$\mathbb{E}[w(1 - e^{-\rho_0 T})] = \frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0}(1 - e^{-(\beta^G + \gamma + \rho_0)s})w.$$

Its partial derivative with respect to $s$ is

$$(\gamma + \rho_0)e^{-(\beta^G + \gamma + \rho_0)s} w > 0.$$

Hence, it is strictly increasing in $s$.

The same argument in Proposition 1.2 also implies that, at public belief $p^G$, the type $G$ agent must be indifferent between censoring or not if a piece of bad news arrives,

$$\frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0}(1 - e^{-(\beta^G + \gamma + \rho_0)s_G})w = \rho_0 c,$$

where $s_G = \frac{\ln[\frac{p^G}{1-p^G} \frac{1-p_{fast}}{p_{fast}}]}{\gamma + \beta^G - \beta^B}$. This indifference condition pins down $p^G$. The type $G$ agent has a strict incentive to censor all bad news when $p_t > p^G$, and has a strict incentive not to censor any news when $p_t < p^G$.

Let $s_B$ be such that

$$\mathbb{E}[-\rho_0 c + \int_0^{s_B \wedge \tau^B} \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu] = 0.$$

Hence, $s_B$ is the time when the public belief drifts from $p^{B\dagger}$ to $p_{fast}$. Thus

$$\frac{\rho_0}{\beta^B + \rho_0}(1 - e^{-(\beta^B+\rho_0)s_B})w = \rho_0 c.$$

Since $c < \underline{c}$, we have

$$\rho_0 c < \frac{\rho_0}{\beta^B + \rho_0} w < \frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0} w.$$

It is easy to see that $\frac{\gamma+\rho_0}{\beta^G+\gamma+\rho_0}(1 - e^{-(\beta^G+\gamma+\rho_0)s})w$ is increasing in $\gamma > 0$, and decreasing in $\beta^G < \beta^B$, thus we have

$$\begin{aligned}\frac{\rho_0}{\beta^B + \rho_0}(1 - e^{-(\beta^B+\rho_0)s_B})w &= \frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0}(1 - e^{-(\beta^G+\gamma+\rho_0)s_G})w \\ &> \frac{\rho_0}{\beta^B + \rho_0}(1 - e^{-(\beta^B+\rho_0)s_G})w.\end{aligned}$$

Hence, $s_B > s_G$, which implies $p^{B\dagger} > p^G$.

Now we show that there exists a $\bar{\beta} < \beta^B$ such that $j(p^G) \leq p_{fast}$ if and only if $\beta^G < \bar{\beta}$.

Note that

$$j^{-1}(p_{fast}) = \frac{p_{fast}\beta^B}{p_{fast}\beta^B + (1 - p_{fast})\beta^G} = \frac{p_{fast}\frac{\beta^B}{\beta^G}}{p_{fast}\frac{\beta^B}{\beta^G} + (1 - p_{fast})},$$

and

$$p^G = \frac{p_{fast}e^{(\gamma+\beta^G-\beta_B)s_G}}{1 - p_{fast} + p_{fast}e^{(\gamma+\beta^G-\beta_B)s_G}}.$$

Hence, $j^{-1}(p_{fast}) \geq p^G$ is equivalent to $\frac{\beta^B}{\beta^G} \geq e^{(\gamma+\beta^G-\beta_B)s_G}$. Since

$$\rho_0 c = \frac{\gamma+\rho_0}{\beta^G+\gamma+\rho_0}(1-e^{-(\beta^G+\gamma+\rho_0)s_G})w,$$

$j^{-1}(p^g) \geq p^*_G$ is also equivalent to

$$\frac{\beta^B}{\beta^G} \geq (\frac{\bar{c}}{\bar{c}-c})^{1-\frac{\beta^B+\rho_0}{\gamma+\beta^G+\rho_0}},$$

where

$$\bar{c} = \frac{\gamma+\rho_0}{\beta^G+\gamma+\rho_0}\frac{w}{\rho_0}.$$

Note that

$$\frac{\mathrm{d}\beta^G(\frac{\bar{c}}{\bar{c}-c})^{1-\frac{\beta^B+\rho_0}{\gamma+\beta^G+\rho_0}}}{\mathrm{d}\beta^G} = (\frac{\bar{c}}{\bar{c}-c})^{1-\frac{\beta^B+\rho_0}{\gamma+\beta^G+\rho_0}}\{1+\frac{\beta^G}{\gamma+\beta^G+\rho_0}[\frac{\rho_0 c(\gamma+\beta^G-\beta^B)}{w(\gamma+\rho_0)-\rho_0 c(\gamma+\beta^G+\rho_0)} + \ln(\frac{\bar{c}}{\bar{c}-c})\frac{\beta^G+\rho_0}{\gamma+\beta^G+\rho_0}]\} > 0.$$

In addition, $\beta^G(\frac{\bar{c}}{\bar{c}-c})^{1-\frac{\beta^B+\rho_0}{\gamma+\beta^G+\rho_0}}$ goes to 0 when $\beta^G$ goes to 0, and it is lager than $\beta^B$ when $\beta^G = \beta^B$. Hence, there exists a $\bar{\beta} \in (0, \beta^B)$, such that when $\beta^G \in (0, \bar{\beta}]$, the condition $j^{-1}(p_{fast}) \geq p^G$ is satisfied.

(3) At last, given the strategies of both types of agent, the evaluator faces the classic bandit problem when $p_t < p^G$, and his best response is a cutoff strategy with the cutoff $p_{fast}$. In addition, when $p_t > p^G$, it is easy to verify that the evaluator has a strict incentive to not dismiss the agent. This completes the proof. □

*Proof of Proposition 1.7.* (1) We first establish that there exists a $\underline{\beta} \in (0, \bar{\beta})$, such that $j(p^{B\dagger}) \leq p_{fast}$ if and only if $\beta^G \in (0, \underline{\beta})$.

Note that

$$p^{B\dagger} = \frac{\frac{p^G}{1-p^G}e^{(\gamma-\beta^B)(s_B-s_G)}}{1+\frac{p^G}{1-p^G}e^{(\gamma-\beta^B)(s_B-s_G)}} = \frac{\frac{p_{fast}}{1-p_{fast}}e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}}{1+\frac{p_{fast}}{1-p_{fast}}e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}} = \frac{p_{fast}e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}}{1-p_{fast}+p_{fast}e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}}.$$

Hence, $j^{-1}(p_{fast}) \geq p^{B\dagger}$ is equivalent to $\frac{\beta^B}{\beta^G} \geq e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}$. Since

$$\rho_0 c = \frac{\gamma+\rho_0}{\beta^G+\gamma+\rho_0}(1-e^{-(\beta^G+\gamma+\rho_0)s_G})w,$$

we have

$$\beta^G s_G = \beta^G \frac{\ln \frac{w(\gamma+\rho_0)}{w(\gamma+\rho_0)-\rho_0 c(\beta^G+\gamma+\rho_0)}}{\beta^G+\gamma+\rho_0}.$$

Clearly, $\beta^G s_G$ is increasing in $\beta^G$.

Consider

$$\frac{\beta^B}{\beta^G} \geq e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}.$$

The left hand side (LHS) is decreasing in $\beta^G$, and the right hand side (RHS) is increasing in $\beta^G$. In addition, $\beta^G e^{(\gamma-\beta^B)s_B}e^{\beta^G s_G}$ goes to 0 when $\beta^G$ goes to 0, and it is larger than $\beta^B$ when $\beta^G = \beta^B$.

Thus, $j^{-1}(p_{fast}) \geq p^{B\dagger}$ if and only if $\beta^G$ is smaller than some threshold $\underline{\beta} > 0$, and $\underline{\beta} < \bar{\beta}$ since $p^{B\dagger} > p^G$.

(2) We show that the evaluator is better off with censorship when $p_0 \in (p^G, p^{B\dagger}]$.

For $p \in [p^G, p^{B\dagger}]$, the HJB equation in the NCB is

$$(\rho_1+p\gamma)U^{\mathbf{0}}(p)+(p\beta^G+(1-p)\beta^B)[U^{\mathbf{0}}(p)-U^{\mathbf{0}}(j(p))]+(\gamma+\beta^G-\beta^B)p(1-p)U^{\mathbf{0}\prime}(p) = p\gamma k(\gamma+\rho_1).$$

In equilibrium, it is

$$(\rho_1+p\gamma)U^{\tilde{\boldsymbol{x}}}(p)+(1-p)\beta^B[U^{\tilde{\boldsymbol{x}}}(p)-m]+(\gamma-\beta^B)p(1-p)U^{\tilde{\boldsymbol{x}}\prime}(p) = p\gamma k(\gamma+\rho_1).$$

For $p \leq p^G$, clearly $U^{\tilde{\boldsymbol{x}}}(p) = U^{\mathbf{0}}(p)$. Hence, at $p = p^G$, comparing the two HJB equations, we have

$$U^{\mathbf{0}\prime}(p^G) < U^{\tilde{\boldsymbol{x}}\prime}(p^G),$$

where $U^{\tilde{\boldsymbol{x}}\prime}(p^G)$ is the right derivative.

Since the value functions are continuously differentiable, the above relation is true for some neighborhood $p \in [p^G, \check{p})$, hence in that neighborhood $U^{\tilde{\boldsymbol{x}}}(p) - U^{\mathbf{0}}(p)$ is increasing in $p$. Thus, $U^{\tilde{\boldsymbol{x}}}(p) > U^{\mathbf{0}}(p)$ for $p \in (p^G, \check{p})$.

To prove by contradiction, suppose there is some $\acute{p} \in (p^G, p^{B\dagger}]$ such that $U^{\tilde{\boldsymbol{x}}}(\acute{p}) \leq U^{\mathbf{0}}(\acute{p})$. Let $\acute{p} = \inf\{p \in (p^G, p^{B\dagger}] : U^{\tilde{\boldsymbol{x}}}(\acute{p}) = U^{\mathbf{0}}(\acute{p})\}$. First, note that it must be that $U^{\mathbf{0}\prime}(\acute{p}) \geq U^{\tilde{\boldsymbol{x}}\prime}(\acute{p})$. Otherwise, if $U^{\mathbf{0}\prime}(\acute{p}) < U^{\tilde{\boldsymbol{x}}\prime}(\acute{p})$, then it is true in a small neighborhood, and $U^{\tilde{\boldsymbol{x}}}(p) - U^{\mathbf{0}}(p)$ is increasing in $p$ in that neighborhood, then it must be that $U^{\tilde{\boldsymbol{x}}}(\acute{p}) > U^{\mathbf{0}}(\acute{p})$. Then

$$\begin{aligned}
&(\acute{p}\beta^G + (1-\acute{p})\beta^B)[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] + (\gamma + \beta^G - \beta^B)\acute{p}(1-\acute{p})U^{\mathbf{0}\prime}(\acute{p}) \\
=&(1-\acute{p})\beta^B[U^{\tilde{\boldsymbol{x}}}(\acute{p}) - m] + (\gamma - \beta^B)\acute{p}(1-\acute{p})U^{\tilde{\boldsymbol{x}}\prime}(\acute{p}) \\
\leq&(1-\acute{p})\beta^B[U^{\tilde{\boldsymbol{x}}}(\acute{p}) - m] + (\gamma - \beta^B)\acute{p}(1-\acute{p})U^{\mathbf{0}\prime}(\acute{p}).
\end{aligned}$$

Hence,

$$(\acute{p}\beta^G + (1-\acute{p})\beta^B)[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] + \beta^G\acute{p}(1-\acute{p})U^{\mathbf{0}\prime}(\acute{p}) - (1-\acute{p})\beta^B[U^{\mathbf{0}}(\acute{p}) - m] \leq 0.$$

Note that if $j(\acute{p}) \leq p_{fast}$, then $U^{\mathbf{0}}(j(\acute{p})) = m$, a contradiction. Hence, it must be $j(\acute{p}) > p_{fast}$. Since $U^{\mathbf{0}}(p)$ is strictly convex in $p \in [p_{fast}, 1]$, we have

$$U^{\mathbf{0}\prime}(\acute{p}) > \frac{U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))}{\acute{p} - j(\acute{p})},$$

and

$$U^{\mathbf{0}}(\acute{p}) - m = U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(p_{fast}) < [U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))]\frac{\acute{p} - p_{fast}}{\acute{p} - j(\acute{p})}.$$

Hence, we have

$$\begin{aligned}
&(\acute{p}\beta^G + (1-\acute{p})\beta^B)[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] + \beta^G\acute{p}(1-\acute{p})U^{\mathbf{0}\prime}(\acute{p}) - (1-\acute{p})\beta^B[U^{\mathbf{0}}(\acute{p}) - m] \\
>&(\acute{p}\beta^G + (1-\acute{p})\beta^B)[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] + \beta^G\acute{p}(1-\acute{p})\frac{U(\acute{p}) - U(j(\acute{p}))}{\acute{p} - j(\acute{p})} \\
&- (1-\acute{p})\beta^B[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))]\frac{\acute{p} - p_{fast}}{\acute{p} - j(\acute{p})} \\
=&(\acute{p}\beta^G + (1-\acute{p})\beta^B + \beta^G\acute{p}(1-\acute{p})\frac{1}{\acute{p} - j(\acute{p})} - (1-\acute{p})\beta^B\frac{\acute{p} - p_{fast}}{\acute{p} - j(\acute{p})})[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] \\
>&(\acute{p}\beta^G + (1-\acute{p})\beta^B + \beta^G\acute{p}(1-\acute{p})\frac{1}{\acute{p} - j(\acute{p})} - (1-\acute{p})\beta^B\frac{\acute{p}}{\acute{p} - j(\acute{p})})[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))] \\
=&(\acute{p}\beta^G + (1-\acute{p})\beta^B + \frac{\acute{p}(1-\acute{p})(\beta^G - \beta^B)}{\acute{p} - j(\acute{p})})[U^{\mathbf{0}}(\acute{p}) - U^{\mathbf{0}}(j(\acute{p}))].
\end{aligned}$$

Since

$$\acute{p} - j(\acute{p}) = \frac{(\beta^B - \beta^G)\acute{p}(1-\acute{p})}{\acute{p}\beta^G + (1-\acute{p})\beta^B},$$

the above equation is strictly larger than 0, a contradiction.

(3) Now we show the type $B$ agent is better off with censorship when $\beta^G \in (0, \underline{\beta}]$ and $p_0 \in (p^G, p^{B\dagger}]$.

In the equilibrium and the NCB, the type $B$ agent does not censor bad news when $p_0 \leq p^{B\dagger}$. In addition, in both cases, she would be dismissed when either a piece of bad news arrives or when the public belief drift down to $p_{fast}$. Since the public belief drifts down slower in the equilibrium than it does in the NCB, the type $B$ agent is better off with censorship.

(4) At last, we show the type $G$ agent is also better off with censorship when $\beta^G \in (0, \underline{\beta}]$ and $p_0 \in (p^G, p^{B\dagger}]$.

In the equilibrium, her payoff when $p_0 \in (p^G, p^{B\dagger}]$ is

$$\int_0^{s^1} \gamma e^{-\gamma\nu}[w - c\int_0^{\nu} \rho_0 e^{-\rho_0 t}\beta^G \,\mathrm{d}t] + e^{-\gamma s^1}[(w - c\beta^G)(1 - e^{-\rho_0 s^1}) + e^{-\rho_0 s^1}\rho_0 c]$$
$$= \frac{\gamma w + \rho_0(w - c\beta^G)}{\gamma + \rho_0}(1 - e^{-(\rho_0+\gamma)s^1}) + e^{-(\rho_0+\gamma)s^1}\rho_0 c,$$

where $s^1 = \frac{\ln[\frac{p_0}{1-p_0}\frac{1-p^G}{p^G}]}{\gamma - \beta^B}$ is the time that belief drifts down from $p_0$ to $p^G$ according to rate $\gamma - \beta^B$.

In the NCB, since $j(p_0) \le p_{fast}$, her payoff is

$$\int_0^{s^0} \beta^G e^{-\beta^G\nu} e^{-\gamma\nu} w(1 - e^{-\rho_0\nu})d\nu + \int_0^{s^0} \gamma e^{-\beta^G\nu} e^{-\gamma\nu} w d\nu$$
$$+ e^{-\beta^G s^0} e^{-\gamma s^0}[w(1 - e^{-\rho_0 s^0}) + e^{-\rho_0 s^0}\rho_0 c]$$
$$= \frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0}(1 - e^{-(\beta^G+\gamma+\rho_0)s^0})w + e^{-(\beta^G+\gamma+\rho_0)s^0}\rho_0 c,$$

where $s^0 = \frac{\ln[\frac{p_0}{1-p_0}\frac{1-p^G}{p^G}]}{\gamma + \beta^G - \beta^B}$ is the time that belief drifts down from $p_0$ to $p^G$ according to rate $\gamma + \beta^G - \beta^B$.

First, since $c < \underline{c}$, we have

$$\frac{\gamma w + \rho_0(w - c\beta^G)}{\gamma + \rho_0} > \frac{\gamma + \rho_0}{\beta^G + \gamma + \rho_0} w.$$

Second, note that

$$e^{-(\rho_0+\gamma)s^1} < e^{-(\beta^G+\gamma+\rho_0)s^0},$$

because

$$(\rho_0+\gamma)s^1 = \ln[\frac{p_0}{1-p_0}\frac{1-p^G}{p^G}]\frac{\rho_0+\gamma}{\gamma-\beta^B} > \ln[\frac{p_0}{1-p_0}\frac{1-p^G}{p^G}]\frac{\beta^G+\gamma+\rho_0}{\gamma+\beta^G-\beta^B} = (\beta^G+\gamma+\rho_0)s^0.$$

Hence, the type $G$ agent is better off with censorship.

□

*Proof of Proposition 1.8.* (1) (Low cost) Given the strategies of both agents, there is no bad news in equilibrium when $p_t > p_{fast}$. We assume the public belief jumps down to 0 after off-path bad news. In addition, the belief drifting process is the same as in Proposition 1.4 when $p_t \geq p_{fast}$.

Given the strategy of the type $G$ agent, the evaluator and the type $B$ agent face the same problem, hence have the same best response as in Proposition 1.4 when $p_t \geq p_{fast}$.

Given the strategy of the evaluator, when $p_t = p_{fast}$, the type $G$ agent's continuation value after censoring one piece of bad news is

$$\mathbb{E}[\int_0^T \rho_0 e^{-\rho_0 \nu} w \,\mathrm{d}\nu - \int_0^{T\wedge\chi} \rho_0 e^{-\rho_0 \nu} c X_\nu^G \,\mathrm{d}N_\nu^G],$$

where

$$T = \begin{cases} \infty, & \chi < \lambda, \\ \lambda, & \chi > \lambda, \end{cases}$$

and $\lambda$ is the arrival time of dismissal induced by $z^* = \frac{w}{c} - \rho_0 - \beta^B$, and $\chi$ is the arrival time of success.

Hence, her continuation value is

$$\begin{aligned} &\mathbb{E}[w(1 - e^{-\rho_0 T}) - \beta^G c(1 - e^{-\rho_0 (T\wedge\chi)})] \\ =& w\frac{\gamma + \rho_0}{z^* + \gamma + \rho_0} - \beta^G c \frac{\rho_0}{z^* + \gamma + \rho_0} \\ =& \frac{w(\gamma + \rho_0) - \beta^G \rho_0 c}{z^* + \gamma + \rho_0}. \end{aligned}$$

Since $z^* = \frac{w}{c} - \rho_0 - \beta^B$, we have

$$\frac{w(\gamma+\rho_0) - \beta^G \rho_0 c}{z^* + \gamma + \rho_0} > \rho_0 \frac{w - \beta^G c}{z^* + \rho_0} = \rho_0 c \frac{w - \beta^G c}{w - \beta^B c} > \rho_0 c.$$

Hence, the type $G$ agent has a strict incentive to censor bad news when $p_t = p_{fast}$. In addition, it is easy to verify that she also has a strict incentive to censor bad news when $p_t > p_{fast}$.

At last, when $p_t < p_{fast}$, we can show that the evaluator has a strict incentive to dismiss the agent, given the strategies of both types of agent. Thus, neither agent has an incentive to censor bad news when $p_t < p_{fast}$.

(2) (Intermediate cost) Given the strategies of both agents, a piece of bad news will make the public belief jump down to $J(p_t, 1, 0) = 0$ when $p_t > p_{slow}$. In the absence of news, the public belief will drift up according to

$$\mathrm{d}p_t = -p_t(1-p_t)(\gamma - \beta^B)\,\mathrm{d}t.$$

The type $B$ agent faces a problem as in Proposition 1.1, since $c > \underline{c}$, so she never censors any bad news.

The evaluator also faces a problem as in Proposition 1.1, thus he uses a cutoff strategy with the cutoff belief $p_{slow}$.

When $p_t > p_{slow}$, if the type $G$ agent censors all bad news before the success, her payoff would be

$$\int_0^\infty \gamma e^{-\gamma\nu}[w - c\int_0^\nu \rho_0 e^{-\rho_0 t}\beta^G\,\mathrm{d}t]\,\mathrm{d}\nu = w - c\beta^G + c\beta^G \frac{\gamma}{\rho_0+\gamma} > \rho_0 c,$$

since $c < \bar{c}$. Hence, the type $G$ agent has a strict incentive to censor bad news.

At last, when $p_t < p_{slow}$, we can show that the evaluator has a strict incentive to dismiss the agent, given the strategies of both types of agent. Thus, neither agent has an incentive to censor bad news when $p_t < p_{slow}$. $\square$

*Proof of Proposition 1.9.* (1) First, we summarize the property of $p^*$ – the strategy of the evaluator in the NCB from the bandit literature. The optimal policy of the evaluator is a cutoff strategy with the cutoff $p^*$. Her value function is continuously differentiable everywhere, with a possible exception at the cutoff $p^*$. In addition, when $p_0 > p^*$, her value function is strictly increasing and strictly convex.

When $\gamma + \beta^G \geq \beta^B$, $p^* = p_{fast}$, and the evaluator's value function is continuously differentiable at $p^*$.

When $\gamma + \beta^G < \beta^B$, $p^* \in (p_{slow}, p_{fast})$, and the evaluator's value function has a kink at $p^*$.

(2) (Low cost) Assume $c < \underline{c}$ and $p_0 > p^*$.

In the NCB, the evaluator can always use the same strategy as in the equilibrium by ignoring bad news and dismissing the agent if no success arrives for some time. However, this strategy is strictly dominated by her optimal strategy in the NCB, which implies she is strictly worse off in the equilibrium when $p_0 > p^*$.

(3) (Intermediate cost) Assume $c \in (\underline{c}, \bar{c})$ and $p_0 > p_{slow}$. Let $U^{\mathbf{0}}(p_0)$ and $U^{\tilde{\boldsymbol{x}}}(p_0)$ be the evaluator's value functions in the NCB and in the equilibrium, respectively.

Clearly, the evaluator is better off in equilibrium when $p_0 \in (p_{slow}, p^*]$ and $U^{\mathbf{0}}(p^*) < U^{\tilde{\boldsymbol{x}}}(p^*)$.

Consider $p_0 \in (p^*, 1)$. To prove by contradiction, suppose there is some $\acute{p} \in (p^*, 1)$ such that $U^{\tilde{\boldsymbol{x}}}(\acute{p}) \leq U^{\mathbf{0}}(\acute{p})$. Let $\acute{p} = \inf\{p \in (p^*, 1) : U^{\tilde{\boldsymbol{x}}}(\acute{p}) = U^{\mathbf{0}}(\acute{p})\}$. First, note that it must be that $U^{\mathbf{0}\prime}(\acute{p}) \geq U^{\tilde{\boldsymbol{x}}\prime}(\acute{p})$. Otherwise, if $U^{\mathbf{0}\prime}(\acute{p}) < U^{\tilde{\boldsymbol{x}}\prime}(\acute{p})$, then it is true in a small neighborhood, and $U^{\tilde{\boldsymbol{x}}}(p) - U^{\mathbf{0}}(p)$ is increasing in $p$ in that neighborhood, then it must be that $U^{\tilde{\boldsymbol{x}}}(\acute{p}) > U^{\mathbf{0}}(\acute{p})$, a contradiction.

Since $U^{\mathbf{0}}(p)$ is strictly convex for $p \in [p^*, 1]$, $U^{\mathbf{0}\prime}(p) > U^{\mathbf{0}\prime}(\acute{p}) \geq U^{\tilde{\boldsymbol{x}}\prime}(\acute{p})$ for $p > \acute{p}$. Also, from the Proof of Proposition 1.1, we know that the value function of the evaluator in the equilibrium is linear when $p_0 > p_{slow}$, hence $U^{\mathbf{0}\prime}(p) > U^{\tilde{\boldsymbol{x}}\prime}(\acute{p}) = U^{\tilde{\boldsymbol{x}}\prime}(p)$ for any $p \in (\acute{p}, 1)$. Hence, $U^{\mathbf{0}}(p) - U^{\tilde{\boldsymbol{x}}}(p)$ is continuous and strictly increasing in $p \in (\acute{p}, 1]$. Thus, $U^{\mathbf{0}}(1) - U^{\tilde{\boldsymbol{x}}}(1) > U^{\mathbf{0}}(\acute{p}) - U^{\tilde{\boldsymbol{x}}}(\acute{p}) = 0$, a contradiction.

□

# Appendix B

# Appendix for Chapter 2

## B.1 Temporary Delay of the Project

In this section, we formalize the idea of delaying the project temporarily. Lemma B.1 will show that it is never optimal for the $H$ type to delay her project temporarily.

Given a contract $C = \{N, W, \boldsymbol{b}\}$, and the agent's belief $\beta_0$ about the principal being $H$, after accepting the contract, the agent would solve a dynamic decision problem, i.e. when to exert efforts. Suppose the agent's optimal action plan is to exert efforts only for periods in $R \subseteq \{1, 2, \cdots, N\}$, then the contract must satisfy the agent's IC constraints[1] for those periods, i.e. for all $1 \leq n \leq \#R$ ($\#R$ is the size of $R$),

$$\sum_{s=n}^{\#R} \Pi_{j=n+1}^{s} \delta_j f_{s-n}(q_n, \beta_n)(q_s \lambda_s b_{r_s} - c)\Delta \geq \sum_{s=n+1}^{\#R} \Pi_{j=n+1}^{s} \delta_j f_{s-n-1}(q_n, \beta_n)(q_{s-1}\lambda_{s-1} b_{r_s} - c)\Delta$$

where $\delta_j = \delta^{r_j - r_{j-1}}$, $r_j$ is the $j$-th minimal item in $R$, and $r_0 = 0$.

[1]The notations are borrowed from the section 2.5.3. The constraints are the agent's IC constraints when the agent believes that the $H$ and $L$ type pool with the prior belief $\beta_0$. The degenerate case $\beta_0 = 0$ (or $\beta_0 = 1$) means the agent believes the principal is the $L$ (or $H$) type. The bonus scheme also needs to satisfy the IC constraints that the agent does not work in the periods $n \in \{1, 2, \cdots, N\} \backslash R$. We omit the requirement of $b_n$ for $n \in \{1, 2, \cdots, N\} \backslash R$, since they are payoff irrelevant in the equilibrium as long as they are low enough to discourage the agent to experiment.

In such a contract, $\#R \leq N$ is the actual number of experiments. Everything remains the same in the IC constraints, except for the discount factor for each experiment. Therefore, we can emphasize the contract in a different way: the principal has to incentivize the agent to work for every experiment, but she can choose how long to suspend the project between two consecutive experiments, or equivalently, choose the discount factors between every two consecutive experiments.

From the perspective of the induced actions, we define the following direct contracts. A *direct contract* is a quadruple, $\mathscr{C} = \{N, W, \boldsymbol{b}, \boldsymbol{\delta}\}$, where $N \in \mathbb{N}_0$ is the number of experiments, $W$ is still the lump-sum time zero transfer from the principal to the agent, $\boldsymbol{b} \in \mathbb{R}^N$ is the vector of bonus scheme for success that *must* incentivize the agent to work for *all* $N$ experiments, and $\boldsymbol{\delta} \in [0, \delta]^N$ is the vector of discount factors between two consecutive experiments, and for $1 \leq n \leq N$, $\delta_n = \delta^k$ for some $k \in \mathbb{N} \cup \{\infty\}$, therefore $0 \leq \delta_n \leq \delta$. The principal can freely choose any discount factors and any number of experiments. However, given the chosen discount factors and the number of experiments, the bonus payments must satisfy the IC constraints for the agent for all $N$ experiments.

With a direct contract $\mathscr{C} = \{N, W, \boldsymbol{b}, \boldsymbol{\delta}\}$, nothing changes except for discount factors. We can easily rewrite the payoffs of all parties, and their IR and IC constraints by properly adjusting their discount factors. For example, given the agent's belief about the principal's type being $H$ is $\beta_0$, his IC constraints become, for $1 \leq n \leq N$

$$\sum_{s=n}^{N} \Pi_{j=n+1}^{s} \delta_j f_{s-n}(q_n, \beta_n)(q_s \lambda_s b_s - c)\Delta \geq \sum_{s=n+1}^{N} \Pi_{j=n+1}^{s} \delta_j f_{s-n-1}(q_n, \beta_n)(q_{s-1} \lambda_{s-1} b_s - c)\Delta.$$

Therefore, we can use direct contracts to analyze the signaling game and the third-party mechanism design problem. However, Lemma B.1 implies that delaying the $H$ type project temporarily never happens in the best equilibrium for the $H$ type. As we will see later, this is also true for the optimal mechanism for the $H$ type.

## B.2 Three Useful Lemmas

We provide three lemmas that we need to prove other results in the Appendix. Lemma B.1 shows that the IC constraints for the agent always bind whenever there are signaling concerns, and temporary delay of the project is not a good way to separate the $H$ type. Lemma B.2 characterizes the recursive relations between two consecutive bonuses when the IC constraints for the agent bind. Lemma B.3 explores the limit of the payoff difference between the two types of principal from the same contract when the period length shrinks, and shows that it is monotonic in the belief about the type of principal when the termination time is less than the efficient termination time for the $H$ type.

### B.2.1 Lemma B.1

*Lemma* B.1. Fix any $N > 1$, let $\boldsymbol{b}^*, \boldsymbol{\delta}^*$ be a solution to the following program:

$$\max_{\boldsymbol{b},\{0\leq\delta_n\leq\delta\}_{n=1}^N} \sum_{n=1}^{N} \Pi_{j=1}^{n}\delta_j[f_{n-1}^H(q_0)q_n^H\lambda^H\Delta - f_{n-1}^L(q_0)q_n^L\lambda^L\Delta](h-b_n)$$

$$\text{s.t. } \sum_{s=n}^{N} \Pi_{j=n+1}^{s}\delta_j f_{s-n}(q_n,\beta_n)(q_s\lambda_s b_s - c)\Delta \geq \sum_{s=n+1}^{N} \Pi_{j=n+1}^{s}\delta_j f_{s-n-1}(q_n,\beta_n)(q_{s-1}\lambda_{s-1}b_s - c)\Delta$$

Let $m^* + 1 = \min\{n : \delta_n^* = 0\}$, or $m^* = N$ if $\{n : \delta_n^* = 0\} = \varnothing$, then for any $1 \leq n \leq m^*$, the above constraints bind and $\delta_n^* = \delta$.

The constraints are the agent's IC constraints. The objective function is the payoff difference between the $H$ and $L$ type when they propose the same contract. To find the best equilibrium for the $H$ type, the lump-sum payment $W$ is pinned down by either a binding IC constraint for the $L$ type as in the BSEH or a binding IR constraint for the agent as in the BPEH. In other words, the $H$ type has no reason to leave rents to the agent, unless she wants to prevent the $L$ type mimicking her. If the $L$ type obtains 0 payoff, then the objective function is equal to the $H$ type's payoff. If the agent obtains 0 payoff, then the objective function represents the part of the $H$ type's payoff that involves the bonus payments.

We show that given any termination date $N$, the IC constraints for the agent must bind for all periods even when arbitrary delay of the project is allowed. Moreover, we also show that temporary delay of the project is never optimal for the $H$ type. Furthermore, the result holds with respect to the following transformation of the objective function: $H \longmapsto \xi H + G(\delta_1, \ldots, \delta_n)$, where $\xi \in \mathbb{R}_{++}$, and $G(\delta_1, \ldots, \delta_n)$ is linear in every $\delta_n$ and independent with every $b_n$ for $1 \leq n \leq N$.

*Proof.* If $m^* = 1$, the statement is trivial. Consider $m^* > 1$. From $m^* + 1$, all terms in the objective function is 0, we can rewrite the program in which the last term is in the $m^*$ period.

Note that $\eta_n := f_{n-1}^H(q_0)q_n^H\lambda^H - f_{n-1}^L(q_0)q_n^L\lambda^L$, and $\eta_n\Delta$ is the difference of the probabilities to succeed for the $n$-th experiment between the $H$ and $L$ type. It is easy to see that when $n \leq n^* = \lfloor \frac{\log\lambda^H - \log\lambda^L}{\log(1-\lambda^L\Delta) - \log(1-\lambda^H\Delta)} \rfloor + 1$, $\eta_n \geq 0$. When $n > n^*$, $\eta_n < 0$.

Therefore, for $1 \leq n \leq n^*$, we would like to set $b_n$ as small as possible. Since $b_n$ is bounded from below, the IC constraints for $1 \leq n \leq n^*$ must bind.

Given the binding constraints for $1 \leq n \leq n^*$, consider $n = n^* + 1$. Note that a change of $b_{n^*+1}$, call it $y_{n^*+1}$, will change all the previous bonuses according to the binding constraints from period 1 to period $n^*$.

We can show that the change of $b_n$ for $1 \leq n \leq n^*$, call it $y_n$, from $y_{n^*+1}$ is

$$y_n = \Pi_{j=n+1}^{n^*+1}\delta_j[\lambda^H - (\lambda^H - \lambda^L)\frac{\lambda^L(1-\beta_{n^*})}{\lambda_{n^*}}]\Delta y_{n^*+1}.$$

Clearly, $\zeta_{n^*} := \lambda^H - (\lambda^H - \lambda^L)\frac{\lambda^L(1-\beta_{n^*})}{\lambda_{n^*}}$ is a positive constant. When $\beta_0 = 1$, it is $\lambda^H$, and when $\beta_0 = 0$, it is $\lambda^L$. When $\beta_0 \in (0, 1)$, it is in between $\lambda^L$ and $\lambda^H$, and converges to $\lambda^L$ when $n^*$ goes to infinity.

Thus, the change of the objective function from $y_{n^*+1}$ is

$$-\sum_{n=1}^{n^*+1}\Pi_{j=1}^{n}\delta_j\eta_n\Delta y_n = -\sum_{n=1}^{n^*}\Pi_{j=1}^{n^*+1}\delta_j\eta_n\Delta\zeta_{n^*}\Delta y_{n^*+1} - \Pi_{j=1}^{n^*+1}\delta_j\eta_{n^*+1}\Delta y_{n^*+1}$$
$$= -\Pi_{j=1}^{n^*+1}\delta_j y_{n^*+1}(\zeta_{n^*}\Delta\sum_{n=1}^{n^*}\eta_n\Delta + \eta_{n^*+1}\Delta) \qquad \text{(B.1)}$$

Note that the distribution on success for $L$ type first-order stochastic dominates that for $H$ type; the $H$ project is more likely to succeed than the $L$ project in the first $s \geq 1$ experiments. Formally, for any $s \geq 1$

$$\phi_s := \sum_{n=1}^{s}\eta_n\Delta = -[f_s^H(q_0) - f_s^L(q_0)] = q_0[(1-\lambda^L\Delta)^s - (1-\lambda^H\Delta)^s] > 0.$$

Therefore, the terms in the parenthesis of equation (B.1) becomes

$$\zeta_{n^*}\Delta q_0[(1-\lambda^L\Delta)^{n^*} - (1-\lambda^H\Delta)^{n^*}] + q_0[(1-\lambda^H\Delta)^{n^*}\lambda^H\Delta - (1-\lambda^L\Delta)^{n^*}\lambda^L\Delta]$$

which is large than $q_0(\lambda^H - \lambda^L)\Delta(1-\lambda^H\Delta)^{n^*} > 0$ since $\zeta_{n^*} \geq \lambda^L$.

Therefore, to maximize the objective function, we would like to make $b_{n^*+1}$ as small as possible. Thus, the IC constraints in period $n^*+1$ also binds.

By induction, we can get a necessary condition for the solution of the program: the IC constraints to incentivize agent to work must bind in each period.

In addition, since every discount factor enter objective function linearly, we must have $\delta_n^* = \delta$ for $1 \leq n \leq m^*$. The reason is the following.

Consider the discount factor in $n+1$ period, $\delta_{n+1}$. It will not affect any $b_s$ for $s \geq n+1$, then $\delta_{n+1}$ enters into the terms linearly beyond period $n+1$

as a discount factor. It will linearly affect every $b_s$ for $s \leq n$, but they are not discounted by $\delta_{n+1}$. Therefore, the discount factor $\delta_{n+1}$ enters objective function linearly.

Thus, the optimal discount factor must be a corner solution, i.e. $\delta_n^* \in \{\delta, 0\}$ for any $1 \leq n \leq m^*$. Since $m^* + 1 = \min\{n : \delta_n^* = 0\}$, we have $\delta_n^* = \delta$ for all $1 \leq n \leq m^*$. □

### B.2.2 Lemma B.2

*Lemma* B.2. Given the binding IC constraints for the agent when he has prior $\beta_0$ about the $H$ type, i.e. for $1 \leq n \leq N$,

$$\sum_{s=n}^{N} \delta^{s-n} f_{s-n}(q_n, \beta_n)(q_s \lambda_s b_s - c)\Delta = \sum_{s=n+1}^{N} \delta^{s-n} f_{s-n-1}(q_n, \beta_n)(q_{s-1}\lambda_{s-1} b_s - c)\Delta.$$

Then

$$q_n \lambda_n b_n - c = \delta(q_n \lambda_n b_{n+1} - c)$$

for $1 \leq n \leq N - 1$, and

$$q_N \lambda_N b_N - c = 0.$$

*Proof.* Immediately, the binding IC constraint for the last period $N$ implies

$$q_N \lambda_N b_N - c = 0.$$

For the IC constraint in period $N - 1$, it gives us

$$q_{N-1}\lambda_{N-1} b_{N-1} - c + \delta f_1(q_{N-1}, \beta_{N-1})(q_N \lambda_N b_N - c) = \delta(q_{N-1}\lambda_{N-1} b_N - c).$$

Since the second term on the LHS is 0, it implies

$$q_{N-1}\lambda_{N-1}b_{N-1} - c = \delta(q_{N-1}\lambda_{N-1}b_N - c).$$

Suppose for all $m+1 \leq n \leq N-1$, we have $q_n\lambda_n b_n - c = \delta(q_n\lambda_n b_{n+1} - c)$. Consider the IC constraint in the period $m$,

$$\sum_{s=m}^{N-1} \delta^{s-m} f_{s-m}(q_m, \beta_m)(q_s\lambda_s b_s - c)\Delta = \sum_{s=m+1}^{N} \delta^{s-m} f_{s-m-1}(q_m, \beta_m)(q_{s-1}\lambda_{s-1}b_s - c)\Delta.$$

Note that the LHS has no terms in the period $N$ since $q_N\lambda_N b_N - c = 0$. Relabeling the terms on the LHS gives us

$$\sum_{s=m+1}^{N} \delta^{s-m-1} f_{s-m-1}(q_m, \beta_m)(q_{s-1}\lambda_{s-1}b_{s-1} - c)\Delta = \sum_{s=m+1}^{N} \delta^{s-m} f_{s-m-1}(q_m, \beta_m)(q_{s-1}\lambda_{s-1}b_s - c)\Delta.$$

Since $q_n\lambda_n b_n - c = \delta(q_n\lambda_n b_{n+1} - c)$ for $m+1 \leq n \leq N-1$, it is clear that the recursive relation remains valid for $n = m$, which concludes this lemma. □

The following observations will be useful for other proofs:

1. By iteration, for $1 \leq n \leq N$

$$b_n = \delta^{N-n}\frac{c}{q_N\lambda_N} + (1-\delta)\sum_{s=n}^{N-1} \delta^{s-n}\frac{c}{q_s\lambda_s};$$

2. For $n < N$, $q_n\lambda_n b_n > c$;

3. For $n < N$, $\delta b_{n+1} \leq b_n \leq b_{n+1}$, and both inequalities are strict when $\delta < 1$.

### B.2.3 Lemma B.3

With a slight abuse of notation, we denote $q_t$, $\lambda_t$, $\eta_t$, $\phi_t$, $f_t^\theta(q)$ and $f_t(q,\beta)$, where $\theta \in \Theta$, as the limit of $q_n$, $\lambda_n$, $\eta_n$, $\phi_n$, $f_n^\theta(q)$ and $f_n(q,\beta)$ when $\Delta \to 0$ and $n\Delta \to t$.

*Lemma* B.3. Let $T = N\Delta$. Define

$$H^{\beta_0}(N) := \sum_{n=1}^{N} \delta^n \eta_n (h - b_n)\Delta,$$

where $b_n = \delta^{N-n}\frac{c}{q_N\lambda_N} + (1-\delta)\sum_{s=n}^{N-1}\delta^{s-n}\frac{c}{q_s\lambda_s}$. Then

$$\lim_{\Delta\to 0} H^{\beta_0}(N) = H_0^{\beta_0}(T) := e^{-\rho T}(h - \frac{c}{q_T\lambda_T})\phi_T + \rho\int_0^T e^{-\rho t}(h - \frac{c}{q_t\lambda_t})\phi_t \,\mathrm{d}t.$$

In addition, $H_0^{\beta_0}(T)$ is strictly increasing in $\beta_0 \in [0,1]$ for any $T \in (0, T_*^H]$.

*Proof.* Note that

$$\sum_{n=1}^{N}\delta^n\eta_n h\Delta = \sum_{n=1}^{N}\delta^n h(\phi_n - \phi_{n-1}) = \delta^N h\phi_N + (1-\delta)\sum_{n=1}^{N-1}\delta^n h\phi_n,$$

$$\begin{aligned}\sum_{n=1}^{N}\delta^n\eta_n b_n\Delta &= \delta^N\frac{c}{q_N\lambda_N}\phi_N + (1-\delta)\sum_{n=1}^{N}\sum_{s=n}^{N-1}\delta^s\eta_n\frac{c}{q_s\lambda_s}\Delta\\ &= \delta^N\frac{c}{q_N\lambda_N}\phi_N + (1-\delta)\sum_{s=1}^{N-1}\sum_{n=1}^{s}\delta^s\eta_n\frac{c}{q_s\lambda_s}\Delta\\ &= \delta^N\frac{c}{q_N\lambda_N}\phi_N + (1-\delta)\sum_{s=1}^{N-1}\delta^s\frac{c}{q_s\lambda_s}\phi_s.\end{aligned}$$

Hence, we have

$$H^{\beta_0}(N) = \delta^N(h - \frac{c}{q_N\lambda_N})\phi_N + \frac{(1-\delta)}{\Delta}\sum_{n=1}^{N-1}\delta^n(h - \frac{c}{q_n\lambda_n})\phi_n\Delta.$$

Let $T = N\Delta$, we have

$$\lim_{\Delta\to 0} \delta^N (h - \frac{c}{q_N\lambda_N})\phi_N = e^{-\rho T}(h - \frac{c}{q_T\lambda_T})\phi_t \text{ and } \lim_{\Delta\to 0}\frac{1-\delta}{\Delta} = \rho.$$

In addition, we can show that for any $t \in [0, T]$,

$$|\delta^n (h - \frac{c}{q_n\lambda_n})\phi_n - e^{-\rho t}(h - \frac{c}{q_t\lambda_t})\phi_t| \underset{\Delta\to 0}{=} \mathcal{O}(\Delta),$$

where $t = n\Delta$. Thus,

$$\lim_{\Delta\to 0}\sum_{n=1}^{N-1} |\delta^n (h - \frac{c}{q_n\lambda_n})\phi_n - e^{-\rho t}(h - \frac{c}{q_t\lambda_t})\phi_t|\Delta = 0.$$

Hence, we have

$$\lim_{\Delta\to 0}\sum_{n=1}^{N-1} \delta^n (h - \frac{c}{q_n\lambda_n})\phi_n\Delta = \lim_{\Delta\to 0}\sum_{n=1}^{N-1} e^{-\rho t}(h - \frac{c}{q_t\lambda_t})\phi_t\Delta = \int_0^T e^{-\rho t}(h - \frac{c}{q_t\lambda_t})\phi_t \,\mathrm{d}t,$$

where the last equality comes from Riemann integral. Therefore, $\lim_{\Delta\to 0} H^{\beta_0}(N) = H_0^{\beta_0}(T)$. Moreover, we have

$$\frac{\mathrm{d}H_0^{\beta_0}(T)}{\mathrm{d}\beta_0} = -[e^{-\rho T}\frac{\mathrm{d}\frac{c}{q_T\lambda_T}}{\mathrm{d}\beta_0}\phi_T + \rho\int_0^T e^{-\rho t}\frac{\mathrm{d}\frac{c}{q_t\lambda_t}}{\mathrm{d}\beta_0}\phi_t \,\mathrm{d}t].$$

Note that for all $0 \le t \le T_*^H$,

$$\frac{\mathrm{d}q_t\lambda_t}{\mathrm{d}\beta_0} = \frac{f_t^H(q_0)f_t^L(q_0)}{f_t^2(q_0,\beta_0)}(q_t^H\lambda^H - q_t^L\lambda^L) > 0,$$

since $q_t^H\lambda^H \ge \frac{c}{h} \ge q_t^L\lambda^L$ and at least one inequality is strict. Thus, for all $0 \le t \le T_*^H$, $\frac{\mathrm{d}\frac{c}{q_t\lambda_t}}{\mathrm{d}\beta_0} < 0$. Therefore, $\frac{\mathrm{d}H_0^{\beta_0}(T)}{\mathrm{d}\beta_0} > 0$ for $T \in (0, T_*^H]$, since $\phi_t > 0$ for $t > 0$. □

Actually, we can show that when $\Delta \to 0$, $n\Delta \to t$ and $N\Delta \to T$, $b_n$ converges to

$$b_t = e^{-\rho(T-t)}\frac{c}{q_T\lambda_T} + \rho\int_t^T e^{-\rho(\tau-t)}\frac{c}{q_\tau\lambda_\tau}\,\mathrm{d}\tau,$$

and $H_0^{\beta_0}(T) = \int_0^T e^{-\rho t}\eta_t(h-b_t)\,\mathrm{d}t$. The latter can be seen by differentiating the RHS,

$$\frac{\mathrm{d}\int_0^T e^{-\rho t}\eta_t(h-b_t)\,\mathrm{d}t}{\mathrm{d}T} = e^{-\rho T}\eta_T(h-b_T) - \int_0^T e^{-\rho t}\eta_t\frac{\mathrm{d}b_t}{\mathrm{d}T}\,\mathrm{d}t = e^{-\rho T}\frac{\mathrm{d}(h-\frac{c}{q_T\lambda_T})\phi_T}{\mathrm{d}T},$$

and applying the fundamental theorem of calculus and integration by parts,

$$\begin{aligned}\int_0^T e^{-\rho t}\eta_t(h-b_t)\,\mathrm{d}t &= \int_0^T e^{-\rho t}\frac{\mathrm{d}(h-\frac{c}{q_t\lambda_t})\phi_t}{\mathrm{d}t}\,\mathrm{d}t \\ &= e^{-\rho T}(h-\frac{c}{q_T\lambda_T})\phi_T + \rho\int_0^T e^{-\rho t}(h-\frac{c}{q_t\lambda_t})\phi_t\,\mathrm{d}t.\end{aligned}$$

Hence, $H_0^{\beta_0}(T) = \int_0^T e^{-\rho t}\eta_t(h-b_t)\,\mathrm{d}t$. Moreover, we have

$$\frac{\mathrm{d}H_0^{\beta_0}(T)}{\mathrm{d}T} = e^{-\rho T}[\eta_T h + \phi_T c - cq_0\frac{(\lambda^H-\lambda^L)\lambda_0 e^{-(\lambda^L+\lambda^H)T}}{q_T\lambda_T(\beta_0\lambda^H e^{-\lambda^H T} + (1-\beta_0)\lambda^L e^{-\lambda^L T})}].$$

## B.3 Proof of Lemma 2.2 and Proposition 2.1

*Proof.* The worst contract, $C^{wt} = (N_*^H, W^{wt}, \boldsymbol{b^{wt}})$, for the $L$ type solves Program I. Since $IR_A^H(N_*^H)$ bind,

$$-W = \sum_{n=1}^{N_*^H}\delta^n f_{n-1}^H(q_0)(q_n^H\lambda^H b_n - c)\Delta,$$

and the objective function in the Program I becomes

$$-\sum_{n=1}^{N_*^H}\delta^n\eta_n(h-b_n)\Delta + V^H(N_*^H).$$

According to Lemma B.1, the solution to Program I must have binding constraints in $IC_A^H(N_*^H)$ for any $1 \leq n \leq N_*^H$, which gives us Lemma 2.2.

Furthermore, from Lemma B.2, for $1 \leq n < N_*^H$,

$$q_n^H \lambda^H b_n - c = \delta(q_n^H \lambda^H b_{n+1} - c),$$

and

$$q_{N_*^H}^H \lambda^H b_{N_*^H} - c = 0.$$

Clearly, $0 < b_n \leq b_{N_*^H} \leq h$, and $q_n^H \lambda^H b_n - c > 0$ for all $n < N_*^H$. It means that the $L$ type can obtain positive payment from the agent before experimentation, i.e. $-W^{wt} \geq 0$, and positive share of profits during experimentation, i.e. $\sum_{n=1}^{N_*^H} \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n)\Delta \geq 0$. Additionally, at least one of the two payoffs is strictly positive. If $-W^{wt} = 0$, then $N_*^H = 1$. Thus, $b_{N_*^H} = c/q_0\lambda^H < h$ and the share of profits during experimentation is strictly positive.

Therefore, by choosing the worst contract for the $L$ type, the $L$ type can obtain a strictly positive payoff, which violates her IC constraint. This gives us Proposition 2.1. □

## B.4 Proof of Lemma 2.3

*Proof.* Obviously, the $L$ type cannot generate a strictly positive payoff. When the agent believes that the principal is the $L$ type, he will either reject the contract which gives the agent a negative payoff, or accept the contract which gives the $L$ type principal a negative payoff. In either case, the $L$ type principal cannot obtain a strictly positive payoff.

Now, consider the $H$ type. We will solve the best contract for the $H$ type if the agent believes that she is a $L$ type but is still willing to accept the contract. The best contract will solve the following Program V:

$$\max_{N,W,\boldsymbol{b}} \Pi^H = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^H(q_0) q_n^H \lambda^H (h - b_n)\Delta$$

$$\text{s.t.} \quad \sum_{s=n}^{N} \delta^{s-n} f_{s-n}^L(q_n^L)(q_s^L \lambda^L b_s - c)\Delta \geq \sum_{s=n+1}^{N} \delta^{s-n} f_{s-n-1}^L(q_n^L)(q_{s-1}^L \lambda^L b_s - c)\Delta$$

$$W + \sum_{n=1}^{N} \delta^n f_{n-1}^L(q_0)(q_n^L \lambda^L b_n - c)\Delta \geq 0$$

The IR constraint for agent must bind, so the program becomes

$$\max_{N,\boldsymbol{b}} \Pi^H = \sum_{n=1}^{N} \delta^n \eta_n (h - b_n)\Delta + V^L(N)$$

$$\text{s.t.} \quad \sum_{s=n}^{N} \delta^{s-n} f_{s-n}^L(q_n^L)(q_s^L \lambda^L b_s - c)\Delta \geq \sum_{s=n+1}^{N} \delta^{s-n} f_{s-n-1}^L(q_n^L)(q_{s-1}^L \lambda^L b_s - c)\Delta$$

According to Lemma B.1 and B.2, the IC constraints must bind, and

$$b_n = \delta^{N-n} \frac{c}{q_N^L \lambda^L} + (1-\delta) \sum_{s=n}^{N-1} \delta^{s-n} \frac{c}{q_s^L \lambda^L}$$

We will show that reducing the number of experiments by 1 makes the $H$ type better off; thus, the termination date of the best contract for the $H$ type is 0 and her maximal profit is 0. Let $b_n^N$ be the bonus in period $n$, and $\Pi^H(N)$ be the profit of the $H$ type, when the termination date is $N$. We can show that for $N \geq 1$

$$b_n^{N+1} - b_n^N = \delta^{N+1-n}(\frac{c}{q_{N+1}^L \lambda^L} - \frac{c}{q_N^L \lambda^L}).$$

Furthermore, for $N \geq 0$,

$$\Pi^H(N) - \Pi^H(N+1) = \delta^{N+1}[(\frac{c}{q_{N+1}^L \lambda^L} - \frac{c}{q_N^L \lambda^L})\phi_N + f_N^H(q_0) q_{N+1}^H \lambda^H (\frac{c}{q_{N+1}^L \lambda^L} - h)\Delta] \geq 0,$$

since $\frac{c}{q_{N+1}^L \lambda^L} - \frac{c}{q_N^L \lambda^L} \geq 0$, $\phi_N \geq 0$ and $\frac{c}{q_{N+1}^L \lambda^L} \geq \frac{c}{q_0 \lambda^L} \geq h$. □

## B.5 Proof of Proposition 2.2

*Proof.* The $H$ type's equilibrium contract, $C^{sep} = \{N^{sep}, W^{sep}, \boldsymbol{b^{sep}}\}$, in the BSEH solves Program II. We will solve the relaxed program without the agent's IR constraint, and verify it later.

First, in the relaxed program, $IC_L^H(N)$ must bind, otherwise we can decrease $W$ without violating any other constraint, and obtain higher payoff.

The binding $IC_L^H(N)$ determines $W = \sum_{n=1}^N \delta^n f_{n-1}^L(q_0) q_n^L \lambda^L (h - b_n)\Delta$. Thus, the objective function becomes

$$\sum_{n=1}^N \delta^n \eta_n (h - b_n)\Delta$$

From Lemma B.1, all IC constraints should bind. Furthermore, Lemma B.2 gives an explicit characterization of the bonus scheme. In addition, Lemma B.3 gives the limit form of the above objective function. Let $\Pi_{sep}^H(T)$ be the limit of the $H$ type's payoff when her termination time $N\Delta$ goes to $T$ and $\Delta$ goes to 0. Thus, we can get[2]

$$\Pi^H_{sep}(T) = V^H_0(T) - V^L_0(T) - c\frac{\lambda^H - \lambda^L}{\lambda^H}[(1-q_0)\frac{1-e^{-(\rho+\lambda^L-\lambda^H)T}}{\rho+\lambda^L-\lambda^H} + q_0\frac{1-e^{-(\rho+\lambda^L)T}}{\rho+\lambda^L}].$$

Take the first derivative with respect to $T$,

$$\frac{\mathrm{d}\Pi^H_{sep}(T)}{\mathrm{d}T} = \frac{q_0}{\lambda^H}(\lambda^H h - c)e^{-(\rho+\lambda^L)T}[\lambda^H e^{(\lambda^L-\lambda^H)T} - (\lambda^H - \lambda^L)\frac{l^H}{l_0}e^{\lambda^H T} - \lambda^L].$$

Note that $[\lambda^H e^{(\lambda^L-\lambda^H)T} - (\lambda^H - \lambda^L)\frac{l^H}{l_0}e^{\lambda^H T} - \lambda^L]$ is strictly decreasing in $T$. At $T = 0$, it is $\frac{q_0}{\lambda^H}(\lambda^H h - c)(\lambda^H - \lambda^L)(1 - \frac{l^H}{l_0}) > 0$, and it goes to negative infinity when $T$ goes to infinity. Therefore, there exists a $T^{sep}$ such that

$$\lambda^H e^{-\lambda^H T^{sep}} - (\lambda^H - \lambda^L)\frac{l^H}{l_0}e^{(\lambda^H-\lambda^L)T^{sep}} - \lambda^L e^{-\lambda^L T^{sep}} = 0.$$

When $T < T^{sep}$, $\Pi^H_{sep}(T)$ is strictly increasing in $T$; when $T > T^{sep}$, $\Pi^H_{sep}(T)$ is strictly decreasing in $T$. $T^{sep}$ maximizes $\Pi^H_{sep}(T)$.

Moreover, since $\frac{l^H}{l_0} = e^{-\lambda^H T^H_*}$, we have

$$\lambda^H e^{-\lambda^H T^H_*} - (\lambda^H - \lambda^L)\frac{l^H}{l_0}e^{(\lambda^H-\lambda^L)T^H_*} - \lambda^L e^{-\lambda^L T^H_*} = \lambda^H(e^{-\lambda^H T^H_*} - e^{-\lambda^L T^H_*}) < 0.$$

Therefore, $T^{sep} < T^H_*$, there is always under experimentation. In addition, since $T^{sep} < T^H_*$ and $b^{sep}_t$ is increasing, we have $b^{sep}_t \leq b^{sep}_{T^{sep}} < h$. Furthermore, $W^{sep} > 0$ since all $b^{sep}_t < h$.

In addition, we can show that $T^{sep}$ is decreasing in $\lambda^L$. When $\lambda^L = 0$, $T^{sep} = \frac{1}{2}T^H_*$, and $W^{sep} = 0$. Therefore, $T^{sep} < \frac{1}{2}T^H_*$ in the main model since $\lambda^L > 0$.

[2]We assume that $\rho + \lambda^L - \lambda^H \neq 0$ so that the denominator is not 0. However, when $\rho+\lambda^L-\lambda^H = 0$, we can obtain the same results. The proofs in Section B.11 of the Appendix also apply, where the term $\rho + \lambda^L - \lambda^H$ is in the denominator.

The last thing is to check the agent's IR constraint. The agent's payoff in the limit is $V_0^H(T^{sep}) - \Pi_{sep}^H(T^{sep})$. Thus, it is equal to

$$
\begin{aligned}
&V_0^L(T^{sep}) + c\frac{\lambda^H - \lambda^L}{\lambda^H}[(1-q_0)\frac{1-e^{-(\rho+\lambda^L-\lambda^H)T^{sep}}}{\rho+\lambda^L-\lambda^H} + q_0\frac{1-e^{-(\rho+\lambda^L)T^{sep}}}{\rho+\lambda^L}] \\
=&\frac{\lambda^L}{\lambda^H}V_0^H(T^{sep}) + c\frac{\lambda^H-\lambda^L}{\lambda^H}(1-q_0)[\frac{1-e^{-(\rho+\lambda^L-\lambda^H)T^{sep}}}{\rho+\lambda^L-\lambda^H} - \frac{1-e^{-\rho T^{sep}}}{\rho}] \\
+&\frac{\lambda^L}{\lambda^H}q_0(\lambda^H h - c)[\frac{1-e^{-(\rho+\lambda^L)T^{sep}}}{\rho+\lambda^L} - \frac{1-e^{-(\rho+\lambda^H)T^{sep}}}{\rho+\lambda^H}] > 0.
\end{aligned}
$$

All terms above are strictly positive. The first term is strictly positive since $0 < T^{sep} < T_*^H$. The other two terms are strictly positive since $\frac{1-e^{-xt}}{x}$ is strictly decreasing when $t > 0$. Therefore, the solution we find is the $H$ type's equilibrium contract in the BSEH. In such an equilibrium, the agent gets rents more than $\frac{\lambda^L}{\lambda^H}V_0^H(T^{sep})$, i.e. $\frac{\lambda^L}{\lambda^H}$ proportion of the total surplus. The $H$ type principal gets payoff less than $(1-\frac{\lambda^L}{\lambda^H})V_0^H(T^{sep})$, i.e. $\frac{\lambda^H-\lambda^L}{\lambda^H}$ proportion of the total surplus. □

## B.6 Proof of Proposition 2.3

*Proof.* (a) First, we show that the BSEH survives the intuitive criterion.

For any contract $C$, let $z$ be a response of the agent, which includes both the acceptance/rejection decision and the action plan conditional on acceptance, and $BR(C, \beta)$ be the set of best responses when his belief about the principal being $H$ type is $\beta \in [0,1]$. Let $\Pi_{sep}^\theta$ be the equilibrium payoff of the $\theta \in \Theta$ type principal in the BSEH, $\Pi^\theta(C, z)$ be the payoff of the $\theta \in \Theta$ type principal when she proposes contract $C$ and the agent's response is $z$, and

$$J(C) = \{\theta \in \Theta : \Pi^{\theta}_{sep} > \max_{z \in \bigcup_{\beta \in [0,1]} BR(C,\beta)} \Pi^{\theta}(C, z)\}$$

be the set of types whose payoffs from deviating to contract $C$ is strictly less than their equilibrium payoffs in the BSEH, for any beliefs that the agent could possibly form. intuitive criterion allows the agent to form beliefs about the types only on $\Theta \backslash J(C)$. When $J(C) = \Theta$, the agent could form any beliefs.

For a contract $C$ that the agent could "reasonably" form belief about the principal being type $L$, i.e $J(C) = \Theta$, $J(C) = \varnothing$ or $J(C) = \{H\}$, the minimal payoff the principal could get is no more than 0, according to Lemma 2.3. Thus, the BSEH cannot fail the intuitive criterion for such contract $C$, since both types could obtain non-negative payoffs in the BSEH and they would not deviate to $C$.

For a contract $C$ that the agent could only "reasonably" form belief about the principal being type $H$, i.e $J(C) = \{L\}$, then the $L$ type could not get positive payoff from deviating to contract $C$ even when the agent believes her type is $H$. However, by definition, the BSEH already gives the highest payoff to the $H$ type, subjected to the $L$ type cannot get strictly positive payoff by proposing the same contract and making the agent believe her type is $H$. Hence, the $H$ type would not deviate to $C$ either. Thus, the BSEH cannot fail the intuitive criterion for such contract $C$.

Therefore, the BSEH must survive the intuitive criterion.

(b) Second, we show that any equilibria that give less equilibrium payoff for the $H$ type than the BSEH fail the intuitive criterion.

For any such equilibrium, let $\bar{\Pi}^{\theta}$ be the equilibrium payoff for $\theta \in \Theta$ type principal. Then we have $\Pi^{H}_{sep} > \bar{\Pi}^{H}$, and $\bar{\Pi}^{L} \geq 0$. Let $\bar{\epsilon} \in (0, \Pi^{H}_{sep} - \bar{\Pi}^{H})$. Consider the following contract $\bar{C}^{H} = \{T^{sep}, W^{sep} + \bar{\epsilon}, \boldsymbol{b^{sep}}\}$. The agent will always accept this contract since $W^{sep} + \bar{\epsilon} > 0$, given any belief about the principal. In addition, for any type of the principal, to experiment in every period is the best action plan for the principal, since the bonus scheme $\boldsymbol{b^{sep}}$ is an increasing sharing rule.

However, the $L$ type will get strictly negative payoff from $\bar{C}^{H}$ even in this best case. Thus, the $L$ type principal would never deviate to that contract $\bar{C}^{H}$, i.e. $L \in J(\bar{C}^{H})$.

On the other hand, the $H$ type can obtain a higher payoff $\Pi^{H}_{sep} - \bar{\epsilon} > \bar{\Pi}^{H}$ if the agent accepts the contract $\bar{C}^{H}$ and works until the termination date, which is the best response of the agent when he believes her type is $H$. Thus, the $H$ type might deviate to the contract $\bar{C}^{H}$, i.e. $H \notin J(\bar{C}^{H})$.

Hence, the agent would believe that it is the $H$ type who deviates to $\bar{C}^{H}$, and the $H$ type can obtain a higher payoff than $\bar{\Pi}^{H}$, given such belief. Therefore, the equilibrium with $\bar{\Pi}^{H} < \Pi^{H}_{sep}$ fails the intuitive criterion. □

## B.7 Proof of Proposition 2.4

*Proof.* The equilibrium contract, $C^{pl} = \{N^{pl}, W^{pl}, \boldsymbol{b^{pl}}\}$, in the BPEH solves Program III. We solve the relaxed program without $IR_L(N)$, and verify it later. In the relaxed program, $IR_A(N)$ must bind, otherwise we can decrease $W$ without violating any other constraint, and obtain higher payoff.

Given the binding $IR_A(N)$, we can determine $W$. Take it into the objective function, then the $H$ type's payoff is

$$(1-\beta_0)\sum_{n=1}^{N}\delta^n\eta_n(h-b_n)\Delta + V(N)$$

From Lemma B.1, all IC constraints should bind. Furthermore, Lemma B.2 gives an explicit characterization of the bonus scheme. In addition, Lemma B.3 gives the limit form of the first part of the objective function. Let $\Pi_{pl}^H(T)$ denote the limit of the $H$ type's payoff when her termination time $N\Delta$ goes to $T$ and $\Delta$ goes to 0. Then,

$$\Pi_{pl}^H(T) = (1-\beta_0)H_0^{\beta_0}(T) + \beta_0 V_0^H(T) + (1-\beta_0)V_0^L(T).$$

Therefore, taking the first derivative with respect to $T$, $\frac{\mathrm{d}\Pi_{pl}^H(T)}{\mathrm{d}T}$ is equal to

$$e^{-\rho T}[-(1-\beta_0)q_0c\frac{(\lambda^H-\lambda^L)\lambda_0 e^{-(\lambda^L+\lambda^H)T}}{q_T\lambda_T(\beta_0\lambda^H e^{-\lambda^H T}+(1-\beta_0)\lambda^L e^{-\lambda^L T})}+q_0(\lambda^H h-c)e^{-\lambda^H T}-(1-q_0)c].$$

Note that $\frac{\mathrm{d}\Pi_{pl}^H(T)}{\mathrm{d}T} > 0$ is equivalent to

$$q_0(\lambda^H h-c) > (1-\beta_0)q_0c\frac{(\lambda^H-\lambda^L)\lambda_0}{q_T\lambda_T(\beta_0\lambda^H e^{(\lambda^L-\lambda^H)T}+(1-\beta_0)\lambda^L)} + (1-q_0)ce^{\lambda^H T}.$$

Clearly, the RHS of above inequality is strictly increasing in $T$, and goes to infinity when $T$ goes to infinity. In addition, when $T = 0$, the RHS equals

$$\begin{aligned}(1-\beta_0)q_0(\lambda^H-\lambda^L)\frac{c}{q_0\lambda_0}+(1-q_0)c &< (1-\beta_0)q_0(\lambda^H-\lambda^L)h+(1-q_0)c\\ &= c-q_0\lambda_0 h+q_0(\lambda^H h-c)\\ &< q_0(\lambda^H h-c),\end{aligned}$$

where both inequities come from the fact that $q_0\lambda_0 h > c$. Hence, there exits a $T^{pl} > 0$, such that $\frac{\mathrm{d}\Pi^H_{pl}(T)}{\mathrm{d}T}|_{T=T^{pl}} = 0$. When $T < T^{pl}$, $\Pi^H_{pl}(T)$ is strictly increasing; when $T > T^{pl}$, $\Pi^H_{pl}(T)$ is strictly decreasing. $T^{pl}$ maximizes $\Pi^H_{pl}(T)$.

To show $T^{pl} < T_*$, we show that $\frac{\mathrm{d}\Pi^H_{pl}(T)}{\mathrm{d}T}|_{T=T_*} < 0$. Since $q_{T_*}\lambda_{T_*}h = c$, we have

$$\beta_0\frac{l_0}{l^H}e^{-\lambda^H T_*}+(1-\beta_0)\frac{l_0}{l^L}e^{-\lambda^L T_*} = 1.$$

Thus, $\frac{\mathrm{d}\Pi^H_{pl}(T)}{\mathrm{d}T}|_{T=T_*}$ is equal to

$$e^{-\rho T}(1-\beta_0)q_0(e^{-\lambda^H T_*}-e^{-\lambda^L T_*})\frac{[\beta_0(\lambda^H-\lambda^L)(\lambda^H h-c)e^{-\lambda^H T_*}+\lambda^L\frac{c}{l_0}]}{(\beta_0\lambda^H e^{-\lambda^H T_*}+(1-\beta_0)\lambda^L e^{-\lambda^L T_*})} < 0.$$

Furthermore, $T^{pl} < T_*$ implies $b_{T^{pl}} < h$, since $q_{T^{pl}}\lambda_{T^{pl}}b_{T^{pl}} = q_{T_*}\lambda_{T_*}h = c$.

Let us explore $W^{pl}$. Note that the agent's expected payoff is 0, i.e.

$$W^{pl}+\int_0^{T^{pl}} e^{-\rho t}f_t(q_0,\beta_0)(q_t\lambda_t b_t-c)\,\mathrm{d}t = 0,$$

and $q_t\lambda_t b_t - c > 0$ for all $t < T^{pl}$. Obviously, $W^{pl} < 0$. This also means the $L$ type must obtain strictly positive expected payoff in the equilibrium, because her payoff before experimentation is strictly positive, i.e. $-W^{pl} > 0$, and her payoff during the experimentation is also strictly positive since $b_t < h$ for all $t$. Therefore, the $L$ type's IR constraint is satisfied, and the contract we solved is indeed the best pooling equilibrium for the $H$ type.

At last, we will show that $\frac{\mathrm{d}\Pi_{pl}^H(T^{pl})}{\mathrm{d}\beta_0} > 0$. Note that

$$\frac{\partial \Pi_{pl}^H(T)}{\partial \beta_0} = (1-\beta_0)\frac{\partial H_0^{\beta_0}(T)}{\partial \beta_0} + V_0^H(T) - V_0^L(T) - H_0^{\beta_0}(T).$$

From Lemma B.3, for $0 < T \le T_*^H$, $(1-\beta_0)\frac{\partial H_0^{\beta_0}(T)}{\partial \beta_0} > 0$, and

$$\begin{aligned} V_0^H(T) - V_0^L(T) - H_0^{\beta_0}(T) &= \int_0^T e^{-\rho t}(\eta_t h + \phi_t c)\,\mathrm{d}t - \int_0^T e^{-\rho t}\eta_t(h - b_t)\,\mathrm{d}t \\ &= \int_0^T e^{-\rho t}(\eta_t b_t + \phi_t c)\,\mathrm{d}t \\ &= e^{-\rho T}\frac{c}{q_T\lambda_T}\phi_T + \rho\int_0^T e^{-\rho t}\frac{c}{q_t\lambda_t}\phi_t\,\mathrm{d}t + \int_0^T e^{-\rho t}\phi_t c\,\mathrm{d}t > 0. \end{aligned}$$

Hence, $\frac{\partial \Pi_{pl}^H(T)}{\partial \beta_0} > 0$ for $0 < T \le T_*^H$. Therefore, according to the Envelope Theorem, $\frac{\mathrm{d}\Pi_{pl}^H(T^{pl})}{\mathrm{d}\beta_0} = \frac{\partial \Pi_{pl}^H(T^{pl})}{\partial \beta_0} > 0$ since $0 < T^{pl} < T_* < T_*^H$. Clearly, when $\beta_0 \to 1$, $T^{pl} \to T_*^H$, and $\Pi_{pl}^H(T^{pl}) \to V_0^H(T_*^H)$. □

## B.8 Proof of Proposition 2.5

*Proof.* Consider the case where $q_0\lambda_0 h > c$, since otherwise the $H$ type must obtain 0 payoff in any pooling equilibria. Let $\Pi_{pl}^\theta$ be the equilibrium payoff of the $\theta \in \Theta$ type principal in the BPEH. Clearly, $\Pi_{pl}^H < V_0^H(T^{pl})$. Let $\check{\epsilon} = V_0^H(T^{pl}) - \Pi_{pl}^H > 0$.

According to Lemma B.3, since $0 < T^{pl} < T_* < T_*^H$, we have

$$\Pi_{pl}^H - \Pi_{pl}^L = \int_0^{T^{pl}} e^{-\rho t}\eta_t(h - b_t^{pl})\,\mathrm{d}t = H_0^{\beta_0}(T^{pl}) < H_0^1(T^{pl}) = \int_0^{T^{pl}} e^{-\rho t}\eta_t(h - \mathring{b}_t)\,\mathrm{d}t,$$

where $\mathring{\boldsymbol{b}}$ are defined by the binding IC constraints for the agent when he believes the principal is the $H$ type, i.e. $\mathring{b}_t = e^{-\rho(T^{pl}-t)} \frac{c}{q^H_{T^{pl}} \lambda^H} + \rho \int_t^{T^{pl}} e^{-\rho(s-t)} \frac{c}{q^H_s \lambda^H} \,\mathrm{d}s$.

Let $\acute{\epsilon} = H_0^1(T^{pl}) - H_0^{\beta_0}(T^{pl}) > 0$, $\mathring{\epsilon} = 1/2 \min\{\acute{\epsilon}, \check{\epsilon}\} > 0$, and $\mathring{W} = \int_0^{T^{pl}} e^{-\rho t} f_t^H(q_0) q_t^H \lambda^H (h - \mathring{b}_t) \,\mathrm{d}t - \Pi^H_{pl} - \mathring{\epsilon}$.

Consider the contract $\mathring{C} = \{T^{pl}, \mathring{W}, \mathring{\boldsymbol{b}}\}$. If the agent believes that it is the $H$ type who proposes $\mathring{C}$, then clearly his IC constraints are satisfied. Moreover, if he accepts the offer, he will obtain a positive payoff, i.e.

$$\mathring{W} + \int_0^{T^{pl}} e^{-\rho t} f_t^H(q_0)(q_t^H \lambda^H \mathring{b}_t - c) \,\mathrm{d}t = V_0^H(T^{pl}) - \Pi^H_{pl} - \mathring{\epsilon} = \check{\epsilon} - \mathring{\epsilon} > 0.$$

Hence, his IR constraint is also satisfied.

When the $H$ type proposes $\mathring{C}$, and the agent believes that she is the $H$ type, her payoff would be higher than $\Pi^H_{pl}$, i.e.

$$\mathring{\Pi}^H := -\mathring{W} + \int_0^{T^{pl}} e^{-\rho t} f_t^H(q_0) q_t^H \lambda^H (h - \mathring{b}_t) \,\mathrm{d}t = \Pi^H_{pl} + \mathring{\epsilon} > \Pi^H_{pl}.$$

When the $L$ type proposes $\mathring{C}$ and the agent accepts the offer, the best action plan for the $L$ type is to experiment on the project until the termination date $T^{pl}$, since the bonus payment is an increasing sharing scheme. Nevertheless, the $L$ type's payoff would be less than $\Pi^L_{pl}$, i.e.

$$\mathring{\Pi}^L := -\mathring{W} + \int_0^{T^{pl}} e^{-\rho t} f_t^L(q_0) q_t^L \lambda^L (h - \mathring{b}_t) \,\mathrm{d}t = \Pi^L_{pl} + \mathring{\epsilon} - \acute{\epsilon} < \Pi^L_{pl}.$$

If the agent rejects the offer, the $L$ type would obtain 0 payoff, which is still strictly less than $\Pi^L_{pl}$.

Thus, the only possible type who proposes $\mathring{C}$ must be the $H$ type, and the $H$ type could obtain a higher payoff from $\mathring{C}$ if the agent believes she is the $H$ type. Therefore, BPEH fails the intuitive criterion. □

## B.9 Proof of Lemma 2.4

*Proof.* Fix any equilibrium $E$ that gives the $L$ type strictly positive payoff. Let $CL$ be the set of equilibrium contracts chosen by the $L$ type, and $CH$ be the set of equilibrium contracts chosen by the $H$ type. Then $CL \subseteq CH$, and the $L$ type can obtain the same payoff by proposing any $C \in CL$, the $H$ type can obtain the same payoff by proposing any $C \in CH$.

For any contract $C \in CL$, let $x^C \in (0, 1]$ be the probability that the $L$ type chooses $C$ in the equilibrium, and $y^C \in (0, 1]$ be the probability that the $H$ type chooses $C$ in the equilibrium.[3] Then $\sum_{C \in CL} y^C \leq \sum_{C \in CL} x^C = 1$.

Clearly, there are only two possible cases.

- $CL = CH$, and $x^C = y^C$ for all $C \in CL = CH$;
- There is at least one contract $C$, such that $x^C > y^C$.

In the first case, for any $C \in CL = CH$, when the principal proposes the contract $C$, the agent's belief about the principal's type is just the prior $\beta_0$. In addition, given this belief and the contract $C$, we know that the $H$'s

[3]For technical simplicity, we assume that every contract over which one principal randomizes, she chooses that contract with strictly positive probability.

payoff in such an equilibrium cannot exceed the payoff she can obtain in the BPEH when the prior is $\beta_0$, by definition.

In the second case, for the contract $C$ such that $x^C > y^C$, the agent's belief about the principal' type $\beta$ is strictly below the prior $\beta_0$ by Bayesian rule. Given such belief and the contract $C$, the $H$ type's payoff cannot exceed the payoff she can obtain in the BPEH when the prior is $\beta$, by definition. From Proposition 2.4, the $H$ type's payoff in the BPEH is increasing in the prior. Thus, we can conclude that the $H$ type's payoff also cannot exceed the payoff she can obtain in the BPEH when the prior is $\beta_0$. □

## B.10 Proof of Lemma 2.5

*Proof.* For the equilibrium such that the $L$ type gets zero payoff, the best contract for the $H$ type, $\hat{C} = \{\hat{N}, \hat{W}, \hat{\boldsymbol{b}}\}$, when the agent's belief about the $H$ type is $\beta_0$ solves the following Program VI:

$$\max_{N,W,\boldsymbol{b}} \Pi^H = -W + \sum_{n=1}^{N} \delta^n f_{n-1}^H(q_0) q_n^H \lambda^H (h - b_n)\Delta$$
$$\text{s.t. } IC_A(N) \ , \ IR_A(N) \ , \ IC_L^H(N)$$

As in the BPEH, when $\beta_0$ is small, i.e. $\frac{c}{q_0\lambda_0} \geq h$, the optimal termination time is 0, and the payoff of the $H$ type is 0. Now, consider the non-trivial case when $\frac{c}{q_0\lambda_0} < h$.

We can still drop the agent's IR constraint, and solve the relaxed program, then verify it later. However, for the purpose to prove this lemma, we don't have to. The reason is that the solution to the relaxed program cannot be smaller than that to the original program, and observe that when $\beta_0 = 1$, the relaxed program is the same program solves the BSEH. What we need to show is just the value function for this relaxed program is increasing in $\beta_0$.

Consider the relaxed program without $IR_A(N)$. Clearly, $IC_L^H(N)$ must bind. Solve for $W$ and take it into the objective function, which is

$$\sum_{n=1}^{N} \delta^n \eta_n (h - b_n)\Delta.$$

By Lemma B.1 and B.2, the agent's IC constraints in $IC_A(N)$ must bind. Let $\hat{\Pi}^H(T)$ denote the limit of the $H$ type's payoff when the termination time $N\Delta$ goes to $T$ and $\Delta$ goes to 0. From Lemma B.3,

$$\hat{\Pi}^H(T) = e^{-\rho T}\phi_T(h - \frac{c}{q_T\lambda_T}) + \rho \int_0^T e^{-\rho t}\phi_t(h - \frac{c}{q_t\lambda_t})\,\mathrm{d}t.$$

For any $T > T^*$, since $q_T\lambda_T h < q_{T_*}\lambda_{T_*}h = c$, we have

$$\hat{\Pi}^H(T^*) - \hat{\Pi}^H(T) = -e^{-\rho T}\phi_T(h - \frac{c}{q_T\lambda_T}) - \rho \int_{T^*}^T e^{-\rho t}\phi_t(h - \frac{c}{q_t\lambda_t})\,\mathrm{d}t > 0.$$

Hence, the principal never wants to extend the project beyond the efficient termination time of the mixed project; we can focus solutions in the compact set $[0, T_*]$. Since the objective function is continuous, there exists a solution $\hat{T} \in [0, T_*]$ to the relaxed program. We can further show that $\hat{T} \in (0, T_*)$ by showing $\frac{\mathrm{d}\hat{\Pi}^H(T)}{\mathrm{d}T}|_{T=0} > 0$ and $\frac{\mathrm{d}\hat{\Pi}^H(T)}{\mathrm{d}T}|_{T=T_*} < 0$.

Lemma B.3 also shows $\frac{\partial\hat{\Pi}^H(T)}{\partial\beta_0} > 0$ for $0 < T \leq T_*^H$. Thus, Envelope Theorem implies $\frac{\mathrm{d}\hat{\Pi}^H(\hat{T})}{\mathrm{d}\beta_0} = \frac{\partial\hat{\Pi}^H(\hat{T})}{\partial\beta_0} > 0$, since $0 < \hat{T} < T_* \leq T_*^H$.

Hence, the value function of the relaxed program is 0 for $\frac{c}{q_0\lambda_0} \leq h$, and then strictly increasing in $\beta_0$ for $\frac{c}{q_0\lambda_0} > h$. Note that when $\beta_0 = 1$, it is the equilibrium payoff for the $H$ type in the BSEH. Therefore, all the equilibrium that gives the $L$ type 0 payoff, cannot give the $H$ type higher payoff than the BSEH. □

## B.11 Proof of Proposition 2.7

In this section, we find the optimal mechanism for the $H$ type principal when $\Delta$ is small. First, we restrict to feasible mechanisms when $N^L = 0$. Then, we focus on feasible mechanisms when $N^L \geq 1$. Last, we compare them when $\Delta$ is small.

*Proof.* (a) Consider the optimal mechanism for the $H$ type when $N^L = 0$, $\mathcal{M}^0 = \{\mathcal{C}^{H0}, \mathcal{C}^{L0}\}$. It will solve the following Program IV:

$$\max_{\mathcal{M}} \Pi^H(\mathcal{M}) = -W^H + q_0 \sum_{n=1}^{N^H} \delta^n (1 - \lambda^H \Delta)^{n-1} \lambda^H (h - b_n^H) \Delta$$

s.t. $IR_H(\mathcal{M})$ , $IR_L(\mathcal{M})$ , $IR_A(\mathcal{M})$ , $IC_H^L(\mathcal{M})$ , $IC_L^H(\mathcal{M})$ , $IC_A(\mathcal{M})$ , and $N^L = 0$

Clearly, $IR_H(\mathcal{M})$ is automatically satisfied since we maximize the $H$ type's payoff. We drop the constraints $IR_L(\mathcal{M})$ and $IC_H^L(\mathcal{M})$, and verify them later. In the relaxed program, $IC_L^H(\mathcal{M})$ and $IR_A(\mathcal{M})$ must bind, otherwise we can adjust $W^H$ and $W^L$ slightly to increase the payoff of the $H$ type.

From the binding $IC_L^H(\mathcal{M})$ and $IR_A(\mathcal{M})$, we can obtain $W^H$ and $W^L$, which in turn pins down

$$\Pi^H = \beta_0 V^H(N^H) + (1-\beta_0)\sum_{n=1}^{N^H}\delta^n\eta_n(h - b_n^H)\Delta,$$

$$\Pi^L = \beta_0 V^H(N^H) - \beta_0\sum_{n=1}^{N^H}\delta^n\eta_n(h - b_n^H)\Delta.$$

For any given $N^H$, according to Lemma B.1 and B.2, the agent's IC constraints in $IC_A(\mathcal{M})$ must also bind, and we can obtain $b_n^H$ for $1 \leq n \leq N^H$.

Let $\Pi_0^\theta(T)$ denote the limit of the $\theta \in \Theta$ type's payoff when $\Delta$ goes to 0, and the termination time of the $H$ type $N^H\Delta$ goes to $T$, where the subscript 0 is reminiscent of the $L$ type's experimenting time. Thus,

$$\Pi_0^H(T) = V_0^H(T) - (1-\beta_0)V_0^L(T) - (1-\beta_0)c\frac{\lambda^H - \lambda^L}{\lambda^H}[(1-q_0)\frac{1 - e^{-(\rho+\lambda^L-\lambda^H)T}}{\rho+\lambda^L-\lambda^H} + q_0\frac{1-e^{-(\rho+\lambda^L)T}}{\rho+\lambda^L}],$$

$$\Pi_0^L(T) = \beta_0 V_0^L(T) + \beta_0 c\frac{\lambda^H - \lambda^L}{\lambda^H}[(1-q_0)\frac{1 - e^{-(\rho+\lambda^L-\lambda^H)T}}{\rho+\lambda^L-\lambda^H} + q_0\frac{1-e^{-(\rho+\lambda^L)T}}{\rho+\lambda^L}].$$

The last thing is to find the optimal termination time for the $H$ type. Take the derivative with respect to $T$, we have

$$\frac{\mathrm{d}\Pi_0^H(T)}{\mathrm{d}T} = (1-\beta_0)\frac{q_0}{\lambda^H}(\lambda^H h - c)e^{-(\rho+\lambda^L)T}\alpha(T),$$

where $\alpha(T) := \frac{1}{1-\beta_0}\lambda^H e^{(\lambda^L-\lambda^H)T} - \frac{\beta_0}{1-\beta_0}\lambda^H\frac{l^H}{l_0}e^{\lambda^L T} - (\lambda^H - \lambda^L)\frac{l^H}{l_0}e^{\lambda^H T} - \lambda^L$ is strictly decreasing in $T$. When $T = 0$, it is $\alpha(0) = (\frac{1}{1-\beta_0}\lambda^H - \lambda^L)(1 - \frac{l^H}{l_0}) > 0$. When $T \to \infty$, it goes to $-\lambda^L$. Hence, there exists a unique solution $T^{H0}$ to the relaxed program, which satisfies $\alpha(T^{H0}) = 0$. We can also show that $\alpha(T^{sep}) > 0 > \alpha(T_*^H)$, thus $T^{sep} < T^{H0} < T_*^H$.

Now, let us check $IR_L(\mathcal{M})$. Note that

$$\frac{\mathrm{d}\Pi_0^L(T)}{\mathrm{d}T} = \beta_0 q_0(\lambda^H h - c)e^{-\rho T}[\frac{\lambda^L}{\lambda^H}e^{-\lambda^L T}(1 - \frac{l^H}{l_0}e^{\lambda^H T}) + \frac{l^H}{l_0}(e^{-(\lambda^L-\lambda^H)T} - 1)].$$

When $T < T_*^H$, $e^{\lambda^H T} < e^{\lambda^H T_*^H} = \frac{l_0}{l^H}$. Hence, $1 - \frac{l^H}{l_0}e^{\lambda^H T} > 0$ for $T < T_*^H$. We also have $e^{-(\lambda^L-\lambda^H)T} - 1 > 0$ for $T > 0$. Thus, $\frac{\mathrm{d}\Pi_0^L(T)}{\mathrm{d}T} > 0$ for $0 \le T \le T_*^H$, i.e. $\Pi_0^L(T)$ is strictly increasing in $T$. Hence, $\Pi_0^L(T^{H0}) > \Pi_0^L(0) = 0$, and $IR_L(\mathcal{M})$ is slack.

We also need to check $IC_H^L(\mathcal{M})$. By choosing $\mathcal{C}^{L0}$, the $H$ type will get $\Pi_0^L(T^{H0})$, thus deviating to $\mathcal{C}^{L0}$ will get

$$\Pi_0^L(T^{H0}) - \Pi_0^H(T^{H0}) = -\int_0^{T^{H0}} e^{-\rho t}\eta_t(h - b_t^H)\,\mathrm{d}t < 0,$$

since $\int_0^{T^{H0}} e^{-\rho t}\eta_t(h - b_t^H)\,\mathrm{d}t = H_0^1(T^{H0}) > 0$. Thus, $IC_H^L(\mathcal{M})$ is also slack.

Therefore, the solution to the relaxed program is the optimal mechanism for the $H$ type when the $L$ type does not experiment. Moreover, $\Pi_0^L(T^{H0}) > 0$ implies $\Pi_0^H(T^{H0}) < V_0^H(T^{H0}) < V_0^H(T_*^H)$. Hence, the payoff of the $H$ type is strictly less than her FIB payoff in the limit $\Delta \to 0$.

(b) Consider the optimal mechanism for the $H$ type when $N^L \ge 1$. As we have shown, for a small $\Delta$, the payoff of the $H$ type has an upper bound $V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}V^L(1) > 0$ when $N^L \ge 1$. We now show the following mechanism $\mathcal{M}^1$ is feasible and achieves this upper bound.

**Recommendations** - The mediator recommends the agent working until $N_*^H$ if the principal reports $H$, but working once if the principal reports $L$.

**Bonus schemes** - Let $\{b_n^H\}_{2\le n\le N_*^H}$ be the bonus transfers such that satisfy the agent's binding IC constraints[4], i.e. for $2 \le n \le N_*^H$,

$$b_n^H = \delta^{N_*^H - n}\frac{c}{q_{N_*^H}^H \lambda^H} + (1-\delta)\sum_{s=n}^{N_*^H-1}\delta^{s-n}\frac{c}{q_s^H\lambda^H}.$$

Fix $\{b_n^H\}_{2\le n\le N_*^H}$, the inequality (2.7) determines an upper bound for $b_1^H$. We choose some $b_1^H$ which is below the upper bound. Fix the chosen $\{b_n^H\}_{1\le n\le N_*^H}$, the inequality (2.6) determines a lower bound for $b_1^L$. Now, choose some $b_1^L$ which is above both the lower bound and $h$. Therefore, the above bonus scheme satisfies both the agent and the $L$ type's IC constraints. Moreover, because $b_1^L \ge h$, if the $H$ type reports she is $L$, she will obtain negative payoff:

$$-W^L + \delta q_0 \lambda^H (h - b_1^L)\Delta = \delta q_0(\lambda^H - \lambda^L)(h - b_1^L)\Delta \le 0.$$

Thus, the above bonus scheme also satisfies the IC constraint of the $H$ type.

**Lump-sum transfers** - Give the chosen bonus schemes, let $W^\theta$ for $\theta \in \Theta$ satisfy equation (2.4) and (2.5) to make the $L$ type and the agent's obtain 0 payoffs. At last, the $H$ type's IR constraint is automatically satisfied.

Hence, $\mathcal{M}^1$ is feasible. Clearly, the $H$ type could achieve her payoff upper bound $V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}V^L(1)$ from $\mathcal{M}^1$, since all the surplus is retained by the $H$ type. When $\Delta \to 0$, her payoff converges to $V_0^H(T_*^H)$, since $\lim_{\Delta\to 0} V^L(1) = 0$.

[4]Whether the IC constraints for the agent bind or not does not affect the result here. However, binding IC constraints are robust in the sense that they also provide the proper incentives for the agent to truthfully report a success when a success cannot be observed by the principal.

(c) We have shown that in the limit $\Delta \to 0$, the $H$ type's payoff in $\mathcal{M}^0$ is strictly below $V_0^H(T_*^H)$, while her payoff in $\mathcal{M}^1$ converges to $V_0^H(T_*^H)$. Therefore, when $\Delta$ is small enough, $\mathcal{M}^1$ is the optimal mechanism for the $H$ type. □

## B.12 Implementation of the FIB

In this section, we examine the implementation problem of the FIB for any fixed $\Delta > 0$. We also allows that the mediator randomizes over recommendations to the agent. Remember, the FIB is that both types of projects are conducted efficiently, and the principal obtains all the surplus of her project. Obviously, Proposition 2.1 implies that the FIB cannot be fully implemented by any mechanism. However, we can show that it can be virtually implemented through random recommendations.

*Proposition* B.1. Fix any $\Delta > 0$, the FIB can be virtually implemented.

*Proof.* Consider the following mechanism $\mathcal{M}^\mu$.

**Recommendations** - The agent is recommended to work until $N_*^H$ if the principal reports $H$. He is recommended to work with probability $\mu$ (and not to work with probability $1 - \mu$) in period 1, where $\mu$ is small but strictly positive, and not to work thereafter, if the principal reports $L$. We have $V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}\mu V^L(1) > 0$ when $\mu$ is close to 0.

**Lump-sum transfers** - Now the $L$ type's lump-sum transfer can also depend on the random event - whether the mediator recommends that the agent work in the first period. Let $W_R^L$ and $W_{NR}^L$ be the lump-sum transfers when the mediator does recommend and does not recommend working in the first period, respectively. Choose the lump-sum transfers as $W_{NR}^L = 0$, $W_R^L = \delta q_0 \lambda^L (h - b_1^L)\Delta$, and

$$W^H = -[V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}\mu V^L(1)] + \sum_{n=1}^{N_*^H} \delta^n f_{n-1}^H(q_0) q_n^H \lambda^H (h - b_n^H)\Delta.$$

Those lump-sum transfers gives the agent and $L$ type binding IR constraints. The $H$ type's IR constraint is automatically satisfied.

**Bonus schemes** - Let $\{b_n^H\}_{2\le n\le N_*^H}$ be the bonus transfers such that satisfy the agent's binding IC constraints for $2 \le n \le N_*^H$. Fix $\{b_n^H\}_{2\le n\le N_*^H}$, the $L$ type's IC constraint is

$$V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}\mu V^L(1) - \sum_{n=1}^{N_*^H} \delta^n \eta_n (h - b_n^H)\Delta \le 0.$$

The above inequality determines an upper bound for $b_1^H$. We choose some $b_1^H$ which is below the upper bound.

Fix the chosen $\{b_n^H\}_{1\le n\le N_*^H}$, the agent's IC constraint for $n = 1$ conditional on being recommended to work is

$$q_0(\beta_0\lambda^H b_1^H + (1-\beta_0)\mu\lambda^L b_1^L) - c \ge \beta_0 q_0 \lambda^H \sum_{s=2}^{N_*^H} \delta^{s-1}(1-\lambda^H\Delta)^{s-2}(\lambda^H b_s^H - c)\Delta.$$

The above inequality determines a lower bound for $b_1^L$. We choose some $b_1^L$ which is above both the lower bound and $h$. Therefore, the above bonus scheme

Table B.1: Comparison between FIB and $\mathcal{M}^\mu$ for the $H$ type principal

| | Termination date | $P$'s expected payoff | $A$'s expected payoff |
|---|---|---|---|
| FIB | $N_*^H$ | $V^H(N_*^H)$ | 0 |
| $\mathcal{M}^\mu$ | $N_*^H$ | $V^H(N_*^H) + \frac{1-\beta_0}{\beta_0}\mu V^L(1)$ | $-\frac{1-\beta_0}{\beta_0}\mu V^L(1)$ |

Table B.2: Comparison between FIB and $\mathcal{M}^\mu$ for the $L$ type principal

| | Termination date | $P$'s expected payoff | $A$'s expected payoff | |
|---|---|---|---|---|
| FIB | 0 | 0 | 0 | |
| $\mathcal{M}^\mu$ | 0 | 0 | 0 | w/ prob. $1-\mu$ |
| | 1 | 0 | $V^L(1)$ | w/ prob. $\mu$ |

satisfies both the agent and the $L$ type's IC constraints. If the $H$ type reports she is $L$, then she will get

$$\mu[-W_R^L + \delta q_0 \lambda^H (h - b_1^L)\Delta] = \mu \delta q_0 (\lambda^H - \lambda^L)(h - b_1^L)\Delta \leq 0.$$

Therefore, the above bonus scheme also satisfies the $H$ type's IC constraint.

Hence, the above mechanism $\mathcal{M}^\mu$ is feasible. Table B.1 and B.2 compare $\mathcal{M}^\mu$ with the FIB, regarding to the termination date of the project, the expected payoff of the principal, and that of the agent. Obviously, the outcome of $\mathcal{M}^\mu$ and the FIB can be arbitrary close as $\mu$ goes to 0. Thus, $\mathcal{M}^\mu$ virtually implements the FIB. □

# Appendix C

# Appendix for Chapter 3

## C.1 Proof of Lemma 3.2

Given the martingale and the evolution of $g_{it}$, we have

$$(p_0 \dot{-} g_{it}) = -\dot{g}_{it} = -\lambda^g \mu_{it}(p_0 - g_{it}).$$

Comparing with $\dot{l}_{it}$, we have

$$\frac{\dot{l}_{it}}{(p_0 \dot{-} g_{it})} = \Delta \frac{l_{it}}{p_0 - g_{it}}.$$

Simplify that, we have

$$\mathrm{d}ln l_{it} = \Delta \mathrm{d}ln(p_0 - g_{it}).$$

Integrate that, we have

$$ln(\frac{l_{it}}{l_0}) = \Delta ln(\frac{p_0 - g_{it}}{p_0 - 0}).$$

Since $\Delta > 0$,

$$g_{it} = p_0(1 - (\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}).$$

From the martingale property,

$$b_{it} = 1 - g_{it} - \frac{p_0 - g_{it}}{p_{it}}$$
$$= 1 - g_{it} - \frac{p_0(\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}}{\frac{l_{it}}{1+l_{it}}}$$
$$= 1 - p_0 - p_0(\frac{l_{it}}{l_0})^{\frac{1}{\Delta}}\frac{1}{l_{it}}$$
$$= (1 - p_0)(1 - (\frac{l_{it}}{l_0})^{\frac{1}{\Delta}-1}).$$

## C.2 Details about Assumption 3.1

Given Lemma 3.3, Assumption 3.1 becomes

$$\frac{g_{it} + (1 - g_{it} - b_{it})\alpha_{ijt}p_{it}}{g_{it} + (1 - g_{it} - b_{it})\alpha_{ijt}} \geq \frac{(1 - g_{it} - b_{it})(1 - \alpha_{ijt})p_{it}}{b_{it} + (1 - g_{it} - b_{it})(1 - \alpha_{ijt})},$$

Given the martingale, we have

$$\frac{g_{it} + (p_{i0} - g_{it})\alpha_{ijt}}{p_{it}g_{it} + (p_{i0} - g_{it})\alpha_{ijt}} \geq \frac{(p_{i0} - g_{it})(1 - \alpha_{ijt})}{p_{it}b_{it} + (p_{i0} - g_{it})(1 - \alpha_{ijt})},$$

It is equivalent (by using the martingale again and the fact that $g_{it} < p_0$) to

$$\alpha_{ijt}(p_0 - p_{it}) \leq p_{it}g_{it}\frac{1 - p_0}{p_0 - g_{it}}.$$

It is clear that the right hand side is weakly positive for any $t \in [0, \infty)$. Since $\Delta > 0$, replace $p_{it}$ with $l_{it}$ and use Lemma 3.2, the inequality is equivalent to

$$\alpha_{ijt}[(\frac{l_{it}}{l_0})^{-1} - 1] \leq (\frac{l_{it}}{l_0})^{-\frac{1}{\Delta}} - 1.$$

Because $\Delta \in (0, 1)$, we have $1 - \frac{1}{\Delta} < 0$ and $\frac{l_{it}}{l_0} \leq 1$, thus $(\frac{l_{it}}{l_0})^{1-\frac{1}{\Delta}} \geq 1$. Therefore, for any $t \in [0, \infty)$,

$$0 \leq (\frac{l_{it}}{l_0})^{-1} - 1 \leq (\frac{l_{it}}{l_0})^{-\frac{1}{\Delta}} - 1.$$

Hence, the inequality holds for any $\alpha_{ijt} \in [0,1]$. That means the Assumption 3.1 imposes no additional restrictions.

## C.3 Proof of Proposition 3.1

First, I will show that both platforms prescribe the Full Transparency policy for both types of consumers is an equilibrium. Then, I will show that's the unique equilibrium.

The Full Transparency policy prescribes $\alpha_{ijt} = 0$ since $\Delta > 0$ and $p_0 < k_{ij}$. It's easy to see that given one platform chooses the Full Transparency policy for both types of consumers, other policies except the Full Transparency policy in the other platform are not implementable when they do not differentiate their products. According to Lemma 3.5, thus the Full Transparency policy is mutual best response to each other. Therefore, they constitute an equilibrium.

Suppose there are another equilibrium strategy profile $\{\hat{\alpha}_{ijt}\}_{i\in\{a,b\},i\in\{a,b\},t\geq 0}$ different from the Full Transparency policy for both types of consumers[1]. Let's assume that $t \geq 0$ is the first time some platform, say platform $a$, uses a non-Full-Transparency policy for some type of consumers, say type $c$ consumers. Thus until time $t$, both platforms prescribe the Full Transparency policy for both types of consumers, i.e. $\hat{\alpha}_{ij\tau} = 0$ for $i \in \{a,b\}$, $i \in \{a,b\}$, $\tau \in [0,t)$,

[1]Technically, $\{\hat{\alpha}_{ijt}\}_{i\in\{a,b\},i\in\{a,b\},t\geq 0}$ should be different from the Full Transparency policy for both types of consumers in some positive measure of time set. Thus the following proof should also consider such time set.

and $\hat{\alpha}_{act} > 0$. Therefore, half of each type of consumers enter each platform until time $t$, and the beliefs evolution until time $t$ are the same for both platforms, i.e. $l_{at} = l_{bt}$. Moreover, both platforms have implementable policy for both types of consumers at that time. According to Lemma 3.4 and 3.5, given $\hat{\alpha}_{act} > 0$, in equilibrium only $\hat{\alpha}_{bct} \in [0, \hat{\alpha}_{act}]$ is implementable for type $c$ consumers in platform $b$. In addition, given $\hat{\alpha}_{bct} \in [0, 1]$, in equilibrium only $\hat{\alpha}_{act} \in [0, \hat{\alpha}_{bct}]$ is implementable for type $c$ consumers in platform $a$. Hence, we must have $\hat{\alpha}_{act} = \hat{\alpha}_{bct} > 0$ and $\hat{z}_{act} = \hat{z}_{bct}$ in equilibrium.

Now consider platform $a$ deviates to an implementable policy at time $t$, $\hat{\alpha}_{act} - \epsilon$, where $\epsilon \in (0, 1/2\hat{\alpha}_{act})$, and keep other actions unchanged. Thus $z_{act} > \hat{z}_{act} = \hat{z}_{bct}$ and platform $a$ attracts all type $c$ consumers and increases her instant payoff at this time due to the increased instant payoff per type $c$ consumer and the increased number of type $c$ consumers. Moreover, $\hat{\mu}_{at} = \rho + 1/2\hat{\alpha}_{act} + 1/2\hat{\alpha}_{adt}$ increases to $\mu_{at} = \rho + \hat{\alpha}_{act} - \epsilon + 1/2\hat{\alpha}_{adt}$. Thus, $l_{a,t+\mathrm{d}t}$ would be smaller than the equilibrium $\hat{l}_{a,t+\mathrm{d}t}$ according to the belief evolution. In addition, according to Lemma 3.4, $z_{a,t+\mathrm{d}t}$ would be larger than the equilibrium $\hat{z}_{a,t+\mathrm{d}t}$, thus relaxes the future constraint. This is a profitable deviation, which leads to a contradiction to the equilibrium. Therefore both platforms prescribe the Full Transparency policy for both types of consumers is the unique equilibrium.

## C.4 Details about $\hat{\alpha}(l_{at})$

The second part of $\hat{\alpha}(l_{at})$ is $\frac{(\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}}-1}{\kappa_1 - l_{at}} l_{at}$, which is the only part that evolves over time. I will show this part is strictly increasing in $t$. Differentiate $\hat{\alpha}(l_{at})$ with respect to $t$, I can get

$$\begin{aligned}
\frac{d\hat{\alpha}(l_{at})}{dt} &= \frac{k_2 - k_1}{1 - k_1} \frac{\frac{1}{\Delta} l_{at} (\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}} + \kappa_1[(1 - \frac{1}{\Delta})(\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}} - 1]}{(\kappa_1 - l_{at})^2} \dot{l}_{at} \\
&= \frac{k_2 - k_1}{1 - k_1} (\frac{l_{at}}{l_0})^{-\frac{1}{\Delta}} \frac{l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]}{\Delta(\kappa_1 - l_{at})^2} \dot{l}_{at} \\
&\propto \{l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]\} \dot{l}_{at}.
\end{aligned}$$

Note that $\dot{l}_{at} < 0$ for any $t$, I will show the item in the brace is also strictly negative. Take derivative, we have

$$\frac{d\{l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]\}}{dl_{at}} = 1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{l_{at}}.$$

Since $\Delta \in (0, 1)$, we have $\frac{1}{\Delta} - 1 > 0$, thus $(l_{at})^{\frac{1}{\Delta}-1}$ is strictly increasing in $l_{at}$ and $l_{at} \in (0, l_0]$. Therefore, $1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{l_{at}} \in [1 - \frac{\kappa_1}{l_0}, 1)$ and is strictly decreasing in $l_{at}$. Note that $1 - \frac{\kappa_1}{l_0} < 0$, thus there exists an $\hat{l}$ such that $1 - (\frac{\hat{l}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{\hat{l}} = 0$, and $1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{l_{at}} > 0$ when $l_{at} \in (0, \hat{l})$, $1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{l_{at}} < 0$ when $l_{at} \in (\hat{l}, l_0]$. Therefore, $l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]$ is strictly increasing when $l_{at} \in (0, \hat{l})$, and strictly decreasing when when $l_{at} \in (\hat{l}, l_0]$. Hence, $l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]$ achieves the maximum when $l_{at} = \hat{l}$.

Note that $1 - (\frac{\hat{l}}{l_0})^{\frac{1}{\Delta}} \frac{\kappa_1}{\hat{l}} = 0$, thus $\hat{l} + \kappa_1[(\Delta - 1) - \Delta(\frac{\hat{l}}{l_0})^{\frac{1}{\Delta}}] = \hat{l} + \kappa_1(\Delta - 1) - \Delta\hat{l} = (\Delta - 1)(\kappa_1 - \hat{l}) < 0$ since $\hat{l} < l_0 < \kappa_1$ and $\Delta < 1$. Hence, I've showed that $l_{at} + \kappa_1[(\Delta - 1) - \Delta(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}]$ is strictly negative, and we can conclude that $\hat{\alpha}(l_{at})$ is strictly increasing in $t$. In addition, $\hat{\alpha}(l_{at}) = 0$ when $t = 0$, and $\hat{\alpha}(l_{at})$ goes to infinity when time goes to infinity.

## C.5 Proof of Lemma 3.6

The Problem A' can be simplified as

$$\sup_{\alpha_{ac}} \int_{t \geq 0} e^{-rt} (l_{act})^{\frac{1}{\Delta}} [\alpha_{act}(1 - \frac{\kappa_1}{l_{act}}) - 1] \, \mathrm{d}t$$
$$\text{s.t.} \quad \dot{l}_{at} = -\Delta \lambda^g (\rho + \alpha_{act}) l_{at}$$
$$\alpha_{act} \in [0, \bar{\alpha}(l_{at})].$$

It turns out it would be easier to solve this problem when we treat the state variable $l$ as "time" and treat time $t(l) = \inf\{t : l_{at} \leq l\}$ as state variable. In addition, this transformation makes that the admissible set of control variable does not depend on state variable, so we can use standard optimal control technique. Then, the problem is equivalent to[2]

$$\sup_{\alpha(l)} \int_0^{l_0} e^{-rt(l)} l^{\frac{1}{\Delta} - 1} [(1 - \frac{\kappa_1}{l}) - \frac{\rho(1 - \frac{k}{l}) + 1}{\rho + \alpha(l)}] \, \mathrm{d}l,$$

s.t

$$t'(l) = -\frac{1}{\Delta \lambda^g (\rho + \alpha(l)) l},$$
$$\alpha(l) \in [0, \bar{\alpha}(l)],$$

and $t(l_0) = 0$.

[2]I omit the subscript "$ac$" and "$a$".

Let $u(l) = \frac{1}{\rho+\alpha(l)}$ be the control variable, then the problem is equivalent to

$$\sup_{\alpha(l)} \int_0^{l_0} e^{-rt(l)} l^{\frac{1}{\Delta}-1}[(1-\frac{\kappa_1}{l}) - (\rho(1-\frac{\kappa_1}{l})+1)u(l)]\,\mathrm{d}l,$$

s.t

$$t'(l) = -\frac{u(l)}{\Delta\lambda^g l},$$
$$u(l) \in [\frac{1}{\rho+\bar{\alpha}(l)}, \frac{1}{\rho}],$$

and $t(l_0) = 0$.

Write Hamiltonian, let $\theta(l)$ be Hamiltonian multiplier,

$$\mathcal{H} = e^{-rt(l)} l^{\frac{1}{\Delta}-1}[(1-\frac{\kappa_1}{l}) - (\rho(1-\frac{\kappa_1}{l})+1)u(l)] - \theta(l)\frac{u(l)}{\Delta\lambda^g l}.$$

The necessary optimality conditions are first

$$\phi(l) \equiv \Delta\lambda^g e^{-rt(l)} l^{\frac{1}{\Delta}}(\rho(1-\frac{\kappa_1}{l})+1) + \theta(l) = 0$$

or $u(l) = \frac{1}{\rho+\bar{\alpha}(l)}$ if $\phi(l) > 0$, $u(l) = \frac{1}{\rho}$ if $\phi(l) < 0$.

And second

$$\theta'(l) = -\frac{\partial\mathcal{H}}{\partial t} = re^{-rt(l)} l^{\frac{1}{\Delta}-1}[(1-\frac{\kappa_1}{l}) - (\rho(1-\frac{\kappa_1}{l})+1)u(l)].$$

Last, the transversality condition $\theta(0) = 0$ from $\lim_{l\to 0}\theta(l)t(l) = 0$.

Let's examine when $\phi(l) = 0$. Note that

$$\begin{aligned}\phi'(l) &= \Delta\lambda^g e^{-rt(l)} l^{\frac{1}{\Delta}-2}[(-rt'(l)l + \frac{1}{\Delta})(\rho(l-\kappa_1)+l) + \rho\kappa_1] + \theta'(l)\\ &= e^{-rt(l)} l^{\frac{1}{\Delta}-2}[(\lambda^g(\rho+1)+r)l - (\lambda^g\rho(1-\Delta)+r)\kappa_1].\end{aligned}$$

Except 0, which cannot be reached in finite time, there is only one $\tilde{l}$ such that $\phi'(\tilde{l}) = 0$. And when $0 < l < \tilde{l}$, $\phi'(l) < 0$, when $l > \tilde{l}$, $\phi'(l) > 0$. Thus $\phi(l)$ is deceasing when $l \in [0, \tilde{l}]$ and increasing when $l \in [\tilde{l}, \infty)$.

It's easy to verify that $\phi(0) = 0$, so except 0, there is only one $l^*$ makes $\phi(l^*) = 0$. In addition, when $0 < l < l^*$, $\phi(l) < 0$, when $l > l^*$, $\phi(l) > 0$. Therefore the optimal policy is "bang-bang" policy: when $0 < l < l^*$, $u(l) = \frac{1}{\rho}$, when $l > l^*$, $u(l) = \frac{1}{\rho + \bar{\alpha}(l)}$. We still need to determine $l^*$.

When $0 < l < l^*$, $u(l) = \frac{1}{\rho}$, thus

$$t'(l) = -\frac{1}{\Delta\lambda^g \rho l},$$

$$\theta'(l) = -\frac{r}{\rho} e^{-rt(l)} l^{\frac{1}{\Delta}-1}.$$

Solve for $t(l)$, get

$$t(l) = C_0 - \frac{lnl}{\Delta\lambda^g \rho},$$

for some finite constant $C_0$. Thus

$$e^{-rt(l)} = e^{-rC_0} l^{\frac{r}{\Delta\lambda^g \rho}},$$

and for some finite constant $C_1$,

$$\theta(l) = -\frac{\Delta\lambda^g r}{r + \lambda^g \rho} e^{-rC_0} l^{\frac{r+\lambda^g \rho}{\Delta\lambda^g \rho}} + C_1.$$

Since $\theta(0) = 0$, $C_1 = 0$.

Thus

$$\begin{aligned}\phi(l) &= \Delta\lambda^g e^{-rC_0} l^{\frac{r}{\Delta\lambda^g \rho}} l^{\frac{1}{\Delta}} (\rho(1 - \frac{\kappa_1}{l}) + 1) - \frac{\Delta\lambda^g r}{r + \lambda^g \rho} e^{-rC_0} l^{\frac{r+\lambda^g \rho}{\Delta\lambda^g \rho}} \\ &= \Delta\lambda^g \rho e^{-rC_0} l^{\frac{r}{\Delta\lambda^g \rho}} l^{\frac{1}{\Delta}} [(1 - \frac{\kappa_1}{l}) + \frac{\lambda^g}{r + \lambda^g \rho}].\end{aligned}$$

$\phi(l^*) = 0$ and $l^* > 0$ implies

$$\frac{\kappa_1}{l^*} = 1 + \frac{\lambda^g}{r + \lambda^g \rho},$$

or

$$l^* = \frac{r + \lambda^g \rho}{r + \lambda^g(\rho + 1)} \kappa_1 < \kappa_1.$$

Therefore, when $l < l^*$, then $\alpha(l) = 0$, when $l > l^*$, then $\alpha(l) = \bar{\alpha}(l)$. In terms of time, if $l_0 \leq l^*$, then $\alpha_t = 0$ for any $t$. And if $l_0 > l^*$, then $\alpha_t = \bar{\alpha}(l_t)$ for $t < t^*$, and $\alpha_t = 0$ for $t > t^*$, where $t^*$ is determined by

$$\begin{cases} t^* = \inf\{t : l_t \leq l^*\}, \\ \dot{l}_t = -\Delta\lambda^g(\rho + \bar{\alpha}(l_t))l_t. \end{cases}$$

## C.6 Proof of Lemma 3.7

The "only if" part is obvious. I will show the "if" part now.

Step 1 Note that if $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, then when time passes to a later time $\bar{t} \geq t^*$, and both platforms' posterior likelihood ratio are $\bar{l} \leq l^*$ when $t = \bar{t}$, given $\alpha_{bct} = \alpha_{bdt} = 0$ for $t \geq \bar{t}$, $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \bar{t}$ is an optimal solution for platform $a$ from $\bar{t}$, i.e. $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \bar{t}$ is an optimal solution for the following problem:

$$
\begin{aligned}
&\sup_{\alpha_{ac},\alpha_{ad}} \int_{t\geq\bar{t}} e^{-r(t-\bar{t})}\{z_{act} + z_{adt}\eta_{adt}\}\,\mathrm{d}t \\
&\quad \text{s.t.} \quad z_{act} \geq v_{bc}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}], \\
&\quad \eta_{adt} = \mathbb{1}(z_{adt} \geq v_{bd}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]), \\
&\quad l_{a\bar{t}} = l_{b\bar{t}} = \bar{l}, \\
&\quad \dot{l}_{bt} = -\Delta\lambda^g \rho l_{bt}, \\
&\quad \dot{l}_{at} = -\Delta\lambda^g(\rho + \alpha_{act} + \alpha_{adt}\eta_{adt})l_{at}, \\
&\quad \alpha_{act}, \alpha_{adt} \in [0,1].
\end{aligned} \tag{Problem B'}
$$

Step 2 Now consider platform $a$'s problem (Problem B")starting from some later time $\tilde{t} > t^*$, at which platform $a$'s posterior is $l^*$ but is bigger than platform $b$'s posterior, i.e. $l_{a\tilde{t}} = l^* > l_{b\tilde{t}}$. In addition, platform $b$ chooses the strategy $\alpha_{bct} = \alpha_{bdt} = 0$ for $t \geq \tilde{t}$. We'll show that if $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, then $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \tilde{t}$ is also the optimal solution to Problem B":

$$
\begin{aligned}
\sup_{\alpha_{ac},\alpha_{ad}} \int_{t\geq\tilde{t}} & e^{-r(t-\tilde{t})}\{z_{act}+z_{adt}\eta_{adt}\}\,\mathrm{d}t \\
\text{s.t.}\quad & z_{act} \geq v_{bc}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}], \\
& \eta_{adt} = \mathbb{1}(z_{adt} \geq v_{bd}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]), \\
& l_{a\tilde{t}} = l^* > l_{b\tilde{t}}, \\
& \dot{l}_{bt} = -\Delta\lambda^g \rho l_{bt}, \\
& \dot{l}_{at} = -\Delta\lambda^g(\rho+\alpha_{act}+\alpha_{adt}\eta_{adt})l_{at}, \\
& \alpha_{act},\alpha_{adt} \in [0,1].
\end{aligned}
\tag{Problem B”}
$$

There are two cases to consider. Case I, suppose $v_{ac}[1-(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}}] \geq v_{bc}[1-(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}}]$, i.e. platform $a$ has an implementable policy for type $c$ consumers at time $\tilde{t}$. Thus we have

$$
\begin{aligned}
v_{ac}-v_{bc} &\geq v_{ac}(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}} - v_{bc}(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}} \\
&\geq [v_{ac}(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}} - v_{bc}(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}}]e^{-\Delta\lambda^g\rho(t-\tilde{t})},
\end{aligned}
$$

for $t \geq \tilde{t}$. That's equivalent to $v_{ac}[1-(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}}e^{-\Delta\lambda^g\rho(t-\tilde{t})}] \geq v_{bc}[1-(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}}e^{-\Delta\lambda^g\rho(t-\tilde{t})}]$, thus $\alpha_{act}=0$ for $t\geq\tilde{t}$ is implementable for Problem B”. If there is another implementable policy could do better than $\alpha_{act}=\alpha_{adt}=0$ for $t\geq\tilde{t}$, it's also implementable in Problem B since $z_{bjt}$ is strictly decreasing in $l_{bt}$. Thus if $\alpha_{act}=\alpha_{adt}=0$ for $t\geq\tilde{t}$ is not an optimal solution for Problem B”, then it's not an optimal solution for Problem B. Therefore, if $\alpha_{act}=\alpha_{adt}=0$ for $t\geq t^*$ is an optimal solution for Problem B, then $\alpha_{act}=\alpha_{adt}=0$ for $t\geq\tilde{t}$ is also the optimal solution to Problem B" in this case.

Case II, suppose $v_{ac}[1-(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}}] < v_{bc}[1-(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}}]$, i.e. platform $a$ has no implementable policy for type $c$ consumers at time $\tilde{t}$. Then there exists some $\hat{t} > \tilde{t}$, at which platform $a$ has an implementable policy for type $c$ consumers at the first time, i.e.

$$\hat{t} = \inf\{t : v_{ac}[1-(\frac{l_{a\tilde{t}}}{l_0})^{\frac{1}{\Delta}} e^{-\Delta\lambda^g \rho(t-\tilde{t})}] \geq v_{bc}[1-(\frac{l_{b\tilde{t}}}{l_0})^{\frac{1}{\Delta}} e^{-\Delta\lambda^g \rho(t-\tilde{t})}]\}.$$

Therefore, $\alpha_{act} = 0$ for $t \geq \hat{t}$ is implementable for Problem B" for the same reason in Case I, and the policy before $\hat{t}$ does not affect anything. If there is another implementable policy could do better than $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \hat{t}$, it's also implementable in Problem B' where $\bar{l} = l_{a\hat{t}}$ and $\bar{t} = \hat{t}$. Thus if $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \hat{t}$ is not an optimal solution for Problem B", then it's not an optimal solution for Problem B' where $\bar{l} = l_{a\hat{t}}$ and $\bar{t} = \hat{t}$. However, as established in Step 1, since $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \hat{t}$ is an optimal solution for Problem B' where $\bar{l} = l_{a\hat{t}}$ and $\bar{t} = \hat{t}$. We get a contradiction.

Therefore, if $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, then $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \tilde{t}$ is also an optimal solution to Problem B".

Step 3 I will show the candidate equilibrium strategy is a mutual best response if $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B.

Suppose platform $b$ chooses $\alpha_{bjt} = \alpha^*_{bjt}$ for $j \in \{c, d\}$ and $t \geq 0$, now I consider platform $a$'s problem.

Before time $t^*$, note that $\alpha_{adt}$ does not matter. Since platform $b$ chooses the maximum implementable experimentation, we have $l_{bt} \leq l_{at}$ for $t \leq t^*$. Therefore, whatever $\{\alpha_{adt}\}_{t\in[0,t^*]}$ it is, we always have $z_{bdt} \geq z_{adt}$ for $t \leq t^*$. Thus, platform $a$ cannot attract any type $d$ consumers before time $t^*$ and type $d$ consumers cannot contribute to the experimentation for platform $a$ before time $t^*$.

Let's consider platform $a$'s strategy for type $c$ consumers before time $t^*$. Suppose platform $a$'s optimal strategy $\{\hat{\alpha}_{ajt}\}$ is not to choose $\alpha_{act} = \alpha^*_{act}$ when $t \leq t^*$, i.e. platform $a$ does not experiment as much as she can, then at time $t^*$, we must have $l_{at^*} > l^* = l_{bt^*}$. In addition, as time passes, there exists some $\bar{t}$ such that $l_{a\bar{t}} = l^* > l_{b\bar{t}}$. Since $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq t^*$ is an optimal solution for Problem B, $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \bar{t}$ is an optimal solution for Problem B" as established in Step 2. Therefore, the best policy for platform $a$ must prescribe $\alpha_{act} = \alpha_{adt} = 0$ for $t \geq \bar{t}$. Given platform $b$'s strategy and platform $a$ prescribe $\hat{\alpha}_{act} = \hat{\alpha}_{adt} = 0$ for $t \geq \bar{t}$, platform $a$ can never attract any type $d$ consumers.

If platform $a$ attracts type $c$ consumers for the entire time, then obviously $\alpha_{act} = \alpha^*_{act}$ when $t \leq t^*$ is the best strategy from Lemma 3.6. If platform $a$ does not attract type $c$ consumers for the entire time, i.e. there exists some positive measure $T \subset \mathbb{R}$, such that $\eta_{act} = 0$ when $t \in T$, then her payoff is even worse-off. Because in this situation, platform $a$'s payoff is strictly less than the payoff in which she attracted type $c$ consumers for the entire time but did not experiment them when $t \in T$. Therefore, platform $a$ must choose $\alpha_{act} = \alpha^*_{act}$ when $t \leq t^*$.

At last, given $\alpha_{act} = \alpha^*_{act}$ when $t \leq t^*$, we come back to Problem B. Therefore, $\{\alpha^*_{ajt}\}_{j\in\{c,d\},t\geq 0}$ is platform $a$'s best response to $\{\alpha^*_{bjt}\}_{j\in\{c,d\},t\geq 0}$, and they constitute an equilibrium.

## C.7 Proof of Lemma 3.9

When $l_0 > l^*$, consider the situation when platform $a$ can experiment all her target type of consumers from time $t^*$, i.e. $\alpha_{act} = 1$ for $t \geq t^*$. If it's infeasible to attract any non-target type of consumers in this situation, then it's never feasible to do so. Therefore, I need to prove that for any $t \geq t^*$, given System A and $\alpha_{act} = 1$,

$$v_{ad}[1 - (\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}] < v_{bd}[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}].$$

That is, for any $t \geq t^*$

$$(1-k_2)[1 - (\frac{l^*}{l_0})^{\frac{1}{\Delta}} e^{-\lambda^g(\rho+1)(t-t^*)}] < (1-k_1)[1 - (\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}} e^{-\lambda^g \rho(t-t^*)}],$$

or equivalently

$$\chi = \frac{k_2 - k_1}{1 - k_1} > (\frac{l^*}{l_0})^{\frac{1}{\Delta}} \frac{e^{-\lambda^g \rho(t-t^*)} - e^{-\lambda^g(\rho+1)(t-t^*)}}{1 - (\frac{l^*}{l_0})^{\frac{1}{\Delta}} e^{-\lambda^g(\rho+1)(t-t^*)}}. \quad (1)$$

Let's consider how the right hand side of the above inequality changes over time. Take derivative,

$$\mathrm{d}\{\frac{e^{-\lambda^g \rho(t-t^*)} - e^{-\lambda^g(\rho+1)(t-t^*)}}{1 - (\frac{l^*}{l_0})^{\frac{1}{\Delta}} e^{-\lambda^g(\rho+1)(t-t^*)}}\}/\mathrm{d}t \propto \Omega \equiv (\rho+1)e^{-\lambda^g(t-t^*)} - \rho - (\frac{l^*}{l_0})^{\frac{1}{\Delta}} e^{-\lambda^g(\rho+1)(t-t^*)}.$$

Note that

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = -\lambda^g(\rho+1)e^{-\lambda^g(t-t^*)}[1-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}e^{-\lambda^g\rho(t-t^*)}] < 0,$$

and $\lim_{t\to t^*}\Omega = 1-(\frac{l^*}{l_0})^{\frac{1}{\Delta}} > 0$, $\lim_{t\to\infty}\Omega = -\rho < 0$, thus there exists some $\check{t} > t^*$ such that when $t = \check{t}$, $\Omega = 0$. In addition, $\Omega > 0$ when $t \in [t^*, \check{t})$, and $\Omega < 0$ when $t \in (\check{t}, \infty)$. Therefore, the right hand side of inequality 1 achieves the maximum when $t = \check{t}$. Take $\Omega = 0$ when $t = \check{t}$, we can get the maximum of the right hand side of inequality 1

$$\frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}e^{-\lambda^g\rho(\check{t}-t^*)}.$$

From $\Omega = 0$ when $t = \check{t}$, we know that

$$\rho+1-\rho e^{\lambda^g(\check{t}-t^*)} = (\frac{l^*}{l_0})^{\frac{1}{\Delta}}e^{-\lambda^g\rho(\check{t}-t^*)} < (\frac{l^*}{l_0})^{\frac{1}{\Delta}},$$

thus

$$e^{-\lambda^g(\check{t}-t^*)} < \frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}.$$

Therefore, the maximum of the right hand side of inequality 1 is strictly less than

$$\frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}[\frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}]^{\rho}.$$

Hence, if $\chi > \frac{1}{1+\rho}(\frac{l^*}{l_0})^{\frac{1}{\Delta}}[\frac{\rho}{1+\rho-(\frac{l^*}{l_0})^{\frac{1}{\Delta}}}]^{\rho}$, then inequality 1 holds for any $t \geq t^*$. Therefore there exists no $t \geq t^*$ such that $v_{ad}[1-(\frac{l_{at}}{l_0})^{\frac{1}{\Delta}}] \geq v_{bd}[1-(\frac{l_{bt}}{l_0})^{\frac{1}{\Delta}}]$ in System A.

When $l_0 \leq l^*$, with very similar argument, we have that if $\chi > \frac{1}{1+\rho}$, then it's infeasible to attract any non-target type of consumers.

# Bibliography

Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2011). Experimentation, patents, and innovation. *American Economic Journal: Microeconomics*, 3(1):37–77.

Acharya, V. V., DeMarzo, P., and Kremer, I. (2011). Endogenous Information Flows and the Clustering of Announcements. *American Economic Review*, 101(7):2955–2979.

Argenton, C. and Prüfer, J. (2012). Search engine competition with network externalities. *Journal of Competition Law and Economics*, 8(1):73–105.

Aumann, R. J. and Maschler, M. (1995). *Repeated games with incomplete information*. MIT press.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817.

Beaudry, P. (1994). Why an informed principal may leave rents to an agent. *International Economic Review*, pages 821–832.

Bedard, N. C. (2016). Contracts in informed-principal problems with moral hazard. *Economic Theory Bulletin*, pages 1–14.

Bergemann, D. and Hege, U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance*, 22(6):703–735.

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

Besley, T. and Prat, A. (2006). Handcuffs for the Grabbing Hand? Media Capture and Government Accountability. *The American Economic Review*, 96(3):720–736.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026.

Bolton, P. and Harris, C. (1999). Strategic experimentation. *Econometrica*, 67(2):349–374.

Che, Y.-K. and Hörner, J. (2018). Recommender Systems as Mechanisms for Social Learning*. *The Quarterly Journal of Economics*, 133(2):871–925.

Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, pages 179–221.

Daughety, A. F. and Reinganum, J. F. (2018). Evidence Suppression by Prosecutors: Violations of the Brady Rule. *SSRN Electronic Journal*.

De Corniere, A. (2013). Search advertising. *Available at SSRN 1967102*.

Dye, R. A. (1985). Disclosure of Nonproprietary Information. *Journal of Accounting Research*, 23(1):123–145.

Dye, R. A. (2017). Optimal disclosure decisions when there are penalties for nondisclosure. *The RAND Journal of Economics*, 48(3):704–732.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic Studies*, 80(4):1422–1458.

Egorov, G., Guriev, S., and Sonin, K. (2009). Why resource-poor dictators allow freer media: A theory and evidence from panel data. *American Political Science Review*, 103(4):645–668.

Eliaz, K. and Spiegler, R. (2011). A simple model of search engine pricing*. *The Economic Journal*, 121(556):F329–F339.

Elitzur, R. and Gavious, A. (2003). Contracting, signaling, and moral hazard: a model of entrepreneurs,'angels,'and venture capitalists. *Journal of Business Venturing*, 18(6):709–725.

Ely, J. C. (2017). Beeps. *American Economic Review*.

Ely, J. C. and Szydlowski, M. (2016). Moving the Goalposts. Technical report, working paper, Northwestern University.

Eraslan, H. and Ozerturk, S. (2017). Information Gatekeeping and Media Bias.

Fryer, R. G. and Holden, R. T. (2012). Multitasking, learning, and incentives: a cautionary tale. Technical Report w17752, National Bureau of Economic Research, Cambridge, MA.

Gehlbach, S. and Sonin, K. (2014). Government control of the media. *Journal of Public Economics*, 118:163–171.

Gratton, G., Holden, R., and Kolotilin, A. (2016). When to Drop a Bombshell. *Discussion Papers*, 13.

Grossman, S. J. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *The Journal of Law & Economics*, 24(3):461–483.

Grubb, M. D. (2011). Developing a reputation for reticence. *Journal of Economics & Management Strategy*, 20(1):225–268.

Guo, Y. (2016). Dynamic delegation of experimentation. *American economic review*.

Guriev, S. M. and Treisman, D. (2018). Informational Autocracy: Theory and Empirics of Modern Authoritarianism. Technical report.

Guttman, I., Kremer, I., and Skrzypacz, A. (2014). Not Only What but Also When: A Theory of Dynamic Voluntary Disclosure. *American Economic Review*, 104(8):2400–2420.

Hagiu, A. and Jullien, B. (2014). Search diversion and platform competition. *International Journal of Industrial Organization*, 33:48–60.

Halac, M., Kartik, N., and Liu, Q. (2014). Contests for experimentation.

Halac, M. C., Liu, Q., and Kartik, N. (2016). Optimal contracts for experimentation. *The Review of Economic Studies*.

Hauser, D. N. (2017). Promotion, Censorship, and Reputation for Quality.

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

Jehiel, P. (2015). On Transparency in Organizations. *The Review of Economic Studies*, 82(2):736–761.

Jung, W.-O. and Kwon, Y. K. (1988). Disclosure When the Market Is Unsure of Information Endowment of Managers. *Journal of Accounting Research*, 26(1):146–153.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Karle, H., Schumacher, H., and Staat, C. (2016). Signaling quality with increased incentives. *European Economic Review*, 85:8–21.

Kartik, N., Lee, F. X., and Suen, W. (2017). A Theorem on Bayesian Updating and Applications to Communication Games.

Kaya, A. (2010). When does it pay to get informed? *International Economic Review*, 51(2):533–551.

Keller, G. and Rady, S. (2010). Strategic experimentation with Poisson bandits. *Theoretical Economics*, 5(2):275–311.

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

Keller, G., Rady, S., and Cripps, M. (2005). Strategic Experimentation with Exponential Bandits. *Econometrica*, 73(1):39–68.

Klees, J. E. and Joines, R. (1996). Occupational health issues in the pharmaceutical research and development process. *Occupational medicine (Philadelphia, Pa.)*, 12(1):5–27.

Laraki, R., Solan, E., and Vieille, N. (2005). Continuous-time games of timing. *Journal of Economic Theory*, 120(2):206–238.

Li, W. and Yang, D. T. (2005). The great leap forward: Anatomy of a central planning disaster. *Journal of Political Economy*, 113(4):840–877.

Lorentzen, P. (2014). China's Strategic Censorship. *American Journal of Political Science*, 58(2):402–414.

Maskin, E. and Tirole, J. (1990). The principal-agent relationship with an informed principal: The case of private values. *Econometrica*, pages 379–409.

Maskin, E. and Tirole, J. (1992). The principal-agent relationship with an informed principal, II: Common values. *Econometrica*, pages 1–42.

Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391.

Moscarini, G. and Squintani, F. (2010). Competitive experimentation with private information: The survivor's curse. *Journal of Economic Theory*, 145(2):639–660.

Myerson, R. B. (1983). Mechanism design by an informed principal. *Econometrica*, pages 1767–1797.

Presman, E. L. (1991). Poisson Version of the Two-Armed Bandit Problem with Discounting. *Theory of Probability & Its Applications*, 35(2):307–317.

Redlicki, J. (2017). What Drives Regimes to Manipulate Information: Criticism, Collective Action, and Coordination.

Renault, J., Solan, E., and Vieille, N. (2017). Optimal dynamic information provision. *Games and Economic Behavior*, 104:329–349.

Riedel, F. and Steg, J.-H. (2017). Subgame-perfect equilibria in stochastic timing games. *Journal of Mathematical Economics*, 72:36–50.

Rodan, G. (2004). *Transparency and Authoritarian Rule in Southeast Asia: Singapore and Malaysia*. Routledge.

Rosenbaum, S. M., Billinger, S., and Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45:181–196.

Shadmehr, M. and Bernhardt, D. (2015). State censorship. *American Economic Journal: Microeconomics*, 7(2):280–307.

Shin, H. S. (2003). Disclosures and Asset Returns. *Econometrica*, 71(1):105–133.

Silvers, R. (2012). The value of information in a principal–agent model with moral hazard: The ex post contracting case. *Games and Economic Behavior*, 74(1):352–365.

Sirowy, L. and Inkeles, A. (1990). The effects of democracy on economic growth and inequality: A review. *Studies in Comparative International Development*, 25(1):126–157.

Smirnov, A. and Starkov, E. (2018). Bad News Turned Good: Reversal Under Censorship.

Smith, L. and Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398.

Stevens, G. A. and Burley, J. (1997). 3,000 raw ideas= 1 commercial success! *Research-Technology Management*, 40(3):16–27.

Taylor, G. (2013). Search quality and revenue cannibalization by competing search engines. *Journal of Economics & Management Strategy*, 22(3):445–467.

Wagner, C., Mylovanov, T., and Tröger, T. (2015). Informed-principal problem with moral hazard, risk neutrality, and no limited liability. *Journal of Economic Theory*, 159:280–289.