Motivating Information Acquisition Under Delegation

Spyros Terovitis†

November 11, 2016

Abstract

We study a model in which a principal delegates a choice between different actions to an agent. The return from each action is unknown but the agent can invest in acquiring (noisy) information before making his choice. The principal would like the agent to invest in information but this investment is unobservable and only the chosen action and its resulting payoff is contractible. We solve for the optimal contract and show that it induces a contrarian bias. As the main application, we explore a setting where an analyst in a brokerage house issues financial recommendations, and argue that our findings are supported by empirical evidence on the behaviour of financial analysts.

Keywords: Information Acquisition, Experts’ Compensation, Financial Recommendation.

*I am grateful to Motty Perry and Ilan Kremer for their support and their very useful comments. Moreover, I would like to thank Herakles Polemarchakis, Dan Bernhardt, Kobi Glazer, Marina Halac, Sinem Hidir, Jacques Cremer, Alvaro Sandroni, and Phil Reny and the participants at the CRETA Conference, the ESSEC Business School Seminar, and the Warwick Economics Workshop. I also wish to thank the Royal Economic Society and the State Scholarships Foundation of Greece for financial support.

†University of Warwick, Department of Economics, Coventry, CV4 7AL, United Kingdom, email: s.terovitis@warwick.ac.uk
1 Introduction

There are many situations where individuals delegate a decision to an expert. For instance, a patient delegates the choice of his medical treatment to his doctor; a car-owner delegates the decision regarding her car repair to a mechanic; a brokerage house delegates the issuance of financial recommendations to an analyst. A key question in these settings is how a principal can incentivize an expert to implement the action which is in the principal’s best interest. What makes answering this question less straightforward is that, in contrast to the standard principal-agent problems, in these settings the principal does not know ex-ante which action is in her best interest. The principal’s desired action depends on the state of the world, which is ex-ante unknown. For instance, a patient would prefer treatment A to treatment B only if the former is more appropriate for his health condition; a car-owner would prefer to repair her car engine only if it is necessary; a brokerage house would prefer its analyst to recommend a short over a long position only when future asset prices fall. This paper addresses the following question: how could a principal incentivize an agent to take a profit-maximizing action when neither he nor the agent knows ex-ante what the profit-maximizing action is?

In this setting, the principal would like the agent to collect information regarding the state of the world, and subsequently take the profit-maximizing action, based on the available information. However, information acquisition is costly, and in many cases unobserved, by the principal. Consequently, the only way to motivate the agent to collect information is through a contract contingent on the implemented action. Such a contract – apart from affecting the agent’s incentives to gather information – also affects his incentives when deciding which action to implement after the information is obtained. We show that the interaction of unobservable information acquisition and the multi-tasking nature of the problem (i.e., information acquisition and implementation of a decision rule), results in a trade-off between the cost of incentivizing information acquisition and the benefit of using the obtained information effectively. This paper shows that the incentive scheme which optimally solves this trade-off should reward contrarian actions, which subsequently leads an agent to adopt such actions more often than the first best.

We develop a model in which: i) a principal delegates the implementation of a binary action to an agent, ii) the action which maximizes the principal’s revenue is state-dependent, iii) the state of the world is ex-ante unknown, iv) the agent can acquire costly information about the state of the world, and v) neither the information acquisition nor the obtained information is

---

1 Other applications include product design, the marketing orientation, and issuing credit ratings.

2 In the product design problem, allocating company’s resources to the production of a new product is the principal’s desired action as long as it succeeds in accommodating the ex-ante unknown consumer’ preference. Similarly, in the market orientation problem, adopting a new marketing technique over a conventional one is the desired action as long as its ex-ante unknown productivity exceeds the productivity of the conventional technique. Finally, in the rating process, a bad rating is preferred over a good as long as the company defaults.
observable to the principal. In our main application, a brokerage house (principal) delegates the issuance of financial recommendations to an analyst (agent). The binary action consists of recommending a short or a long position. Finally, the ex-ante unknown state of the world refers to assets’ future prices.

Our model differs from the typical principal-agent set-up in three dimensions. First, neither the principal nor the agent knows ex-ante the profit maximizing action. Second, neither the action of learning nor the learning outcome is observable to the principal. Finally, the agent does not hold any private information ex-ante. In fact, the agent finds it optimal to incentivize the agent to gather private information. In this environment, we show that the derivation of the optimal compensation contract consists of two parts: i) the characterization of the most cost-efficient contract which implements a given decision rule, and ii) the characterization of the decision rule which maximizes the principal’s profits.

The first set of findings refers to the optimal compensation contract. We find that the promised payment is positive only when the agent implements the ex-post revenue-maximizing action. Furthermore, the optimal contract promises a premium when the agent ‘goes against the flow’, i.e., takes an action which differs from the action the principal would implement, had he not hired an agent. In turn, this premium leads to over-implementation against the flow, with the extent of this bias being increasing in the cost of acquiring information and the prior beliefs. Applying these findings to the financial recommendation example, the analyst’s compensation is positive as long as he recommends a short (long) position and the price eventually falls (rises). In addition, when prior beliefs indicate that the asset is undervalued, the analyst recommends a short position more often than the first best, and the other way around. The bias against the flow is consistent with the empirical evidence on analysts’ recommendations provided in Bernhardt, Wan, and Xiao (2016). Also, given that analysts’ recommendations reflect their forecast about the asset price, this bias is in line with Pierdzioch, Rülke, and Stadtmann (2013), Laster, Bennett, and Geoum (1997), and Ehrbeck and Waldmann (1996), who provide evidence that analysts issue contrarian forecasts.

The second set of findings refers to the effect of the optimal compensation contract on the informational role of the implemented action. We show that the informational role of the implemented action differs depending on which action the principal would implement, had he not hired an agent. In particular, we find that the conditional probability that an implemented action is revenue-maximizing given that it is against the flow, is lower than the corresponding probability when there is no agency problem. An against-the-flow action might be driven by the premium of the corresponding payment, rather than the agent’s belief that this action is revenue-maximizing. Following a similar logic, the conditional probability that an implemented action is revenue-maximizing given that it is following the flow, is higher than the corresponding probability when there is no agency problem. This finding relates to the result that the agent is willing to forgo the premium that an against-the-flow action entails, only if he is very confident
that this action is not revenue-maximizing. In the financial recommendation setting, when an asset is considered over-valued: i) the conditional probability that its price will fall given a recommendation of a short position is lower than the corresponding probability when there is no agency problem, and ii) the conditional probability that its price will increase given a recommendation of a long position, is higher than the corresponding probability when there is no agency problem. Thus, the recommendation of a short position is a weak indicator that the asset price will fall, whereas the recommendation of a long position is a strong indicator that the asset price will rise. The opposite findings hold when the asset is considered undervalued.

The previous mechanism could provide a theoretical foundation for the empirical evidence indicating that investors over-react and/or under-react to financial analysts’ recommendations or forecasts (Elgers, Lo, and Pfeiffer Jr 2001, Elliot, Philbrick, and Wiedman 1995, Elliot et al. 1995, Mendenhall 1991, Sloan 1996). The rationale is that analysts’ recommendations are indicative about their private information, and thus can be used by investors to infer the future price of an asset. Our model predicts that rational investors should under-react when the analyst recommends a contrarian position, and over-react otherwise.

The bias against the flow is a key finding, which plays a critical role in the implications of the optimal contract. The intuition behind this bias relies on the interaction of unobservable information acquisition and the multi-tasking nature of the problem. To incentivize the agent to acquire unobservable information, the principal should promise a positive payment only if the ex-post revenue-maximizing action is implemented. However, the compensation contract also shapes the agent’s incentives when considering which action to implement, which in turn, determines the principal’s expected revenue. A critical property is that under the optimal contract, the agent’s utility coincides with his outside option. Also, for the payments which maximize the principal’s revenue, the agent’s outside option is to follow the flow without acquiring information. The principal can thus lower the cost of incentivizing information acquisition by simply worsening the agent’s outside option. This is achieved by lowering the payment when the flow is followed. However, changing the ratio of payments comes at the cost of decreasing the principal’s expected revenue: for the new ratio of payments, the expected revenue is no longer maximized. The optimal contract thus solves the inherent trade-off between the cost of incentivizing learning and the benefit of implementing a decision rule as close as possible to the revenue-maximizing one. This trade-off lies at the heart of our paper. We show that for small deviations from the first best, the decrease in the expected cost dominates the decrease in the expected revenue.

In Appendix A and B, we explore two extensions of the benchmark setting. Appendix A analyzes the case where the state of the world is imperfectly observed. Appendix B examines the case where the principal allocates the task of information acquisition and the task of implementing a decision rule to two different individuals. We show that the optimal contract in these cases has the same features as the optimal contract in the benchmark model, and the
main finding of bias against the flow remains.

In Section 6 we discuss four alternative environments which share the same characteristics as the benchmark setting, and provide the main implications of the optimal contract in each setting. First, in a portfolio allocation problem where a fund manager invests in a risky or a safe asset, we argue that the optimal contract leads to under-investment or over-investment in the risky asset. Second, in a product design problem, we argue that the optimal contract leads to product features which are less likely to accommodate future demand, compared to the first best. Third, we consider the rating process in credit rating agencies. We claim that the optimal contract implies more frequent good ratings (rating inflation) than the first best when the market is pessimistic about the company’s creditworthiness, and vice versa. This finding could support extreme prior beliefs about a company’s creditworthiness. Finally, in an environment where the principal aims to motivate innovation, we claim that the optimal contract can capture both under-implementation and over-implementation of innovative strategies.

The outline of the paper is as follows. Section 2 discusses the related literature. Section 3 describes the benchmark model. Section 4 characterizes the optimal compensation contract. Section 5 explores the implications of the optimal compensation contract. Section 6 discusses alternative applications and concludes.

2 Related Literature

Contracting and information acquisition

Similarly to our model, Lewis and Sappington (1997), Gromb and Martimort (2007), Lambert (1986), Chade and Kovrijnykh (2016), and Inderst and Ottaviani (2009), explore the optimal contracting problem in a setting where a principal delegates information acquisition to an agent.

Lewis and Sappington (1997) consider a set-up where an agent is incentivized to acquire information about the state of the world before choosing an unobservable level of cost-reducing effort. A critical difference from this paper is that, in our setting, the principal’s desired action is state-dependent. This characteristic is responsible for the inherent trade-off explained in the introduction. In addition, as opposed to Lewis and Sappington (1997), we find that allocating the tasks to two different agents, is not preferred by the principal.

The papers which consider environments which are closer to ours is Lambert (1986) and Gromb and Martimort (2007). Gromb and Martimort (2007) characterize the most cost-efficient contract of incentivizing an analyst – first, to acquire a costly binary signal about a project’s quality, and second, to truthfully reveal the signal realization. We differ from this paper by endogenizing the decision rule that the principal prefers the agent to implement. This element gives rise to the aforementioned trade-off, which results in the bias in the implemented action. Such a trade-off, thus, bias, does not appear in Gromb and Martimort (2007). Lambert (1986) explores the compensation contract which incentivizes the manager to obtain information before
deciding to invest in a risky or a safe asset. As opposed to [Lambert (1986)], we are able to formally characterize a simple condition which determines the direction of the bias compared to the first best. The simplicity of this condition allows us to generate testable implications about the informational role of the implemented action, and perform a series of comparative statics. Also, we explore a more general set-up than [Lambert (1986)] by focusing on information acquisition regarding the state of the world, rather than a particular action. Hence, our set-up could be used to explore alternative environments, like the ones discussed in Section 6.

A recent paper which allows for delegation of information acquisition is [Chade and Kovrijnykh (2016)]. The main focus of the two papers differ; we focus on the implemented action, whereas [Chade and Kovrijnykh (2016)] focus on the quality of information available. Another critical difference is that in [Chade and Kovrijnykh (2016)] the signal realization (but not its precision) is observed by all parties, and the agents cannot misreport it. In contrast, we explore a set-up where the principal cannot observe the realized signal, and as a result, the contract cannot be contingent on that. Also, our paper is close to [Inderst and Ottaviani (2009)] who consider a set-up where a principal delegates a purchase recommendation to a seller representative. We differ from [Inderst and Ottaviani (2009)] because the principal can never observe whether the agent actually acquired information. This is a key departure, which critically affects not only the optimal compensation contract but also its main implications. The literature on delegation also includes, among others, [Demski and Sappington (1987)], [Garicano and Santos (2001)], [Malcomson (2009)], [Szalay (2005)], [Eső and Szentes (2007)].

**Information asymmetry, investment decision and analysts’ reports**

This work also relates to the literature which recognizes that information asymmetry might distort manager’s investment decision or financial analysts’ forecasts. This strand of the literature, which starts with the seminal work of [Scharfstein and Stein (1990)], and it includes [Ottaviani and Sørensen (2006)], [Zwiebel (1995)], and [Levy (2004)], highlights the impact of publicly available information on financial decisions. In particular, [Scharfstein and Stein (1990)] and [Ottaviani and Sørensen (2006)] provide models which give rise to managers mimicking the decisions of other managers (herding behavior), whereas, [Zwiebel (1995)], and [Levy (2004)] develop models where anti-herding behavior in the analysts’ forecasts arises. The existing empirical evidence on the emergence of herding or anti-herding behavior is conflicting. For instance, [Trueman (1990)] and [Hong, Kubik, and Solomon (2000)] provide evidence which indicates herding behavior. In contrast, [Bernhardt and Kutsoati (2001)] and [Bernhardt, Campello, and Kutsoati (2006)] find strong evidence that support anti-herding behavior. This finding is supported by [Laster et al. (1997)], [Ehrbeck and Waldmann (1996)], and [Pierdzioch et al. (2013)]. In these papers, the main mechanism which leads to distortion compared to the first best is the agent’s career concerns; the manager (analyst) uses his investment decision (forecast) as an instrument to signal his type, rather than to maximize the principal’s profits.
We differ from this literature in two important directions. First, we are interested in the optimal compensation of an agent, a topic that has been overlooked in the previous papers. Second, we recognize that assuming that agents are endowed with a private signal is a significant shortcoming. Thus, we are interested in exploring not only how to incentivize an agent to take a decision, or an analyst to disclose his private information truthfully, but also how to optimally incentivize the agent to acquire costly information before the action is taken. Also, the source of information asymmetry differs; our environment is not characterized by information asymmetry regarding the ability of the manager, but by ex-ante moral hazard (information acquisition is a hidden action) and ex-post asymmetric information (the signal realization is the agent’s private information). Hence, the mechanism that lead to the bias differs; the bias is a consequence of the principal’s motive incentivize information acquisition at the lowest cost.

Optimality of relative performance pay
An implication of our analysis is that under the optimal contract, both the agent’s compensation and the bias in the agent’s decision depend on the prior beliefs. Since we do not impose any restriction on how prior beliefs are formed, our set-up can allow for the case where prior beliefs capture actions of other agents, as long as these actions uncover some information about the state of the world. Based on this remark, our work relates to Holmstrom (1979), Holmstrom (1982), Lazear and Rosen (1979), McConnell, Diamond, and Verrecchia (1982) which highlight the benefit of evaluating agents on the basis of their relative performances in environments where agents’ performance is affected by common shocks. We differ from this literature in two directions. First, our set-up explores the case where before taking a decision, the agent observes the actions of other agents. The second variation refers to the benefit of relative pay. In our model, the benefit of a relative pay comes from worsening the agent’s outside option of not acquiring information, whereas in Holmstrom (1979) the benefit stems from the fact that relative performance is a more informative indicator about whether an unobservable action is taken.

Contracting under multi-tasking
Finally, this paper pertains on the literature which highlights the misallocation consequences of the optimal contract. For instance, Holmstrom and Milgrom (1991) explore a setting, where an agent allocates his effort across multiple tasks, and the principal observes a performance measure for each of these tasks. The nature of this multi-tasking problem differs from Holmstrom and Milgrom (1991). In our set-up, multi-tasking emerges endogenously: the principal is only interested in information acquisition to the extent that it improves the information set, based on which the agent takes an action. Inderst and Ottaviani (2009), and Lambert (1986) also find a resulting bias in one of the tasks.
3 Benchmark Model

**Environment:** We consider a setting where a risk-neutral principal delegates a decision, denoted by $d$, to a risk-neutral agent. Decision $d$ refers to implementing one of the two available actions, $b$ or $g$. The generated revenue of the implemented action is received by the principal, and its value depends on the state of the world, $\theta \in \{B, G\}$. In particular, the realized revenue, denoted by $R(d, \theta)$, is given by:

$$R(d = b, \theta = B) = R(d = g, \theta = G) = v$$  \hspace{1cm} (1)

$$R(d = b, \theta = G) = R(d = g, \theta = B) = -v$$  \hspace{1cm} (2)

The first observation derived from (1) and (2) is that the principal’s preferred action is state-dependent. The second observation is that the revenue structure is symmetric: i) the gain (loss) of implementing the action which matches (does not match) the state is independent of the realized state, and ii) the gain and the loss have the same magnitude. The assumption of symmetry improves the tractability of the model, however, it could be relaxed without affecting the main findings qualitatively.

We focus on the case where the realized state of the world is ex-ante unknown to the principal and the agent, who, however, have common prior beliefs about it. In particular, they expect that the state is good ($\theta = G$) with probability $p$, and bad ($\theta = B$) with probability $1 - p$. Thus, the value of $p$ incorporates any publicly available information regarding the state of the world, whereas we do not impose any structure on how these beliefs are formed.

**Information Acquisition Technology:** Before an action is implemented, the agent can acquire a signal about the state of the world. The agent’s private information consists of the signal realization and signal acquisition, the latter of which incurs a cost $c$. This cost can be interpreted as a utility loss due to effort of acquiring a signal. Conditional on obtaining information, the agent receives a noisy signal $s \in [0, 1]$ about the state of the world. The signal $s$ is distributed according to a continuous density function $f_\theta(.)$, with a distribution function $F_\theta(.)$, where $\theta = \{B, G\}$. After observing the signal realization $s'$, the agent updates his beliefs as follows:

$$Pr(\theta = G|s = s') = \frac{f_G(s')p}{f_G(s')p + f_B(s')(1 - p)} = 1 - Pr(\theta = B|s = s')$$

**Assumption 1:** Monotone Likelihood Ratio Property (MLRP)

*For any signal realization $s' \in [0, 1]$, the ratio $\frac{f_G(s')}{f_B(s')} \text{ is increasing in } s'.$*

---

3This characteristic relates to the nature of the delegation problem – the principal, as opposed to the agent, does not have expertise in inferring the unknown state of the world. Thus, the principal is unaware of the evidence the agent needs to collect to infer the unknown state of the world.
A direct consequence of Assumption 1 is that the probability that the state is $\theta = G$ is increasing in the signal realization. Besides, given that $f_G(s)$ and $f_B(s)$ represent probability density functions, Assumption 1 implies that $f_G(s)$ and $f_B(s)$ satisfy the single-crossing condition.

**Preferences and Actions:** The principal designs the agent’s compensation contract, which is denoted by $\hat{W}$ and specified below, in order to maximize his expected profits:

$$\mathbb{E} \Pi = \mathbb{E}[R|\hat{W}] - \mathbb{E}[C|\hat{W}]$$  \hspace{1cm} (3)

where $\mathbb{E}[R|\hat{W}]$ stands for the expected revenue of the implemented action, whereas $\mathbb{E}[C|\hat{W}]$ stands for the agent’s expected compensation, given a contract $\hat{W}$. Note that the compensation contract affects the expected revenue, $\mathbb{E} R$, through shaping the agent’s incentives when deciding which action to implement.

The agent faces two kinds of decisions: the information acquisition decision, and the decision to implement action $g$ or $b$. The objective of the agent is to maximize his expected utility:

$$\mathbb{E} V = \mathbb{E}[C|\hat{W}] - \mathbb{1}c$$  \hspace{1cm} (4)

where $\mathbb{1}$ equals 1 if information is obtained, and zero otherwise.

**Set of available contracts:** We allow for contracts contingent on the implemented action, $d$, and the state realization, $\theta$, i.e., $\hat{W} : d \times \theta \mapsto \mathbb{R}^+$, where $d = \{b, g\}$ and $\theta = \{B, G\}$. Thus, the contract is characterized by the following quadruple:

$$\hat{W} = \{w_{bB}, w_{bG}, w_{gB}, w_{gG}\}$$

where $w_{bG}$ ($w_{bB}$) denotes the payment when the agent implements action $b$ and the state is revealed to be $G$ ($B$). Likewise, $w_{gG}$ ($w_{gB}$) denotes the payment when the agent implements action $g$ and the state is revealed to be $G$ ($B$). For the sake of tractability, we assume that the agent is protected by limited liability, i.e., payment $w_{d\theta}$ is non-negative. In section 4.2, we show that this assumption can be relaxed without affecting the main findings qualitatively.

In Appendix A, we explore the case where it is not feasible for the principal to offer contracts contingent on the realized state of the world. We show that as long as there is a public signal, which is revealed after the action is implemented and is informative about the realized state of the world, the main findings go through. In Appendix B, we allow for contracts contingent on: i) messages sent by the agent to the principal, and ii) the realized state of the world. We show that the optimal contract is effectively the same, independently of whether it is contingent on the implemented action or the agent’s messages.
**Timing:** The sequence of events is:

1. The principal offers a compensation contract $\hat{W}$.
2. The agent decides whether to acquire information.
3. If information is obtained, the agent observes the signal realization.
4. The agent chooses the implemented action.
5. The state is realized, and the contract is executed.

**Financial Recommendations Example**

Consider a brokerage house that issues financial recommendations. Equity-holders of a brokerage house (principal) delegate financial recommendations to a financial analyst (agent), who has two options: recommending a “long” ($d = g$) position or recommending a “short” position ($d = b$). The expected revenue of each recommendation/position depends on the price movement of the asset, $P_1^\theta - P_0$, which in turn, depends on the ex-ante unknown quality of the asset (state of the world). State $\theta = G$ and $\theta = B$ stand for the case where the asset is of good and bad quality, respectively. In order to replicate (1) and (2), we assume that if the asset is of good quality, its price is expected to increase to $P_1^G = P_0 + v$, whereas if the asset is of bad quality, its price is expected to decrease to $P_1^B = P_0 - v$. This assumption captures the idea that the asset price moves towards its fundamental value.

A difference from the benchmark case is that the expected revenue generated by the recommendation accrues to the customers of the brokerage house, rather than the brokerage house itself. However, in a highly competitive industry such as financial advice, one would expect that the objective of brokerage houses is to generate revenue for their customers, which would in turn, allow these firms to increase their commission. We show in the Appendix that the main findings would be qualitatively similar if the brokerage house’s objective were to maximize the the expected probability of issuing the “correct” recommendation.

### 4 Optimal Compensation Contract

In this section, we characterize the optimal compensation contract, focusing on the case where the principal finds it optimal to incentivize the agent to acquire information. In section 4.7, we characterize the conditions when incentivizing information acquisition is optimal.

Given that information acquisition is unobservable, the only way to motivate the agent to acquire information is through a contract contingent on the implemented action. However, the compensation contract affects not only the agent’s incentives to acquire information, but also his decision regarding the implemented action after the information is obtained. Thus,
following [Grossman and Hart (1983)], there are two parts to the derivation of the optimal compensation contract consists: i) the characterization of the most cost-efficient contract which implements a given decision rule (mapping from the signal realization to an action), and ii) the characterization of the decision rule which maximizes the principal’s profits, given the cost-efficient contract which implements this decision rule.

Our analysis highlights the inherent trade-off between the cost of incentivizing information acquisition and the benefit of implementing a decision rule as close as possible to the revenue-maximizing one. The goal of this section is to show that the contract which optimally solves this trade-off gives rise to over-implementation against the flow. In other words, the agent takes the opposite action to the one that the principal would take had he not hired an agent, more often than the first best.

**Characterization of the optimal compensation contract: 4-step process**

The derivation of the optimal contract can be analyzed in 4 steps:

1. Characterization of the decision rule, $\mathcal{DR} : s \mapsto d$, which maximizes principal’s profits.

2. Characterization of the contract which minimizes the expected compensation of the agent, subject to the constraint that a given decision rule $\mathcal{DR}$ is implemented. This step effectively derives the principal’s expected cost of implementing a decision rule $\mathcal{DR}$, which we denote as $E C(\mathcal{DR})$.

3. Derivation of the expected revenue of implementing a decision rule $\mathcal{DR}$, $E R(\mathcal{DR})$.

4. Characterization of the optimal decision rule $\mathcal{DR}^*$, i.e., the decision rule $\mathcal{DR}$ which maximizes profits, $E \Pi(\mathcal{DR}) = E R(\mathcal{DR}) - E C(\mathcal{DR})$.

Thus, the optimal compensation contract is the constrained-optimal compensation contract of step 2, subject to the constraint that $\mathcal{DR} \equiv \mathcal{DR}^*$, where $\mathcal{DR}^*$ is characterized in step 4.

### 4.1 Step 1: Decision Rule $\mathcal{DR}$: cut-off $\hat{s}$.

In this step, we seek the mapping from the signal realization to the implemented action, i.e. $\mathcal{DR} : s \mapsto d$, which maximizes principal’s profits. Lemma 1 captures that the principal’s desired action is characterized by a monotonic relation.

**Lemma 1**: Given any compensation contract which implies non-negative expected profits: i) if the principal prefers the agent to take action $g$ for $s = s'$, then he also prefers the agent to take the same action for any $s > s'$, ii) if action $b$ is preferred for $s = s'$, this also holds true for any $s < s'$, and iii) there always exists a unique $s$, denoted by $\hat{s}$, where the principal is indifferent between the agent taking action $g$ or $b$.

**Proof.** See Appendix C.
Lemma 1 is an implication of the MLRP, according to which the higher the signal realization, the higher the probability that \( g \) is the profit-maximizing action. A direct consequence of Lemma 1 is that the decision rule is characterized by a threshold \( \hat{s} \), such as:

\[
\mathcal{DR} = \begin{cases} 
  d = g & \text{if } s > \hat{s} \\
  d = b & \text{if } s < \hat{s}
\end{cases}
\]

Hence, a particular decision rule corresponds to a particular threshold \( \hat{s} \), and vice versa.

### 4.2 Step 2: Constrained-optimal contract that implements cut-off \( \hat{s} \)

Here we characterize the constrained-optimal compensation contract, i.e., the compensation contract which minimizes the expected compensation of the agent, subject to the constraint that cut-off \( \hat{s} \) is implemented.

Suppose that the principal designs a contract such that the agent acquires information, and then implements \( \hat{s} \). The optimal contract should satisfy two sets of constraints: the agent should prefer acquiring a signal, and given that a signal is obtained, the agent should prefer implementing \( \hat{s} \), to any other mapping from the signal realization to the implemented action.

**Constraints for implementing \( \mathcal{DR} \)**

Implementing \( \hat{s} \) implies that the following two sets of constraints are satisfied. The first set of constraints guarantees that the agent prefers action \( g \) for any \( s' \in [\hat{s}, 1] \).

\[
E[V[d = g|s']] \geq E[V[d = b|s']] \implies Pr(\theta = G|s')[w_{gG} - w_{bG}] \geq Pr(\theta = B|s')[w_{bB} - w_{gB}].
\]

The second set of constraints guarantees that the agent prefers action \( b \) for any \( s'' \in [0, \hat{s}] \).

\[
E[V[d = g|s'']] \leq E[V[d = b|s'']] \implies Pr(\theta = G|s'')[w_{gG} - w_{bG}] \leq Pr(\theta = B|s'')[w_{bB} - w_{gB}].
\]

Since \( Pr(\theta = G|\hat{s}) \) is increasing in \( \hat{s} \), the previous constraints are not mutually exclusive if:

\[
w_{gG} - w_{bG} \geq 0 \quad (5)
\]

\[
w_{bB} - w_{gB} \geq 0 \quad (6)
\]

Relations (5) and (6) imply that payment which corresponds to the ex-post revenue-maximizing action should not be lower than the payment which corresponds to the opposite action. As
long as (5) and (6) hold, the previous two sets of constraints reduce to:

$$Pr(\theta = G | \hat{s}) [w_G - w_{bG}] = Pr(\theta = B | \hat{s}) [w_{bB} - w_{gB}] .$$  

(7)

That is, for \( s = \hat{s} \), the agent is indifferent between taking action \( g \) and action \( b \).

**Information acquisition constraints**

Relation (8) provides the agent’s expected utility under his outside option of taking an action without acquiring information.

$$EV[\text{no signal}] = \max \{ EV[\text{no signal} & d = g], EV[\text{no signal} & d = b] \}$$  

(8)

where the expected utility of each action is given by:

$$EV[\text{no signal} & d = g] = p \times w_G + (1 - p) \times w_{gB}$$  

$$EV[\text{no signal} & d = b] = p \times w_{bG} + (1 - p) \times w_{bB}.$$  

Note that the agent has no private information, thus, his beliefs about the state of the world equal his prior. In the derivation of his expected utility when acquiring information, the agent takes into account the decision rule that he anticipates following after the signal is obtained, i.e., taking action \( g \) for signals \( s > \hat{s} \) and taking action \( b \) otherwise. The expected utility equals the expected compensation reduced by the cost of acquiring information, \( c \):

$$EV[\text{signal} | \hat{s}] = -c + \int_0^{\hat{s}} EC[d = b | s] f(s) ds + \int_{\hat{s}}^1 EC[d = g | s] f(s) ds$$  

(9)

where \( f(s) = pf_G(s) + (1 - p)f_B(s) \). The expected compensation of each action is given by:

$$EC[d = g | s] = Pr(G | s) \times w_G + Pr(B | s) \times w_B$$  

(10)

$$EC[d = b | s] = Pr(G | s) \times w_{bG} + Pr(B | s) \times w_{bB}.$$  

(11)

By substituting (10) and (11) into (9), and applying Bayes rule, we obtain:

$$EV[\text{signal} | \hat{s}] = -c + pF_G(\hat{s})w_{bG} + (1 - p)F_B(\hat{s})w_{bB} + p(1 - F_G(\hat{s}))w_G + (1 - p)(1 - F_B(\hat{s}))w_{gB}$$  

(12)

Hence, the information acquisition constraints become:

$$EV[\text{signal} | \hat{s}] \geq EV[\text{no signal} & d = g] \implies$$

$$pF_G(\hat{s})(w_{bG} - w_{gG}) + (1 - p)F_B(\hat{s})(w_{bB} - w_{gB}) \geq c$$  

(13)

$$EV[\text{signal} | \hat{s}] \geq EV[\text{no signal} & d = b] \implies$$

$$p(1 - F_G(\hat{s}))(w_{gG} - w_{bG}) + (1 - p)(1 - F_B(\hat{s}))(w_{gB} - w_{bB}) \geq c.$$  

(14)
Cost Minimization Problem
Following the previous analysis, the cost minimization problem of incentivizing information acquisition, given a decision rule which is characterized by a threshold \( \hat{s} \), is:

Minimize \( w_bG, w_bB, w_gG, w_gB \) \[ EC(\hat{s}) \]  
\[ pF_G(\hat{s})(w_{bG} - w_{gG}) + (1-p)F_B(\hat{s})(w_{bB} - w_{gB}) \geq c \]  
\[ p(1 - F_G(\hat{s}))(w_{gG} - w_{bG}) + (1-p)(1 - F_B(\hat{s}))(w_{gB} - w_{bB}) \geq c \]  
\[ f_G(\hat{s})p(\hat{s})(w_{gG} - w_{bG}) = f_B(\hat{s})(1-p)[w_{bB} - w_{gB}] \]  
\[ w_{gG} - w_{bG} \geq 0 \]  
\[ w_{bB} - w_{gB} \geq 0 \]  
\[ w_{bG} \geq 0, \ w_{bB} \geq 0, \ w_{gG} \geq 0, \ w_{gB} \geq 0 \]  

where \( EC(\hat{s}) \) is given by:

\[ \mathbb{E}C(\hat{s}) = pF_G(\hat{s})w_{bG} + (1-p)F_B(\hat{s})w_{bB} + p(1 - F_G(\hat{s}))w_{gG} + (1-p)(1 - F_B(\hat{s}))w_{gB} \]  

Proposition 1 provides the optimal contract, subject to the constraint that decision rule is characterized by threshold \( \hat{s} \).

**Proposition 1:** Constrained Optimal Contract.
If \( \hat{s} \leq \hat{s}_{\text{min}} \), the constrained optimal contract is given by:

\[ w^*_b(\hat{s}) = \frac{f_G(\hat{s})}{(1-p)(F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s}))}c \]  
\[ w^*_g(\hat{s}) = \frac{f_B(\hat{s})}{p(F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s}))}c \]  
\[ w^*_b(\hat{s}) = w^*_g(\hat{s}) = 0. \]

If \( \hat{s} \geq \hat{s}_{\text{min}} \), the constrained optimal contract is given by:

\[ w^*_b(\hat{s}) = \frac{f_G(\hat{s})}{(1-p)[(1 - F_G(\hat{s}))f_B(\hat{s}) - (1 - F_B(\hat{s}))f_G(\hat{s})]}c \]  
\[ w^*_g(\hat{s}) = \frac{f_B(\hat{s})}{p[(1 - F_G(\hat{s}))f_B(\hat{s}) - (1 - F_B(\hat{s}))f_G(\hat{s})]}c \]  
\[ w^*_g(\hat{s}) = w^*_b(\hat{s}) = 0, \]  

where \( \hat{s}_{\text{min}} \) solves \( f_B(\hat{s}_{\text{min}}) = f_G(\hat{s}_{\text{min}}) \).

*Proof.* See Appendix C.
Three points are relevant to the intuition for the constrained optimal contract.

1st Remark: Note that the incentive constraints in the minimization problem are expressed in terms of the difference between the payment which corresponds to the revenue-maximizing action and the payment which corresponds to the opposite action; That is, $w_{bB} - w_{gB}$, if the realized state of the world is $B$, and $w_{gG} - w_{bG}$, if the realized state of the world is $G$. Following this remark, we can show that under the optimal contract, the payment when the agent does not take the revenue-maximizing action is zero, i.e., $w_{gB}^* = w_{bG}^* = 0$. For $w_{gB}^* = 0$, the intuition is that by decreasing the payment which corresponds to action $g$ ($w_{gB}$) and action $b$ ($w_{bB}$) by the same amount, the principal decreases the expected compensation of the agent without affecting the incentive constraints. Such a deviation is feasible until $w_{gB}$ hits its low bound. Similar intuition applies for the optimality of $w_{bG}^* = 0$.

2nd Remark: By relation (15), the relative pay $w_{bB}/w_{gG}$ should be equal to $Pr(\theta = G|\hat{s})/Pr(\theta = B|\hat{s})$, which is determined by the choice of $\hat{s}$. Note that $w_{bB}/w_{gG}$ is increasing in $\hat{s}$. The intuition behind this remark is that as $\hat{s}$ increases, the probability that the agent attributes to the state being $G$, for a signal realization $s = \hat{s}$, increases as well. Thus, the relative payment that the manager requires to take action $b$ for $s = \hat{s}$, increases with $\hat{s}$, in order to compensate for the lower probability that this payment will be realized. Similar intuition applies when $\hat{s}$ decreases; the main difference is that as $\hat{s}$ decreases, the manager requires a higher premium for going long.

3rd Remark: The last remark relates to the information acquisition constraints. It can be shown that if $s < \hat{s}_{min}$, (14) becomes redundant, and under the optimal contract, (13) binds. The opposite holds when $s > \hat{s}_{min}$. The intuition is the following. First, note that the take-it-or-leave-it nature of the contract implies that under the optimal contract, the agents’ utility when acquiring information and implementing $\hat{s}$ should be equal to his outside option. Suppose now that the principal aims to implement a low value of $\hat{s}$. For this value of $\hat{s}$, the agent attributes a high probability to the state being $B$, and, following the second remark, requires a relatively high payment $w_{gG}/w_{bB}$ to take action $g$. For this ratio of payments, and conditional on not acquiring information, the agent’s utility when taking action $g$ exceeds his utility when taking action $b$. Hence (13) is more restrictive than (14).

The following lemma explores how the optimal payment $w_{bG}^*(\hat{s})$ and $w_{gG}^*(\hat{s})$ relates to $\hat{s}$.

---

4It is worth highlighting that relaxing the agent’s limited liability assumption would not affect the optimal contract qualitatively. This is captured by the first remark. Hence, if the agent has an initial wealth of $\bar{w}$, and $\bar{w}$ is common knowledge, the payments of the optimal contract would be: $w_{bG}^*(\hat{s}) = -\bar{w}$, $w_{gB}^*(\hat{s}) = -\bar{w}$, $w_{bB}^*(\hat{s}) = w_{bB}^*(\hat{s}) - \bar{w}$, and $w_{gG}^*(\hat{s}) = w_{gG}^*(\hat{s}) - \bar{w}$. 

14
Lemma 2: Relationship between optimal payments and $\hat{s}$.

The optimal payments $w^*_b(\hat{s})$ and $w^*_g(\hat{s})$ are:

(i) decreasing in $\hat{s}$, for $\hat{s} \leq \hat{s}_{\text{min}}$.

(ii) increasing in $\hat{s}$, for $\hat{s} \geq \hat{s}_{\text{min}}$.

(iii) minimized for $\hat{s}_{\text{min}}$ such as $f_G(\hat{s}) = f_B(\hat{s})$.

(iv) convex in $\hat{s}$.

Proof. See Appendix C.

![Figure 1: Optimal payments, $w^*_b(\hat{s})$ and $w^*_g(\hat{s})$, for $p = 0.5$, $f_G(s) = 2s$, $f_B(s) = 2(1 - s)$.](image)

Critical for the underlying mechanism in Lemma 2 is the role of the payments as opportunity cost. In fact, when the agent takes action $b$ and the revenue-maximizing action is $g$, the opportunity cost of taking the “wrong” action coincides with $w_g^*$. Likewise, when the agent takes action $g$ and the revenue-maximizing action is $b$, the opportunity cost coincides with $w_b^*$.

In order to capture the main mechanism, we explore the case where: i) the principal considers switching from the optimal contract which implements $\hat{s} = \hat{s}_{\text{min}}$ to the optimal contract which implements $\hat{s} = \hat{s}' > \hat{s}_{\text{min}}$, and ii) $p = 0.5$. For $p = 0.5$, the ratio $w_b^*/w_g^*$ that implements $\hat{s}_{\text{min}}$ equals one. Note that increasing $\hat{s}$ leads to a higher value of $\Pr(\theta = G | \hat{s}) / \Pr(\theta = B | \hat{s})$. Thus, by (15), the increase in $\hat{s}$ should be accompanied by an increase in the ratio $w_b^*/w_g^*$. This implies that the principal, to prevent the agent from taking action $g$ for $s \in [\hat{s}_{\text{min}}, \hat{s}']$, should increase $w_b^*$. However, the increase in $w_b^*$ increases the agent’s utility when taking action $b$ without acquiring information, which now exceeds the agent’s utility when acquiring information. The only way the principal can prevent the agent from taking action $b$ without acquiring information is by increasing the opportunity cost of taking the wrong action, which is achieved by increasing $w_g^*$. Hence, a deviation from $\hat{s} = \hat{s}_{\text{min}}$ to $\hat{s} = \hat{s}' > \hat{s}_{\text{min}}$ increases both $w_b^*$ and $w_g^*$.

---

5We show in Appendix C that decreasing $w_g^*$ instead of increasing $w_b^*$ is not feasible, because this would contradict with the optimality of the initial contract which implements $\hat{s} = \hat{s}_{\text{min}}$.

6Similar intuition applies when the principal switches from $\hat{s} = \hat{s}_{\text{min}}$ to $\hat{s} = \hat{s}'' < \hat{s}_{\text{min}}$. 

15
From the constrained-optimal to the optimal contract.
Recall that the optimal contract is the constrained-optimal contract that implements a threshold \( \hat{s} = \hat{s}^* \), where \( \hat{s}^* \) is the value of \( \hat{s} \) which maximizes the expected profits of the principal:

\[
\mathbb{E} \Pi(\hat{s}) = \mathbb{E} R(\hat{s}) - \mathbb{E} C(\hat{s}).
\]  

(20)

where \( \mathbb{E} C(\hat{s}) \) denotes the expected compensation cost of implementing \( \hat{s} \) (derived in subsection 4.3), and \( \mathbb{E} R(\hat{s}) \) denotes the expected revenue of implementing \( \hat{s} \) (derived in subsection 4.4).

4.3 Expected compensation cost of implementing \( \hat{s} \), \( \mathbb{E} C(\hat{s}) \).
The constrained optimal contract enables us to derive the expected cost of implementing \( \hat{s} \), \( \mathbb{E} C(\hat{s}) \), which is provided in Corollary 1.

Corollary 1: Expected cost of implementing \( \hat{s} \), \( \mathbb{E} C(\hat{s}) \).
If \( \hat{s} \leq \hat{s}_{\text{min}} \), the expected cost is given by:

\[
\mathbb{E} C(\hat{s}) \equiv \mathbb{E} C(\hat{s})^- = \left[ \frac{F_B(\hat{s})f_G(\hat{s}) + (1 - F_G(\hat{s}))f_B(\hat{s})}{F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s})} \right] c
\]

If \( \hat{s} \geq \hat{s}_{\text{min}} \), the expected cost is given by:

\[
\mathbb{E} C(\hat{s}) \equiv \mathbb{E} C(\hat{s})^+ = \left[ \frac{F_B(\hat{s})f_G(\hat{s}) + (1 - F_G(\hat{s}))f_B(\hat{s})}{(1 - F_G(\hat{s}))f_B(\hat{s}) - (1 - F_B(\hat{s}))f_G(\hat{s})} \right] c
\]

The first remark derived from Corollary 1 is that \( \mathbb{E} C(\hat{s}) \) is linearly dependent on the cost of acquiring information, \( c \). The second remark is about the behavior of \( \mathbb{E} C(\hat{s}) \) with respect to threshold \( \hat{s} \), which is summarized in Lemma 3.

Lemma 3: Relationship between \( \mathbb{E} C(\hat{s}) \) and \( \hat{s} \).

(i) For each \( \hat{s} \in [0, \hat{s}_{\text{min}}) \), \( \mathbb{E} C(\hat{s}) \) is decreasing in \( \hat{s} \).

(ii) For each \( \hat{s} \in (\hat{s}_{\text{min}}, 1] \), \( \mathbb{E} C(\hat{s}) \) is increasing in \( \hat{s} \).

(iii) \( \mathbb{E} C(\hat{s}) \) is minimized for \( \hat{s}_{\text{min}} \) such as \( f_G(\hat{s}) = f_B(\hat{s}) \).

(iv) \( \mathbb{E} C(\hat{s}) \) is convex in \( \hat{s} \).

Proof. See Appendix.

The intuition behind Lemma 3 is identical to the intuition in Lemma 2. Before we proceed to the next section, it is worth providing the rationale behind part three of Lemma 3, as it
plays a critical role in understanding the main features of the optimal contract. First, note that \( EC(\hat{s}) \) does not depend on \( p \), although \( p \) affects the optimal payments. This is because the impact of \( p \) on each optimal payment is offset by the probability that this payment will be realized. Intuitively, the publicly available information, which is captured by \( p \), is internalized and therefore irrelevant to the cost of incentivizing information acquisition: what matters is the agent’s private information. As a result, the cost is minimized when the agent implements action \( g \) only for signal realizations \( s \) which are more likely to correspond to state \( G \) than state \( B \), i.e., for \( s \) such as \( f_G(s) > f_B(s) \), denoted as \( \hat{s}_{min} \). Thus the cost-minimizing decision rule is implementing action \( g \) for \( s \in [0, \hat{s}_{min}) \) and action \( b \) for \( s \in (\hat{s}_{min}, 1] \).

The last step is to derive the expected revenue of implementing \( \hat{s} \), which combined with the expected cost of implementing \( \hat{s} \), will allow us to characterize the optimal threshold, \( \hat{s}^* \).

\[ \begin{align*}
\text{Figure 2: Expected cost of implementing } \hat{s}, \ E\!\!C(\hat{s}), \\
\text{for } f_G(s) = 2s, f_B(s) = 2(1-s).
\end{align*} \]

### 4.4 Step 3: Expected revenue of implementing \( \hat{s} \), \( E\!\!R(\hat{s}) \)

Suppose that the principal offers a contract which implements \( \hat{s} \). When forming his beliefs about the expected revenue, he anticipates the agent to take action \( g \) when the signal is greater than \( \hat{s} \), and to take action \( b \) otherwise. Hence, the expected revenue of implementing \( s = \hat{s} \) is:

\[
E\!\!R(\hat{s}) = \int_0^{\hat{s}} E\!\!R[d = b|s]f(s)ds + \int_{\hat{s}}^1 E\!\!R[d = g|s]f(s)ds
\]

where \( f(s) = pf_G(s) + (1-p)f_B(s) \). Also, the principal’s expected revenue when the agent implements action \( b \) and \( g \), given a signal \( s \), is given by:

\[
E\!\!R[d = b|s] = Pr(\theta = G|s)R(d = b, \theta = G) + Pr(\theta = B|s)R(d = b, \theta = B) \\
E\!\!R[d = g|s] = Pr(\theta = G|s)R(d = g, \theta = G) + Pr(\theta = B|s)R(d = g, \theta = B)
\]
By substituting (22) and (23) into (21), and applying the Bayes rule, we obtain:

$$E_R(\hat{s}) = v\{p(1 - 2F_G(\hat{s})) + (1 - p)(2F_B(\hat{s}) - 1)\} \quad (24)$$

Lemma 4 provides the relationship between the expected revenue and threshold $\hat{s}$.

**Lemma 4** Relationship between $E_R$ and $\hat{s}$.

(i) For each $\hat{s} \in [0, \hat{s}^{FB}]$, $E_R(\hat{s})$ is increasing and concave in $\hat{s}$.

(ii) For each $\hat{s} \in (\hat{s}^{FB}, 1]$, $E_R(\hat{s})$ is decreasing and concave in $\hat{s}$.

(iii) The expected revenue $E_R(\hat{s})$ is single-peaked at $\hat{s} = \hat{s}^{FB}$, where $\hat{s}^{FB}$ solves:

$$\frac{f_G(\hat{s})}{f_B(\hat{s})} = \frac{(1 - p)}{p} \quad (25)$$

**Proof.** See Appendix C.

![Figure 3: Expected revenue of implementing $\hat{s}$, $E_R(\hat{s})$, for $f_G(s) = 2s$ and $f_B(s) = 2(1 - s)$.](image)

1st Remark: It is worth highlighting that $\hat{s}^{FB}$ corresponds to the signal realization for which the expected utility of implementing action $g$ and action $b$ are equal. Given the symmetry of the revenue structure, $\hat{s}^{FB}$ is the signal realization which is associated with posterior beliefs equal to 0.5. We show in the Appendix that the signal realization which solves (25) coincides with the threshold that the principal would implement if there were no agency problem.

2nd Remark: The concavity of $E_R(\hat{s})$ is an implication of the MLRP. Recall that by applying Bayes rule, $\frac{Pr(\theta = G|s)}{Pr(\theta = B|s)}$ equals $\frac{f_G(s)}{f_B(s)} \times \frac{p}{1 - p}$. Hence, the lower the value of $\hat{s}$ compared to $\hat{s}^{FB}$, the higher the expected distortion of taking action $g$ in the area $s \in (\hat{s}, \hat{s}^{FB})$, and the more likely
that \( b \) is the revenue-maximizing action. Similar intuition applies for deviations to \( \hat{s} > \hat{s}^{FB} \).

3rd Remark: A clear implication of (25) is that the peak of \( \mathbb{E} R(\hat{s}) \) moves towards lower \( \hat{s} \) as the value of \( p \) increases. The intuition is straightforward: the higher the prior belief \( p \), the higher the prior probability that the principal attributes to action \( g \) being revenue-maximizing. Thus, for high values of \( p \), the principal prefers the agent to take action \( g \) unless he receives strong evidence that \( b \) is the revenue-maximizing action, i.e., a very low signal. The opposite holds for low values of \( p \).

4.5 Step 4: Optimal threshold \( \hat{s}^* \)

The goal of this section is to characterize the optimal threshold \( \hat{s}^* \), which maximizes the expected profit of the principal, \( \mathbb{E} \Pi(\hat{s}) \), where,

\[
\mathbb{E} \Pi(\hat{s}) = \mathbb{E} R(\hat{s}) - \mathbb{E} C(\hat{s})
\]

By Lemma 3, the expected cost \( \mathbb{E} C(\hat{s}) \) is minimized for \( \hat{s} = \hat{s}_{min} \), such as:

\[
\frac{f_G(\hat{s})}{f_B(\hat{s})} = 1 \tag{26}
\]

Also, by Lemma 4, the expected revenue \( \mathbb{E} R(\hat{s}) \) is maximized for \( \hat{s} = \hat{s}^{FB} \), such as:

\[
\frac{f_G(\hat{s})}{f_B(\hat{s})} = \frac{(1 - p)}{p} \tag{27}
\]

Thus, as long as the principal prefers one of the two actions, had he not hired an agent (i.e., when \( p \neq 0.5 \)), the value of \( \hat{s} \) which maximizes the expected revenue, \( \hat{s}^{FB} \), differs from the value which minimizes the expected cost, \( \hat{s}_{min} \). This finding highlights that the optimal value of \( \hat{s} \) relies on an inherent trade-off between minimizing the cost of incentivizing information acquisition, and the benefit of implementing a threshold \( \hat{s} \), which is as close as possible to the revenue-maximizing threshold, \( \hat{s}^{FB} \).

**Optimality condition for \( \hat{s} \).**

The principal offers a contract that corresponds to a threshold \( \hat{s} \) which optimally resolves this trade-off. Thus, for a threshold \( \hat{s} \) to be optimal, the following condition needs to hold:

\[
\frac{\partial \mathbb{E} \Pi(\hat{s})}{\partial \hat{s}} = \frac{\partial \mathbb{E} R(\hat{s})}{\partial \hat{s}} - \frac{\partial \mathbb{E} C(\hat{s})}{\partial \hat{s}} = 0 \tag{28}
\]

\(^8\)Relaxing the symmetry assumption implies that the critical value of \( p \) is the one which solves \( \frac{f_G(\hat{s})}{f_B(\hat{s})} = 1 \), which effectively is the value of \( p \) which solves \( \frac{1-p}{p} = \frac{R(d=b, \theta=B) - R(d=g, \theta=B)}{R(d=g, \theta=G) - R(d=g, \theta=G)} \).
Condition (28) captures the idea that for the optimal value of \( \hat{s} \), denoted by \( \hat{s}^* \), the expected marginal revenue \( (\partial E R(\hat{s})/\partial \hat{s}) \) should equal the expected marginal cost \( (\partial E C(\hat{s})/\partial \hat{s}) \).

4.6 Optimal compensation contract

The optimal compensation contract \( \hat{W}^*(\hat{s}^*) \) is the constrained-optimal compensation contract of Proposition 1, subject to the constraint that \( \hat{s} \equiv \hat{s}^* \).

Proposition 2: Optimal Contract, \( \hat{W}^*(\hat{s}^*) \)

If \( \hat{s}^F \leq \hat{s}_{\text{min}} \) (i.e., \( p \geq \hat{p} = 0.5 \)), the optimal value of \( \hat{s} \), \( \hat{s}^* \) solves:

\[
2v\{-pf_G(\hat{s}) + (1 - p)f_B(\hat{s})\} = c \frac{f_B(\hat{s})[f'_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})f'_G(\hat{s})]}{(f_B(\hat{s})f_G(\hat{s})f_B(\hat{s})f_G(\hat{s}))^2}
\]

(29)

and the optimal payments are given by:

\[
w^*_{bb}(\hat{s}^*) = \frac{f_G(\hat{s}^*)}{(1 - p)(f_B(\hat{s}^*)f_G(\hat{s}^*) - f_G(\hat{s}^*)f_B(\hat{s}^*))}c
\]

\[
w^*_{gg}(\hat{s}^*) = \frac{f_B(\hat{s}^*)}{p(f_B(\hat{s}^*)f_G(\hat{s}^*) - f_G(\hat{s}^*)f_B(\hat{s}^*))}c
\]

\[
w^*_{gb}(\hat{s}^*) = w^*_{bg}(\hat{s}^*) = 0.
\]

If \( \hat{s}^F \geq \hat{s}_{\text{min}} \) (i.e., \( p \leq \hat{p} = 0.5 \)), the optimal value of \( \hat{s} \), \( \hat{s}^* \) solves:

\[
2v\{-pf_G(\hat{s}) + (1 - p)f_B(\hat{s})\} = -c \frac{(f_B(\hat{s})f'_G(\hat{s}) - f'_B(\hat{s})f_G(\hat{s}))(-1 + F_G(\hat{s}))}{[f_G(\hat{s}) - f_B(\hat{s})f_G(\hat{s}) + f_B(\hat{s})(-1 + F_G(\hat{s}))]^2}
\]

(30)

and the optimal payments are given by:

\[
w^*_{bb}(\hat{s}^*) = \frac{f_G(\hat{s}^*)}{(1 - p)[(1 - F_G(\hat{s}^*))f_B(\hat{s}^*) - (1 - F_B(\hat{s}^*))f_G(\hat{s}^*)]}c
\]

\[
w^*_{gg}(\hat{s}^*) = \frac{f_B(\hat{s}^*)}{p[(1 - F_G(\hat{s}^*))f_B(\hat{s}^*) - (1 - F_B(\hat{s}^*))f_G(\hat{s}^*)]}c
\]

\[
w^*_{gb}(\hat{s}^*) = w^*_{bg}(\hat{s}^*) = 0.
\]

Proof. See Appendix C.

It is worth mentioning that the critical value, \( p = 0.5 \) relies on the assumption of the symmetric revenue structure. For a more general revenue structure, the critical value of \( p \),
denoted as \( \hat{p} \), solves:

\[
\begin{align*}
 pR(d = g, \theta = G) + (1 - p)R(d = g, \theta = B) = pR(d = b, \theta = G) + (1 - p)R(d = b, \theta = B)
\end{align*}
\]

Thus, as long as \( p > \hat{p} \), \( \hat{s}^{FB} < \hat{s}_{\text{min}} \), whereas if \( p < \hat{p} \), \( \hat{s}^{FB} > \hat{s}_{\text{min}} \), where \( \hat{p} \) is given by:

\[
\hat{p} = \frac{(R(d = b, \theta = B) - R(d = g, \theta = B))}{(R(d = g, \theta = G) - R(d = b, \theta = G)) + (R(d = b, \theta = B) - R(d = g, \theta = B))}.
\]

### 4.7 When is incentivizing information acquisition profitable?

Recall that the previous analysis refers to the case where the principal finds it optimal to incentivize the agent to acquire information. However, the principal has the option of offering a contract which does not require the agent to obtain information. If the agent does not hold any private information, the principal would prefer the agent to take action \( g \) if \( p > 0.5 \), and action \( b \) otherwise. We can show that the principal could implement this decision rule by offering a contract, denoted \( \hat{W}' \), which pays an arbitrarily small amount \( \tilde{\eta} \to 0 \), if the agent follows the aforementioned strategy, and zero otherwise. Lemma 7 characterizes the condition which needs to hold for it to be optimal for the principal to incentivize information acquisition.

**Lemma 7:**

The principal finds it optimal to incentivize the agent to acquire information as long as:

\[
2v\{(1 - p)F_B(\hat{s}^*) - pF_G(\hat{s}^*) - 2p\} \geq \frac{F_B(\hat{s}^*)f_G(\hat{s}^*) + (1 - F_G(\hat{s}^*))f_B(\hat{s}^*)}{F_B(\hat{s}^*)f_G(\hat{s}^*) - F_G(\hat{s}^*)f_B(\hat{s}^*)}c 
\]

\[
2v\{(1 - p)F_B(\hat{s}^*) - pF_G(\hat{s}^*) - 2(1 - p)\} \geq \frac{F_B(\hat{s}^*)f_G(\hat{s}^*) + (1 - F_G(\hat{s}^*))f_B(\hat{s}^*)}{(1 - F_G(\hat{s}^*))f_B(\hat{s}^*) - (1 - F_B(\hat{s}^*))f_G(\hat{s}^*)}c
\]

where (33) (34) corresponds to the case where \( p > 0.5 \) \( p < 0.5 \). Note that \( 2v \) reflects the benefit of taking the “right” over the “wrong” action. Note also that the term in the curly brackets is the expected increase in the probability of taking the “right” action due to the signal acquisition. Thus, the LHS captures the principal’s expected benefit of acquiring information, whereas the RHS is the expected compensation cost, derived in Corollary 1.

---

If the realized state is good and the agent takes action \( g \), principal’s net return is \( v \). In contrast, if the agent takes action \( b \), principal’s net return is \( -v \). Likewise, if the realized state is bad and the agent takes action \( b \), principal’s net return is \( v \). In contrast, if the agent takes action \( g \), principal’s net return is \( -v \). Thus, independently of the state of the world, the net benefit of taking the right over the wrong action is \( 2v \).
5 Implications of the Optimal Contract

In this section, we explore the main implications of the optimal contract. We focus on the most interesting case where the principal aims to incentivize the agent to acquire information.

5.1 Premium against the flow & bias in the implemented action

Proposition 3 explores the relation between the optimal payment when the agent “goes against the flow” and the optimal payment when the agent “follows the flow”. By “following the flow” we define the case where the agent takes the same action as the one that the principal would take, had he not hired an agent. Thus, “following the flow” means implementing the action which corresponds to the highest ex-ante expected revenue, i.e., action $g$ when $p > 0.5$, and action $b$ otherwise.

**Proposition 3:** Premium for going against the flow.

Under the optimal contract:

(i) If $\hat{s}_{FB} < \hat{s}_{min}$ (i.e., $p > \hat{\rho} = 0.5$), then $w_{bB}(\hat{s}^*) > w_{gG}(\hat{s}^*)$.

(ii) If $\hat{s}_{FB} > \hat{s}_{min}$ (i.e., $p < \hat{\rho} = 0.5$), then $w_{gG}(\hat{s}^*) > w_{bB}(\hat{s}^*)$.

**Proof.** See Appendix C.

Hence, under the optimal contract, there is a premium for going against the flow. Proposition 4 explores how $\hat{s}^*$ relates to: i) the optimal value of $\hat{s}$ if there is no agency problem (first best), $\hat{s}_{FB}$, and ii) the value which minimizes the cost of incentivizing information acquisition, $\hat{s}_{min}$.

**Proposition 4:** Bias in the implemented action.

Under the optimal contract:

(i) If $\hat{s}_{FB} < \hat{s}_{min}$ (i.e., $p > \hat{\rho} = 0.5$), then $\hat{s}_{FB} < \hat{s}^* \leq \hat{s}_{min}$.

(ii) If $\hat{s}_{FB} > \hat{s}_{min}$ (i.e., $p < \hat{\rho} = 0.5$), then $\hat{s}_{FB} > \hat{s}^* \geq \hat{s}_{min}$.

**Proof.** See Appendix C.

The proof of Proposition 3 and Proposition 4 can be captured in Figure 4, which illustrates Lemma 3 and Lemma 4 for $f_G = 2s$ and $f_B = 2(1 - s)$. Panel A of Figure 4 presents the case where $p > 0.5$. Note that for these values of $p$, a threshold $\hat{s}' \in [0, \hat{s}_{FB}]$ cannot be optimal, because switching to $\hat{s}'' = \hat{s}' + \eta$, where $\eta$ is a small positive number, decreases the expected compensation cost and increases the expected revenue. Likewise, a threshold $\hat{s}' \in [\hat{s}_{min}, 1]$ cannot be optimal, because switching to $\hat{s}'' = \hat{s}' - \eta$ decreases the expected compensation cost and increases the expected revenue. Thus, for $p > 0.5$, $\hat{s}^* \in (\hat{s}_{FB}, \hat{s}_{min}]$. Similar intuition applies for the case where $p < 0.5$, which is illustrated in Panel B.
Panel A: Case where \(p > 0.5\), flow $\rightarrow$ action \(g\). Panel B: Case where \(p < 0.5\), flow $\rightarrow$ action \(b\).

Figure 4: Bias against the flow. Case where \(f_G(s) = 2s\), \(f_B(s) = 2(1-s)\).

Proposition 4 captures the key feature of the optimal contract: the agent goes against the flow more often than the first best. The intuition behind the premium and the bias for going against the flow relies on the interaction of unobservable information acquisition and the multitasking nature of the problem. In order to incentivize the agent to acquire information, the principal should promise a positive payment only if the ex-post profit maximizing action is taken. However, the compensation contract – apart from affecting the incentive of the agent to acquire information – also affects his implemented action, which in turn, determines the principal’s expected revenue.

A critical property of the optimal contract is that the agent’s utility should coincide with his outside option, i.e.,

\[
-c + (1-p)F_B(\hat{s})w_{\text{gB}}^* + p(1-F_G(\hat{s}))w_{\text{gG}}^* = \max\{ EV[\text{signal} | \hat{s}], (1-p)w_{\text{gB}}^*, (1-p)w_{\text{gG}}^* \} \tag{35}
\]

Also, in this setting, the expected revenue is maximized for \(w_{gG}^* = w_{bB}^*\). For such payments, the outside option of the agent is to follow the flow without acquiring information. Note that for \(p > 0.5\), by decreasing the payment when the flow is followed \((w_{gG}^*)\), the RHS of (35) decreases more than the LHS, which in turn allows the principal to decrease also \(w_{bB}^*\) without violating this condition. Thus, the principal, by worsening the agent’s outside option (through lowering \(w_{gG}^*)\), can incentivizing information acquisition at a lower cost. However, changing the ratio of payments comes at the cost of decreasing the principal’s expected revenue; for a ratio of payments different than one, the first best is no longer implemented. The optimal contract thus solves the inherent trade-off between the cost of incentivizing learning and benefit of implementing a threshold as close as possible to the revenue-maximizing one. We show that for small deviations from the first best, the decrease in the expected cost is greater than the decrease in the expected revenue.
Following the intuition provided in the previous paragraph, it is worth highlighting that what matters for the direction of the bias is not the whether state \( B \) is ex-ante more or less likely than state \( G \); the direction of the bias depends on which action the principal would implement, had he not hired an agent. In other words, what matters is the relation between the principal’s ex-ante expected revenue of each action rather than the ex-ante probability that each action is revenue-maximizing.\(^{10}\) For instance, suppose that \( p \) is close to one. As long as implementing action \( b \) generates sufficiently high revenue such as \( \hat{s}^{FB} > \hat{s}_{\text{min}} \), the optimal contract would imply a premium and over-implementation of the action \( g \), although it is ex-ante more likely to be revenue-maximizing. Of course, assuming a symmetric revenue structure implies that these two interpretations are equal, however, this is not generally true. The rationale is that if the principal has a strict preference for action \( g \) over action \( b \) (either because this action is more likely be profit-maximizing or because it corresponds to a higher revenue), he would prefer the agent to implement action \( g \) even when observing a signal which is more likely to correspond to state \( B \) than state \( G \), i.e., \( f_B(\hat{s}^{FB}) > f_G(\hat{s}^{FB}) \). Hence, \( \hat{s}^{FB} \) is smaller than the value of \( \hat{s} \) which minimizes the cost of incentivizing information acquisition, i.e., \( \hat{s} \) such as \( f_B(\hat{s}) = f_G(\hat{s}) \). Thus, the principal finds it optimal to “scarify” some revenue by switching towards a higher implementation threshold, which corresponds to a lower compensation cost.

Corollary 2 explores the main implications of the bias in the implemented decision on: i) the probability of the agent taking action \( b \) or \( g \), ii) the probability that the implemented action is revenue-maximizing, and iii) the posterior beliefs about the state of the world.

**Corollary 2: Implications of the Optimal Contract compared to First Best.**

If \( \hat{s}^{FB} < \hat{s}_{\text{min}} \), i.e., \( p > \hat{p} = 0.5 \) (\( \hat{s}^{FB} > \hat{s}_{\text{min}} \), i.e., \( p < \hat{p} = 0.5 \)), compared to the case where there is no agency problem, under the optimal contract:

(i) The agent is more (less) likely to take action \( b \).

(ii) Conditional that action \( b \) is implemented, it is less (more) likely to be revenue-maximizing.

(iii) Conditional that action \( g \) is implemented, it is more (less) likely to be revenue-maximizing.

(iv) Conditional on the implemented action, the beliefs about the state being good are higher (lower).

*Proof. See Appendix C.*

The proof and the main intuition of Corollary 2 can be captured in Figure 5, where the solid line depicts the probability that the state is good given a signal realization \( s \), \( Pr(\theta = G|s) \), for the case where \( f_G = 2s \) and \( f_B = 2(1-s) \). To avoid any unnecessary repetition, we only present the case where \( p > 0.5 \) (illustrated in Panel A); similar intuition applies for the case where \( p < 0.5 \) (illustrated in Panel B).

\(^{10}\)This is perfectly captured in relation \( 31 \), and the discussion that follows Proposition 2.
Recall that due the symmetric revenue structure, under the first best, action $b$ is implemented for $s$ such as $Pr(\theta = G|s) \leq 0.5$, and action $g$ is implemented otherwise. Thus, for $p > 0.5$, under the first best, $b$ is implemented in area $D$, and $g$ in areas $E$ and $F$. In contrast, under the optimal contract, the premium for going against the flow results in the agent implementing action $b$ not only in area $D$ but also in area $E$.

Part one follows directly from the last remark of the previous paragraph. The intuition behind part two is that under the optimal contract, the agent is tempted by the premium to take action $b$, even for signal realizations for which it is more likely that action $g$ is revenue-maximizing. This is captured in Panel A, where the agent implements action $b$ also in area $E$, for which the corresponding signals indicate that it is more likely that the state is good. Thus conditional on action $b$ being implemented, the probability of this action being revenue-maximizing is lower than in the case where the first best threshold is implemented.

The rationale for part three follows a similar logic: the agent is willing to implement action $g$ and forgo the premium that action $b$ entails, only if he holds strong evidence that the actual state is $G$. This is captured in Panel A where the agent implements action $g$ only in area $F$, instead of areas $E$ and $F$. Note that compared to area $E$, signal realizations in area $F$ correspond to a higher probability that the action state is good. Thus, conditional on action $g$ being implemented, the probability that this action is revenue-maximizing is higher than in the case where the first best threshold is implemented. Finally, part four is a direct consequence of part two and three.

![Figure 5: Informational role of implemented actions](image)

**5.2 Informational role of implemented action**

In this section we show that the informational value of the implemented action differs, depending on whether the agent goes against the flow, or follows the flow. This finding relates to Corollary 2. Recall that if following the flow means taking action $g$, the conditional probability that the state is good given that action $b$ is implemented is lower than the corresponding probability when there is no agency problem. Thus, action $b$ is a weak indicator that the actual state is $B$.  

25
Following a similar logic, action $g$ is a *strong indicator* that the actual state is $G^{11}$. In contrast, if following the flow means taking action $b$ (for $p < 0.5$), action $b$ is a *strong indicator* that the actual state is bad, whereas action $g$ is a *weak indicator* that the actual state is good.

Another implication of Corollary 2 is that the resulting bias in the implemented action can sustain extreme beliefs about the state of the world. For instance, if the ex-ante prior probability that the state is good is high (i.e., $p$ has a high value), then the posterior beliefs that the state is good, conditional on action $g$ and $b$, are higher compared to the case where the first best threshold is implemented. Intuitively, this is because a high value of $p$ leads to a high premium when action $b$ is implemented, which in turn, tempts the agent to take this action. As a result, after observing action $b$, the beliefs about the probability that the state is good are not radically downgraded. This is because the action $b$ might be chosen not because of a low signal realization, but due to its high premium. Similarly, after observing action $g$, the beliefs about the probability that the state is good are radically upgraded. This relies on the fact that if the agent is willing to forgo the premium that action $b$ entails, it must be that he attributes a high probability on the state being good. Following a similar reasoning, if the prior probability that the state is good is low, then the posterior beliefs that the state is good, conditional on both action $g$ and $b$, are lower compared to the first best.

### 5.3 Financial recommendation setting

Recall that action $g$ corresponds to recommending a long position, whereas action $b$ corresponds to recommending a short position. Also, the state of the world refers to the future asset price: a good state corresponds to a price increase, and a bad state corresponds to a price decrease. Hence, the revenue-maximizing action is recommending a short position when the state is bad, and a long position if the state is good. Finally, the prior beliefs, which can be interpreted as market beliefs, indicate that the asset is undervalued if $p > 0.5$, and overvalued if $p < 0.5$, where an asset is considered undervalued when the expected future price is high than the current price, and vice versa.

Following the aforementioned analogy, and under the optimal contract, the promised payment is positive only: i) when the analyst recommends a long position, and the price eventually increases, and ii) when the analyst recommends a short position, and the price eventually decreases. In all other cases, the payment is zero. Besides, when the prior beliefs indicate that the asset is overvalued, the payment which corresponds to a recommendation of a long position exceeds the payment which corresponds to a recommendation of a short position, and vice versa. This premium leads to *over-investment against the flow*: if the asset is considered undervalued, a short position is recommended more often than the first best, and vice versa.

---

11 This is because the conditional probability that the state is good given that action $g$ is implemented exceeds the corresponding probability, when there is no agency problem.
**Informational role of recommendations**

Note that in practice, and similarly to our model, recommendations of financial analysts reflect their private information, and thus can be used by market participants to infer the future price of assets. A direct consequence of Section 5.2 that the informational role of each recommendation differs depending on whether the asset is overvalued or undervalued. Translating the implications of Corollary 2 to this setting, if the asset is considered undervalued, the conditional probability that the asset price will drop given a recommendation of short position, is lower than the corresponding probability when there is no agency problem. In other words, a recommendation of a short position is a weak indicator that the asset price will eventually drop. Following a similar logic, a recommendation of a long position is a strong indicator that the asset price will eventually increase.\(^{12}\) By contrast, if the asset is considered overvalued, a recommendation of a short position is a strong indicator that the price will drop, whereas a recommendation of a long position is a weak indicator that the price will increase.

**Link with empirical evidence and empirical predictions**

The first prediction of our model is that the analyst issues recommendations against the flow more often than the first best. This prediction is in line with Bernhardt et al. (2016). Bernhardt et al. (2016) explore a setting where analysts follow a three-tier recommendation policy: “sell”, “hold”, “buy”. Extending the logic of our paper to this setting, and interpreting recommendation “hold” as following the flow, we would expect analysts to revise their recommendations away from “hold” more often than they revise to “hold”. Consistently with that, Bernhardt et al. (2016) provide evidence that analysts revise away from “hold” more quickly than the revise to hold, i.e. they spend less time ‘at the prior’.

Recall that recommendations are based on the analyst’s beliefs about the future price. Thus, recommendations reflect the analyst’s forecast about future prices. Earlier, we showed that our model predicts that when prior beliefs indicate that the asset is under-valued, the conditional probability that the asset price increases given a long position, is lower than the corresponding probability when the first best threshold is implemented. Allowing for the limitations that our binary set-up imposes, this finding is consistent with Bernhardt et al. (2006), who show that the conditional probability that a forecast exceeds realized earnings, given that the forecast exceeds the consensus forecast, is lower than the unconditional probability.

Also, Corollary 2 and the mechanism presented in the section 5.2 could provide a theoretical foundation for the empirical evidence which indicates that investors over-react and/or under-react to financial analysts’ forecasts and recommendations (Elgers et al., 2001, Elliot et al., 1995, Elliot et al., 1995, Mendenhall, 1991, Sloan, 1996). Again, the underlying intuition relates to the observation that the recommendation is indicative about the analyst’s private

\(^{12}\) This is because the conditional probability that the asset price will increase given a recommendation of a long position exceeds the corresponding probability, when there is no agency problem.
information regarding the future price of the asset. Following Corollary 2, our model predicts that rational investors should under-react when an analyst issues a contrarian recommendation, and over-react otherwise. The reason is that a contrarian recommendation is a weak indicator that the price will move against the flow, whereas a non-contrarian recommendation is a strong indicator that the price will move with the flow.

5.4 Impact of information acquisition cost, \( c \).

Proposition 5 summarizes the impact of the cost of acquiring information \( c \) on the optimal contract and its main implications. We consider values of \( c \) for which the principal finds it optimal to incentivize the agent to acquire information.

**Proposition 5:** Impact of an increase in \( c \).

If \( \hat{s}^{FB} < \hat{s}_{min} \) i.e., \( p > \hat{p} = 0.5 \) (\( \hat{s}^{FB} > \hat{s}_{min} \) i.e., \( p < \hat{p} = 0.5 \)), an increase in the cost of information acquisition from \( c \) to \( c' \), leads to:

(i) An increase (decrease) in the optimal threshold from \( \hat{s}^* \) to \( \hat{s}'^* \).

(ii) A higher bias \(|\hat{s}'^* - \hat{s}^{FB}| > |\hat{s}^* - \hat{s}^{FB}|\).

(iii) A lower expected probability of taking the revenue-maximizing action.

**Proof.** See Appendix C.

![Figure 6: Impact of an increase in the information acquisition cost from \( c \) to \( c' \).](image)

We provide the intuition for Proposition 5 for the case where \( \hat{s}^{FB} < \hat{s}_{min} \). Similar intuition applies for the case where \( \hat{s}^{FB} > \hat{s}_{min} \). The findings of Proposition 5 depend on the behavior of \( \mathbb{E}C(\hat{s}) \) as \( c \) increases. By Lemma 3, \( \mathbb{E}C(\hat{s}) \) is convex, decreasing in \( \hat{s} \) for \( \hat{s} \in [0, \hat{s}_{min}) \), increasing in \( \hat{s} \) for \( \hat{s} \in (\hat{s}_{min}, 1] \) and linearly dependent in \( c \). Thus, an increase in cost \( c \) shifts the entire \( \mathbb{E}C(\hat{s}) \) curve upwards, and leads to a steeper-sloped U-shape. This is captured in Figure 6, where \( \hat{s} \) is depicted in the horizontal axis, the red line represents the \( \mathbb{E}R(\hat{s}) \), the green line represents \( \mathbb{E}C(\hat{s}, c) \) and the blue line represents \( \mathbb{E}C(\hat{s}, c') \).
Recall that when $\hat{s}^{FB} < \hat{s}_{\text{min}}$, a deviation from $\hat{s}^{FB}$ to $\hat{s} = \hat{s}^{FB} + \eta$ has a positive and a negative effect. On the one hand, it decreases the expected compensation cost, but on the other hand, it decreases the expected revenue. Under the optimal value, $\hat{s}^*$, the two opposite forces offset each other, i.e., the benefit of decreasing the expected cost coincides with the loss of decreasing the expected revenue. Notice that a steeper-sloped U-shape for $EC(\hat{s})$ strengthens the incentive to increase $\hat{s}$. This is because for a given deviation $\eta$ from the first best, the reduction in the expected cost is higher the steeper the slope of the U-shape. This can be seen in Figure 6, where the distance $A'B'$ captures the reduction in the expected cost before the increase in $c$, whereas the distance $AB$ captures the corresponding reduction after the increase in $c$. Thus, $\hat{s}'' > \hat{s}^*$, which combining with the fact that $\hat{s}^{FB}$ is unaffected by the change in $c$, leads to a higher bias compared to the first best.

Finally, part three stems from the monotonic relation between the expected revenue and the probability of achieving the revenue-maximizing action. It is worth highlighting that the lower probability of taking the right action is not a consequence of the fact that the agent does not acquire information. The rationale is that the distortion in the decision rule increases, which in turn, diminishes the information which is incorporated into the implemented action.

5.5 Impact of prior beliefs, $p$.

Proposition 6 summarizes the impact of prior beliefs $p$ on the implemented action. In particular, we explore the case where prior beliefs become more extreme, i.e. $|p - 0.5|$ increases. We denote as $\hat{s}^*$ and $\hat{s}''$ the equilibrium value of $\hat{s}$ before and after the change in $p$. Similarly, $\hat{s}^{FB}$ and $\hat{s}''^{FB}$ correspond to the first best value of $\hat{s}$ before and after the change in $p$. We consider values of $p$ for which incentivizing information acquisition is optimal.

**Proposition 6:** Impact of more extreme prior beliefs.

If $\hat{s}^{FB} < \hat{s}_{\text{min}}$ i.e., $p > \hat{p} = 0.5$ ($\hat{s}^{FB} > \hat{s}_{\text{min}}$ i.e., $p < \hat{p} = 0.5$), an increase (decrease) in beliefs $p$ leads to:

(i) A decrease (increase) in the optimal threshold from $\hat{s}^*$ to $\hat{s}''$.

(ii) Higher bias $|\hat{s}'' - \hat{s}^{FB}| > |\hat{s}^* - \hat{s}^{FB}|$, when the signal structure is linear.

**Proof.** See Appendix C.

The intuition behind part one is straightforward: an increase (decrease) in $p$ decreases (increases) $\hat{s}^{FB}$ without affecting the threshold which minimizes the expected cost, $\hat{s}_{\text{min}}$. Hence, the second best $\hat{s}^*$ moves away from $\hat{s}_{\text{min}}$ towards $\hat{s}''^{FB}$.

Part two is less straightforward. We have shown that a deviation from the first best by a given $\eta$ implies a loss due to the lower expected revenue, and a gain due to the lower expected

\[13\text{Recall that Proposition 5 refers to the case where incentivizing information acquisition is optimal.}\]
compensation cost. As the prior beliefs and the corresponding value of $\hat{s}^{FB}$ become more extreme, the convexity of the expected compensation cost implies that the gain due to the lower expected cost increases. However, the decrease in the expected revenue depends on the particular distribution of $f_{θ}(s)$.

In Proposition 6, we focus on the linear signaling structure, i.e., $f_{θ}(\hat{s})'$ is assumed to be constant. The advantage of this signaling structure is that the loss of a given deviation from the first best is independent of the value of $p$. Hence, as prior beliefs become more extreme, for a given deviation $η$, the decrease in the cost is larger, whereas the decrease in the revenue is unaffected. As a result, more extreme prior beliefs correspond to a higher distortion in the implemented action, compared to the first best.

6 Concluding Remarks & Further Discussion

This paper examines the case where a principal delegates a decision to an agent. We explore an environment where the return from the implemented action depends on the ex-ante unknown state of the world, and the information acquisition by the agent is unobservable and unverifiable. We show that the optimal contract which incentivizes the agent to acquire information promises a positive payment only if the agent’s decision is proven correct. Also, we find that a key feature of the optimal contract is that the agent takes contrarian actions more often than the first best. Furthermore, we find that both the direction and the extent of the distortion in the implemented decision relates to prior beliefs. Besides, we show how these beliefs affect the informational role of the implemented action. Finally, we show in Appendix A and B that the main findings are robust to: i) a set-up where the state of the world is imperfectly observed, and ii) a set-up where the principal allocates the tasks of information acquisition and a decision implementation to two different agents.

In this section, we make the connection between the benchmark setting and alternative environments. Recall that there are three critical characteristics in the benchmark model: i) delegation of a decision by a principal to an agent, ii) the revenue-maximizing action is state-dependent, and, iii) the agent can acquire costly and private information about the ex-ante unknown state of the world. In what follows, we discuss four environments which share, to some extent, these three characteristics, and present the implications of the optimal contract in each setting. Table 2 provides the analogy between the benchmark case and the environments we discuss.

Credit Rating Agencies

We argue in the Online Appendix that the analysis presented in the previous sections can also be applied to the rating process in credit rating agencies (CRAs). In practice, the rating process is characterized by delegation of the evaluation of a company’s creditworthiness to an
analyst, who can obtain information before issuing a rating. Consistently with the literature on credit rating agencies, which recognizes reputation concerns as the main objective of CRAs, we assume that the principal’s objective is to maximize the probability of issuing a rating which corresponds to the actual type. In our setting, this objective translates into giving a good rating to a creditworthy company, and a bad rating otherwise, where the creditworthiness of a company corresponds to the ex-ante unknown state of the world.

In this environment, the optimal contract implies a premium for a rating against the flow. This premium gives rise to more frequent good ratings (rating inflation) compared to the first best, when the company is ex-ante less likely to be of good type, and vice versa. Moreover, if the company is ex-ante more likely to be of good type, then a bad rating is less (more) likely to be correct, whereas a good rating is more (less) likely to be correct. Also the resulting bias can sustain extreme prior beliefs about a company’s type. Finally, we endogenize the company’s borrowing interest rates after the rating is issued. We find that conditional on the rating, the interest rate is lower (higher) than the first best when prior beliefs indicate that the company is of good (bad) type. This finding could explain the emergence of long periods of low or high interest rates.

Product Design
The main role of a product manager is to analyze the market and decide about those product features which will accommodate the ex-ante unknown future demand. In this environment, the optimal contract pays the product manager only if he designs a product which eventually accommodates the demand. Also, the optimal contract leads to over-investment in products and product features which are ex-ante less likely to accommodate the demand. Besides, conditional on adopting such a product, it is less likely than the first best that the demand will be actually addressed. In contrast, conditional on adopting a product which is ex-ante more likely to accommodate the demand, it is more likely than the first best that it will succeed in accommodating the demand. Thus, this model would predict an excessive supply of products which have a low ex-ante probability of accommodating the demand, and in turn are very likely not to accommodate the demand.

Portfolio Allocation
Another application is a portfolio allocation problem where the manager considers investing either in a risky project or a safe project. The return of the safe project is fixed, whereas the return of the risky project depends on its quality, which is unknown to both the equity holder and the manager. The net return of the risky project is positive if its type is good, and negative if the type is bad. This environment is similar to Lambert (1986).

In this environment, the optimal contract pays the manager only in the case he invests on a safe (risky) asset, and the type of the risky asset is revealed to be bad (good). In any other
case, the payment to the agent is zero. Besides, if the prior indicates that the NPV of the risky asset is higher (lower) than the NPV of the safe asset, then the optimal compensation contract gives rise to under-investment (over-investment) to the risky asset.

The resulting over-investment against the flow has implications for the riskiness of the portfolio, captured by its variance. In particular, the riskiness is lower than the first best when the ex-ante NPV of the risky asset exceeds the NPV of the safe asset, and the other way around.

**Incentivizing Innovation**

Similarly to [Manso (2011)](Manso2011), the benchmark model can also be used to characterize the optimal compensation contracts which incentivize an agent to innovate. For instance, consider an environment where two strategies/techniques are available: a conventional one, which leads to a fixed return $R$, and an innovative one, which returns $R_S > R$, if it is successful, and $R_F < R$ otherwise. By nature, the return on the innovative strategy is ex-ante unknown to the agents, who, however, hold prior beliefs about the return being $R_S$, denoted by $p$. The main difficulty which arises in this set-up is that the principal and the agent cannot observe whether the innovative strategy is successful, unless this strategy is implemented. This difficulty is addressed in Appendix A.

In this environment, the optimal contract pays the agent only when he adopts an innovative strategy which is successful or a conventional strategy which is supported by the public signal. Besides, if prior beliefs indicate that the NPV of the innovative strategy is higher (lower) than the NPV of the conventional strategy, the optimal compensation contract results in under-implementation (over-implementation) of the innovative strategy.
References


Lazear, E. P. and S. Rosen (1979): “Rank-order tournaments as optimum labor contracts.”


Appendix A: Imperfect State Realization

The previous analysis relies on the assumption that the state of the world is observable and contractible. There are many situations, however, where the state of the world is imperfectly revealed. For instance, in the financial-recommendation setting, suppose that the asset that the analyst evaluates is a stock of Company “Z”. In that case, rather than observing the realized quality of the asset, which in this example corresponds to the profitability of Company “Z”, the principal and the agent might observe an imperfect public signal about the profitability of Company “Z” such as earnings forecasts or credit rating announcements. Another case that the state of the world might not be contractible is when its realization takes place after the contract is terminated.

The goal of this section is two-fold. First, to derive the optimal compensation contract in an alternative set-up where the payments can only be contingent on an imperfect signal about the state of the world. Second, to examine whether the main features and implications of the optimal contract are similar to the ones in the benchmark model.

We explore an environment where the principal and the agent have access to a public signal, which is revealed after the decision is taken, and imperfectly reveals the actual state of the world. In particular, we allow for a binary public signal i.e., $\sigma = \{\tilde{B}, \tilde{G}\}$, where:

\[
Pr(\sigma = \tilde{G} | \theta = G) = q_G \\
Pr(\sigma = \tilde{G} | \theta = B) = q_B
\]

with $q_G > q_B$. The intuition is straightforward; compared to the case where the state is $\theta = B$, when the state is $\theta = G$, it is more likely that the public signal will be $\sigma = \tilde{G}$. For instance, a good credit rating is more likely when Company “Z” is of good rather than bad quality. Likewise, compared to the case where the state is $\theta = G$, when the state is $\theta = B$, it is more likely that the public signal will be $\sigma = \tilde{B}$. Note that the benchmark model can be thought as a special case of the imperfect state realization case, if we restrict the values of $q_G$ and $q_B$ to be equal to one and zero, respectively.

The difference in the compensation contract is that instead of the payments being contingent on the realized state, $\theta$, they are contingent on the realized public signal, $\sigma$, i.e., $\hat{W}'' : d \times \sigma \mapsto \mathbb{R}^+$, where $d = \{b, g\}$ and $\sigma = \{\tilde{B}, \tilde{G}\}$. The contract is characterized by the following quadruple:

\[
\hat{W}'' = \{w_{b\tilde{G}}, w_{b\tilde{B}}, w_{g\tilde{G}}, w_{g\tilde{B}}\}
\]

In order to derive the optimal compensation contract, we follow the same four-step process we followed in the case where the state of the world is contractible.
A.1 Step 1: Decision Rule $\mathcal{DR}$

We show in Appendix C that Lemma 1 extends to the case where the state of the world is imperfectly observed. Hence, similarly to the case where the state of the world is contractible, the decision rule that the principal aims to implement is given by a threshold $\hat{s}$, such as:

$$
\mathcal{DR} = \begin{cases} 
  d = g & \text{if } s > \hat{s} \\
  d = b & \text{if } s < \hat{s}
\end{cases}
$$

Proof. See Lemma 1B in Appendix C.

A.2 Step 2: Constrained-optimal contract that implements $\hat{s}$.

In this step, we derive the most cost-efficient contract of implementing a given threshold $\hat{s}$.

**Proposition 1B: Constrained Optimal Contract**

For $\hat{s} \in [0, \hat{s}_{\text{min}}]$ the constrained optimal contract is given by:

$$
\begin{align*}
  w_{b_B}^*(\hat{s}) &= \frac{f_B(\hat{s})(1-p)q_B + f_G(\hat{s})pq_G}{(f_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})F_G(\hat{s}))(-1+p)p(q_B - q_G)}c \\
  w_{g_G}^*(\hat{s}) &= \frac{f_B(\hat{s})(-1+p)(-1+q_B) - f_G(\hat{s})p(-1+q_G)}{(f_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})F_G(\hat{s}))(-1+p)p(q_B - q_G)}c \\
  w_{g_B}^*(\hat{s}) = w_{b_G}^*(\hat{s}) &= 0.
\end{align*}
$$

For $\hat{s} \in [\hat{s}_{\text{min}}, 1]$ the constrained optimal contract is given by:

$$
\begin{align*}
  w_{b_B}^*(\hat{s}) &= \frac{f_B(\hat{s})(-1+p)q_B - f_G(\hat{s})pq_G}{(f_B(\hat{s})(-1 + F_G(\hat{s})) + f_G(\hat{s})(1 - F_B(\hat{s})))(-1+p)p(q_B - q_G)}c \\
  w_{g_G}^*(\hat{s}) &= \frac{f_B(\hat{s})(-1+p)(1-q_B) + f_G(\hat{s})p(-1+q_G)}{((1 - F_B(\hat{s}))f_G(\hat{s}) + f_B(\hat{s})(-1 + F_G(\hat{s})))(-1+p)p(q_B - q_G)}c \\
  w_{g_B}^*(\hat{s}) = w_{b_G}^*(\hat{s}) &= 0.
\end{align*}
$$

where $\hat{s}_{\text{min}}$ solves $f_G(\hat{s}_{\text{min}}) = f_B(\hat{s}_{\text{min}})$.

Proof. See Appendix C.

Proposition 1B is the analog of Proposition 1. The proof and the intuition is identical to ones in Proposition 1. Also, we show in Appendix C that the relationship between the optimal payments and $\hat{s}$, which is captured in Lemma 2, extends to this set-up. Lemma 3 also extends to this set-up. Thus, $E_C(\hat{s})$ is U-shaped, and minimized for $f_G(\hat{s}) = f_B(\hat{s})$. 

37
A.3 Step 3: Expected revenue $\mathbb{E} R(\hat{s})$

A critical remark is that the expected revenue of a given threshold $\hat{s}$ is determined by the actual state of the world, rather than the public signal about the state of the world. Hence, the second step is identical to the corresponding step in the benchmark model. The behavior of $\mathbb{E} R(\hat{s})$ with respect to the implemented threshold $\hat{s}$ is given by Lemma 4.

A.4 Step 4: Optimal Threshold $\hat{s}^*$

The optimal value of $\hat{s}$ denoted as $\hat{s}^*$ equates the expected marginal revenue $\left(\frac{\partial \mathbb{E} R(\hat{s})}{\partial \hat{s}}\right)$ with the expected marginal cost $\left(\frac{\partial \mathbb{E} C(\hat{s})}{\partial \hat{s}}\right)$, where $\mathbb{E} R(\hat{s})$ is defined in Lemma 2 whereas $\mathbb{E} C(\hat{s})$ is defined in Lemma 3B.

A.5 Optimal Compensation Contract

The optimal contract is given by the constrained optimal contract of Proposition 1B, subject to the constraint the $\hat{s}$ satisfies the optimality condition which is provided in Lemma A.4.

Proposition 2B: Optimality Contract, $\hat{W}^*(\hat{s}^*)$

If $\hat{s}^{FB} < \hat{s}_{\text{min}}$ (i.e., $p > \hat{p} = 0.5$), the optimal value $\hat{s}$, $\hat{s}^*$ solves:

$$2v\{-pf_G(\hat{s}) + (1-p)f_B(\hat{s})\} = \left[\frac{((-1+p)q_B - pq_G)((-1+p)(-1+q_B)F_B(\hat{s}) - p(-1+q_G)F_G(\hat{s}))\gamma(\hat{s})}{[F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s})(-1+p)p(q_B - q_G)]^2}\right]_c \quad (36)$$

and the optimal payments are given by:

$$w_{bB}^*(\hat{s}^*) = \frac{f_B(\hat{s}^*)(1-p)q_B + f_G(\hat{s}^*)pq_G}{(F_B(\hat{s}^*)f_G(\hat{s}^*) - f_B(\hat{s}^*)F_G(\hat{s}^*))(-1+p)p(q_B - q_G)}_c$$

$$w_{gG}^*(\hat{s}^*) = \frac{f_B(\hat{s}^*)(-1+q_B) - f_G(\hat{s}^*)p(-1+q_G)}{(F_B(\hat{s}^*)f_G(\hat{s}^*) - f_B(\hat{s}^*)F_G(\hat{s}^*))(-1+p)p(q_B - q_G)}_c$$

$$w_{gB}^*(\hat{s}^*) = w_{bG}^*(\hat{s}^*) = 0$$

If $\hat{s}^{FB} > \hat{s}_{\text{min}}$ (i.e., $p < \hat{p} = 0.5$), the optimal value $\hat{s}$, $\hat{s}^*$ solves:

$$2v\{-pf_G(\hat{s}) + (1-p)f_B(\hat{s})\} = \left[\frac{(1 - (1-p)q_B - pq_G)((1-p)q_B(1 - F_B(\hat{s})) + pq_G(1 - F_G(\hat{s})))\gamma(\hat{s})}{[((-1 + F_B(\hat{s})f_G(\hat{s}) + (1 - F_G(\hat{s}))f_B(\hat{s})(-1+p)p(q_B - q_G))]}_c \quad (37)$$
and the optimal payments are given by:

\[ w^*_b(s^*) = \frac{f_B(s^*)(-1 + p)q_B - f_G(s^*)p q_G}{(f_B(s^*)(-1 + F_G(s^*))) + f_G(s^*)(1 - F_B(s^*)))(1 - p)p(q_B - q_G)^c} \]

\[ w^*_g(s^*) = \frac{f_B(s^*)(-1 + p)(1 - q_B) + f_G(s^*)p(-1 + q_G)}{((1 - F_B(s^*))f_G(s^*) + f_B(s^*)(-1 + F_G(s^*)))(1 - p)p(q_B - q_G)} \]

\[ \gamma(s) \equiv (f_G(s)f'_B(s) - f_B(s)f'_G(s)) \]

where \( \gamma(s) \equiv (f_G(s)f'_B(s) - f_B(s)f'_G(s)) \).

Proof. See Appendix C.

\( \square \)

### A.6 Bias in the investment decision

A crucial remark of the previous analysis is that \( \mathbb{E} C(s) \) has similar behavior (U-shaped, and minimized for \( f_G(s) = f_B(s) \)) independently of whether the principal and the agent observe the state of the world or a public signal which is informative about the state of the world. Recall also that \( \mathbb{E} R(s) \) is unaffected by whether the state of the world or a public signal, is observed.

Hence, the analysis of how the optimal threshold \( s^* \) relates to the first best threshold \( s^{FB} \) is qualitatively the same with the benchmark model.

**Proposition 4B:** Bias in the implemented decision.

**Under the optimal contract:**

(i) If \( s^{FB} < \hat{s}_{min} \) (i.e., \( p > \hat{p} = 0.5 \)), then \( s^{FB} < s^* \leq \hat{s}_{min} \)

(ii) If \( s^{FB} > \hat{s}_{min} \) (i.e., \( p < \hat{p} = 0.5 \)), then \( s^{FB} > s^* \geq \hat{s}_{min} \)

Proof. See Appendix C.

Hence, the main finding that under the optimal contract, that the agent adopts contrarian actions more often than the first best, remains.

### A.6.1 Impact of \( q_G \) and \( q_B \) on the implemented decision

Proposition 7 explores the impact of the informativeness of the public signal, which is captured by \( q_G \) and \( q_B \), on the implemented decision and the emerging bias. Recall that \( q_G \) and \( q_B \) is the probability of observing a good public signal when the state is good and bad, respectively.

**Proposition 7:** Impact of \( q_G \) and \( q_B \) on the implemented decision.

(i) If \( s^{FB} < \hat{s}_{min} \) (i.e., \( p > \hat{p} = 0.5 \)), \( s^* \) is increasing in \( q_B \) and decreasing in \( q_G \).
(ii) If $\hat{s}^{FB} > \hat{s}_{min}$ (i.e., $p < \hat{p} = 0.5$), $\hat{s}^*$ is decreasing in $q_B$ and increasing in $q_G$.

(iii) The bias $|\hat{s}^* - \hat{s}^{FB}|$ is decreasing in $q_B$ and increasing in $q_G$.

Proof. See Appendix C.

Proposition 7 relies on Lemma 6 and on the observation that the expected revenue of implementing $\hat{s}$ is independent of $q_G$ or $q_B$. Thus, $q_G$ and $q_B$ affect $\hat{s}^*$ through their impact on expected cost $\mathbb{E}C(\hat{s})$. An implication of Lemma 6 is that a decrease in $q_G$ and/or an increase in $q_B$ moves the graphical illustration of $\mathbb{E}C(\hat{s})$ upwards, and leads to a steeper-sloped U-shape. The intuition behind Lemma 6 relies on the fact that as the difference $q_G - q_B$ decreases, the quality of monitoring worsens, and as a result, the principal can only incentivize the agent to acquire information through higher payments. The underlying intuition in Proposition 7 is similar to the intuition in Proposition 5. Namely, for a given deviation $\eta$ from the first best $\hat{s}^{FB}$, a higher $q_B$ and/or a lower $q_G$ leads to a larger reduction in the expected cost, which in turn, strengthens the incentive to deviate. Proposition 6 is captured in Panel A and Panel B of Figure 7.

![Figure 6: Impact of Better Monitoring, for $f_G = 2s$, and $f_B = 2(1 - s)$.]

**Appendix B: Allocating tasks to different agents**

In Section 4, we argue that as long as the agent is not indifferent between action $g$ and $b$ in the absence of a signal (i.e., $p \neq 0.5$), there is a trade-off because the threshold $\hat{s}$ which minimizes the cost of incentivizing the agent to acquire information differs from the threshold which maximizes the expected revenue. Two questions which arise naturally are whether allocating the tasks to two different individuals would increase principal’s profits, and whether the optimal contract would exhibit the key feature of the “bias against the flow”.

40
To this end, we allow for two agents: a researcher and a manager, each of them performing one task. The researcher acquires information and sends a message to his manager. The manager receives the researcher’s message and subsequently implements one of the two available actions. In this environment, the researcher’s contract is contingent on the realized state of the world \((\theta)\) and his message, which we denote as \(\tilde{s}\), i.e., \(\tilde{W}_A : \tilde{s} \times \theta \mapsto \mathbb{R}^+\). Regarding the manager’s compensation contract, we focus on contracts contingent on the researcher’s message and the manager’s implemented action, i.e., \(\tilde{W}_M : \tilde{s} \times d \mapsto \mathbb{R}^+\).

**Lemma 8:** Equilibrium Messages.

*Under the optimal contract, the researcher sends a message \(s\) if his private signal belongs in partition \([0, \hat{s}]\), and a message \(\bar{s}\) if his private signal belongs in partition \((\hat{s}, 1]\).*

**Proof.** Recall that what matters for taking the revenue-maximizing action is whether \(\Pr(\theta = G|s)\) is greater or lower than a threshold value, which in the symmetric case equals one-half. This is because the principal’s best respond depends solely on that. Following the spirit of Lemma 1, for the incentives of the principal and the researcher to be aligned, it is sufficient for the principal to offer a contract where two messages are issued in equilibrium. Message \(s\), that corresponds to partition \([0, \hat{s}]\), and message \(\bar{s}\), that corresponds to partition \((\hat{s}, 1]\), where \(\hat{s} \in (0, 1)\). It is never optimal for the principal to offer a contract where more than two messages are issued in equilibrium; allowing for more messages would never lead to higher revenue, whereas it would increase the cost of incentivizing information acquisition. This is because more messages would imply a higher number of incentive compatibility constraints. \(\square\)

Lemma 9 states that the researcher’s optimal contract in this environment would coincide with the optimal contract of the agent in the benchmark model, characterized in Proposition 2.

**Lemma 9:** Researcher’s optimal contract.

\[
 w(s, B)^* \equiv w_{bB}(\hat{s}^*), \quad w(\tilde{s}, G)^* \equiv w_{dG}(\hat{s}^*), \quad w(s, G)^* \equiv w_{gG}(\hat{s}^*) \quad \text{and} \quad w(\tilde{s}, B)^* \equiv w_{bB}(\hat{s}^*)
\]

**Proof.** Lemma 9 is a consequence of Lemma 8. The main difference is that instead of incentivizing the agent to acquire information and follow a decision rule, which is characterized by a threshold \(\hat{s}\), the principal incentivizes a researcher to acquire information and follow a message rule, which is characterized by the same threshold \(\hat{s}\). \(\square\)

The main message of Lemma 9 is that the optimal compensation contract, and thus the cost of incentivizing an agent to acquire information, is the same independently of whether this agent is in charge of implementing an action. Thus, the feature of the “bias against the flow” emerges even when the principal allocates the tasks to two different individuals. To complete the analysis of this environment, we should also characterize the optimal compensation contract
Lemma 10: Manager’s optimal contract.
\[ w(s, S)^* = w(\bar{s}, L)^* = \kappa \text{ and } w(s, L)^* = w(\bar{s}, S)^* = 0, \text{ where } \kappa \to 0. \]

Proof. Since the manager has no private information, the incentives of the principal and the manager are aligned if the principal offers an arbitrarily small payment \( \kappa \) when the manager takes the revenue-maximizing action, and zero otherwise.

To conclude, we show in this section that, as long as information acquisition is unobservable, delegating the tasks to different agents leads to the same implemented decision, and it does not improve the profitability of the principal: the principal also incurs the cost of incentivizing the manager to implement the revenue-maximizing decision.

Appendix C: Proofs

C.1 First Best

We start by analyzing the case where there is no agency problem. First, we derive the expected profit of the principal when no information is acquired. Second, we derive the expected profit for the case where principal acquires information. Finally, we characterize the condition such as information acquisition is optimal.

Case where no information is acquired: In the absence of information acquisition the expected revenue of action \( b \) and \( g \) is given by:

\[
\begin{align*}
\mathbb{E} \Pi[d = b] &= pR(d = b, \theta = G) + (1 - p)R(d = b, \theta = B) = (1 - 2p)v \\
\mathbb{E} \Pi[d = g] &= pR(d = g, \theta = G) + (1 - p)R(d = g, \theta = B) = (2p - 1)v
\end{align*}
\]

Thus, action \( b \) (\( g \)) is revenue-maximizing when \( p < 0.5 \) (\( p > 0.5 \)).

\[
\mathbb{E} \Pi[\text{no signal}] = \max\{(2p - 1), (1 - 2p)\} \times v
\]

Case where information is acquired: First, we characterize the optimal action for each signal realization, \( s \). The principal takes action \( g \) if:

\[
\mathbb{E} \Pi[d = g|s] > \mathbb{E} \Pi[d = b|s] \iff [Pr(\theta = G|s) - Pr(\theta = B|s)]v > [-Pr(\theta = G|s) + Pr(\theta = B|s)]v
\]
whereas he takes action b otherwise. Since \( Pr(\theta = G | s) \) is strictly increasing in \( s \), there is a unique signal realization \( s = \hat{s}^{FB} \) such that \( E\Pi[d = g|\hat{s}^{FB}] = E\Pi[d = b|\hat{s}^{FB}] \). This is true for \( \hat{s}^{FB} \) such that:

\[
\frac{f_G(\hat{s})}{f_B(\hat{s})} = \frac{(1 - p)}{p} \tag{42}
\]

Hence, the expected profit of acquiring information is given by:

\[
E\Pi[signal|\hat{s}^{FB}] = \int_0^{\hat{s}^{FB}} E\Pi[d = b|s]f(s)ds + \int_{\hat{s}^{FB}}^1 E\Pi[d = g|s]f(s)ds - c \tag{43}
\]

In section 4.4 we showed that the expected revenue of acquiring a signal, given \( \hat{s} \), is:

\[
E R(\hat{s}) = v\{p(1 - 2F_G(\hat{s})) + (1 - p)(2F_B(\hat{s}) - 1)\}
\]

Thus, the expected profit of acquiring a signal, given the first best threshold \( \hat{s}^{FB} \), is:

\[
E\Pi[signal|\hat{s}^{FB}] = v\{p(1 - 2F_G(\hat{s}^{FB})) + (1 - p)(2F_B(\hat{s}^{FB}) - 1)\} - c \tag{44}
\]

As a result, the principal acquires information as long as:

\[
E\Pi[signal|\hat{s}^{FB}] \geq E\Pi[no\ signal]
\]

which is satisfied for values of \( c \) such as:

\[
c \leq [p(1 - 2F_G(\hat{s}^{FB})) + (1 - p)(2F_B(\hat{s}^{FB}) - 1) - \max\{(2p - 1), (1 - 2p)\}] \times v \tag{45}
\]

where the LHS of (45) captures the cost of acquiring information, whereas the RHS captures the expected benefit of acquiring information net of the principal’s outside option.

### C.2 Useful Lemmas

We provide Lemma A.1 and Lemma A.2, which are going to be useful for the remaining proofs.

#### C.2.1 Lemma A.1

**Lemma A.1:** For each \( \hat{s} \in S \), the following relation holds:

\[
\gamma_1 \equiv f'_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})f'_G(\hat{s}) < 0
\]

**Proof.** Crucial for the determination of the sign of \( \gamma_1 \) is \( f'_G(\hat{s}) \) and \( f'_B(\hat{s}) \), where \( f'_G(\hat{s}) = \frac{\partial f_G(\hat{s})}{\partial \hat{s}} \) and \( f'_B(\hat{s}) = \frac{\partial f_B(\hat{s})}{\partial \hat{s}} \). Although we have made no assumption about \( f'_G(\hat{s}) \) and \( f'_B(\hat{s}) \), they are

\[\text{The implicit assumption w.l.o.g. is that if the principal is indifferent, he goes short.}\]
indirectly constrained by MLRP. For instance, a direct consequence of MLRP is that \( f_G'(\hat{s}) > f_B'(\hat{s}) \), otherwise \( \frac{f_G(\hat{s})}{f_B(\hat{s})} \) is decreasing in \( \hat{s} \). Thus, we explore the sign of \( \gamma_1 \) for all possible relations between \( f_G'(\hat{s}) \) and \( f_B'(\hat{s}) \), by taking into consideration the constraints imposed by MLRP.

(i) Case 1: \( f_G'(\hat{s}) > f_B'(\hat{s}) > 0 \)
For this relation, \( \gamma_1 < 0 \), since \( f_B(\hat{s}) > f_G(\hat{s}) \) and \( f_G'(\hat{s}) > f_B'(\hat{s}) > 0 \).

(ii) Case 2: \( f_G'(\hat{s}) > 0 > f_B'(\hat{s}) \)
For this relation, \( \gamma_1 < 0 \), since its first term is negative and its second term is positive.

(iii) Case 3: \( 0 > f_G'(\hat{s}) > f_B'(\hat{s}) \)
For this relation, \( \gamma_1 < 0 \). The proof is less straightforward since the sign of \( \gamma_1 \) depends on the relation between ratio \( \frac{f_G'(\hat{s})}{f_G(\hat{s})} \) and \( \frac{f_B'(\hat{s})}{f_B(\hat{s})} \). We show that \( \gamma_1 < 0 \) by following the method of contradiction. Suppose that \( \gamma_1 > 0 \), i.e.,
\[
f_B'(\hat{s})f_G(\hat{s}) - f_B(\hat{s})f_G'(\hat{s}) > 0 \tag{46}
\]
By dividing the LHS of (46) by \( f_G(\hat{s})f_B(\hat{s}) \) and rearranging, we obtain:
\[
\frac{f_B'(\hat{s})}{f_B(\hat{s})} > \frac{f_G'(\hat{s})}{f_G(\hat{s})} \tag{47}
\]
where \( f_B'(\hat{s}) = \lim_{h \to 0} \frac{f_B(\hat{s}_0 + h) - f_B(\hat{s}_0)}{h} \) and \( f_G'(\hat{s}) = \lim_{h \to 0} \frac{f_G(\hat{s}_0 + h) - f_G(\hat{s}_0)}{h} \).
Hence, (47) implies the following relation:
\[
\lim_{h \to 0} \frac{f_B(\hat{s}_0 + h) - f_B(\hat{s}_0)}{f_B(\hat{s}_0)} > \lim_{h \to 0} \frac{f_G(\hat{s}_0 + h) - f_G(\hat{s}_0)}{f_G(\hat{s}_0)} \tag{48}
\]
Thus, relation (48) implies that the percentage decrease in the \( f_B(\hat{s}) \) as \( \hat{s} \) moves from \( \hat{s}_0 \) to \( \hat{s}_1 = \hat{s}_0 + h \) is smaller than the percentage decrease in \( f_G(\hat{s}) \) as \( \hat{s} \) moves from \( \hat{s}_0 \) to \( \hat{s}_1 = \hat{s}_0 + h \). If this is true, then relation (48) gives rise to following relation
\[
\frac{f_G(\hat{s}_0)}{f_B(\hat{s}_0)} > \frac{f_G(\hat{s}_1)}{f_B(\hat{s}_1)} \tag{49}
\]
However, relation (49) contradicts with the MLRP assumption, according to which, \( \frac{f_G(\hat{s})}{f_B(\hat{s})} \) is increasing in \( \hat{s} \). Hence, \( \gamma_1 < 0 \).

Finally, for the sake of completeness, we list the remaining cases of the relationship between \( f_G'(\hat{s}) \) and \( f_B'(\hat{s}) \), which are irrelevant, since they violate the MLRP assumption.

(iv) Case 4: \( f_B'(\hat{s}) > f_G'(\hat{s}) > 0 \)
The MLRP is violated since for this relation, the ratio \( \frac{f_G(\hat{s})}{f_B(\hat{s})} \) is decreasing in \( \hat{s} \).
(v) Case 5: $f_B(\hat{s}) > 0 > f_G(\hat{s})$

The MLRP is violated since for this relation, the ratio $\frac{f_G(\hat{s})}{f_B(\hat{s})}$ is decreasing in $\hat{s}$.

(vi) Case 6: $0 > f_B(\hat{s}) > f_G(\hat{s})$

The MLRP is violated since for this relation, the ratio $\frac{f_G(\hat{s})}{f_B(\hat{s})}$ is decreasing in $\hat{s}$.

Thus, for the relevant cases (cases i - iii), $\gamma_1$ is negative independently of the value of $\hat{s}$.

C.2.2 Lemma A.2

**Lemma A.2:** The monotone likelihood ratio property assumption implies that:

$$\frac{1 - F_G(\hat{s})}{1 - F_B(\hat{s})} \geq \frac{f_G(\hat{s})}{f_B(\hat{s})} \geq \frac{F_G(\hat{s})}{F_B(\hat{s})}.$$  

**Proof.** By the definition of the MLRP, for $s_1 \geq s_0$, the following condition holds:

$$\frac{f_G(s_1)}{f_B(s_1)} \geq \frac{f_G(s_0)}{f_B(s_0)}$$

which can be rearranged as:

$$f_G(s_1)f_B(s_0) \geq f_G(s_0)f_B(s_1) \quad (50)$$

Integrating both sides of (50) over $s_0$ from the lower bound of the distribution to $s_1$:

$$\int_0^{s_1} f_G(s_1)f_B(s_0)ds_0 \geq \int_0^{s_1} f_G(s_1)f_B(s_0)ds_0$$

which leads to:

$$\frac{f_G(s_1)}{f_B(s_1)} \geq \frac{F_G(s_1)}{F_B(s_1)} \quad (51)$$

Next, we integrate both sides of (50) over $s_1$ from $s_0$ to the upper bound of the distribution:

$$\int_{s_0}^1 f_G(s_1)f_B(s_0)ds_0 \geq \int_{s_0}^1 f_G(s_1)f_B(s_0)ds_0$$

which leads to:

$$\frac{f_G(s_0)}{f_B(s_0)} \leq \frac{1 - F_G(s_0)}{1 - F_B(s_0)} \quad (52)$$

Notice that $s_0$ and $s_1$ in (51) and (52) are arbitrary. Thus, combining these equation bu letting $s_0 = s_1 = \hat{s}$, we obtain:

$$\frac{1 - F_G(\hat{s})}{1 - F_B(\hat{s})} \geq \frac{f_G(\hat{s})}{f_B(\hat{s})} \geq \frac{F_G(\hat{s})}{F_B(\hat{s})}.$$  

\[\blacksquare\]
C.3 Benchmark Model

C.3.1 Proof of Lemma 1 (Monotonic decision rule)

The proof of Lemma 1 is identical to the proof of Lemma 1B once we set \( q_G = 1, q_B = 0 \), and we replace \( w_{b\tilde{B}} \) with \( w_{bB} \), \( w_{bG} \) with \( w_{gB} \), \( w_{g\tilde{B}} \) with \( w_{gB} \), and \( w_{gG} \) with \( w_{gG} \).

C.3.2 Proof of Proposition 1

The proof of Proposition 1 is identical to the proof of Proposition 1B once we set \( q_G = 1, q_B = 0 \), and we replace \( w_{b\tilde{B}} \) with \( w_{bB} \), \( w_{bG} \) with \( w_{gB} \), \( w_{g\tilde{B}} \) with \( w_{gB} \), and \( w_{gG} \) with \( w_{gG} \).

C.3.3 Proof of Lemma 2

The proof of Lemma 2 is identical to the proof of Lemma 2B once we set \( q_G = 1, q_B = 0 \), and we replace \( w_{b\tilde{B}} \) with \( w_{bB} \), \( w_{bG} \) with \( w_{gB} \), \( w_{g\tilde{B}} \) with \( w_{gB} \), and \( w_{gG} \) with \( w_{gG} \).

C.3.4 Proof of Lemma 4

In section 4.4 we showed that the expected revenue from acquiring a signal, given \( \hat{s} \), is:

\[
\mathbb{E} R(\hat{s}) = v\{p(1 - 2F_G(\hat{s})) + (1 - p)(2F_B(\hat{s}) - 1)\}.
\]

Thus, the first order condition is given by:

\[
\frac{\partial \mathbb{E} R(\hat{s})}{\partial (\hat{s})} = 2v\{f_B(\hat{s})(1 - p) - f_G(\hat{s})p\} = 0 \tag{53}
\]

which is satisfied for a unique (due to MLRP) \( \hat{s} = \hat{s}^{FB} \), which solves:

\[
\frac{f_G(\hat{s})}{f_B(\hat{s})} = \frac{(1 - p)}{p}. \tag{54}
\]

The second derivative of \( \mathbb{E} R(\hat{s}) \) is given by:

\[
\frac{\partial^2 \mathbb{E} R(\hat{s})}{\partial (\hat{s})^2} = f'_B(\hat{s})(1 - p) - f'_G(\hat{s})p. \tag{55}
\]

For the second order condition to be satisfied, it must hold:

\[
\left. \frac{\partial^2 ER(\hat{s})}{\partial (\hat{s})^2} \right|_{(1-p)=\frac{f_G(\hat{s}^{FB})}{f_B(\hat{s}^{FB})}} = p \left[ \frac{f'_B(\hat{s}^{FB})f_G(\hat{s}^{FB})}{f_B(\hat{s}^{FB})} - f'_G(\hat{s}^{FB}) \right] < 0 \tag{56}
\]
which implies that the following should hold:

\[ f'_B(s^{FB}) f'_G(s^{FB}) - f'_G(s^{FB}) f_B(s^{FB}) < 0. \]  

(57)

In Lemma A.1 we proved that \( f'_B(s) f'_G(s) - f'_G(s) f_B(s) < 0 \) for any \( s \). Thus, (57) holds.

We now explore the behavior of \( E R(s) \) for \( s > s^{FB} \) and \( s < s^{FB} \). We start by analysing the case where \( s' > s^{FB} \). For \( s' > s^{FB} \), and given (54), the MLRP implies that:

\[ \frac{f_G(s')}{f_B(s')} > \frac{f_G(s^{FB})}{f_B(s^{FB})} = \frac{(1 - p)}{p}. \]

Thus, \( \frac{f_G(s')}{f_B(s')} \) can be expressed as:

\[ \frac{f_G(s')}{f_B(s')} = \frac{(1 - p)}{p} + \eta(s') \]  

(58)

where \( \eta(s') > 0 \) and increasing in \( s' \) by the MLRP. Also, by rearranging the first derivative of the expected revenue with respect to \( s \), we obtain:

\[ \frac{\partial E R(s')}{\partial s} / f_B(s) = v\{(1 - p) - \frac{f_G(s')}{f_B(s')} p\} \]  

(59)

and, by substituting (58) into (59):

\[ \frac{\partial E R(s')}{\partial s} / f_B(s) = v\{(1 - p) - \left( \frac{(1 - p)}{p} + \eta(s') \right) p\} = -v\eta(s') p < 0. \]  

(60)

Hence, for any \( s' > s^{FB} \), \( E R(s) \) is decreasing and concave in \( s' \), since \( \eta(s') \) is increasing in \( s' \).

Second, we explore the case where \( s < s^{FB} \). For \( s' > s^{FB} \), the MLRP implies that:

\[ \frac{f_G(s')}{f_B(s')} \]  

where \( f_G(s') / f_B(s') \) can be expressed as:

\[ \frac{f_G(s')}{f_B(s')} = \frac{(1 - p)}{p} - \eta(s') \]  

(61)

where \( \eta(s') > 0 \) and decreasing in \( s' \) by the MLRP. Also, by substituting (61) into (59):

\[ \frac{\partial E R(s')}{\partial s} / f_B(s) = v\{(1 - p) - \left( \frac{(1 - p)}{p} - \eta(s') \right) p\} = v\eta(s') p > 0. \]  

(62)
Hence, for any \( \hat{s}' > \hat{s}^{FB} \), \( \mathbb{E} R(\hat{s}) \) is increasing and concave in \( \hat{s} \) since \( \eta'(\hat{s}') \) is decreasing in \( \hat{s}' \).

**C.3.5 Proof of Proposition 2**

The proof of Proposition 2 is identical to the proof of Proposition 2B once we set \( q_G = 1 \), \( q_B = 0 \), and we replace \( w_{\tilde{b}B} \) with \( w_{bB} \), \( w_{\tilde{b}G} \) with \( w_{bG} \), \( w_{\tilde{g}B} \) with \( w_{gB} \), and \( w_{\tilde{g}G} \) with \( w_{gG} \).

**C.3.6 Proof of Proposition 3**

Notice that by (15), the ratio of the payments is given by:

\[
\frac{w_{bB}(\hat{s}^*)}{w_{gG}(\hat{s}^*)} = \frac{f_G(\hat{s}^*)}{f_B(\hat{s}^*)} \cdot \frac{1-p}{p}.
\]

In addition, under the first best, the following condition should hold:

\[
\frac{f_G(\hat{s}^{FB})}{f_B(\hat{s}^{FB})} = \frac{1-p}{p}.
\]

*Part one:* By Proposition 4, if \( p > 0.5 \), \( \hat{s}^* > \hat{s}^{FB} \), which combined with MLRP implies:

\[
\frac{f_G(\hat{s}^*)}{f_B(\hat{s}^*)} > \frac{1-p}{p} \equiv \frac{f_G(\hat{s}^{FB})}{f_B(\hat{s}^{FB})}.
\]

Thus,

\[
\frac{w_{bB}(\hat{s}^*)}{w_{gG}(\hat{s}^*)} > 1 \implies w_{bB}(\hat{s}^*) > w_{gG}(\hat{s}^*)
\]

*Part two:* By Proposition 4, if \( p < 0.5 \), \( \hat{s}^* < \hat{s}^{FB} \), which combined with MLRP implies:

\[
\frac{f_G(\hat{s}^*)}{f_B(\hat{s}^*)} < \frac{1-p}{p} \equiv \frac{f_G(\hat{s}^{FB})}{f_B(\hat{s}^{FB})}.
\]

Thus,

\[
\frac{w_{bB}(\hat{s}^*)}{w_{gG}(\hat{s}^*)} < 1 \implies w_{bB}(\hat{s}^*) < w_{gG}(\hat{s}^*)
\]

**C.3.7 Proof of Proposition 4**

The proof of Proposition 4 is identical to the proof of Proposition 4B once we set \( q_G = 1 \), \( q_B = 0 \), and we replace \( w_{\tilde{b}B} \) with \( w_{bB} \), \( w_{\tilde{b}G} \) with \( w_{bG} \), \( w_{\tilde{g}B} \) with \( w_{gB} \), and \( w_{\tilde{g}G} \) with \( w_{gG} \).
C.3.8 Proof of Corollary 2

We focus on the case where \( p > 0.5 \). Similar intuition applies for the case where \( p < 0.5 \). Recall that by Proposition 4, for \( p > 0.5 \), \( \hat{s}^* > \hat{s}^{FB} \), whereas for \( p < 0.5 \), \( \hat{s}^* < \hat{s}^{FB} \).

**Part one:** Recall that for a given \( \hat{s} \), the expected probability that action \( b \) is chosen, is:

\[
Pr(d = b|\hat{s}) = pPr(s \leq \hat{s}|\theta = G) + (1 - p)Pr(s \leq \hat{s}|\theta = B)
= pF_G(\hat{s}) + (1 - p)F_B(\hat{s})
\]

where \( F_G(\hat{s}) \) and \( F_B(\hat{s}) \) are increasing in \( \hat{s} \). Thus, as \( \hat{s}^* > \hat{s}^{FB} \),

\[
Pr(d = b|\hat{s} = \hat{s}^*) > Pr(d = b|\hat{s} = \hat{s}^{FB})
\]

**Part two:** The probability that action \( b \) is revenue-maximizing is given by:

\[
Pr(\theta = B|d = b) = \frac{Pr(d = b|\theta = B)(1 - p)}{Pr(d = b|\theta = B)(1 - p) + Pr(d = b|\theta = G)p} = \frac{F_B(\hat{s})(1 - p)}{F_B(\hat{s})(1 - p) + F_G(\hat{s})p}
\]

Differentiating \( Pr(\theta = B|d = b) \) with respect to \( \hat{s} \), we obtain:

\[
\frac{\partial Pr(\theta = B|d = b)}{\partial \hat{s}} = \frac{(-1 + p)p[F_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})f_G(\hat{s})]}{[F_B(\hat{s})(1 + p) - F_G(\hat{s})p]^2}
\]

where, by Lemma A.1, is negative for any value of \( \hat{s} \). Recall that, if \( p > 0.5 \), \( \hat{s}^* > \hat{s}^{FB} \), thus:

\[
Pr(\theta = B|d = b, \hat{s} = \hat{s}^*) < Pr(\theta = B|d = b, \hat{s} = \hat{s}^{FB}) \quad (64)
\]

whereas, if \( p < 0.5 \), \( \hat{s}^* < \hat{s}^{FB} \), thus:

\[
Pr(\theta = B|d = b, \hat{s} = \hat{s}^*) > Pr(\theta = B|d = b, \hat{s} = \hat{s}^{FB}) \quad (65)
\]

**Part three:** The probability that action \( g \) is revenue-maximizing, is given by:

\[
Pr(\theta = G|d = g) = \frac{Pr(d = g|\theta = G)p}{Pr(d = g|\theta = G)p + Pr(d = g|\theta = B)(1 - p)}
= \frac{(1 - F_G(\hat{s}))p}{(1 - F_G(\hat{s}))p + (1 - F_B(\hat{s}))(1 - p)} \quad (66)
\]

Differentiating \( Pr(\theta = G|d = g) \) with respect to \( \hat{s} \), we obtain:

\[
\frac{\partial Pr(\theta = G|d = g)}{\partial \hat{s}} = \frac{(-1 + p)p[(1 - F_B(\hat{s}))f_G(\hat{s}) - f_B(\hat{s})(1 - F_G(\hat{s}))]}{[F_B(\hat{s})(1 + p) - F_G(\hat{s})p + 1]^2}
\]
where, by Lemma A.1, it is positive for any value of  \( \hat{s} \). Recall that, if  \( p > 0.5 \),  \( \hat{s}^* > \hat{s}^{FB} \), thus:

\[
Pr(\theta = G | d = g, \hat{s} = \hat{s}^*) > Pr(\theta = G | d = g, \hat{s} = \hat{s}^{FB})
\]

(67)

whereas, if  \( p < 0.5 \),  \( \hat{s}^* < \hat{s}^{FB} \), thus:

\[
Pr(\theta = G | d = g, \hat{s} = \hat{s}^*) < Pr(\theta = G | d = g, \hat{s} = \hat{s}^{FB})
\]

(68)

**Part four**: This part is a direct consequence of part two and three, and in particular, of relations (64) and (67) if  \( p > 0.5 \), and (65) and (68) for  \( p < 0.5 \).

### C.3.9 Proof of Proposition 5

**Part one**: In this part we show that if  \( p > 0.5 \) (\( p < 0.5 \)), an increase (decrease) in the cost of acquiring information from  \( c \) to  \( c' \) increases (decreases) the optimal threshold from  \( \hat{s}^* \) to  \( \hat{s}'^* \). Similar intuition applies for the case where  \( p < 0.5 \). First, we explore the case where the cost of acquiring information is  \( c \). Given that the principal chooses  \( \hat{s}^* \), it must be the case that his expected profit of implementing  \( \hat{s}'^* \) is not higher than the expected profit of implementing  \( \hat{s}^* \):

\[
E \Pi(\hat{s}'^*) \leq E \Pi(\hat{s}^* ) \implies \underbrace{E R(\hat{s}'^*) - E R(\hat{s}^*)}_D R \leq \underbrace{E C'(\hat{s}'^*) - E C'(\hat{s}^*)}_C D'
\]

(69)

where  \( E C'(.) \) denotes the expected compensation before the increase in the cost of acquiring information, and  \( E R(.) \) denotes the expected revenue.

Second, we explore the case where the cost of acquiring information is  \( c' > c \). Given that the principal chooses  \( \hat{s}'^* \), it must be the case that his expected profit of implementing  \( \hat{s}'^* \) is higher than the expected profit of implementing  \( \hat{s}^* \), i.e.,

\[
E \Pi(\hat{s}'^*) \geq E \Pi(\hat{s}^* ) \implies \underbrace{E R(\hat{s}'^*) - E R(\hat{s}^*)}_D R \geq \underbrace{E C''(\hat{s}'^*) - E C''(\hat{s}^*)}_C D''
\]

(70)

where  \( E C''(.) \) denotes the expected compensation after the increase in the cost of acquiring information. We prove part one by the method of contradiction. Suppose that  \( \hat{s}'^* < \hat{s}^* \). Then, by Lemma 2,  \( DR > 0 \) and by Lemma 4,  \( DC'' > 0 \) and  \( DC > 0 \). Relations (69) and (70) are not mutually exclusive as long as:

\[
DC \geq DR \geq DC'' \implies E C'(\hat{s}'^*) - E C'(\hat{s}^*) \geq E C''(\hat{s}'^*) - E C''(\hat{s}^*).
\]

(71)
Notice that (71) never holds because by $EC(\hat{s})$ is convex in $\hat{s}$ and linearly dependent on the cost of acquiring information (Lemma 3). Thus, an increase in $c$ leads to a steeper-sloped U-shape of $EC(\hat{s})$. As a result, (69) and (70) are mutually exclusive, hence, the initial hypothesis that $\hat{s}'^{*} < \hat{s}^{*}$ is not valid. Thus, $\hat{s}'^{*} > \hat{s}^{*}$.

**Part two:** Part one implies that as $c$ increases, $\hat{s}^{*}$ moves towards $\hat{s}_{\text{min}}$. Recall that $\hat{s}^{FB}$ is not related to $c$. Thus, the bias $|\hat{s}^{*} - \hat{s}^{FB}|$ is increasing in $c$.

**Part three:** The expected precision is denoted as $\Phi(\hat{s})$, where:

$$
\Phi(\hat{s}) = Pr(d = b|\hat{s})Pr(\theta = B|d = b) + Pr(d = g|\hat{s})Pr(\theta = G|d = g) \\
= F_{B}(\hat{s})(1 - p) + (1 - F_{G}(\hat{s}))p
$$

(72)

Recall that the expected return is given by:

$$
\mathbb{E}R(\hat{s}) = v\{p(1 - 2F_{G}(\hat{s})) + (1 - p)(2F_{B}(\hat{s}) - 1)\}.
$$

Thus, we can express $\Phi(\hat{s})$ as a function $\mathbb{E}R(\hat{s})$,

$$
\Phi(\hat{s}) = \frac{\mathbb{E}R(\hat{s})}{2v} + p.
$$

(73)

Thus, an increase in $c$ increases the bias $|\hat{s}^{*} - \hat{s}^{FB}|$, which decreases $\mathbb{E}R(\hat{s})$ and $\Phi(\hat{s})$.

**C.3.10 Proof of Proposition 6**

**Part one:** Note that $EC(\hat{s})$ is not a function of $p$. Thus, only $\mathbb{E}R(\hat{s})$ is affected by a change in $p$. Lemma A.3 explores the impact of a change in $p$ on $\mathbb{E}R(\hat{s})$.

**Lemma A.3:** Relationship between $p$ and $\mathbb{E}R(\hat{s})$.

(i) $\frac{\partial \mathbb{E}R(\hat{s})}{\partial \hat{s}}$ is decreasing in $p$.

(ii) $\frac{\partial^2 \mathbb{E}R(\hat{s})}{\partial \hat{s}^2}$ is not a function of $p$ when the signaling structure is linear.

**Proof.**

$$
\frac{\partial \mathbb{E}R(\hat{s})}{\partial \hat{s}} = 2v\{-pf_{G}(\hat{s}) + (1 - p)f_{B}(\hat{s})\} < 0
$$

$$
\frac{\partial \left(\frac{\partial \mathbb{E}R(\hat{s})}{\partial \hat{s}}\right)}{\partial p} = 2v\{-f_{G}(\hat{s}) - f_{B}(\hat{s})\} < 0
$$

since both $f_{G}(\hat{s})$ and $f_{B}(\hat{s})$ are positive.
\[
\frac{\partial^2 \mathbb{E} R(\hat{s})}{\partial \hat{s}^2} = 2v \{-p(f'_G(\hat{s}) + f'_B(\hat{s})) + f'_B(\hat{s})\}
\]

Note that for linear signaling structure, it holds that \(f'_G(\hat{s}) + f'_B(\hat{s}) = 0\).

Returning to the proof of the first part of Proposition 6, we denote \(\frac{\partial \mathbb{E} R(\hat{s})}{\partial \hat{s}}\) as \(\mathbb{E} MR(\hat{s})\), and \(\frac{\partial \mathbb{E} C(\hat{s})}{\partial \hat{s}}\) as \(\mathbb{E} MC(\hat{s})\). Given that \(\mathbb{E} MR(\hat{s})\) decreases in \(p\), the intersection of \(\mathbb{E} MR(\hat{s})\) with \(\mathbb{E} MC(\hat{s})\), which defines the equilibrium value \(\hat{s}^*\), moves towards lower values of \(\hat{s}\) as \(p\) increases.

![Figure 7: Impact of more extreme prior beliefs](image)

**Part two:** From Proposition 4, we know that if \(p > 0.5\), \(\hat{s}^{FB} < \hat{s}^* \leq \hat{s}_{min}\). For those values of \(\hat{s}^*\), and conditional on a linear signaling structure, the following four properties hold: i) \(\mathbb{E} MR(\hat{s})\) is linear in \(\hat{s} \in [0,1]\), ii) the slope of \(\mathbb{E} MR(\hat{s})\) is independent of \(p\) iii) \(\mathbb{E} MR(\hat{s})\) is negative for \(\hat{s} > \hat{s}^{FB}\) and positive otherwise, iv) \(\mathbb{E} MC(\hat{s})\) is increasing, and v) concave for \(\hat{s} \in [0,\hat{s}_{min}]\). These properties are captured in the Figure 8, where the blue and red line represent \(\mathbb{E} MR(\hat{s})\) before and after the increase in \(p\), whereas the red curve depicts the \(\mathbb{E} MC(\hat{s})\). Point \(A\) (\(B\)) denotes the first best \(\hat{s}\) before (after) the increase \(p\), whereas point \(A^*\) (\(B^*\)) denotes the equilibrium value of \(\hat{s}\) before (after) the increase \(p\). Thus, the distance \(AA^*\) is the bias before the increase in \(p\), whereas the distance \(BB^*\) is the bias after the increase in \(p\). Given that the aforementioned properties hold, it is always the case that the distance \(BB^*\) is greater than the distance \(AA^*\).

A more intuitive approach is that for a given deviation \(\eta\) from the first best, the drop in \(\mathbb{E} R(\hat{s})\), is independent of the value of \(\hat{s}^{FB}\), thus the value of \(p\), whereas the drop in \(\mathbb{E} C(\hat{s})\) is decreasing in \(p\). Hence, the incentive to deviate is stronger, the lower the value of \(p\) is.
C.4 Imperfect State Realization Case

C.4.1 Proof of Lemma 1B (Monotonic decision rule)

The expected profit of the principal when the agent takes action $b$ or $g$, given that the latter observes a signal $s$, is given by:

$$
\mathbb{E} \Pi[d = b|s] = Pr(\sigma = \tilde{G}|s)(-v - w_{bG}q_G - w_{b\tilde{B}}(1 - q_G)) + Pr(\sigma = \tilde{B}|s)(v - w_{b\tilde{B}}(1 - q_B) - w_{b\tilde{G}}q_{\tilde{B}})
$$

$$
\mathbb{E} \Pi[d = g|s] = Pr(\sigma = \tilde{G}|s)(v - w_{g\tilde{G}}q_G - w_{g\tilde{B}}(1 - q_B)) + Pr(\sigma = \tilde{B}|s)(-v - w_{g\tilde{G}}q_B - w_{g\tilde{B}}(1 - q_B))
$$

given that $Pr(\sigma = \tilde{B}|s) = 1 - Pr(\sigma = \tilde{G}|s)$, we can rearrange the previous relations as:

$$
\mathbb{E} \Pi[d = b|s] = \left(-2v - w_{b\tilde{G}}(q_G - q_B) + w_{b\tilde{B}}(q_G - q_B)\right) Pr(\sigma = \tilde{G}|s) + (v - w_{b\tilde{B}}(1 - q_B) - w_{b\tilde{G}}q_{\tilde{B}})
$$

$$
\mathbb{E} \Pi[d = g|s] = \left(2v - w_{g\tilde{G}}(q_G - q_B) + w_{g\tilde{B}}(q_G - q_B)\right) Pr(\sigma = \tilde{G}|s) + (-v - w_{g\tilde{B}}(1 - q_B) - w_{g\tilde{G}}q_B)
$$

Suppose that for $s = s'$ the principal prefers the agent to take action $g$. Then, it must hold:

$$
\mathbb{E} \Pi[d = g|s'] \geq \mathbb{E} \Pi[d = b|s'] \implies \quad A_L Pr(\sigma = \tilde{G}|s') + B_L \geq A_S Pr(\sigma = \tilde{G}|s') + B_S
$$

(74)

where,

$$
A_L - A_S = 4v - w_{g\tilde{G}}(q_G - q_B) + w_{g\tilde{B}}(q_G - q_B) + w_{bG}(q_G - q_B) - w_{b\tilde{B}}(q_G - q_B)
$$

$$
B_S - B_L = 2v - w_{b\tilde{B}}(1 - q_B) + w_{g\tilde{B}}(1 - q_B) - w_{b\tilde{G}}q_B + w_{g\tilde{G}}q_B
$$

(75)

A useful remark is that the principal’s net gain of taking the right action over taking the wrong action is $2v$ where the right action is the one which is supported by the public signal, i.e. $d = g$ if $\sigma = \tilde{G}$, and $d = b$ if $\sigma = \tilde{B}$. Hence, the maximum value that a payment can reach is $2v$. A payment higher than $2v$ would imply negative profits for the principal, as a result, the principal would not find profitable to incentivize the agent to acquire information. This remark implies that even in the extreme case where $w_{g\tilde{B}}$ and $w_{g\tilde{G}}$ reach their minimum value and $w_{b\tilde{B}}$ and $w_{b\tilde{G}}$ reach their maximum values, $B_S - B_L$ is positive. Hence, under the optimal contract, $B_S - B_L$, is always positive. Following the same intuition, it can be shown that $A_L - A_S$ is also

\footnote{If the realized state is good and the agent takes action $g$, principal’s net return is $v$. In contrast, if the agent takes action $b$, principal’s net return is $-v$. Likewise, if the realized state is bad and the agent takes action $b$, principal’s net return is $v$. In contrast, if the agent takes action $g$, principal’s net return is $v$. Thus, independently of the state of the world, the net benefit of taking the right over the wrong action is $2v.$}
positive. Thus, (74) can be expressed as:

\[ A_L - A_S \geq \frac{B_S - B_L}{Pr(\sigma = \bar{G}|s)}. \] (76)

Besides, given that \( Pr(\sigma = \bar{G}|s) \) is increasing in \( s \), if relation (76) holds for \( s = s' \), then it also holds for any \( s \geq s' \), i.e., \( \mathbb{E} \Pi[d = g|s'] - \mathbb{E} \Pi[d = b|s'] \) is increasing in \( s \). Also, by MLRP, there exists a unique \( s \), denoted as \( \hat{s} \), such as (76) binds. Thus, the decision rule that the principal prefers the agent to implement, is such as the latter goes short for \( s \leq \hat{s} \) and long otherwise.

C.4.2 Proposition 1B: Minimization Problem when state imperfectly observed

Constraints for implementing \( \mathcal{DR} \)

Similarly to the analysis of section 4.2, the agent implements \( \mathcal{DR} \) as long as his preferences are aligned with the preferences of the principal. This alignment pins down to the following three conditions, which are the analogue of (5), (6) and (7) in the benchmark model.

\[ DG = w_{g\bar{G}} - w_{b\bar{G}} \geq 0 \] (77)
\[ DB = w_{b\bar{B}} - w_{g\bar{B}} \geq 0 \] (78)
\[ Pr(\sigma = \bar{G}|\hat{s})[w_{g\bar{G}} - w_{b\bar{G}}] = Pr(\sigma = \bar{B}|\hat{s})[w_{b\bar{B}} - w_{g\bar{B}}] \] (79)

Information acquisition constraints

When the agent does not acquire information, his utility given the action is:

\[ \mathbb{E} V[d = g] = \{pq_G + (1-p)q_B\} w_{g\bar{G}} + \{p(1-q_G) + (1-p)(1-q_B)\} w_{b\bar{B}} \]
\[ \mathbb{E} V[d = b] = \{pq_G + (1-p)q_B\} w_{b\bar{G}} + \{p(1-q_G) + (1-p)(1-q_B)\} w_{g\bar{B}} \]

Hence, the agent’s expected utility if no information is acquired is given by:

\[ \mathbb{E} V[\text{no signal}] = \max\{\mathbb{E} V[d = g], \mathbb{E} V[d = b]\} \]

The agent’s expected utility of acquiring information, given a threshold \( \hat{s} \), is given by:

\[ \mathbb{E} V[\text{signal}|\hat{s}] = \int_0^{\hat{s}} \mathbb{E} C[d = b|s] f(s) ds + \int_{\hat{s}}^1 \mathbb{E} C[d = g|s] f(s) ds - c. \] (80)

where the expected compensation of action \( g \) and \( b \) given a signal realization \( s \), is:

\[ \mathbb{E} C[d = g|s] = Pr(\sigma = \bar{G}|s) \times w_{g\bar{G}} + Pr(\sigma = \bar{B}|s) \times w_{g\bar{B}} \] (81)
\[ \mathbb{E} C[d = b|s] = Pr(\sigma = \bar{G}|s) \times w_{b\bar{G}} + Pr(\sigma = \bar{B}|s) \times w_{b\bar{B}} \] (82)
Hence, the information acquisition constraints lead to:

\[
\mathbb{E}V[\text{signal} | \hat{s}] = -c + \\
\int_0^1 \left\{ \left( \frac{f_G(s)pq_G}{f(s)} + \frac{f_B(s)(1-p)q_B}{f(s)} \right)w_{bG} + \left( \frac{f_B(s)(1-p)(1-q_B)}{f(s)} + \frac{f_G(s)(1-q_G)}{f(s)} \right)w_{bB} \right\} f(s) ds + \\
\int_1 \left\{ \left( \frac{f_G(s)pq_G}{f(s)} + \frac{f_B(s)(1-p)q_B}{f(s)} \right)w_{gG} + \left( \frac{f_B(s)(1-p)(1-q_B)}{f(s)} + \frac{f_G(s)(1-q_G)}{f(s)} \right)w_{gB} \right\} f(s) ds
\]

which simplifies to:

\[
\mathbb{E}V[\text{signal} | \hat{s}] = -c + \\
\{ pq_G F_G(\hat{s}) + (1-p)q_B F_B(\hat{s}) \} w_{bG} + \{ (1-p)(1-q_B) F_B(\hat{s}) + p(1-q_G) F_G(\hat{s}) \} w_{bB} + \\
\{ pq_G (1-F_G(\hat{s})) + (1-p)q_B (1-F_B(\hat{s})) \} w_{gG} + \{ (1-p)(1-q_B)(1-F_B(\hat{s})) + p(1-q_G)(1-F_G(\hat{s})) \} w_{gB}
\]

Hence, the information acquisition constraints lead to:

\[
\mathbb{E}V[\text{signal} | \hat{s}] \geq \mathbb{E}V[d = g] \implies \\
\{ -pq_G F_G(\hat{s}) - (1-p)q_B F_B(\hat{s}) \} (w_{gG} - w_{bG}) + \\
\{ (1-p)(1-q_B) F_B(\hat{s}) + F_G(\hat{s}) (1-q_G) \} (w_{bB} - w_{gB}) \geq c
\]

\[
\mathbb{E}V[\text{signal} | \hat{s}] \geq \mathbb{E}V[d = b] \implies \\
\{ pq_G (1-F_G(\hat{s})) + (1-p)q_B (1-F_B(\hat{s})) \} (w_{gG} - w_{bG}) + \\
\{ -(1-p)(1-q_B)(1-F_B(\hat{s})) - p(1-q_G)(1-F_G(\hat{s})) \} (w_{bB} - w_{gB}) \geq c
\]

**Cost Minimization Problem**

Following the previous analysis, the cost minimization problem of incentivizing information acquisition, given an decision rule characterized by a threshold \( \hat{s} \), pins down to:

\[
\text{Minimize} \quad \mathbb{E}C(\hat{s})
\]

subject to \( \{ 77 \}, \{ 78 \}, \{ 79 \}, \{ 85 \}, \{ 86 \}, \) \( w_{bG} \geq 0, \) \( w_{bB} \geq 0, \) \( w_{gG} \geq 0, \) \( w_{gB} \geq 0, \) where \( \mathbb{E}C(\hat{s}) \) is:

\[
\mathbb{E}C(\hat{s}) = \\
\{ pq_G F_G(\hat{s}) + (1-p)q_B F_B(\hat{s}) \} w_{bG} + \{ (1-p)(1-q_B) F_B(\hat{s}) + p(1-q_G) F_G(\hat{s}) \} w_{bB} + \\
\{ pq_G (1-F_G(\hat{s})) + (1-p)q_B (1-F_B(\hat{s})) \} w_{gG} + \{ (1-p)(1-q_B)(1-F_B(\hat{s})) + p(1-q_G)(1-F_G(\hat{s})) \} w_{gB}
\]
C.4.3 Proof of Proposition 1B

Proof that \( w^*_{bg}(\hat{s}) = 0 \).

Substituting constraints (77) and (78) into (79) implies:

\[
DB = \frac{\{pq_g f_G(\hat{s}) + (1 - p)q_B f_B(\hat{s})\}}{\{p(1 - q_g) f_G(\hat{s}) + (1 - p)(1 - q_B) f_B(\hat{s})\}} DG
\]  
(88)

By substituting (88) into (85) and rearranging, we obtain:

\[
DG \geq \frac{f_B(\hat{s})(-1 + p)(-1 + q_B) - f_G(\hat{s})p(-1 + q_B)}{(-1 - F_B(\hat{s})) f_G(\hat{s}) + f_B(\hat{s})(1 - F_G(\hat{s}))) (1 + p)p(q_B - q_G)} c
\]  
(89)

where the RHS of (89) is positive, as an implication of Lemma A.2.

Substituting (88) into (86) and rearranging, we obtain:

\[
DG \geq \frac{f_B(\hat{s})(-1 + p)(-1 + q_B) - f_G(\hat{s})p(-1 + q_B)}{(-1 - F_B(\hat{s})) f_G(\hat{s}) + f_B(\hat{s})(1 - F_G(\hat{s}))) (1 + p)p(q_B - q_G)} c
\]  
(90)

where the RHS of (90) is positive, as an implication of Lemma A.2.

The previous analysis shows that incentive constraints of the maximization problem are satisfied as long as \( DG \) satisfies (89) and (90). Hence, the principal would never find it optimal to offer a contract where \( w_{bg}^* \) is positive, because, by decreasing \( w_{gG}^* \) and \( w_{bG}^* \) by the same amount, the incentive constraints are unaffected, and the expected compensation is lower.

Proof that \( w^*_{gb}(\hat{s}) = 0 \).

Substituting constraints (78) and (77) into (79) implies:

\[
DG = \frac{\{p(1 - q_g) f_G(\hat{s}) + (1 - p)(1 - q_B) f_B(\hat{s})\}}{\{pq_g f_G(\hat{s}) + (1 - p)q_B f_B(\hat{s})\}} DB
\]  
(91)

By substituting (91) into (85) and rearranging, we obtain:

\[
DB \geq \frac{f_B(\hat{s})(1 - p)q_B + f_G(\hat{s})pq_G}{(F_B(\hat{s}) f_G(\hat{s}) - f_B(\hat{s}) F_G(\hat{s}))(1 + p)p(q_B - q_G)} c
\]  
(92)

where the RHS of (92) is positive, as an implication of Lemma A.2.

By substituting (91) into (86) and rearranging, we obtain:

\[
DB \geq \frac{f_B(\hat{s})(1 - p)q_B + f_G(\hat{s})pq_G}{(f_B(\hat{s})(1 - F_G(\hat{s})) - (1 - F_B(\hat{s})) f_G(\hat{s}))(1 + p)p(q_B - q_G)} c
\]  
(93)
where the RHS of (93) is positive, as an implication of Lemma A.2.

The previous analysis shows that incentive constraints of the maximization problem are satisfied as long as $DB$ satisfies (91) and (92). Hence, the principal would never find it optimal to offer a contract where $w_{gB}$ is positive, because, by decreasing both $w_{bB}$ and $w_{gB}$ by the same amount, the incentive constraints are unaffected, and the expected payment is lower. Following the previous analysis, the minimization problem pins down to:

\[
\text{Minimize } \mathbb{E}C(\hat{s}) \quad \text{s.t.}
\]

\[
\begin{align*}
\{-pqGF_G(\hat{s}) - (1-p)qBF_B(\hat{s})\}w_{g\hat{G}} + \{(1-p)(1-qB)F_B(\hat{s}) + F_G(\hat{s})p(1-qG)\}w_{b\hat{B}} &\geq c \quad (94) \\
\{pqG(1-F_G(\hat{s})) + (1-p)qB(1-F_B(\hat{s}))\}w_{gG} - \\
\{(1-p)(1-qB)(1-F_B(\hat{s})) + p(1-qG)(1-F_G(\hat{s}))\}w_{bB} &\geq c \quad (95) \\
\{pqGf_G(\hat{s}) + (1-p)qBf_B(\hat{s})\}w_{g\hat{G}} = \{p(1-qG)f_G(\hat{s}) + (1-p)(1-qB)f_B(\hat{s})\}w_{b\hat{B}} \\
& \quad w_{b\hat{B}} \geq 0, \ w_{g\hat{G}} \geq 0 \quad (96)
\end{align*}
\]

**Redundant constraints**

After substituting (96) into (94) and (95), simple algebra implies that (94) is redundant by (95) as long as:

\[
f_G(\hat{s}) \geq f_B(\hat{s})
\]

which holds for $s \in [\hat{s}_{min}, 1]$. Consequently, for $s \in [0, \hat{s}_{min}]$, (95) is redundant by (94).

**Case where $s \in [\hat{s}_{min}, 1]$.**

In this case, (94) becomes redundant. By substituting (96) into (95) and rearranging, we obtain:

\[
w_{g\hat{G}} \geq \frac{f_B(\hat{s})(-1 + p)(1 - qB) + f_G(\hat{s})p(-1 + qG)}{((1 - F_B(\hat{s}))f_G(\hat{s}) + f_B(\hat{s})(-1 + F_G(\hat{s})))(-1 + p)p(qB - qG)}c.
\]

Note that under the optimal contract, (95) binds, otherwise the principal could increase his profit by decreasing $w_{g\hat{G}}$ until (95) is binding. Thus,

\[
w_{g\hat{G}}^*(\hat{s}) = \frac{f_B(\hat{s})(-1 + p)(1 - qB) + f_G(\hat{s})p(-1 + qG)}{((1 - F_B(\hat{s}))f_G(\hat{s}) + f_B(\hat{s})(-1 + F_G(\hat{s})))(-1 + p)p(qB - qG)}c.
\]

The last step is to substitute $w_{g\hat{G}}^*(\hat{s})$ into (96) to derive the optimal value of $w_{b\hat{B}}^*(\hat{s})$ which is:

\[
w_{b\hat{B}}^*(\hat{s}) = \frac{f_B(\hat{s})(-1 + p)qB - f_G(\hat{s})pqG}{(f_B(\hat{s})(-1 + F_G(\hat{s})) + f_G(\hat{s})(1 - F_B(\hat{s})))(-1 + p)p(qB - qG)}c.
\]
Note that under the optimal contract, (94) binds, otherwise the principal could increase his profit by decreasing \( w_gG \) until (94) is binding. Thus,

\[
\frac{\partial w_{gG}(\hat{s})}{\partial \hat{s}} = \frac{f_B(\hat{s})(-1 + p)(1 + q_B) - f_G(\hat{s})(-1 + q_B)}{(f_B(\hat{s}) f_G(\hat{s}) - f_B(\hat{s}) F_G(\hat{s}))(1 + p)(q_B - q_G)}
\]

The last step is to substitute \( w_{gG}^*(\hat{s}) \) into (96) to derive the optimal value of \( w_{bB}^*(\hat{s}) \), which is:

\[
w_{bB}^*(\hat{s}) = \frac{f_B(\hat{s})(1 - p)q_B + f_G(\hat{s})pq_G}{(f_B(\hat{s}) f_G(\hat{s}) - f_B(\hat{s}) F_G(\hat{s}))(1 + p)(q_B - q_G)}
\]

**C.4.4 Proof of Lemma 2B**

**Part one:** The derivative of \( w_{bB}(\hat{s}) \) and \( w_{gG}(\hat{s}) \) with respect to \( \hat{s} \) (for \( \hat{s} \in [0, \hat{s}_{min}] \)), is given by:

\[
\frac{\partial w_{bB}(\hat{s})}{\partial \hat{s}} = c \frac{(pq_G f_G(\hat{s}) + (1 - p)q_B f_B(\hat{s})) \gamma_1(\hat{s})}{(-1 + p)(q_B - q_G)(f_B(\hat{s} f_G(\hat{s}) - f_B(\hat{s}) F_G(\hat{s}))^2}
\]

\[
\frac{\partial w_{gG}(\hat{s})}{\partial \hat{s}} = -c \frac{((1 - p)(-1 + q_B) F_B + p(-1 + q_G) F_G(\hat{s})) \gamma_1(\hat{s})}{p(-1 + p)(q_B - q_G)(f_B(\hat{s} f_G(\hat{s}) - f_B(\hat{s}) F_G(\hat{s}))^2}
\]

where \( \gamma_1(\hat{s}) \equiv [f_B'(\hat{s}) f_G(\hat{s}) - f_B(\hat{s}) f_G'(\hat{s})] \). Recall that \( q_G \geq q_B \). Also, we show in Lemma A.1 that \( \gamma_1 < 0 \). Hence, both \( w_{bB}(\hat{s}) \) and \( w_{gG}(\hat{s}) \) are decreasing in \( \hat{s} \) for \( \hat{s} \in [0, \hat{s}_{min}] \).

**Part two:** The derivative of \( w_{bB}(\hat{s}) \) and \( w_{gG}(\hat{s}) \) with respect to \( \hat{s} \) (for \( \hat{s} \in [\hat{s}_{min}, 1] \)), is given by:

\[
\frac{\partial w_{bB}(\hat{s})}{\partial \hat{s}} = c \frac{(pq_G(1 - F_G(\hat{s})) + (1 - p)(F_B(\hat{s}) - 1))(-\gamma_1(\hat{s}))}{(-1 + p)(q_B - q_G)((1 - F_G(\hat{s})) f_B(\hat{s}) - f_G(\hat{s})(1 - F_B(\hat{s}))^2}
\]

\[
\frac{\partial w_{gG}(\hat{s})}{\partial \hat{s}} = -c \frac{-K_1}{p(-1 + p)(q_B - q_G)((1 - F_G(\hat{s})) f_B(\hat{s}) + f_G(\hat{s})(1 - F_B(\hat{s})))^2}
\]

where \( \frac{\partial w_{bB}(\hat{s})}{\partial \hat{s}} \geq 0 \), given that \( q_G \geq q_B \) and \( \gamma_1 < 0 \). Defining the sign of \( \frac{\partial w_{gG}(\hat{s})}{\partial \hat{s}} \) is less straightforward. Notice that \( \partial K_1/\partial q_G < 0 \). Hence, in order to show that \( K_1 \) is negative, it suffices to show that \( K_1 < 0 \) for the minimum value of \( q_G \), i.e., \( q_G = q_B \). Replacing \( q_G \) with \( q_B \) in \( K_1 \) leads to negative \( K_1 \), thus, \( K_1 < 0 \) for any admissible value of \( q_G \). Also, by Lemma A.1, \( \gamma_1 < 0 \), hence, both \( w_{bB}(\hat{s}) \) and \( w_{gG}(\hat{s}) \) are increasing in \( \hat{s} \) for \( \hat{s} \in [\hat{s}_{min}, 1] \).
Part three: Part three is a direct consequence of part one and two.

Part four: We first provide the intuition behind part one, which is going to act as a stepping stone for part four. The intuition behind part one (and part two) builds on three remarks. First, the agent is willing to implement \( \hat{s} \), as long as:

\[
Pr(\sigma = \tilde{G}|\hat{s})/Pr(\sigma = \tilde{B}|\hat{s}) = w_{b\tilde{B}}/w_{g\tilde{G}}. \tag{99}
\]

Second, under the optimal contract, the agent’s expected utility when acquiring information should be equal to his outside option of not acquiring information, given by:

\[
max\{EV(\text{no signal} \& d = g), EV(\text{no signal} \& d = b)\} =
max\{(pqG + (1-p)qB)w_{g\tilde{G}}, (p(1-qG) + (1-p)(1-qB))w_{b\tilde{B}}\}
\]

Third, the opportunity cost of not taking the revenue-maximizing action, as indicated by the public signal, is \( w_{g\tilde{G}} \), when the agent takes action \( b \) and the revenue-maximizing action is \( g \), and \( w_{b\tilde{B}} \), when the agent takes action \( g \) and the right action is \( b \).

Given these remarks, in order to shed light on the intuition behind part one, we explore the ramifications of a deviation in the implementation threshold from \( \hat{s}_{min} \) to \( \hat{s}' < \hat{s}_{min} \). Suppose that \( p = 0.5 \) and that the principal aims to implement \( \hat{s} = \hat{s}_{min} \). For \( \hat{s} = \hat{s}_{min} \), the agent’s outside option is equal to \( 0.5(qG + qb)w_{g\tilde{G}}(\hat{s}_{min}) \) and the cost in case of taking the wrong action is \( w_{b\tilde{B}}(\hat{s}_{min}) \). Suppose now that the principal aims to implement a lower threshold, \( \hat{s}' < \hat{s}_{min} \), for which \( Pr(\sigma = \tilde{G}|\hat{s}')/Pr(\sigma = \tilde{B}|\hat{s}') < Pr(\sigma = \tilde{G}|\hat{s}_{min})/Pr(\sigma = \tilde{B}|\hat{s}_{min}) \). For relation (99) to be satisfied, it must be that:

\[
w_{b\tilde{B}}(\hat{s}')/w_{g\tilde{G}}(\hat{s}') = Pr(\sigma = \tilde{G}|\hat{s}')/Pr(\sigma = \tilde{B}|\hat{s}').
\]

Note that \( Pr(\sigma = \tilde{G}|\hat{s}')/Pr(\sigma = \tilde{B}|\hat{s}') < w_{b\tilde{B}}(\hat{s}_{min})/w_{g\tilde{G}}(\hat{s}_{min}) \). There are four ways to achieve this.

(i) \( w_{g\tilde{G}}(\hat{s}') > w_{g\tilde{G}}(\hat{s}_{min}) \& w_{b\tilde{B}}(\hat{s}') = w_{b\tilde{B}}(\hat{s}_{min}) \): In this case, the agent’s outside option improves, whereas the cost of taking the wrong action is unaffected. Thus, the outside option now exceeds the agent’s utility when he acquires information. Hence, this case is not feasible.

(ii) \( w_{b\tilde{B}}(\hat{s}') < w_{b\tilde{B}}(\hat{s}_{min}) \& w_{g\tilde{G}}(\hat{s}') = w_{g\tilde{G}}(\hat{s}_{min}) \): In this case, the agent’s utility from acquiring information drops, given that the reward when the agent takes action \( b \) drops, whereas

\[16\]Similar intuition applies for part two.
the agent’s outside option is unaffected. Thus, agent’s utility when acquiring information falls below the agent’s outside option. Hence, this case is not feasible.

(iii) \( w_{gG}(\hat{s}') < w_{gG}(\hat{s}_{min}) \) \& \( w_{bB}(\hat{s}') < w_{bB}(\hat{s}_{min}) \): This case is not feasible. If this case were feasible, it would imply that the principal when implementing \( \hat{s}_{min} \) would be able to increase his profits by decreasing both \( w_{gG}(\hat{s}_{min}) \) and \( w_{gG}(\hat{s}_{min}) \) without violating (99). However, such a profitable deviation would contradict with the initial hypothesis that \( w_{gG}(\hat{s}_{min}) \) and \( w_{gG}(\hat{s}_{min}) \) are the optimal payments.

(iv) \( w_{gG}(\hat{s}') > w_{gG}(\hat{s}_{min}) \) \& \( w_{bB}(\hat{s}') < w_{bB}(\hat{s}_{min}) \): In this case, both the agent’s outside option and the cost of taking the wrong action increase. For \( w_{gG}(\hat{s}') \) small enough and for \( w_{bB}(\hat{s}') \) large enough, the two opposing forces cancel each other. Hence, this is the only feasible case.

Summarizing the previous analysis, we show that as the value of \( \hat{s} \) decreases, the value of \( \frac{f_G(\hat{s})}{f_B(\hat{s})} \), and effectively of \( \frac{Pr(\sigma = G|\hat{s})}{Pr(\sigma = B|\hat{s})} \), decreases as well. Hence, the principal, in order to incentivize the agent to take action \( g \) for \( s \in (\hat{s}, 1] \), should offer a high relative payment \( \frac{w_{bB}}{w_{gG}} \). This can only happen if the principal increases \( w_{gG} \). This increase, however, improves the agent’s outside option, and tempts him to take action \( g \) without acquiring costly information. Hence, in order to prevent the agent from taking action \( g \) without acquiring information, the principal should increase the opportunity cost of not taking the revenue-maximizing action, i.e., offer a higher payment \( w_{bB} \).

Following the reasoning of the previous paragraph, the convexity of \( w^*_L(\hat{s}) \) and \( w^*_S(\hat{s}) \) can be seen in the following example. Suppose that we start from a case where the principal implements \( \hat{s} = 0.5 \), and the optimal payments are denoted by \( w^*_L(0.5) \) and \( w^*_S(0.5) \). Suppose now that the principal aims to implement \( \hat{s} = 0.49 \). In order to incentivize the agent to take action \( g \) for \( s \in (0.49, 0.50) \), the principal should increase \( w_{gG} \) from \( w^*_L(0.5) \) to \( w^*_G \), such as \( \frac{w'_{gG}}{w''_{gG}(0.5)} = \frac{Pr(\sigma = B|s=0.49)}{Pr(\sigma = G|s=0.49)} \). At the same time, the increase in \( w_{gG} \) triggers the agent to take action \( g \) without acquiring information. Thus, in order to prevent the agent from taking action \( g \) without acquiring information, the principal should also increase \( w_{bB} \) to \( w'_{bB} \). However, after this increase, \( \frac{w'_{gG}}{w'_{bB}} < \frac{Pr(\sigma = B|s=0.49)}{Pr(\sigma = G|s=0.49)} \), thus the principal should also increase \( w'_{gG} \) to \( w''_{gG} \), which will in turn, trigger another increase in \( w''_{bB} \) to \( w''_{bB} \) and so on. Notice that each loop leads to smaller increases in \( w_{gG} \) and \( w_{bB} \) because the opportunity cost of not taking the revenue-maximizing action \( w_{bB} \) becomes very high. This process converges to the equilibrium, where the equilibrium payments are denoted by \( w^*_L(0.49) \) and \( w^*_S(0.49) \).

Suppose now that the principal aims to implement \( \hat{s} = 0.48 \). In order to incentivize the agent to take action \( g \) for \( s \in (0.48, 0.49) \), the principal should increase \( w_{gG} \) from \( w^*_L(0.49) \) to \( w^*_{gG} \), such as \( \frac{w^*_{gG}}{w'_{gG}(0.49)} = \frac{Pr(\sigma = B|\hat{s}=0.48)}{Pr(\sigma = G|\hat{s}=0.48)} \). Recall that \( \frac{f_G(\hat{s})}{f_B(\hat{s})} \), and effectively \( \frac{Pr(\sigma = G|\hat{s})}{Pr(\sigma = B|\hat{s})} \), is increasing in \( \hat{s} \) due to the MLRP assumption. Hence, the combination of MLRP and the finding that
$w^*_{sb}(0.49) > w^*_{sb}(0.50)$ implies that the increase $w^+_{gG} - w^*_Lg(0.49)$ will be higher than the increase $w'_{gG} - w^*_Lg(0.5)$. This implies that the incentive to take action $g$ without acquiring information is stronger than before, thus, $w_{sb}$ should increase significantly, which in turn, leads to a higher $w_{gG}$, such as the principal has incentive to take action $g$ for $s \in (0.48, 0.49)$. Thus, the equilibrium payments should increase more when the principal deviates from $\hat{s} = 0.49$ to $\hat{s} = 0.48$, compared to the deviation from $\hat{s} = 0.5$ to $\hat{s} = 0.49$, i.e., $w^*_{Lg}(0.49) - w^*_{Lg}(0.50) > w^*_{Lg}(0.48) - w^*_{Lg}(0.49)$, and $w^*_{sb}(0.49) - w^*_{sb}(0.50) > w^*_{sb}(0.48) - w^*_{sb}(0.49)$.

The last paragraphs refer to the part of Lemma 2B which explores deviations to lower values of $\hat{s}$. Similar intuition applies for the case where the principal considers shifting from implementing $\hat{s} = \hat{s}_{min}$ to implementing $\hat{s} = \hat{s}'' > \hat{s}_{min}$.

### C.4.5 Expected cost of implementing $\hat{s}$

Corollary 1B derives $\mathbb{E}C(\hat{s})$ which arises from Proposition 1B. Lemma 3 extends to this set-up.

**Corollary 1B:** Expected compensation cost of implementing $\hat{s}$, $\mathbb{E}C(\hat{s})$

For $\hat{s} \in [0, \hat{s}_{min}]$, the expected compensation cost is given by:

$$
\mathbb{E}C(\hat{s}) \equiv \mathbb{E}C(\hat{s})^- = 
\begin{align*}
(f_B(\hat{s})(-1+p)(-1+q_B) - p(-1+q_G)f_G(\hat{s}))&((-1+p)q_B(-1+F_B(\hat{s})) - pq_G(-1+F_G(\hat{s}))) \\
+ (f_G(\hat{s})pq_G + q_B(1-p)f_B(\hat{s}))&((-1+p)(-1+q_B)F_B(\hat{s}) - p(-1+q_G)F_G(\hat{s})) \\
(F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s}))&(-1+p)p(q_B - q_G)
\end{align*}
\right] \ c
\tag{100}
$$

For $\hat{s} \in [\hat{s}_{min}, 1]$, the expected compensation cost is given by:

$$
\mathbb{E}C(\hat{s}) \equiv \mathbb{E}C(\hat{s})^+ = 
\begin{align*}
(pf_G(\hat{s})(-1+q_G)(q_B - pq_B + pq_G) + (1-p)(q_B - q_G)F_B(\hat{s})) \\
+ (-1+p)f_B(\hat{s})((-1+q_B)((-1+p)q_B - pq_G) + p(q_B - q_G)F_G(\hat{s}))) \\
((-1+F_G(\hat{s})f_B(\hat{s}) + (1-F_B(\hat{s}))f_G(\hat{s}))(-1+p)p(q_B - q_G))
\end{align*}
\right] \ c
\tag{101}
$$

### C.4.6 Proof of Lemma 3

One can easily notice the similarity of Lemma 3 with Lemma 2 – the behavior of $\mathbb{E}C(\hat{s})$ with respect to $\hat{s}$ follows a similar pattern with the behavior of $w_{hB}(\hat{s})$ and $w_{gG}(\hat{s})$ with respect to $\hat{s}$. This relates to the following remark. Under the optimal contract, the expected utility of the risk-neutral agent when implementing $\hat{s}$: i) coincides with his expected compensation reduced
by the cost of acquiring information, \( c \), and ii) equals his outside option, which is given by:

\[
\mathbb{E} V(\hat{s}) = \max \{ \mathbb{E} V(\text{no signal} \& d = g), \mathbb{E} V(\text{no signal} \& d = b) \} = \max \{ p \times w^*_g(\hat{s}), (1 - p) \times w^*_b(\hat{s}) \}.
\] (102)

Hence, following Proposition 1, the expected cost \( \mathbb{E} C(\hat{s}) \) for \( \hat{s} < \hat{s}_{\text{min}} \) equals:

\[
\mathbb{E} C(\hat{s}) = p \times w^*_g(\hat{s}) + c
\] (103)

whereas, the expected cost \( \mathbb{E} C(\hat{s}) \) for \( \hat{s} > \hat{s}_{\text{min}} \) equals:

\[
\mathbb{E} C(\hat{s}) = (1 - p) \times w^*_b(\hat{s}) + c.
\] (104)

Thus, the intuition behind the relation of \( \mathbb{E} C(\hat{s}) \) with \( \hat{s} \) is identical to the underlying intuition in Lemma 2.

C.4.7 Proof of Proposition 2B

The optimal contract consists of two parts: i) the optimal payment scheme, given \( \hat{s} \), and, ii) an optimality condition for \( \hat{s} \). The first part is characterized in Proposition 1B. The optimality condition for \( \hat{s} \), given by (28), is characterized in Lemma A.4.

Lemma A.4: Optimality Conditions

For \( \hat{s} \leq \hat{s}_{\text{min}} \), optimal value \( \hat{s} \), \( \hat{s}^* \) solves:

\[
2v\left\{-pf_G(\hat{s}) + (1 - p)f_B(\hat{s})\right\} = \left[\frac{((-1 + p)q_B - pq_G)((-1 + p)(-1 + q_B)F_B(\hat{s}) - p(-1 + q_G)F_G(\hat{s}))\gamma_1(\hat{s})}{[F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s})(-1 + p)p(q_B - q_G)]^2} \right]c \] (105)

\[\frac{\partial \mathbb{E} R(\hat{s})}{\partial \hat{s}} \]

\[\frac{\partial \mathbb{E} C(\hat{s})}{\partial \hat{s}}\]
For \( \hat{s} \geq \hat{s}_{\min} \), optimal value \( \hat{s}, \hat{s}^* \) solves:

\[
2v\{ -pf_G(\hat{s}) + (1 - p)f_B(\hat{s}) \} = \partial \mathbb{E} R(\hat{s})/\partial \hat{s}
\]

\[
- \left[ \frac{(1 - (1 - p)q_B - pq_G)(1 - p)q_B(1 - F_B(\hat{s})) + pq_G(1 - F_G(\hat{s})))\gamma_1(\hat{s})}{((-1 + F_B(\hat{s}))f_g(\hat{s}) + (1 - F_G(\hat{s}))f_B(\hat{s})(-1 + p)p(q_B - q_G))^2} \right] c
\]

(106)

where \( \gamma_1(\hat{s}) \equiv [f'_B(\hat{s})f_G(\hat{s}) - f_B(\hat{s})f'_G(\hat{s})] \).

We show in Proposition 4B, that for \( p > 0.5 \) (\( p < 0.5 \)), it is never optimal to choose a threshold \( \hat{s} \in (\hat{s}_{\min}, 1) \) (\( \hat{s} \in [0, \hat{s}_{\min}] \)). Hence, for \( p \geq 0.5 \), the optimality condition is (105), whereas, for \( p \leq 0.5 \), the optimality condition is (106). The combination of this observation with Lemma A.4 and Proposition 1B, leads to Proposition 2B.

C.4.8 Proof of Proposition 4B

Proposition 4B emerges naturally from Lemma 3B and Lemma 4. For instance, if \( p > 0.5 \), we know that \( \hat{s}^{FB} < \hat{s}_{\min} \). Also, \( \hat{s}^* \) cannot belong to \( [0, \hat{s}^{FB}] \) as by switching to a higher threshold the expected revenue increases, and the expected cost decreases. Similarly, \( \hat{s}^* \) cannot belong to \( [\hat{s}_{\min}, 1] \) as by switching to a lower threshold the expected revenue increases, and the expected cost decreases. Hence, \( \hat{s}^* \in (\hat{s}^{FB}, \hat{s}_{\min}) \). A more formal proof is provided below.

Recall that the optimality conditions for the case where \( p \geq 0.5 \) and \( p \leq 0.5 \) are given by:

\[
\partial \mathbb{E} R(\hat{s})/\partial \hat{s} = \partial \mathbb{E} C^-(\hat{s})/\partial \hat{s}
\]

\[
\partial \mathbb{E} R(\hat{s})/\partial \hat{s} = \partial \mathbb{E} C^+(\hat{s})/\partial \hat{s}
\]

By Lemma 3B, \( \partial \mathbb{E} C^-(\hat{s})/\partial \hat{s} \) is negative, and \( \partial \mathbb{E} C^+(\hat{s})/\partial \hat{s} \) is positive. Also, by Lemma 4, \( \partial \mathbb{E} R(\hat{s})/\partial \hat{s} \) is positive for \( s < \hat{s}^{FB} \) and negative for \( s > \hat{s}^{FB} \). Hence, for \( p \geq 0.5 \), for an equilibrium to exist it must be that \( \partial \mathbb{E} R(\hat{s})/\partial \hat{s} = \mathbb{E} C^-(\hat{s})/\partial \hat{s} \). Since \( \partial \mathbb{E} C^-(\hat{s})/\partial \hat{s} \) is negative, then \( \partial \mathbb{E} R(\hat{s})/\partial \hat{s} \) has to be negative, which is true only if \( s > \hat{s}^{FB} \). Similar intuition applies for the case where \( p \leq 0.5 \).

C.4.9 Lemma 6

Lemma 6 explores the relationship between \( \mathbb{E} C(\hat{s}) \) and \( q_G \) and \( q_B \).

**Lemma 6**: Relationship between \( \mathbb{E} C(\hat{s}) \) and \( q_G, q_B \).
Figure 8: Bias in the investment decision

(i) $\mathbb{E} C(\hat{s})$ is decreasing in $q_B$.

(ii) $\mathbb{E} C(\hat{s})$ is increasing in $q_B$.

(iii) $\frac{\partial \mathbb{E} C(\hat{s})}{\partial q_G}$ is increasing in $q_G$ for $\hat{s} \in [0, \hat{s}_{\text{min}})$ and decreasing in $q_G$ for $\hat{s} \in (\hat{s}_{\text{min}}, 1]$.

(iv) $\frac{\partial \mathbb{E} C(\hat{s})}{\partial q_G}$ is decreasing in $q_B$ for $\hat{s} \in [0, \hat{s}_{\text{min}})$ and increasing in $q_B$ for $\hat{s} \in (\hat{s}_{\text{min}}, 1]$.

**Proof.** Part one: We explore the derivative of $\mathbb{E} C(\hat{s})$ with respect to $q_G$ and $q_B$, first for $\hat{s} \in [0, \hat{s}_{\text{min}}]$, and second for $\hat{s} \in [\hat{s}_{\text{min}}, 1]$.

\[
\frac{\partial \mathbb{E} C(\hat{s})^-}{\partial q_G} = c \frac{f_B(\hat{s})(-1 + p)(-1 + q_B)q_B + f_G(\hat{s})p(q_B(1 - q_B) + p(q_G - q_B)^2)}{(F_B(\hat{s})f_G(\hat{s}) - F_G(\hat{s})f_B(\hat{s}))(1 + p)p(q_B - q_G)^2}
\]

\[
\frac{\partial \mathbb{E} C(\hat{s})^+}{\partial q_G} = c \frac{f_B(\hat{s})(-1 + p)(-1 + q_B)q_B + f_G(\hat{s})p(q_B(1 - q_B) + p(q_G - q_B)^2)}{f_G(\hat{s})(1 - F_B(\hat{s})) - f_B(\hat{s})(1 - F_G(\hat{s}))(1 + p)p(q_B - q_G)^2}
\]

which are negative as an implication of Lemma A.2.

\[
\frac{\partial \mathbb{E} C(\hat{s})^-}{\partial q_B} = -c \frac{f_B(\hat{s})(-1 + p)(-1 + q_B)q_B + f_G(\hat{s})\overline{p}(q_B + (-1 + p)q_B^2 - 2pq_Bq_G + pq_G^2)}{f_G(\hat{s})(1 - F_B(\hat{s})) - f_B(\hat{s})(1 - F_G(\hat{s}))(1 + p)p(q_B - q_G)^2}
\]

which is positive as an implication of Lemma A.2. Note that $e_1$ is increasing in $q_B$. Hence, it is easy to show that for the maximum feasible value of $q_B$, i.e., $q_G$, $e_1$ is negative.
Part three: First, we analyze the case where \( \hat{s} \in [0, \hat{s}_{\min}) \).

\[
\frac{\partial E\text{MC}(\hat{s})^-}{\partial q_G} = -c \gamma(\hat{s})(-1 + p)(-1 + q_B)q_B^2 \left[ -p(1 + q_B) + \frac{(1 + p)q_B^2}{-2pq_Bq_G + pq_G^2} \right] \frac{L_1}{(F_B(\hat{s})_G(\hat{s}) - f_B(\hat{s})F_G(\hat{s}))^2} \]

where \( \gamma(\hat{s}) \equiv (f_B(\hat{s})F_G(\hat{s}) - f_B(\hat{s})F_G(\hat{s})). \) Note that the denominator is always negative. Also, by Lemma A.1, \( \gamma(\hat{s}) > 0 \). Notice that \( \partial L_1/\partial q_G > 0 \). Hence, if \( L_1 \) is positive for the minimum feasible value of \( q_G \), i.e. \( q_G = q_B \), then \( L_1 \) is positive for any feasible value of \( q_G \), where:

\[
L_1|_{q_G=qb} = (F_B(\hat{s})(-1 + p) - pF_G(\hat{s}))(-1 + q_B)q_G > 0
\]

Hence, \( \frac{\partial E\text{MC}(\hat{s})^-}{\partial q_G} > 0 \). We now analyze the case where \( \hat{s} \in (\hat{s}_{\min}, 1] \):

\[
\frac{\partial E\text{MC}(\hat{s})^+}{\partial q_G} = \frac{\gamma(\hat{s})(-1 + p)(-1 + q_B)q_B^2 + p(1 + q_B)q_B^2 + p(1 + p)q_B^2 - 2pq_Bq_G + pq_G^2}{((1 - F_B(\hat{s}))F_G(\hat{s}) - f_B(\hat{s})(1 - F_G(\hat{s})))^2} \frac{L_2}{(1 - p)(q_B - q_G)^2}
\]

where the denominator is always negative. Also, by Lemma A.1, \( \gamma(\hat{s}) > 0 \). Notice that \( \partial L_2/\partial q_G < 0 \). Hence, if \( L_2 \) is negative for the minimum feasible value of \( q_G \), i.e. \( q_G = q_B \), then \( L_2 \) is negative for any feasible value of \( q_G \), where:

\[
L_2|_{q_G=qb} = (1 + F_B(\hat{s})(-1 + p) - F_G(\hat{s})q_B)(-1 + q_G)q_B < 0
\]

Hence, \( \frac{\partial E\text{MC}(\hat{s})^+}{\partial q_G} < 0 \).

Part four: First, we analyze the case where \( \hat{s} \in [0, \hat{s}_{\min}) \).

\[
\frac{\partial E\text{MC}(\hat{s})^-}{\partial q_B} = \gamma(\hat{s}) \left[ \frac{((1 + p)(1 + p)q_B^2 - 2(-1 + p)q_Bq_G + q_G(-1 + p)q_G)}{(F_B(\hat{s})F_G(\hat{s}) - f_B(\hat{s})F_G(\hat{s}))^2} \frac{L_3}{(1 - p)(q_B - q_G)^2}
\]

where the denominator is always negative. Also, by Lemma A.1, \( \gamma(\hat{s}) > 0 \). Notice that \( \partial L_3/\partial q_B < 0 \). Hence, if \( L_3 \) is positive for the maximum feasible value of \( q_B \), i.e. \( q_B = q_G \), then \( L_3 \) is positive for any feasible value of \( q_B \), where:

\[
L_3|_{q_B=qb} = (F_B(\hat{s})(-1 + p) - F_G(\hat{s}))(-1 + q_G)q_G > 0
\]

Hence, \( \frac{\partial E\text{MC}(\hat{s})^-}{\partial q_G} < 0 \). We now analyze the case where \( \hat{s} \in (\hat{s}_{\min}, 1] \).

\[
\frac{\partial E\text{MC}(\hat{s})^+}{\partial q_B} = -c \frac{\gamma L_4}{\kappa}
\]

(108)

65
where
\[ \gamma \equiv f_B(\hat{s})f'_G(\hat{s}) - f'_B(\hat{s})f_G(\hat{s}) \]  
(109)

\[ L_4 \equiv (-q_B^2 + 2pq_B - p^2q_B - q_G + 2q_Bq_G - 4pq_Bq_G + 2p^2q_Bq_G + 2pq_G - p^2q_G + (-1 + p)((-1 + p)q_B^2 - 2(-1 + p)q_Bq_G + q_G(-1 + pq_G))f_B(\hat{s}) - p(-1 + q_G)q_Gf_G(\hat{s}) \]  
(110)
\[ \kappa \equiv ((1 - F_B(\hat{s}))f_G(\hat{s}) - f_B(\hat{s})(1 - F_G(\hat{s})))^2(-1 + p)p(q_B - q_G)^2 \]  
(111)

where the denominator is always negative. Also, by Lemma A.1, \( \gamma > 0 \). Notice that \( \partial L_4 / \partial q_B > 0 \). Hence, if \( L_4 \) is negative for the maximum value of \( q_B \), i.e. \( q_B = q_G \), then \( L_4 \) is negative for any feasible value of \( q_B \), where:

\[ L_4|_{q_B=q_G} = (1 + F_B(\hat{s})(-1 + p) - F_G(\hat{s})p)(-1 + q_B)q_B < 0 \]

Hence, \( \frac{\partial EM C(\hat{s})^+}{\partial q_G} > 0 \).

\[ \square \]

C.4.10 Proof of Proposition 7

The proof relies on five remarks: i) the expected return of implementing \( \hat{s} \), \( \mathbb{E} R(\hat{s}) \), is independent of \( q_G \) and \( q_B \), ii) for \( p > 0.5 \), \( \hat{s}^* \in (\hat{s}^{FB}, \hat{s}_{min}] \), and for \( p < 0.5 \), \( \hat{s}^* \in [\hat{s}_{min}, \hat{s}^{FB}) \), iii) by Lemma 4, the expected marginal revenue \( \mathbb{E} MR(\hat{s}) \equiv \partial \mathbb{E} R(\hat{s}) / \partial \hat{s} \) is positive for \( \hat{s} < \hat{s}^{FB} \) and negative for \( \hat{s} > \hat{s}^{FB} \), iv) \( \frac{\partial EM C(\hat{s})}{\partial \hat{s}} \) is increasing in \( q_G \) for \( \hat{s} \in [0, \hat{s}_{min}] \), and decreasing in \( q_G \), for \( \hat{s} \in (\hat{s}_{min}, 1] \), v) \( \frac{\partial EM C(\hat{s})}{\partial \hat{s}} \) is decreasing in \( q_B \) for \( \hat{s} \in [0, \hat{s}_{min}] \), and increasing in \( q_B \) for \( \hat{s} \in (\hat{s}_{min}, 1] \)

**Part one:** For \( p > 0.5 \) where \( \hat{s}^* \in (\hat{s}^{FB}, \hat{s}_{min}] \), an increase in \( q_G \) leads to a greater (still negative) expected marginal cost, thus the intersection with the expected marginal revenue corresponds to lower values of \( \hat{s} \). Similarly, for \( p < 0.5 \) where \( \hat{s}^* \in [\hat{s}_{min}, \hat{s}^{FB}) \), an increase in \( q_G \) leads to a lower (still positive) expected marginal cost, thus the intersection with the expected marginal revenue corresponds to higher values of \( \hat{s} \).

**Part two:** For \( p > 0.5 \) where \( \hat{s}^* \in (\hat{s}^{FB}, \hat{s}_{min}] \), an increase in \( q_B \) leads to a lower (still negative) expected marginal cost, thus the intersection with the expected marginal corresponds to higher values of \( \hat{s} \). Similarly, for \( p < 0.5 \) where \( \hat{s}^* \in [\hat{s}_{min}, \hat{s}^{FB}) \), an increase in \( q_G \) leads to a higher (still positive) expected marginal cost, thus the intersection with the expected marginal corresponds to lower values of \( \hat{s} \).

**Part three:** Part three is a direct implication of part one and two.
Table 1: Analogy

<table>
<thead>
<tr>
<th>Setting</th>
<th>Portfolio I</th>
<th>CRA</th>
<th>Product Design</th>
<th>Innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal</td>
<td>Equity holders</td>
<td>Manager</td>
<td>Equity holders</td>
<td>Venture Capitalist</td>
</tr>
<tr>
<td>Agent</td>
<td>Manager</td>
<td>Analyst</td>
<td>Manager</td>
<td>Researcher</td>
</tr>
<tr>
<td>State</td>
<td>Type of risky</td>
<td>Creditworthiness</td>
<td>Demand</td>
<td>Quality of innovation</td>
</tr>
<tr>
<td>Action</td>
<td>Risky/safe</td>
<td>Good/bad rating</td>
<td>Features</td>
<td>Conventional/innovation</td>
</tr>
<tr>
<td>Objective</td>
<td>Profits</td>
<td>Rating precision</td>
<td>Profits</td>
<td>Profits</td>
</tr>
<tr>
<td>Implications compared to the First Best</td>
<td>Over (under) investment if NPV of risky lower (higher) than NPV of safe asset.</td>
<td>Rating deflation (inflation) if ex-ante more (less) likely the asset type to be good.</td>
<td>Over-implementation to features which are ex-ante less likely to succeed in accommodating demand.</td>
<td>Under(over)-implementation of the innovative strategy when its NPV is higher (lower) than the NPV of conventional strategy.</td>
</tr>
</tbody>
</table>