Indicator Choice in Pay-for-Performance

Majid Mahzoon, Ali Shourideh, Ariel Zetlin-Jones
Carnegie Mellon University*

December 26, 2022

Abstract

We study the classic principal-agent model when the signal observed by the principal is chosen by the agent. We fully characterize the optimal information structure from an agent's perspective in a general moral hazard setting. Unlike standard information design, the full distribution of beliefs is relevant to the principal. We show that the problem can be mapped into a geometrical game between the principal and the agent in the space of likelihood ratios. We use this representation result to show that coarse contracts are sufficient: The agent can achieve her best with binary signals. Additionally, we characterize conditions under which the agent is able to extract all the surplus and implement the efficient allocation under full information. Finally, we show that when effort and performance are one-dimensional, under a general class of models, threshold signals are optimal.

1 Introduction

The use of pay-for-performance contracting is a cornerstone of modern employment contracts. Executive compensation is often indexed in part to company per-

*Emails: mmahzoon@andrew.cmu.edu, ashourid@andrew.cmu.edu, azj@andrew.cmu.edu

formance metrics including growth in the company's stock price. Employment contracts for top athletes frequently involve performance bonuses for specific outcomes such as goals scored for soccer players or the number of touch-downs for players in the NFL. When employment contracts feature such performance pay incentives, employers and employees must agree to a set of performance indicators during contract negotiations.[1] Given this observation, what are the incentives of employees and employers to negotiate on performance indicators? If the employees have a role in choosing the performance metrics, what metrics would they choose? Finally, how does the choice of indicators interact with the ultimate productive efficiency of the firm?

In this paper, we answer these questions by considering the problem of indicator design in the textbook moral hazard problem with limited liability. More specifically, we consider the standard principal-agent problem of Holmström (1979) in which an agent has quasi-linear preferences and must be paid a non-negative wage. The *performance technology* is one that maps costly effort, e, by the agent into a distribution of some performance measure x. Before the principal offers a compensation contract, the agent chooses an *indicator*, a possibly random signal s of x where the principal can only offer a contract that is contingent on the indicator, s. Once the indicator is chosen, the principal and the agent play the textbook moral hazard game.[2]

In this environment, one might conjecture that more information would lead to more efficient outcomes. While this is true under certain circumstances – see the informativeness principle of Holmström (1979), and its extension by Chaigneau et al. (2019) – information can often be detrimental to the agent. To see the intu-

[1] Bebchuk and Fried (2004) discuss the various issues with the negotiations process between the CEOs and boards and possible issues arising from choosing particular performance indicators, i.e., vesting stocks. In soccer, news outlets often describe the process in which players and soccer clubs agree on what performance measure to use. See for example this article in Daily Mail which describes several soccer players in the Premier League negotiating over the relevant performance indicator as a base for pay.

[2] Throughout the paper, we assume that the agent commits to the indicator s while she cannot commit to the effort level e. A justification for this assumption is that employment contracts are often enforced by courts and thus hard to break.

ition for this observation, suppose that the performance technology is degenerate at $x = e$ and thus the agent can use her effort as the indicator. Then, revealing all information by choosing $s = x = e$ would give the principal the ability to fully capture all the surplus generated by the effort. On the other hand, choosing a fully uninformative indicator leads to no surplus for the agent as the principal is unable to incentivize the agent to contribute effort when signals are fully uninformative. This points to a trade-off in the problem of indicator design by the agent.

In this model, we show three main results: first, we provide a geometric interpretation of the indicator design game in the space of likelihoods; the probability of a particular performance/signal realization for an arbitrary effort relative to the effort that the agent would like to implement. Under this interpretation, the agent's indicator design problem is equivalent to a geometric game where the agent chooses a convex set (of likelihood ratios) and the principal chooses a point within that set. Our geometric interpretation provides a tractable formulation to understand how the agent's choice of indicator influences the resulting compensation scheme offered by the principal. Second, we provide conditions under which the agent is able to choose the indicator in such a way to implement the first-best efficient effort and capture all the surplus created by her effort. Finally, we consider a specific case where performance measure x has a continuous distribution in real numbers and show that under certain conditions optimal indicator structure takes the form of monotone or hump-shaped thresholds signals.

The reason that the agent is sometimes able to capture all of the surplus and implement the efficient outcome can be easily understood when performance technology is fully informative, i.e., $x = e$. In this case, suppose the agent chooses an indicator with two realizations: high and low. By choosing the probability of high and low indicators appropriately for every $x = e$, the agent is able to force the principal's hand. Note that if the principal wants to implement a given effort level $\hat{e}$, he must compensate the agent following a realization of the high signal. The amount of compensation the principal must deliver is increasing as the cost of the desired effort level $\hat{e}$ rises relative to the cost of any other effort and is increasing in the likelihood of observing a high signal from some other effort relative to that of observing a high signal from effort $\hat{e}$. By choosing the signal probabilities appropriately, the agent is then able to make the off-path likelihood of the high signal sufficiently large so that the principal must give up all the surplus in order to implement the desired effort level $\hat{e}$.

In general, our geometric interpretation allows us to show that the agent can implement any desired effort level with a "coarse" information structure that has at most two signal realizations. Additionally, this interpretation allows us to provide conditions under which even when the performance technology is stochastic, it is possible for the agent to capture all the surplus and implement the first-best outcome. Our sufficient condition amounts to checking whether a particular point in the likelihood space belongs to convex hull of the likelihood functions implied by the performance technology. This type of sufficient condition is easy to evaluate using existing convex hull algorithms.

Beyond our technical contributions, our results shed light on the debate on the efficiency of incentive contracts in executive compensation. Bebchuk and Fried (2004) have argued that the standard agency model of shareholder maximization is at odds with the data since negotiations often happen between the CEO and the board whose incentives are not necessarily aligned with that of the shareholders – in fact they claim that CEOs and compensation committees often trade favors at a cost to shareholders. In our model, and consistent with Bebchuk and Fried (2004)'s interpretation, we endow the agent with full bargaining power over the choice of performance pay indicators. While the agent captures all of the gains from trade under this assumption, her choices also maximize total surplus. In contrast, standard models of moral hazard with exogenous performance technology often feature inefficiencies in the sense that providing incentives to the agent often entails reductions in total surplus. In this sense, we find that optimal design of performance pay indicators may help to reduce inefficiencies within the firm. Further investigation of the data on negotiations and choice of indicators would be a good test of our theory.

1.1 Related Literature

Our paper is related to several strands of the literature on contracting and information design.

With respect to the moral hazard literature, our key innovation is to consider the problem of choosing indicators whereupon contracts are based on showing that this can remove all inefficiencies. In the classical model, Innes (1990) and Poblete and Spulber (2012) consider moral hazard with limited liability and a risk-neutral agent, and show that simple debt contracts are optimal. Carroll (2015) as well as Walton and Carroll (2022) consider the moral hazard problem with limited liability, where the principal has non-Bayesian uncertainty about the production technology and wishes to maximize her worst-case payoff. [3]

While often information design incentives are ignored in moral hazard, Holmström (1979) and later Chaigneau et al. (2019) are exceptions. They investigate the comparative statics of changing the performance technology on the payoffs; by the informativeness principle, the more informative the output is about the effort, the lower wage the principal needs to pay.

Perhaps, the closest paper to ours is that of Garrett et al. (2020). They consider a model in which the agent can design the performance technology and cost function – they refer to this as technology design – and show that the agent-optimal design involves only binary distributions. In contrast, in our model, the performance technology is fixed and the agent chooses an indicator of this performance, i.e., an information structure to garble the principal's observations. Moreover, our paper has a technical contribution by reformulating the problem in terms of likelihood ratios.

Georgiadis and Szentes (2020) study moral hazard with limited liability and a risk-averse agent where the principal continuously observes signals about the

[3] It is needless to say that a rather large body of work has considered models with risk-averse agents and risk-neutral principals.

agent's effort at a constant marginal cost. They show that the principal-optimal information-acquisition strategy is a two-threshold policy. Barron et al. (2020) consider an agent who can costlessly add mean-preserving noise to the output. This is also the case in our model; however, in our setting, the noisy output is just used for contracting purposes and does not change the principal's payoff compared to the original output. Moreover, in our model, first the agent chooses the information structure, then the principal offers the contract, and finally the agent chooses her effort, whereas in Barron et al. (2020), first the principal offers a contract and then the agent chooses the effort and the information structure.

Another related strand of the moral hazard literature focuses on career concerns introduced by Holmström (1979) and changes in information structure observed by both the principal and the agent. In this context Holmström (1999) shows that noisy performance signals are beneficial for incentives. Similarly Dewatripont et al. (1999) show that more informative signals in the sense of Blackwell do not necessarily increase incentives. In contrast in our model, indicator design or information design can be used as a tool to reshuffle surplus between the agent and the principal while achieving efficiency.

In the context of team production and moral hazard, Halac et al. (2021) show that the principal can benefit from private contract offers by leveraging rank uncertainty: Each agent is informed only of her own bonus and a ranking distribution; each agent's bonus makes work dominant if higher-rank agents work. Interestingly, in our setting, it is the agent that can use uncertainty about performance for the principal to improve efficiency.

We also contribute to the growing literature on incentives in Bayesian Persuasion. Several papers, including Boleslavsky and Kim (2018), Rosar (2017), Perez-Richet and Skreta (2022), Ball (2019), Saeedi and Shourideh (2020), and Zapechelnyuk (2020) have considered the effect of incentives in the Bayesian persuasion problem where a third party designs an information structure, and a "sender" determines the distribution of the underlying state by exerting a costly effort. From a technical perspective, our problem is different from this class of models. This is partly due to the fact that for the ex-post incentive to exert effort by the agent (after the choice of indicator and contract), the distribution of the signals – or alternatively in the language of Kamenica and Gentzkow (2011), the distribution of posteriors – off the equilibrium path is also relevant. By casting the problem in the space of likelihood functions – as opposed to beliefs as in Kamenica and Gentzkow (2011) – we can characterize its solution using the geometric game. This technique can be used in other information design problems in which on- and off-path beliefs are involved.

The rest of the paper is organized as follows: Section 2 describes the key insight about full surplus extraction and first-best implementation in a simple example. Section 3 describes the basic model. Section 4 describes the geometric interpretation of the indicator choice game between the principal and the agent. Section 6 describes optimality of threshold signals. Section 7 concludes.

2 A Simple Example

In this section, we use a basic environment to illustrate the main mechanisms at work in our model. Consider the basic textbook model of moral hazard. A principal (he) is hiring an agent (she) to perform a task whose output $x \in \{0, 1\}$ is collected by the principal. The agent chooses how much effort to put in to perform the task. She can either choose the low effort e_L or a costly high effort e_H whose cost is given by $c > 0$.

Choosing the high effort leads to output $x = 1$ with certainty while choosing the low effort leads to output $x = 1$ with probability $p < 1$ and $x = 0$ with probability $1 - p$. We assume that the total surplus under high effort $1 - c$ is higher than that under low effort p and thus it is efficient to implement the high effort.

In this standard principal-agent model with moral hazard, principal observes the output but not the effort of the agent. He can compensate the agent for each output realization but cannot make these payments negative, i.e., he is subject to limited liability. If the principal sets wage w when output is high, then the agent's incentive compatibility constraint is

$$w - c \geq p \cdot w.$$

Hence, as long as $w \geq \frac{c}{1-p}$, the principal can implement the high effort. If the principal is to choose $w = \frac{c}{1-p}$, then his payoff is $1 - \frac{c}{1-p} > 0$ while the agent's payoff is $\frac{c}{1-p} - c = \frac{p}{1-p}c > 0$. Moreover, for the principal to prefer implementing the high effort to low, we must have that $1 - \frac{c}{1-p} \geq p$ or $c \leq (1-p)^2$.

Now suppose that the agent can control principal's information about the output. Specifically, suppose that before the contracting stage, the agent can design a device that can potentially hide the output of the project. More specifically, suppose that the agent can choose an information structure or a Blackwell experiment that probabilistically maps output $x \in \{0, 1\}$ to a signal $S = \{L, H\}$ which is observed by the principal. The principal observes the signal $s = H$ with probability π_H when $x = 1$ is realized and observes $s = H$ with probability π_L if $x = 0$ is realized, where $\pi_L < \pi_H$.

Since the principal can only observe the signal designed by the agent, he will compensate her only when $s = H$. If this compensation is w, then the agent's incentive compatibility constraint is

$$\pi_H \cdot w - c \geq (\pi_H p + (1-p)\, \pi_L) \cdot w.$$

Hence, the principal is able to implement high effort when

$$w \geq \frac{c}{(\pi_H - \pi_L)(1-p)}$$

When minimizing the wage, the expected cost of compensating the agent for the principal is $\pi_H w = \frac{c}{\left(1 - \frac{\pi_L}{\pi_H}\right)(1-p)}$. Therefore, as long as $p \leq 1 - \frac{c}{\left(1 - \frac{\pi_L}{\pi_H}\right)(1-p)}$, the principal finds it profitable to implement the high effort. This inequality can be rewritten as

$$\frac{c}{(1-p)^2} \leq 1 - \frac{\pi_L}{\pi_H}.$$

This, in turn, implies that when $\frac{c}{(1-p)^2} \leq 1$, the agent can find a signal structure (π_L, π_H) such that the above holds with equality. Under such an information structure, the payoff of the principal is p, what he can achieve without any costly effort, and the agent captures the rest of the surplus, $1 - p - c$. In other words, giving the agent the ability to choose an information structure enables her to guarantee the highest value of the surplus under an efficient level of effort. Intuitively, the change of the information structure allows the agent to induce an arbitrary high value of the wage by increasing the likelihood π_L/π_H, and capture the entire surplus.

The above example illustrates that agent's freedom to choose the information structure, based on which she will be paid, can be extremely powerful. A few natural questions arise: When can the agent capture the efficient level of surplus? What information structure should be used by the agent to achieve her desired outcome? In what follows, we provide a characterization of the optimal signal structure by the agent as well as her ability to extract surplus from the principal.

3 Model

Our general model builds upon the textbook moral hazard problem. Consider a principal employing an agent to perform a task whose output is represented by $x \in X$, where X is finite. The agent chooses effort $e \in E = \{e_1, \cdots, e_m\}$ to perform the task, where E is finite. The agent's effort choice induces a probability distribution $f(x|e)$ over the outcome space X, where $\sum_x f(x|e) = 1, \forall e \in E$. We refer to $f(\cdot|\cdot)$ as the performance technology. Effort is costly to the agent; the cost of putting effort e is given by $c(e)$ for some real-valued function $c : E \to \mathbb{R}_+$.

Throughout the analysis, we assume that $e_1 \in E$ represents the effort with the lowest cost; for simplicity, let $c(e_1) = 0$. The principal's payoff from realization of output x is given by $g(x)$ for some real-valued function $g : X \to \mathbb{R}$.

The principal cannot observe the agent's effort and thus cannot offer a contract contingent on the agent's effort; he can only offer contracts contingent on observable outcomes.

The point of departure from the textbook moral hazard model is that the agent may influence the principal's information about the output by choosing an information structure (S, π). Here, S is a signal space and $\pi(\cdot|x) : X \to \Delta(S)$ is a stochastic mapping from the output space X to the signal space, where $\sum_s \pi(s|x) = 1,\ \forall x \in X$. The principal only observes the signal $s \in S$ generated from this information structure and can thus only offer a contract contingent on this signal realization. Therefore, the principal's choice of contract can be represented by a real-valued function $w : S \to \mathbb{R}_+$ where $w(s)$ is the wage paid to the agent when signal s is realized. One interpretation of this contractual restriction is that the task output is not observable to the principal but the agent may verifiably disclose information about the output. An alternative interpretation is that the output is observable but the agent has the option to choose the performance measure based on which she will be paid.

Note that as in the simple example, we have assumed that agent enjoys limited liability, i.e., the contract offered to her by the principal guarantees a non-negative wage regardless of the effort she puts in. The principal's payoff u_P is equal to the payoff from output less the wages paid to the agent. The agent's payoff u_A is equal to the wage she receives from the principal minus the cost of her effort. Both the principal and the agent are assumed to maximize expected utility. Notice that principal can always implement the zero-cost effort, i.e. e_1, by offering $w(s) = 0,\ \forall s \in S$. Therefore, his outside option is to implement e_1 and obtain $\underline{u}_P = \sum_x g(x) f(x|e_1)$.

The timing of the game is as follows:

- Agent chooses an information structure (S, π). To ease the exposition, we assume the signal space S is finite and simply represent the information structure by π. [4]

[4] Later, we show this restriction to finite signal spaces is without loss of generality.

- Observing the information structure (S, π) chosen by the agent, principal offers the agent a contract $w : S \to \mathbb{R}_+$ contingent on the realized signal.

- Observing the contract w offered by the principal (and the information structure (S, π) she has chosen), the agent chooses how much effort e to exert.

- Given the effort e chosen by the agent, output x is realized according to $f(x|e)$ and then signal $s \in S$ is realized according to $\pi(s|x)$.

- Payoffs are realized where agent's payoff is $u_A = w(s) - c(e)$, and the principal's payoff is $u_P = g(x) - w(s)$.

To summarize, the game is played sequentially in three stages. In the first stage, the agent chooses the information structure π to (partially) inform principal about the realized output. In the second stage, the principal offers the agent a contract w contingent on the signal realization. In the last stage, the agent chooses her effort e.

The equilibrium concept is the standard subgame perfect equilibrium (SPE). The agent's strategy $(\pi, \sigma_e(\pi, w))$ consists of a choice of information structure π and a choice of effort $\sigma_e(\pi, w)$ as a function of π and the contract w offered by the principal. The principal's strategy $\sigma_w(\pi)$ is the contract he offers to the agent as a function of the information structure π chosen by the agent.

Definition 1. An SPE of the game $\sigma^* = (\pi^*, \sigma_e^*(\pi, w), \sigma_w^*(\pi))$ consists of a strategy for the agent $(\pi^*, \sigma_e^*(\pi, w))$ and a strategy for the principal $\sigma_w^*(\pi)$ such that:

- $\sigma_e^*(\pi, w)$ maximizes the agent's expected utility for every information structure π and every principal's choice of contract w.

- $\sigma_w^*(\pi)$ maximizes the principal's expected utility for every agent's choice of information structure π given agent's equilibrium effort strategy $\sigma_e^*(\pi, w)$.

- π^* maximizes the agent's expected utility given principal's equilibrium strategy $\sigma_w^*(\pi)$ and her own equilibrium effort strategy $\sigma_e^*(\pi, w)$.

We let $p(\cdot|e) : E \to \Delta(S)$ represent the resulting stochastic mapping from the effort space E to the signal space S induced by a given information structure (S, π) and the underlying probability distribution of outcomes give effort. Note, for any level of effort e and any signal realization s, $p(s|e) = \sum_x f(x|e)\pi(s|x)$ and $\sum_s p(s|e) = 1,\ \forall e \in E$. Conditional on a given information structure (S, π), both the principal and the agent use the stochastic mapping p to evaluate their expected payoffs.

We now formulate the problems the principal and the agent solve beginning with the agent's choice of effort. The agent's expected utility is $U_A = \sum_s p(s|e)w(s) - c(e)$. Therefore, the agent's problem in the last stage of the game (where π and w have previously been chosen in preceding stages) is

$$\max_e \sum_s p(s|e)w(s) - c(e). \tag{IC-A}$$

The principal's expected utility is $U_P = \sum_x g(x)f(x|e) - \sum_s p(s|e)w(s)$ if he chooses to implement effort e. It is convenient to define $\mathbb{E}[g(x)|e] = \sum_x g(x)f(x|e)$. For a given desired effort level e, the principal chooses a wage schedule that solves

$$\begin{aligned} \min_w \ & \sum_s p(s|e)w(s) \text{ s.t.} \\ & \sum_s p(s|e)w(s) - c(e) \geq \sum_s p(s|\hat{e})w(s) - c(\hat{e}),\ \forall \hat{e} \in E, \\ & w(s) \geq 0,\ \forall s \in S. \end{aligned} \tag{1}$$

The constraints represent the agent's incentive compatibility and a set of limited liability constraints. Notice that we have not imposed a participation constraint for the agent. This is because the assumption $c(e_1) = 0$ together with limited liability implies that setting $e = e_1$ guarantees a non-negative payoff for the agent. Let $W(e, \pi)$ represent the optimal value in (1). This is the minimum expected wage the principal must pay the agent to implement effort e given the information structure π chosen by the agent.

Given $W(e, \pi)$, the problem of the principal is to choose an effort level to maximize her expected utility, or,

$$\max_{e} \mathbb{E}\left[g\left(x\right)|e\right] - W(e, \pi). \tag{2}$$

It is possible to express the above as an incentive compatibility constraint and thus write the agent's problem in the first stage of the game as

$$\max_{e,\pi} W(e, \pi) - c(e)$$

subject to the principal's incentive compatibility constraint

$$\mathbb{E}\left[g\left(x\right)|e\right] - W(e, \pi) \geq \mathbb{E}\left[g\left(x\right)|\hat{e}\right] - W(\hat{e}, \pi),\ \forall \hat{e} \in E. \tag{IC-P}$$

3.1 Remarks on the Environment

It is useful to discuss various interpretations of the model as well as our key assumptions.

Performance Measure as Information Structure

In the textbook model of moral hazard, the principal cannot observe the agent's effort. He therefore uses some imperfect signal of effort to incentivize the agent. If he can observe the output, he offers an output-contingent contract; this makes the output the performance measure for the agent. In our model, the principal does not observe the output but does observe a signal, which may be correlated with effort, and he offers a signal-contingent contract. As a result, the signal is the relevant performance measure for the agent. By choosing the information structure, the agent influences the set of observables that will ultimately dictate her compensation, which we interpret as the agent choosing her performance measure.

We make the assumption that the principal cannot offer output-contingent contracts. As we have described, while a conventional interpretation of this assumption is that the principal cannot observe output directly, an alternative interpretation is that this restriction arises during the negotiations between the agent and the principal in choosing contractual performance measures.

Commitment

We assume that the agent commits to an information structure in the first stage of the game. Our interpretation of the information structure as a contractual performance measure provides a natural justification of the commitment assumption. When signing a contract, all parties are aware of and agree on the probabilistic nature of the chosen performance measure as a function of the output. The contractible nature of the performance measure makes the commitment assumption necessary.

Comparison to the Literature on Bayesian Persuasion

As in the Bayesian persuasion literature, we can write the problem in terms of the distribution of posteriors induced by the information structure. In the Bayesian persuasion, every choice of information structure induces a distribution of posteriors, where the whole distribution matters: not only the support, but also the probabilities. In our setting, every choice of information structure induces a set of distributions of posteriors, one for each choice of effort. These distributions are related through their supports: given the support of one distribution, the supports of the others are pinned down. As we will discuss in section 4, the key sufficient statistic about the choice of information structure is the distribution of the likelihood ratios and these are determined by the supports of the above distributions. Therefore, in our model, only the support of the distribution of posteriors matters.

4 A Geometric Analysis of the Game

We now characterize the equilibrium outcomes of the game. To do so, we first describe the set of effort levels that are implementable by some information structure. We then show a "coarse"-ness result. That is, we show that it is without loss of generality to restrict the agent's choice of information structures to binary structures, where the set of signals has only two discrete points. Using such information structures, we derive sufficient conditions such that the first-best level of effort is implementable. When these sufficient conditions are satisfied, we show that the agent chooses an information structure that implements the first-best effort level and extracts the entire surplus.

While it is possible to work with zero probability events and define likelihood ratios – by describing how division by 0 is defined – in order to avoid complications, we make the following assumption:

Assumption 1. *The performance technology is full support, i.e.,* $\forall x \in X, \forall e \in E, f(x|e) > 0$.

This assumption implies that all the likelihood ratios below are well-defined.

Implementable Effort. To characterize the set of implementable effort levels, we use backward induction and first re-cast the problem of the principal geometrically. Specifically, we show that the likelihood ratios for each signal realization s for any effort level e relative to the desired, implementable level e^* are sufficient statistics to solve the principal's problem. In other words, we argue that any desired effort level to implement e^* and any information structure (S, π) give rise to a (geometric) space of possible likelihood ratios and that the principal's optimal choice of compensation schemes may be reduced to choosing a point in this space of likelihood ratios.

More formally, we define an implementable effort level e^* as follows:

Definition 2. An effort level e^* is *implementable* if there exists an information structure (S, π) such that e^* is a solution to the principal's problem (2).

Given this definition, an implementable effort level e^* must satisfy the two incentive compatibility constraints in (IC-P) and (IC-A) for the principal and the agent where the agent's incentive compatibility constraint must hold for all possible histories including those following a deviation by the principal that involves recommending an alternative level of effort.

Let e^* represent some effort level the agent would like to implement (in the first stage of the game). To motivate the relevance of likelihood ratios, consider the agent's interim incentive compatibility constraint when the principal only pays the agent following a signal realization s:

$$p\left(s|e^*\right) w\left(s\right) - c\left(e^*\right) \geq p\left(s|\hat{e}\right) w\left(s\right) - c\left(\hat{e}\right), \forall \hat{e}.$$

These constraints imply that for any effort level $\hat{e}$ where signal s is less likely than under e^*, i.e., $p\left(s|\hat{e}\right) < p\left(s|e^*\right)$, the wage must satisfy

$$w\left(s\right) \geq \frac{c\left(e^*\right) - c\left(\hat{e}\right)}{p\left(s|e^*\right) - p\left(s|\hat{e}\right)}, \tag{3}$$

and if s is more likely under $\hat{e}$ than under e^* then

$$\frac{c\left(\hat{e}\right) - c\left(e^*\right)}{p\left(s|\hat{e}\right) - p\left(s|e^*\right)} \geq w\left(s\right). \tag{4}$$

The first set of constraints (3) imply that the expected wage must satisfy

$$p\left(s|e^*\right) w\left(s\right) \geq \max_{\hat{e}:p(s|e^*)>p(s|\hat{e})} \frac{c\left(e^*\right) - c\left(\hat{e}\right)}{1 - \frac{p(s|\hat{e})}{p(s|e^*)}}.$$

In other words, the expected cost of implementing e^* for the principal (and its implicit benefit for the agent) is determined by the likelihood of state s. The second set of constraints (4) place an upper bound on the wages the principal may deliver while respecting incentives of the agent. As we show below, this restriction can also be formulated in terms of likelihood ratios.

The above illustration only holds under the assumption that the principal only compensates the agent following a single signal s. We now show a more general version of this analysis for arbitrary compensation schemes. To this end, consider an arbitrary wage schedule $w(s)$ chosen by the principal – for any effort $e_i \in E$ chosen by the principal on- and off- equilibrium path . We may write the expected wage paid to the worker as

$$\begin{aligned}
\sum_s w(s)\, p(s|e_i) &= \sum_s w(s)\, p(s|e^*)\, \frac{p(s|e_i)}{p(s|e^*)} \\
&= \sum_s w(s)\, p(s|e^*) \cdot \sum_s \frac{w(s)\, p(s|e^*)}{\sum_s w(s)\, p(s|e^*)} \frac{p(s|e_i)}{p(s|e^*)} \\
&= \sum_s w(s)\, p(s|e^*) \cdot \sum_s \alpha_s \frac{p(s|e_i)}{p(s|e^*)},
\end{aligned}$$

where $\sum_s \alpha_s = 1$. Since the weights α_s do not depend on e_i, we may write the agent's interim incentive compatibility constraint as

$$\sum_s w(s)\, p(s|e^*) \cdot \sum_s \alpha_s \frac{p(s|e_i)}{p(s|e^*)} - c(e_i) \geq \sum_s w(s)\, p(s|e^*) \cdot \sum_s \alpha_s \frac{p(s|e_j)}{p(s|e^*)} - c(e_j).$$

Therefore, if we define $\ell_i = 1 - \sum_s \alpha_s p(s|e_i) / p(s|e^*)$, this incentive constraint may be written as

$$\sum_s w(s)\, p(s|e^*) \cdot [\ell_j - \ell_i] \geq c(e_i) - c(e_j), \forall j = 1, \cdots, |E|. \tag{5}$$

Writing the incentive constraint in this manner reveals that the choice of likelihood ratios, ℓ_i, is sufficient to characterize payoffs of the principal and the agent and that the choice of contract $w(s)$ by the principal may be decomposed into a choice of $\sum_s w(s)\, p(s|e^*)$ as well as the choice of $\{\alpha_s\}$ or alternatively ℓ_i. In other words, if we represent the likelihood ratio ℓ_i profile with the following vector

$$\mathbf{l} = \left(\ell_1, \cdots, \ell_{|E|}\right),$$

then $\mathbf{l} \in$ convex hull $\left(\left\{1-\frac{p(s|e_1)}{p(s|e^*)}, \cdots, 1-\frac{p\left(s|e_{|E|}\right)}{p(s|e^*)}\right\}_{s\in S}\right) = \text{co}\left(\mathbf{p}\right)$. Thus the choice of the principal can be summarized by the overall level of the compensation $\sum_s p\left(s|e^*\right) w\left(s\right) = \overline{w}$ together with an element of the set co $\left(\mathbf{p}\right)$. We thus have the following lemma:

Lemma 1. *Consider an implementable effort e^* and its associated information structure $\mathbf{p} = \{p\left(\cdot|\cdot\right)\}$. Then the cost to the principal from implementing any effort is given by*

$$W(e_i, \mathbf{p}) = \min_{\ell \in co(\mathbf{p}) \cap \Omega_i} (1-\ell_i) \cdot \max_{i:\ell_j \geq \ell_i} \frac{c(e_i)-c(e_j)}{\ell_j - \ell_i}, \tag{6}$$

$$\Omega_i = \left\{\ell \in \mathbb{R}^{|E|} : \max_{j:\ell_j \geq \ell_i} \frac{c\left(e_i\right)-c\left(e_j\right)}{\ell_j - \ell_i} \leq \min_{j:\ell_j < \ell_i} \frac{c\left(e_i\right)-c\left(e_j\right)}{\ell_j - \ell_i}\right\}. \tag{7}$$

The above lemma states that the problem of the principal can be reduced to choosing an effort level as well as a point in the convex hull of likelihood ratios – its intersection with the convex cone in (7). Note that not all likelihood ratios ℓ_i are feasible in the sense that there may be no compensation scheme that satisfies the agent's interim incentive compatibility constraints. As in our motivating example above, the likelihood ratios must permit a wage $w(s)$ that lies between the lower and upper bounds in (7). The convex cone Ω_i defines the set of likelihood ratios that admit incentive compatible compensation schemes.

These results reveal that by choosing an information structure, the agent effectively determines the convex hull co $\left(\mathbf{p}\right)$ and then the principal chooses a point that is in the intersection of co $\left(\mathbf{p}\right)$ and the convex cone Ω_i. This result by itself does not make the analysis more tractable as it does not immediately describe the set of feasible convex hulls co$(\mathbf{p})$. The following proposition provides such a characterization. Note that, similar to co $\left(\mathbf{p}\right)$, the convex hull co $\left(\mathbf{f}\right)$ is the convex hull created by the points $\left\{1-\frac{f(x|e_1)}{f(x|e^*)}, \cdots, 1-\frac{f\left(x|e_{|E|}\right)}{f(x|e^*)}\right\}_{x\in X}$.

Proposition 1. *For any information structure (S, π) with $|S| < \infty$, its associated co $(\mathbf{p})$ is a subset of co $(\mathbf{f})$ that contains the origin $\mathbf{0} = (0, \cdots, 0)$. Additionally, for any convex subset C of co $(\mathbf{f})$ that contains the origin and has a finite set of extreme points, there exists an information structure (S, π) such that co $(\mathbf{p}) = C$.*

Proof. Let (S, π) be an information structure. Then,

$$\frac{p(s|e_i)}{p(s|e^*)} = \frac{\sum_x f(x|e_i)\pi(s|x)}{\sum_x f(x|e^*)\pi(s|x)} = \sum_x \frac{f(x|e^*)\pi(s|x)}{\sum_x f(x|e^*)\pi(s|x)} \frac{f(x|e_i)}{f(x|e^*)}.$$
$$= \sum_x \beta_s(x) \frac{f(x|e_i)}{f(x|e^*)}$$

where $\sum_x \beta_s(x) = 1$; $\beta_s(x)$ is the principal's posterior probability of x after observing s. The above implies that

$$\left(1 - \frac{p(s|e_1)}{p(s|e^*)}, \cdots, 1 - \frac{p(s|e_{|E|})}{p(s|e^*)}\right) = \sum_x \beta_s(x) \left(1 - \frac{f(x|e_1)}{f(x|e^*)}, \cdots, 1 - \frac{f(x|e_{|E|})}{f(x|e^*)}\right)$$

Hence, the left hand side of the above is a member of co $(\mathbf{f})$ for all $s \in S$. As a result co $(\mathbf{p}) \subset$ co $(\mathbf{f})$. Moreover, we have

$$\sum_s p(s|e^*) \left(1 - \frac{p(s|e_i)}{p(s|e^*)}\right) = 1 - \sum_s p(s|e_i) = 0, \forall i.$$

This implies that the convex set co$(\mathbf{p})$ includes the origin.

Now consider an arbitrary convex set C that contains the origin. Let S be the set of extreme points of C with each of its member being of the form $\mathbf{z} = (z_1, \cdots, z_{|E|})$. Then, since $\mathbf{0} \in C$ by Caratheodory theorem – see Rockafellar (1970) –, there must exist $\{\tau_{\mathbf{z}}\}_{\mathbf{z} \in S}$ such that $\sum_{\mathbf{z}} \tau_{\mathbf{z}} = 1$ and

$$0 = \sum_{\mathbf{z} \in S} \tau_{\mathbf{z}} z_i \tag{8}$$

by definition of co $(\mathbf{f})$, we must have that $z_i \leq 1$ for all $\mathbf{z} \in C$. Moreover, since $\mathbf{z} \in \text{co}(\mathbf{f})$, there must exist a subset $Y \subset X$ whose members are linearly independent together with $\beta_{\mathbf{z}}(x)$ such that

$$\sum_{x \in Y} \beta_{\mathbf{z}}(x) = 1, \beta_{\mathbf{z}}(x) \geq 0, z_i = \sum_{x \in Y} \beta_{\mathbf{z}}(x) \left[1 - \frac{f(x|e_i)}{f(x|e^*)}\right].$$

Replacing the above in (8) leads to

$$0 = \sum_{x \in Y} \sum_{\mathbf{z} \in S} \tau_{\mathbf{z}} \beta_{\mathbf{z}}(x) \left[1 - \frac{f(x|e_i)}{f(x|e^*)}\right]$$

Since the points in Y are linearly independent and we also know that

$$0 = \sum_{x \in Y} f(x|e^*) \left[1 - \frac{f(x|e_i)}{f(x|e^*)}\right]$$

we must have that

$$f(x|e^*) = \sum_{\mathbf{z}} \tau_{\mathbf{z}} \beta_{\mathbf{z}}(x)$$

Let us define

$$\pi(\mathbf{z}|x) = \frac{\beta_{\mathbf{z}}(x) \tau_{\mathbf{z}}}{\sum_{\mathbf{z} \in S} \beta_{\mathbf{z}}(x) \tau_{\mathbf{z}}} = \frac{\beta_{\mathbf{z}}(x) \tau_{\mathbf{z}}}{f(x|e^*)}$$

Then under $\pi(\mathbf{z}|x)$, we have

$$\begin{aligned}
\frac{p(\mathbf{z}|e)}{p(\mathbf{z}|e^*)} &= \frac{\sum_x \frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{f(x|e^*)} f(x|e)}{\sum_x \frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{f(x|e^*)} f(x|e^*)} \\
&= \frac{\sum_x \beta_{\mathbf{z}}(x) \tau_{\mathbf{z}} \frac{f(x|e)}{f(x|e^*)}}{\tau_{\mathbf{z}} \sum_x \beta_{\mathbf{z}}(x)} \\
&= \sum_x \beta_{\mathbf{z}}(x) \frac{f(x|e)}{f(x|e^*)} = \mathbf{z}
\end{aligned}$$

which concludes the proof. □

Proposition 1 implies that any convex subset of $\text{co}(\mathbf{f})$ that contains the origin can be chosen by the agent as co $(\mathbf{p})$. This implies that instead of a choice of $\mathbf{p}$, we can focus on a choice of a convex subset of co $(\mathbf{f})$ that contains the origin. Thus the problem of the best agent-optimal information structure that implements e^* can be thought of as the equilibrium of the following game:

- **Stage 1.** The agent chooses a finite number of points L inside the convex set co $(\mathbf{f})$ such that the convex hull of these points conv (L) includes the origin.

- **Stage 2.** Principal chooses an effort level $e_i \in E$ and a point $\ell \in$ conv $(L) \cap \Omega_i$ to maximize $\mathbb{E}\left[g\left(x\right)|e_i\right] - \left(1-\ell_i\right) \cdot \max_{j:\ell_j>\ell_i} \frac{c(e_i)-c(e_j)}{\ell_j-\ell_i}$.

- **Stage 3.** The choice of principal in stage 2 coincides with e^*.

The following example illustrates the construction of co $(\mathbf{f})$ and equilibrium response of the principal.

Example 1. Suppose that $E = \{e_1, e_2, e_3\}$, where $c\left(e_1\right) = 0$, $c\left(e_2\right) = 0.1$, and $c\left(e_3\right) = 0.3$. The performance technology is given by $X = \{x_1 = 0, x_2 = 1, x_3 = 2\}$ and $f\left(x|e\right)$ is

$$f = \begin{bmatrix} 0.35 & 0.50 & 0.15 \\ 0.05 & 0.50 & 0.45 \\ 0.10 & 0.15 & 0.75 \end{bmatrix},$$

where $f_{ij} = f\left(x_j|e_i\right)$. This yields

$$\mathbb{E}\left[x|e_1\right] = 0.8,\ \mathbb{E}\left[x|e_2\right] = 1.4,\ \mathbb{E}\left[x|e_3\right] = 1.65.$$

Therefore, the first-best is to implement effort e_3. If we set $e^* = e_3$, then since $1 - f\left(x|e_3\right)/f\left(x|e^*\right) = 0$, we can embed the set co $(\mathbf{f})$ in $\mathbb{R}^2$. The gray area in Figure 1 represents this set with a_i being the point associated with $x_i \in X$.

To understand the geometry of the game between the principal and the agent, consider a signal structure with 4 realizations given by $S = \{s_1, s_2, s_3, s_4\}$ and

$$\pi = \begin{bmatrix} 0.5 & 0.2 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.05 & 0.05 & 0.1 & 0.8 \end{bmatrix},$$

where $\pi_{ij} = \pi(s_j|x_i)$. We can use $p(s|e) = \sum_{x\in X} \pi(s|x) f(x|e)$ to construct the likelihood ratios. The green area in Figure 1 illustrates the convex set $\text{co}(\mathbf{p})$.

To better illustrate the incentives in the geometric game between the principal and the agent, suppose the agent reveals x to the principal. As we have illustrated above, the choice of contract by the principal is equivalent to the choice of a point in the convex set $\text{co}(\mathbf{p})$.

To think about the incentives of the principal, if the principal is to implement e_3, he chooses $\mathbf{l} \in \text{co}(\mathbf{f})$ to minimize its cost given by

$$\max\left\{\frac{c(e_3) - c(e_2)}{\ell_2}, \frac{c(e_3) - c(e_1)}{\ell_1}\right\} = \max\left\{\frac{0.2}{\ell_2}, \frac{0.3}{\ell_1}\right\}$$

The upper-contour sets associated with the above cost function are convex cones in the shape of positive orthants – the shaded blue area highlighted in the right panel of Figure 1. For this example, the above cost function is minimized at a_3. On the other hand, if the principal is to implement e_2, he chooses $\mathbf{l} \in \text{co}(\mathbf{f})$ to minimize the cost given by $(1 - \ell_2)\frac{0.1}{\ell_1 - \ell_2}$ when $\ell_1 \geq \ell_2$. This cost is minimized at a_3. For e_3 to be implementable the payoff associated with e_2 – implied by the principal choosing a_3 – should be lower than the payoff associated with e_3. The set of likelihood ratios that satisfy the latter are highlighted by the blue shaded area. Note that a_3 is not contained in this area. As a result, by choosing to reveal x to the principal, the agent is unable to implement e_3. One can show that e_2 is implementable under full revelation of x.

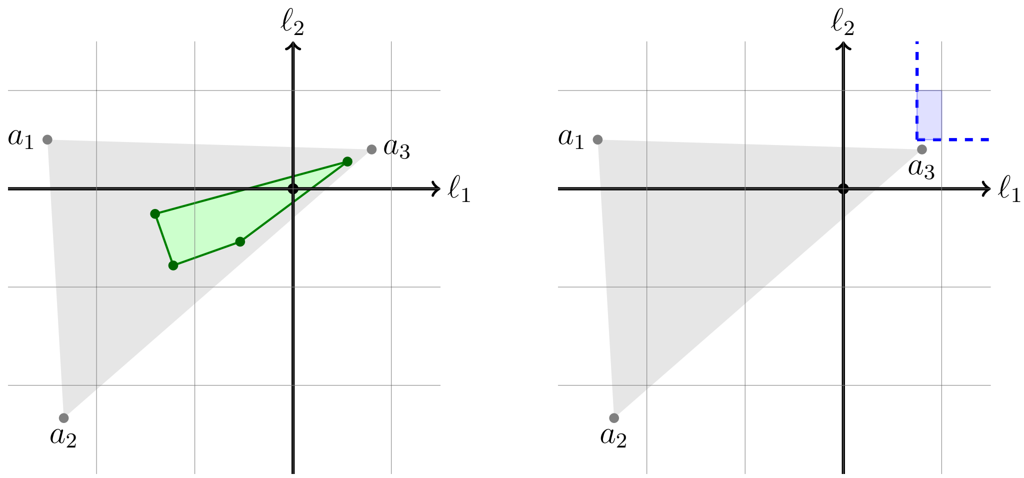

Figure 1: The geometric representation of information structures in the space of likelihood ratios. The left panel shows an example with 4 signals; the right panel shows the incentives of the principal when the agent reveals x.

4.1 Binary Information Structures

We now exploit our geometric interpretation to show that we may restrict the agent to choose information structures with at most two signals without loss of generality.

Proposition 2. *If e^* is implementable by some information structure (S, π) and delivers expected wage $W(e^*, \pi)$, to the agent, then e^* is also implementable by a binary information structure $\left(\hat{S}, \hat{\pi}\right)$ with $|S| = 2$ and $W(e^*, \pi) = W(e^*, \hat{\pi})$.*

Proof. Suppose the effort level e^* is implementable by (S, π) and consider the optimization problem in (6). Let $\mathbf{l}^* \in \text{co}(\mathbf{p})$ be the optimal choice for the principal when choosing e^*. Given the definition of $\text{co}(\mathbf{p})$ and the geometric game described above, there must exist $\{\alpha_s\}_{s \in S}$ such that $\alpha_s \geq 0$ and $\sum_{s \in S} \alpha_s = 1$ and

$$\ell_i^* = 1 - \sum_{s \in S} \alpha_s \frac{p(s|e_i)}{p(s|e^*)}, \forall e_i \in E$$

Let the information structure $\left(\hat{S}, \hat{\pi}\right)$ be defined as $\hat{S} = \{L, H\}$ and

$$\hat{\pi}\left(H|x\right) = \sum_{s} \beta_s \pi\left(s|x\right), \hat{\pi}\left(L|x\right) = 1 - \hat{\pi}\left(H|x\right)$$

where

$$\beta_s = \frac{\alpha_s / p\left(s|e^*\right)}{\sum_s \alpha_s / p\left(s|e^*\right)}$$

The probability function given this information structure satisfies

$$\begin{aligned} 1 - \frac{\hat{p}\left(H|e_i\right)}{\hat{p}\left(H|e^*\right)} &= 1 - \frac{\sum_x \hat{\pi}\left(H|x\right) f\left(x|e_i\right)}{\sum_x \hat{\pi}\left(H|x\right) f\left(x|e^*\right)}, \\ &= 1 - \frac{\sum_{s \in S} \frac{\alpha_s}{p(s|e^*)} \sum_x \pi\left(s|x\right) f\left(x|e_i\right)}{\sum_{s \in S} \frac{\alpha_s}{p(s|e^*)} \sum_x \pi\left(s|x\right) f\left(x|e^*\right)} \\ &= 1 - \frac{\sum_{s \in S} \frac{\alpha_s}{p(s|e^*)} p\left(s|e_i\right)}{\sum_{s \in S} \frac{\alpha_s}{p(s|e^*)} p\left(s|e^*\right)} \\ &= 1 - \sum_{s \in S} \frac{\alpha_s}{p\left(s|e^*\right)} p\left(s|e_i\right) = \ell_i^* \end{aligned}$$

The above implies that if the principal is to choose e^*, $\mathbf{l}^*$ is feasible under the new information structure $\left(\hat{S}, \hat{\pi}\right)$, i.e., $\mathbf{l}^* \in \text{co}\left(\hat{\mathbf{p}}\right)$. Hence, in order to establish our claim, it is sufficient to show that for any alternative $e_i \neq e^*$, $W\left(e_i, \hat{\pi}\right) \geq W\left(e_i, \pi\right)$. To show this, it is sufficient to show that

$$\left(1 - \frac{\hat{p}\left(L|e_1\right)}{\hat{p}\left(L|e^*\right)}, \cdots, 1 - \frac{\hat{p}\left(L|e_{|E|}\right)}{\hat{p}\left(L|e^*\right)}\right) \in \text{co}\left(\mathbf{p}\right)$$

This would imply that $\text{co}\left(\hat{\mathbf{p}}\right)$ is a subset of $\text{co}\left(\mathbf{p}\right)$ since $\text{co}\left(\hat{\mathbf{p}}\right)$ is the line that connects the above point to $\mathbf{l}^*$.

We have

$$\begin{aligned}
1-\frac{\hat{p}\left(L|e_i\right)}{\hat{p}\left(L|e^*\right)} &= 1-\frac{\sum_x\left(1-\hat{\pi}\left(H|x\right)\right)f\left(x|e_i\right)}{\sum_x\left(1-\hat{\pi}\left(H|x\right)\right)f\left(x|e^*\right)}\\
&= 1-\frac{\sum_x\sum_s\left(1-\beta_s\right)\pi\left(s|x\right)f\left(x|e_i\right)}{\sum_x\sum_s\left(1-\beta_s\right)\pi\left(s|x\right)f\left(x|e^*\right)}\\
&= 1-\frac{\sum_s\left(1-\beta_s\right)p\left(s|e_i\right)}{\sum_s\left(1-\beta_s\right)p\left(s|e^*\right)}\\
&= 1-\sum_s\frac{\left(1-\beta_s\right)p\left(s|e^*\right)}{\sum_s\left(1-\beta_s\right)p\left(s|e^*\right)}\frac{p\left(s|e_i\right)}{p\left(s|e^*\right)}
\end{aligned}$$

which establishes the claim. This concludes the proof. □

We can describe the intuition behind the above proof graphically. Consider an effort e^* and suppose that its associated information structure is as is depicted in Figure 2. The green area represents the convex hull of an arbitrary information structure, $\text{co}\left(\hat{\mathbf{p}}\right)$. The point b is the point of optimality for the principal in $\text{co}\left(\mathbf{p}\right)$ if he were to implement e^*. The red line represents the new information structure $\hat{\pi}$ – since there are only two signal realizations this has to be a line. If the principal is to implement e^*, since $b\in\text{co}\left(\hat{\mathbf{p}}\right)$ and b is chosen under π, b remains optimal. Moreover, since $\text{co}\left(\hat{\mathbf{p}}\right)\subset\text{co}\left(\mathbf{p}\right)$, for any other effort $e_i\neq e^*$, minimized cost under $\hat{\pi}$ must be at least as high as the minimized cost under π. This, in turn implies that e^* is implementable under $\hat{\pi}$ and implements the same outcome as π.

The above observations imply that if principal prefers to implement effort e^* under the information structure π, he would prefer the same under the binary information structure $\hat{\pi}$ and the expected wage he has to pay to achieve this is the same under both information structures. In what follows, we will use this result to characterize optimal information structures.

Example 1 (Continued). Recall Example 1 in which we argued that the principal – under full revelation – chooses to implement e_2 and thus his choice is inefficient (since total surplus under e_2 is lower than that of e_3.)

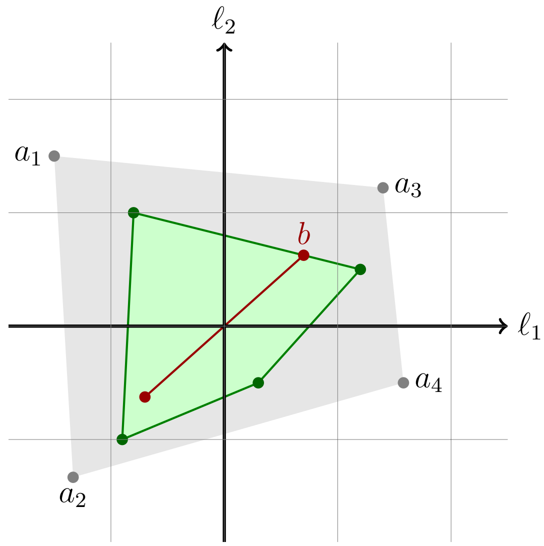

Figure 2: The intuition for the construction of a two-point signal

It is still remain to be seen whether e_3 is implementable and how well can the agent do by choosing an information structure. By Proposition 2, we can focus on information structures that only have two support points. Note that since the output under e_1 is $\mathbb{E}\left[x|e_1\right] = 0.8$ and its cost is 0, the principal can always guarantee $\mathbb{E}\left[x|e_1\right] = 0.8$.

Now consider the point $\hat{\ell}$ satisfying

$$\frac{0.2}{\ell_2} = \frac{0.3}{\ell_1} = 0.85$$

At this point, the payoff to the principal of implementing each effort level is given by

$$\begin{aligned} e_3 &: 1.65 - \max\left\{\frac{0.2}{\ell_2}, \frac{0.3}{\ell_1}\right\} = 0.8 \\ e_2 &: 1.40 - (1-\ell_2)\cdot\frac{c(e_2)-c(e_1)}{\ell_1-\ell_2} = 1.40 - 0.65 = 0.75 \\ e_1 &: 0.8 - 0 = 0.8 \end{aligned}$$

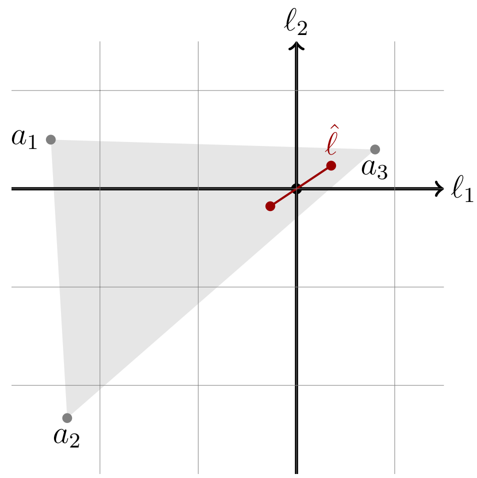

Figure 3: Agent optimal information structure in Example 1

This implies that at $\hat{\ell}$, e_3 maximizes the payoff of the principal. Thus if $\hat{\ell} \in \text{co}\left(\mathbf{f}\right)$, we can choose an information structure that implements e_3. In this example, this is indeed the case. Figure 3 depicts the point $\hat{\ell}$ as well as the two point information structure that implements it; the red line going through $\hat{\ell}$. If we set $S = \{L, H\}$ and

$$\pi\left(H|x\right) = \begin{cases} \frac{27}{83} & x = x_1 \\ \frac{15}{83} & x = x_2 \\ \frac{41}{83} & x = x_3 \end{cases}, \pi\left(L|x\right) = 1 - \pi\left(H|x\right)$$

it can be readily checked that (π, S) is associated with the one depicted in Figure 3.

The above example illustrates the agent's power in the choice of the information structure provides. By choosing the aforementioned information structure, not only the agent is able to implement e_3, i.e., the efficient level of effort, but also able to capture all the surplus. In what follows, we show that under some conditions on the performance technology, this is always possible.

5 Efficient Surplus Extraction

In this section, we provide sufficient conditions on the performance technology $f(\cdot|\cdot)$ and cost function function $c(\cdot)$ so that the agent is able to implement first best effort and fully extract all the surplus.

Let e^* be the first-best level of effort that satisfies

$$e^* \in \arg\max_{e \in E} \mathbb{E}\left[g(x)|e\right] - c(e).$$

Suppose that the agent wishes to implement e^* and capture all the surplus given by $\mathbb{E}\left[g(x)|e^*\right] - c(e^*) - \mathbb{E}\left[g(x)|e_1\right]$. For this to occur, we need to choose $\text{co}(\mathbf{p})$ and $\ell^* \in \text{co}(\mathbf{p})$ that satisfies

$$\begin{aligned}\mathbb{E}\left[g(x)|e_1\right] &= \mathbb{E}\left[g(x)|e^*\right] - \max_{i:\ell_i^* \geq 0} \frac{c(e^*) - c(e_i)}{\ell_i^*} \\ &\geq \max_{e_j \in E, \ell \in \text{co}(\mathbf{p}) \cap \Omega_j} \mathbb{E}\left[g(x)|e_j\right] - (1 - \ell_j) \max_{i:\ell_i > \ell_j} \frac{c(e_j) - c(e_i)}{\ell_i - \ell_j}\end{aligned}$$

Consider the likelihood vector ℓ^* that satisfies the following properties

$$\Delta_y = \mathbb{E}\left[g(x)|e^*\right] - \mathbb{E}\left[g(x)|e_1\right] = \frac{c(e^*) - c(e_i)}{\ell_i^*}, \forall i$$

In words, this is a point in which the agent is indifferent between all efforts – see the reformulation of (IC-A) in (5). Moreover, the principal's payoff is his outside option and thus should the principal choose to implement e^* and chooses $\ell^* \in \text{co}(\mathbf{p})$, agent captures all the surplus. In the following proposition, we show that if $\ell^* \in \text{co}(\mathbf{f})$, we can construct an information structure – and its associated $\text{co}(\mathbf{p})$ – in which ℓ^* is the best choice for the principal and e^* is the best level of effort.

Theorem 1. *Suppose that $\ell^* \in co(\mathbf{f})$. Then e^* is implementable and then there exists an information structure (π, S) for which agent's payoff is $u_A = \mathbb{E}[g(x)|e^*] - \mathbb{E}[g(x)|e_1] - c(e^*)$ and $u_P = \mathbb{E}[g(x)|e^*]$, i.e., the agent can capture all the surplus.*

Proof. First, we note that by Lemma 2 in the Appendix, $\mathbf{0} \in \text{co}(\mathbf{f})$ is an interior point of this convex set where interiority is defined in appropriatley defined subspace of $\mathbb{R}^m$. We define the following convex hull (and by Proposition 1 its associated information structure):

$$\text{co}(\mathbf{p}) = \{\lambda\ell^* + (1-\lambda)(-\alpha\ell^*) : \forall\lambda \in [0,1]\}$$

where in the above $\alpha > 0$ is such that $-\alpha\ell^* \in \text{co}(\mathbf{p})$. Such an α always exists due to $\mathbf{0}$ being an interior point. Since co $(\mathbf{p})$ is a line through ℓ^* and the origin, all of its members must satisfy

$$\forall i, j, \frac{\ell_i}{\ell_j} = \frac{c(e^*) - c(e_i)}{c(e^*) - c(e_j)} = \frac{\ell_i^*}{\ell_j^*} \rightarrow \frac{\ell_i}{\ell_1} = \frac{c(e^*) - c(e_i)}{c(e^*)} \tag{9}$$

This implies that that the cost of choosing any member of co $(\mathbf{p})$ is given by

$$\begin{aligned}(1-\ell_j)\max_{\ell_i \le \ell_j} \frac{c(e_j) - c(e_i)}{\ell_i - \ell_j} &= \left(1 - \frac{c(e^*) - c(e_j)}{c(e^*)}\ell_1\right)\frac{c(e^*)}{\ell_1} \\ &= \left(1 - \frac{c(e^*) - c(e_j)}{c(e^*)}\ell_1\right)\frac{c(e^*)}{\ell_1} \\ &= \frac{c(e^*)}{\ell_1} + c(e_j) - c(e^*)\end{aligned}$$

Since choice of ℓ under e_j must be a member of the cone Ω_j, we must have $\ell_j \le \ell_1$ and this combined with (9) implies that $\ell_1 \ge 0$. Thus the above expression is maximized at ℓ^*. Hence, the highest payoff of the principal from choosing e_j is given by

$$\mathbb{E}\left[g\left(x\right)|e_j\right]-c\left(e_j\right)-\frac{c\left(e^*\right)}{\ell_1^*}+c\left(e^*\right)=$$
$$\mathbb{E}\left[g\left(x\right)|e_j\right]-c\left(e_j\right)+c\left(e^*\right)-\mathbb{E}\left[g\left(x\right)|e^*\right]+\mathbb{E}\left[g(x)|e_1\right]$$

The above is maximized at $e_j = e^*$ which implies that e^* can be implemented with the chosen information structure. Moreover, the payoff of the principal is given by $\mathbb{E}\left[g\left(x\right)|e_1\right]$. This concludes the proof. □

The proof of the above proposition emphasizes the power of the agent when $\ell^* \in \text{co}\left(\mathbf{f}\right)$. By being able to control the information structure, the agent can control the wage that is needed for the principal to implement his desired effort. By choosing ℓ^*, the agent is forcing the principal to have to fully compensate the agent for her cost of effort. This implies that the payoff of the principal becomes total surplus shifted by a constant. Hence, it is optimal to choose the first best level of effort.

While Theorem 1 provides sufficient conditions for implementability of first best effort and full surplus extraction, it is not immediately evident what it imposes on the structure of the model. In what follows, we try to shed light on this.

Almost Perfect Performance Technology Suppose that the performance technology satisfies the following property

$$\begin{aligned} X = E \subset \mathbb{R}_+, e_1 = 0 \\ f\left(e_j|e_j\right) = 1-\left(m-1\right)\varepsilon \\ f\left(e_i|e_j\right) = \varepsilon, \forall i \neq j \end{aligned}$$

where $1/\left(m-1\right) > \varepsilon > 0$. In words, the above performance technology puts probability ε on $e_i \neq e_j$ if e_j chosen. As ε converges to 0, it converges to a setting where observing $x \in X$ fully reveals e, i.e., *a perfect performance technology.*

Note that for ε small enough, first best level of effort is given by

$$e_j \in \arg\max_{e \in E} e - c(e)$$

Suppose that in the above $e_j > e_1 = 0$. Moreover, as ε converges to 0, the point ℓ^* converges to

$$\ell_i^* = \frac{c(e^*) - c(e_i)}{e^*}$$

Finally, note that for any ε, the set $\text{co}(\mathbf{f})$ is the convex hull of the following likelihood ratios

$$1 - \frac{f(e_k|e_i)}{f(e_k|e_j)} = \begin{cases} 1 - \frac{\varepsilon}{1-(m-1)\varepsilon} & k = j, k \neq i \\ 0 & k = i = j \\ 0 & k \neq j, k \neq i \\ m - \frac{1}{\varepsilon} & k \neq j, k = i \end{cases}$$

As ε converges to 0, the above set gets larger and the point associated with $k = j$ converges to

$$\left(\underbrace{1, \cdots, 1}_{j-1\text{-times}}, 0, 1, \cdots, 1\right)$$

while the points associated with $k \neq j$ converges to

$$\left(\underbrace{0, \cdots, 0}_{k-1\text{-times}}, -\infty, 0, \cdots, 0\right)$$

This implies that as ε converges to 0, $\text{co}(\mathbf{f}) \to (-\infty, 1]^{j-1} \times \{0\} \times (-\infty, 1]^{m-j}$. This is depicted in Figure 4. Since $\ell_i^* \leq \frac{c(e_j)}{e_j} < 1$, for ε small enough, we must have that $\ell^* \in \text{co}(\mathbf{f})$. We thus have the following corollary to Theorem 1:

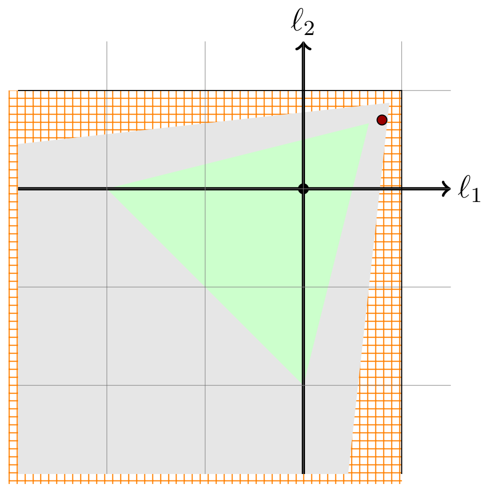

Figure 4: Almost Perfect Performance Technology for $m = 3$. The grey and green shaded areas represent $\text{co}(\mathbf{f})$ when as ε becomes smaller, respectively. As $\varepsilon \to 0$, $\text{co}(\mathbf{f})$ converges to the shaded quarter-space south-west of $(1, 1)$.

Corollary 1. *Let $f(\cdot|\cdot)$ be a almost perfect performance technology satisfying*

$$X = E \subset \mathbb{R}_+, e_1 = 0$$
$$f(e_j|e_j) \leq 1 - (m-1)\varepsilon$$
$$f(e_i|e_j) \geq \varepsilon, \forall i \neq j$$

Then there exists $\bar{\varepsilon}$ such that for all $\varepsilon \leq \bar{\varepsilon}$, such that the agent can implement first level of effort and capture all the surplus.

6 Continuous Effort and Output

In this section, we consider a version of the model from Section 3 where the effort space and the output space are continuous. Applying a first-order approach, we derive sufficient conditions such that the optimal indicator structure takes the form of monotone or hump-shaped threshold signals. In other words, we derive conditions such that the optimal information structure is characterized by an indicator with at most two thresholds.

Consider a principal employing an agent to perform a task whose output is represented by a real number $x \in X = [0,1]$. The agent chooses effort $e \in E = [0,1]$ to perform the task. The agent's effort choice induces a probability distribution $f(x|e)$ over the output space X, whose density function $f(x|e)$ is assumed to be differentiable with respect to e for every $x \in X$. Effort is costly to the agent; the cost of effort e is given by $c(e)$ for some real-valued function $c : E \to \mathbb{R}_+$ where $c'(e) \geq 0,\ c''(e) > 0,\ \forall e \in E$. For simplicity, let $c(0) = 0$. The timing and payoff functions of the game are otherwise the same as in the model in Section 3. To proceed, we assume that the first-order approach is valid for all the optimization problems faced by the agent and the principal.

We characterize this continuous version of the game in the same manner as the discrete case. Given the result in Proposition 2, one can use an approximation argument to show that optimal signal is binary; $s = L, H$. As a result, let $p(e) : E \to [0,1]$ represent the probability of $s = H$ induced by a given information structure (S, π) and the underlying probability distribution of outcomes given effort $f(x|e)$. Note, for any level of effort e, $p(e) = \int_0^1 f(x|e)\, \pi(H|x)\, dx$. Differentiating the stochastic mapping p with respect to e, yields

$$p'(e) = \int_0^1 f_e(x|e)\, \pi(H|x)\, dx,\ \forall e \in E$$

The agent's problem in the last stage of the game given an information structure and a compensation scheme is

$$\max_{e \in E} p(e)\, w(H) + [1 - p(e)]\, w(L) - c(e)$$

The first-order condition characterizing the agent's optimal effort is

$$p'(e)\, [w(H) - w(L)] = c'(e).$$

For any desired level of effort, the principal chooses a compensation scheme that solves

$$\begin{aligned} \min_{w} \; & p\left(e\right) w\left(H\right) + \left(1 - p\left(e\right)\right) w\left(L\right) \\ \text{s.t} & p'\left(e\right)\left[w\left(H\right) - w\left(L\right)\right] = c'\left(e\right), \\ & w\left(H\right), w\left(L\right) \geq 0. \end{aligned} \tag{10}$$

If $\lambda(e, \pi)$ denotes the Lagrange multiplier on the agent's incentive compatibility constraint in 10, then

$$\lambda(e, \pi) = \min\left\{\frac{p\left(e\right)}{p'\left(e\right)}, -\frac{1 - p\left(e\right)}{p'\left(e\right)}\right\}.$$

Exactly as in the discrete case, the above reveals that the agent's choice of information structure, which induces a set of likelihood ratios dependent on the stochastic mapping p determines the principal's shadow cost of incentivizing a given effort level.

If we define the expected payoff to the principal of a given effort level $\mathbb{E}\left[g(x)|e\right] = \int_X g(x) f(x|e) dx$, then the principal's optimal choice of effort solves

$$\max_{e} \mathbb{E}\left[g\left(x\right)|e\right] - \lambda\left(e, \pi\right) c'\left(e\right)$$

with associated optimality condition

$$\frac{\partial \mathbb{E}\left[g(x)|e\right]}{\partial e} - \frac{\partial \lambda(e, \pi)}{\partial e} c'(e) - \lambda(e, \pi) c''(e) = 0.$$

Using the above sequence of optimality conditions, we obtain the following optimization problem that describes the agent's choice of effort and information structure in the first stage of the game:

$$
\begin{aligned}
&\min_{e,\pi} \lambda(e,\pi) - c(e) \text{ s.t.} \\
&\quad \frac{\partial \mathbb{E}[g(x)|e]}{\partial e} - \frac{\partial \lambda(e,\pi)}{\partial e} c'(e) - \lambda(e,\pi) c''(e) = 0, \\
&\quad \frac{1}{\lambda(e,\pi)} = \max \left\{ \frac{\int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx}, -\frac{1 - \int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx} \right\}.
\end{aligned}
\tag{11}
$$

The following assumption allows us to provide a sharp characterization of optimal information structure:

Assumption 2. *Given any effort $e \in E$, the likelihood $\frac{f_e(x|e)}{f(x|e)}$ is strictly monotone in output x and its derivative $\frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)}$ is a convex function of the likelihood $\frac{f_e(x|e)}{f(x|e)}$.*

Several distribution functions satisfy Assumption 2. Examples include power distributions: $f(x|e) = ex^{e-1}$, and truncated exponential distributions: $f(x|e) = \frac{d}{dx}\frac{e^x - 1}{e-1}$.

Our main characterization of optimal information structure is as follows:

Proposition 3. *Suppose that Assumption 2 holds. Then, the equilibrium information structure is characterized by at most two thresholds in the output space. If the equilibrium information structure has a single threshold, say x^*, then $\pi(H|x) = 1$ if and only if $x \geq x^*$. If the equilibrium information structure has two thresholds, say (x_1^*, x_2^*) then $\pi(H|x) = 1$ if and only if $x \in [x_1^*, x_2^*]$.*

Given an effort choice, the agent's optimal choice of information structure requires the agent to choose the probability of being paid at output $\tilde{x}$, i.e., $\pi(\tilde{x})$, such that there is no net marginal benefit from a marginal change in this probability. Generally, it is not possible to satisfy this condition for all output levels. The agent ends up choosing the extreme probabilities for output $\tilde{x}$, depending on whether her net marginal benefit is increasing or decreasing in the probability of being paid at that output level. The assumptions of the Proposition ensure that the output intervals at which the agent uses either of the extreme probabilities are structured such that the optimal information structure takes the form of monotone or hump-shaped thresholds signals.

Example 2. Let $f(x|e) = 3ex^{3e-1}$. As mentioned earlier, this distribution satisfies Assumption 2. If the effort cost is $c(e) = \frac{e^2}{2}$,the equilibrium information structure is characterized by $x^* = 0.45$ where $\pi(H|x) = 1$ if and only if $x \geq x^*$. The principal pays the agent $w = 0.1048$ if he receives the high signal; otherwise, he pays nothing. This choice implements effort $e^* = 0.2725$ yielding the following payoffs: $U_P = 0.3450$, $U_A = 0.0677$. For comparison, the first-best effort is $e = 0.4708$. In this case, and opposite to that of section 5, the agent is unable to implement efficient outcome.

7 Conclusion

In this paper, we have developed the theoretical tool in the design of contracts in principal-agent settings. Part of the design of contract is what indicators should be used as contingencies for payments. Despite the importance of this question and its relevance for analysis of contracting decisions, this part of the contracting procedure is less explored.

Our paper has two broad implications. First, our methodology of thinking about likelihood ratios can be applied to other settings in which considerations of communication off the equilibrium path are important. In our setup, unlike other models of communication and information design, the design of information structure for the equilibrium level of effort affects communication off-path and affects the ability of the principal and agent to capture the surplus. This can arise in other settings with strategic information transmission and our method can be useful for that.

Second, our paper ties the choice of indicators in contracting to the bargaining power of the parties. In the textbook moral hazard problem, principal makes a take-it-or-leave it offer. As a result, a version of informativeness principle often holds; principal wishes to use all the information available, if possible, to use as contingency for payments. In contrast, in our model, agent has different incentives for choice of indicators. There are some casual observations that are in line with this explanation. For example, contracts in NFL are often extremely detailed and payments to football players are highly contingent on various measures of individual and team outcomes. In contrast, contracts found in the English Premier League are not as detailed. They are often contingent on very coarse personal outcomes such as number of goals scored reaching a particular threshold. In light of our theory, the level of competition in English/European soccer (in the form of increased player bargaining power) compared to a lack thereof in NFL could be behind this observation. Future work can hopefully shed light on the importance of this channel in the data.

References

BALL, I. (2019): "Scoring strategic agents," *arXiv preprint arXiv:1909.01888*. 6

BARRON, D., G. GEORGIADIS, AND J. SWINKELS (2020): "Optimal contracts with a risk-taking agent," *Theoretical Economics*, 15, 715–761. 6

BEBCHUK, L. A. AND J. M. FRIED (2004): *Pay without performance: The unfulfilled promise of executive compensation*, Harvard University Press. 2, 4

BOLESLAVSKY, R. AND K. KIM (2018): "Bayesian persuasion and moral hazard," *Available at SSRN 2913669*. 6

CARROLL, G. (2015): "Robustness and linear contracts," *American Economic Review*, 105, 536–563. 5

CHAIGNEAU, P., A. EDMANS, , AND D. GOTTLIEB (2019): "The informativeness principle without the first-order approach," *Games and Economic Behavior*, 113, 743–755. 2, 5

DEWATRIPONT, M., I. JEWITT, AND J. TIROLE (1999): "The economics of career concerns, part I: Comparing information structures," *The Review of Economic Studies*, 66, 183–198. 6

GARRETT, D., G. GEORGIADIS, A. SMOLIN, AND B. SZENTES (2020): "Optimal Technology Design," *Available at SSRN 3720594*. 5

GEORGIADIS, G. AND B. SZENTES (2020): "Optimal monitoring design," *Econometrica*, 88, 2075–2107. 5

HALAC, M., E. LIPNOWSKI, AND D. RAPPOPORT (2021): "Rank uncertainty in organizations," *American Economic Review*, 111, 757–86. 6

HOLMSTRÖM, B. (1979): "Moral hazard and observability," *The Bell journal of economics*, 74–91. 2, 5, 6

——— (1999): "Managerial incentive problems: A dynamic perspective," *The review of Economic studies*, 66, 169–182. 6

Innes, R. D. (1990): "Limited liability and incentive contracting with ex-ante action choices," *Journal of economic theory*, 52, 45–67. 5

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615. 7

Perez-Richet, E. and V. Skreta (2022): "Test design under falsification," *Econometrica*, 90, 1109–1142. 6

Poblete, J. and D. Spulber (2012): "The form of incentive contracts: agency with moral hazard, risk neutrality, and limited liability," *The RAND Journal of Economics*, 40, 215–234. 5

Rockafellar, R. T. (1970): *Convex analysis*, vol. 18, Princeton university press. 19

Rosar, F. (2017): "Test design under voluntary participation," *Games and Economic Behavior*, 104, 632–655. 6

Saeedi, M. and A. Shourideh (2020): "Optimal rating design," *arXiv preprint arXiv:2008.09529*. 6

Walton, D. and G. Carroll (2022): "A General Framework for Robust Contracting Models," *Econometrica*. 5

Zapechelnyuk, A. (2020): "Optimal quality certification," *American Economic Review: Insights*, 2, 161–176. 6

8 Appendix

Lemma 2. *Let S represent the lowest-dimensional linear subspace in $\mathbb{R}^{|E|}$ that contains $co\,(\mathbf{f})$. If $f(\cdot|e^*)$ is full-support, then the origin is an interior point of $co\,(\mathbf{f})$ with respect to S.*

Proof. Notice that the origin can be written as a convex combination of the points defining $co\,(\mathbf{f})$ with weights $f(x|e^*)$:

$$\sum_x f(x|e^*)\left(1-\frac{f(x|e_i)}{f(x|e^*)}\right)=\sum_x f(x|e^*)-\sum_x f(x|e_i)=1-1=0,\quad \forall i,$$

where $f(x|e^*)>0, \quad \forall x\in X$ because of the full-support assumption. Therefore, the origin is always included in the convex set co $(\mathbf{f})$. Suppose by contradiction that the origin is not an interior point of co $(\mathbf{f})$ with respect to S. By the supporting hyperplane theorem, there exists a hyperplane in the linear subspace S that contains the origin and co $(\mathbf{f})$ is entirely contained in one of the two closed half-spaces bounded by the hyperplane. However, since $f(x|e^*)>0, \quad \forall x\in X$, this is possible only if co $(\mathbf{f})$ is entirely contained in this hyperplane. This is a contradiction because S is the lowest-dimensional linear subspace in $\mathbb{R}^{|E|}$ that contains $co\,(\mathbf{f})$.

Proof of Proposition 3 : Without loss of generality, for any desired implementable effort level e, we impose $\int_X f_e(x|e)\pi(x)dx\geq 0$. Consequently, we have

$$\lambda(e,\pi)=\frac{\int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx}\geq 0.$$

We may then write the agent's problem in (11) as

$$\begin{aligned}
&\max_{e,\pi} \lambda(e,\pi)c'(e) - c(e) \text{ s.t.} \\
&\quad \frac{\partial E[g(x)|e]}{\partial e} - \frac{\partial \lambda(e,\pi)}{\partial e}c'(e) - \lambda(e,\pi)c''(e) = 0, \\
&\quad \lambda(e,\pi) = \frac{\int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx}, \\
&\quad \lambda(e,\pi) \geq 0,\ 0 \leq \pi(x) \leq 1,\ \forall x \in X.
\end{aligned} \tag{12}$$

We write the Lagrangian corresponding to (12), momentarily ignoring the inequality constraints:

$$\mathcal{L}(e,\pi,\eta) = \lambda(e,\pi)c'(e) - c(e) + \eta\left[\frac{\partial \mathbb{E}[g(x)|e]}{\partial e} - \frac{\partial \lambda(e,\pi)}{\partial e}c'(e) - \lambda(e,\pi)c''(e)\right].$$

For every output $\tilde{x} \in X$, the agent's optimal information structure in (12) must satisfy

$$\begin{aligned}
&\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = 0, \text{ if } 0 < \pi(\tilde{x}) < 1, \\
&\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} \geq 0, \text{ if } \pi(\tilde{x}) = 1, \\
&\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} \leq 0, \text{ if } \pi(\tilde{x}) = 0.
\end{aligned} \tag{13}$$

With some algebra, we get

$$\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = f(\tilde{x}|e)\left[A_1(e,\pi,\eta) - A_2(e,\pi,\eta)\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)} + A_3(e,\pi,\eta)\frac{f_{ee}(\tilde{x}|e)}{f(\tilde{x}|e)}\right] \tag{14}$$

where A_1, A_2, A_3 are some functions independent of $\tilde{x}$. Note that since

$$\frac{f_{ee}(x|e)}{f(x|e)} = \frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)} + \left(\frac{f_e(x|e)}{f(x|e)}\right)^2,$$

we can write (14) as

$$\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = f(\tilde{x}|e)\Bigg[A_1(e,\pi,\eta) - A_2(e,\pi,\eta)\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)} + A_3(e,\pi,\eta)\left(\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)}\right)^2 + A_3(e,\pi,\eta)\frac{\partial}{\partial e}\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)}\Bigg]. \quad (15)$$

Since $\frac{f_e(x|e)}{f(x|e)}$ is monotone, the function in (15) inherits the curvature of $\frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)}$. That is, $\frac{1}{f(x|e)}\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ is a convex or concave function of $\frac{f_e(x|e)}{f(x|e)}$ depending on the sign of $A_3(e,\pi,\eta)$. As a result, the sign of $\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ changes at most twice over the interval X.

Note that if $\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ is always positive or always negative (for a given effort level e) then the information structure is fully uninformative. Such information structures cannot be incentive compatible as they provide no incentives for the agent to conduct any effort level $e > \underline{e}$. Consequently, the equilibrium information structure has either one or two thresholds. In either case, it follows immediately from (13) that if the equilibrium information structure has a single threshold, say x^*, then $s(x) = 1$ if and only if $x \geq x^*$. If the equilibrium information structure has two thresholds, say (x_1^*, x_2^*), then $s(x) = 1$ if and only if $x \in [x_1^*, x_2^*]$. □