Buying Opinions

Mark Whitmeyer*　　Kun Zhang†

July 27, 2023

Abstract

A principal hires an agent to acquire soft information about an unknown state. Even though neither *how* the agent learns nor *what* the agent discovers are contractible, we show the principal is unconstrained as to what information the agent can be induced to acquire and report honestly. When the agent is risk neutral, and a) is not asked to learn too much, b) can acquire information sufficiently cheaply, or c) can face sufficiently large penalties, the principal can attain the first-best outcome. We discuss the effect of risk aversion (on the part of the agent) and characterize the second-best contracts.

Keywords: Moral hazard, Information acquisition, Rational inattention, Information design

JEL Classifications: D81; D82; D83; D86

---

*Arizona State University, Email: mark.whitmeyer@gmail.com

†Arizona State University, Email: kunzhang@asu.edu

We thank Brian Albrecht, Hector Chade, Vasudha Jain, Andreas Kleiner, Alejandro Manelli, Teemu Pekkarinen, Ludvig Sinander, Eddie Schlee, Ina Taneva, Can Urgun, Han Wang, Joseph Whitmeyer, Thomas Wiseman, and Hanzhe Zhang for their advice. We also appreciate the useful feedback from conference audiences at AMES, ESEM, MWET, NASMES, and Stony Brook 2022; and seminar audiences at ASU, Arizona, CUHK-HKU-HKUST, Fordham, Princeton, and Warwick.

1 Introduction

People buy advice: investors pay for stock picks, politicians and executives in firms employ advisors, and bettors at the race track ask for winners. In some situations this advice can be backed up with hard, verifiable evidence; whereas in others advice is merely cheap talk and honesty is supported only by the advisor's incentives to be truthful. This paper studies the latter situation: we analyze a contracting problem in which a principal hires an agent to acquire unverifiable evidence, which cannot be credibly disclosed or contracted upon.[1]

In our model, it is costly for the agent to acquire information, and he has significant freedom in his learning: he may choose any distribution over posterior beliefs whose mean is the prior. Although the evidence an agent acquires is non-contractible, in our main analysis, we assume that the true state is. Under this assumption, we begin by observing that any contract induces a decision problem for an agent. This allows us to show that the principal can implement any feasible learning: she can write a contract such that the agent is willing to learn precisely as desired *and* report honestly. That is, the agency problem does not impede the principal's ability to acquire information. This result stands in stark contrast to the classical setting (Holmström (1979)), wherein not all effort levels can be implemented in the second-best world.

Next, we show that the required incentives for the agent's learning produce a number of conditions whose structure allows us to simplify the principal's problem. For any state, each message contingent transfer in that state can be written as the difference between the transfer paid in that state for a "benchmark message" and a constant that depends only on exogenous values and the posteriors themselves. Not only do the relative incentives completely pin down the agent's

[1]This is the key difference between this paper and Rappoport and Somma (2017), who explore a similar problem but specify that evidence is observable and contractible.

optimal learning, but the converse is also true: the agent's optimal learning specifies the relative incentives.

We solve for the cheapest contracts that induce the agent to acquire the desired information *and* report his findings truthfully. As in the classical moral hazard environment, there is a natural benchmark in our model: the first-best problem in which learning (our analog of effort) is observable and contractible. We show that when the agent is risk neutral and negative transfers are allowed, any distribution over posteriors can be implemented at the first-best cost, even in our main setting with hidden learning and unverifiable evidence. Moreover, this holds even if the agent may exit the relationship after acquiring information, which renders the "selling the project to the agent" contract generically ineffective. This highlights another essential difference between the canonical setting and ours. In the classical setting, the possibility of an interim exit allows the agent to accrue rents.

If negative transfers are forbidden (limited liability) and the outside option is sufficiently low, the principal cannot efficiently acquire information through the agent. Nevertheless, we show that optimal incentives take simple forms in a number of cases. We provide a full characterization of the optimal contract when the agent's outside option is sufficiently small. There, it is only the limited liability constraint that binds, which allows us to pin down the optimal contract for any desired distribution over posteriors. We also fully characterize optimal implementation with an arbitrary outside option and limited liability in the binary-state case when the agent is risk neutral. In particular, implementation is efficient if and only if the agent is not asked to learn too much (in relation to her cost of acquiring information and outside option).

We also show that the agent's risk aversion introduces inefficiencies: providing incentives for the agent to learn requires that he be exposed to risk, which is surplus destroying when the agent is risk averse. Similar to the classical setting, we establish that with only an *ex ante* participation constraint, the agent gets zero rents. On the other hand, the possibility of an interim exit generically grants the agent surplus: there, the principal trades off conceding rents with better risk sharing.

We finish this section by reconciling our main setting's theoretical predictions with what we see in practice and also discuss related literature. After that, Section 2 lays out the model before Section 3 states the principal's problem and discusses the first-best benchmark. Section 4 presents some preliminary results, and Sections 5 and 6 contain the main results in the absence and presence of limited liability constraints, respectively. We wrap things up in Section 7.

1.1 Buying Opinions in Practice

It seems indisputable that one essential assumption of our model–the softness and non-contractibility of the agent's findings–captures reality in some settings. Consider, for instance, talent scouting in sports. Although teams have hard evidence about prospects (goals scored, batting average, shooting percentage etc...), they nevertheless send scouts to obtain soft (unquantifiable) information: a report from an FC Barcelona scout describes a player's running style, balance, and control, among other things (Vidal (2019)). The scout writes about the player's positioning, "Excellent. It is undoubtedly his best quality. He is always where he should be..."

On the other hand, our assumption that the true state is contractible is more difficult to justify. Several comments are in order. First, if the true state is observable to the principal but not contractible, then as long as the principal has deep pockets, Kleiner and Whitmeyer (In Progress) show that the contracts we construct in this paper can be approximated by a mediated protocol, in which the agent occasionally acquires close to full information and the principal is penalized if caught in a (probable) lie.

Second, it may be that even the principal does not observe the realized state.

If the principal has no other source of information about the state other than the agent then the situation is hopeless: the agent must be provided with differential incentives in order to learn and report honestly. However, it is reasonable in many cases that the principal does have access to information about the state other than that provided by the agent. For example, in professional baseball, once a prospect is identified by one of its scouts, a team will bring the player in to its training facility and evaluate him further directly.

The outcome of this sort of evaluation is subjective and private; however, the ultimate decision by the principal is not. Given this, one common contract is one that conditions the agent's reward on the principal's action. Baseball scouts receive bonuses if a prospect they recommended makes it to a team. Headhunters are rewarded if they identify a candidate that is hired–it is common for firms to pay recruiters a fraction of a hired worker's salary (crucially, *if they are hired*, i.e., make it past the firm's final screening). In the supplemental appendix, we explore an example in which the principal obtains a private signal and the agent's payment is conditioned on the principal's action. There, we point out that i. not all distributions over posteriors may be implementable; ii. the agent may accrue greater rents; and iii. the principal's behavior in her decision problem may be distorted, a new variety of inefficiency.

1.2 Related Literature

Our study belongs to the literature on delegated expertise, pioneered by Lambert (1986), Demski and Sappington (1987) and Osband (1989), in which a principal hires an agent to collect payoff relevant information. The central theme of this literature is incentive design for effective information acquisition and communication.

There are five recent papers that are close to ours: the already referenced Rappoport and Somma (2017), who also study contracting for flexible information acquisition, but where the posterior generated by the agent's choice of distribution is verifiable and contractible;[2] Zermeño (2011); Sharma et al. (2020); Clark and Reggiani (2021); and Müller-Itten et al. (2021). Zermeño (2011) and Clark and Reggiani (2021) both explore contracting environments in which both information acquisition and decision making are delegated to an agent. Zermeño's focus is the interaction between the variables on which the transfer schemes can depend and whether contracts specify transfer scheme menus. Clark and Reggiani (2021) show that any Pareto-optimal contract can be decomposed into a fraction of output, a state-dependent transfer, and an optimal distortion.

Sharma et al. (2020) and Müller-Itten et al. (2021) are especially related. The former explores a two-state version of our environment with a risk-neutral agent and limited liability (and a low outside option). Our Proposition 6.2 is; therefore, a generalization of their elegant characterization result (Theorem 1). The latter work introduces a strikingly useful object–the ignorance equivalent–for studying rational inattention problems. As an application, they show that a principal can efficiently acquire information through a risk-neutral agent with only an *ex ante* participation constraint.

Carroll (2019) studies a robust contracting problem in which the principal has limited knowledge about how the agent can learn and evaluates each possible contract by its worst-case guarantee. In Häfner and Taylor (2022) the agent acquires information to help the principal decide how much she should invest in a project. The distribution over posteriors and its cost are primitives of the model, and the agent's report of the realized posterior is unverifiable. Their focus is on finding the optimal contract–which can depend on the report and the outcome of the project–

[2] Bizzotto et al. (2020) consider a similar problem. However, they only allow the agent to deviate to a "default" distribution, instead of any Bayes-plausible distribution. Also related is Yoder (2022), who generalizes the two-state (risk-neutral agent) environment of Rappoport and Somma (2017) by incorporating a screening problem: the agent's marginal cost of acquiring information (κ) is his private information. Wang (2023) subsequently generalizes this to n states.

that motivates the agent to conduct the experiment and report truthfully.[3] Deb et al. (2018) study the problem of a principal that designs a dynamic mechanism (without transfers) to identify a competent forecaster.

Gromb and Martimort (2007) consider a problem of delegated expertise with two agents, where the agents may collude among themselves or with the principal. In their model, the state space is binary, and the agent is restricted to a fixed message space containing two signals whose meaning is common knowledge. Their one-agent/one-signal case is similar to our model: the agent is risk neutral and protected by limited liability, the compensation can be conditioned on both the report and the realized state, and incentives must be provided for the agent to gather information and report truthfully. Chade and Kovrijnykh (2016) study a dynamic model of contracting for information acquisition in a two state-two (fixed) signals environment. The more effort the agent exerts, the more informative the signal he acquires. They assume that the realized signals are contractible, but the true state is not.

Since in our model every contract induces a decision problem with a posterior separable cost of the agent, our work is naturally related to the rational inattention literature pioneered by Sims (1998, 2003). To analyze the agent's problem, we use insights from Caplin et al. (2022). Maćkowiak et al. (Forthcoming) provides an excellent review of this literature that covers both theory and applications.

Our principal also needs to elicit information from the agent, which connects our paper to the belief elicitation literature. Indeed, our transfer scheme is a scoring rule. The most important distinction is that the beliefs in our work are endogenously determined through the agent's learning. While the papers in that literature study what scoring rules induce truthful reporting (Gneiting and Raftery (2007) and Schlag et al. (2015) are good surveys) and what properties of a state dis-

[3] Terovitis (2018) tackles a similar problem. In his framework, the outcome is deterministically pinned down by the action and state, and the decision is delegated to the agent.

tribution can be elicited (see Lambert (2019) and references therein), our focus is on deriving incentive contracts that induce the agent to learn *and* report truthfully.

Finally, because we study the motivation of an agent to acquire costly and unverifiable information, our work also connects to the moral hazard literature. In the canonical moral hazard problem (see, for example, Mirrlees (1999), Holmström (1979), Grossman and Hart (1983), and Holmström and Milgrom (1987)), the agent is impelled to exert costly effort that yields some (distribution over) output; whereas in ours, he must be coerced into choosing a much more complicated object (a particular probability distribution) then reporting honestly.

2 The Model

There is an unknown state of the world $\theta \in \Theta$, where $|\Theta| = n < \infty$; and both principal and agent share a common full support prior $\mu \in \Delta(\Theta)$. The principal (she) has a continuous (reduced-form) payoff function over posteriors $\mathbf{x} \in \Delta(\Theta)$, $V(\mathbf{x})$.[4] The principal cannot acquire information herself but instead must rely on the assistance of an agent (he), who acquires information by conducting a costly experiment. As shown in Kamenica and Gentzkow (2011), this is equivalent to him choosing a distribution over posteriors $F \in \Delta\Delta(\Theta)$ that is Bayes-plausible: $\mathbb{E}_F[\mathbf{x}] = \mu$. The agent's cost of acquiring F, denoted by $C(F)$, is posterior separable à la Caplin et al. (2022); that is,

$$C(F) = \kappa \int_{\Delta(\Theta)} c(\mathbf{x})\, dF(\mathbf{x}) \ ,$$

where $\kappa > 0$ is a scaling parameter, $c\colon \Delta(\Theta) \to \mathbb{R}_+$ is a strictly convex and twice continuously differentiable function bounded on the interior of $\Delta(\Theta)$, and $c(\mu) = 0$.[5]

[4]This could correspond, for example, to a decision problem faced by the principal.

[5]This class of information costs includes the entropy-based cost function (see e.g. Sims (1998, 2003), and Matějka and McKay (2015)); the log-likelihood cost of Pomatto et al. (2023); the

After acquiring information, the agent sends a message to the principal. The true state is eventually observable to both parties after the principal's value is realized and can be contracted upon. A contract specifies the set of messages available to the agent, and a transfer paid to the agent which can be contingent on both the realized state and the message sent. Formally, the principal proposes a pair (M, t) consisting of a compact set of messages M available to the agent, and a transfer $t\colon M \times \Theta \to \mathbb{R}$ ($t\colon M \times \Theta \to \mathbb{R}_+$ when the agent is protected by limited liability). We assume the principal's payoff is quasi-linear in the transfer. The agent's payoff is additive separable in his utility from the transfer and the cost of acquiring information, and he values the transfer according to a continuously differentiable, concave, and strictly increasing function v, with $v(0) = 0$. To ease presentation, transfer t is expressed in utils.

We further assume the agent has access to an outside option of value $v_0 \geq 0$, and that there are two chances for him to leave with his outside option: he can choose not to accept the contract, or walk away after acquiring information by reporting nothing. In doing so, the agent sends the "null message," $\varnothing$.[6]

Unless otherwise noted, we assume throughout that the principal suffers a penalty that is strictly greater than v_0 if the agent takes his outside option.[7] This ensures that it is not optimal for the principal to offer the agent a contract in which he ever takes his outside option. In the discussion following Theorem 4.3, we argue that our framework can accommodate "shoot the messenger" contracts–in which the agent is asked to exit the relationship with positive probability on path–with ease.

neighborhood-based cost function studied by Hébert and Woodford (2021); and the quadratic (posterior variance) cost function.

[6]More generally, the agent's outside option could derive from some salvage value for information that is an arbitrary upper semicontinuous function of the posterior $p(\mathbf{x})$. In the supplementary appendix, we explain that this does not alter our analysis in a meaningful way.

[7]This captures the cost of hiring a new agent, for example.

The timing of the game is as follows:

(i) The principal proposes a contract (M, t);

(ii) If the agent does not accept, the game ends, and the agent and principal receive v_0 and $V(\mu)$, respectively; otherwise the agent chooses a Bayes-plausible distribution F, from which a posterior $\mathbf{x} \in \Delta(\Theta)$ is drawn and privately observed by the agent;

(iii) The agent chooses whether to report. If he reports, he sends a message $m \in M$; and if he does not report, he takes his outside option v_0 (and the principal observes the null message);[8]

(iv) Payoffs accrue: given belief $\mathbf{x}(m)$ induced by message m, the principal gets $V(\mathbf{x}(m)) - \mathbb{E}_{\mathbf{x}(m)} v^{-1}(t(m,\theta))$ and the agent $\mathbb{E}_{\mathbf{x}(m)} t(m,\theta) - c(F)$.

3 The Principal's Problem

3.1 The First Best Benchmark

Denote the set of Bayes-plausible distributions over posteriors by $\mathcal{F}(\mu)$. It is a convex and compact subset of $\Delta\Delta(\Theta)$. If the principal controlled the information acquisition, she would solve

$$\max_{F \in \mathcal{F}(\mu)} \int (V - \kappa c)\, dF \ ,$$

which is a linear functional of F, guaranteeing the existence of a maximizer. In our context, "first best" refers to the situation where the principal can observe the distribution over posteriors chosen by the agent, so the principal can specify transfer $t \colon \Delta\Delta(\Theta) \to \mathbb{R}_+$. When the distribution is observable, the following contract implements any distribution F and is optimal: the principal pays the agent precisely

[8] If the optimal contract is such that the null message, $\varnothing$, is off-path (as it will be provided the principal incurs a sufficiently large cost from the agent exiting the relationship), the principal obtains $V(\mu)$. If $\varnothing$ is on-path, corresponding to posterior $\mathbf{x}'$, the principal gets $V(\mathbf{x}')$.

the amount that makes him indifferent between learning and walking away with his outside option if and only if the agent acquires F. Otherwise, the principal pays the agent nothing. Evidently, the transfer is never strictly negative, and the agent is willing to acquire F. Therefore, at the first best, the principal's cost of acquiring information is $v^{-1}(C(F)+v_0)$.

3.2 The Contracting Problem

A contract must guarantee that the agent chooses the right distribution and reports honestly. Without loss of generality, every message contained in the message space M–except for the null message $\varnothing$–uniquely identifies a posterior in the support of F,[9] and hence $M = \text{supp}(F) \cup \{\varnothing\}$. We say that a distribution F is *implementable* if choosing F and reporting truthfully is an optimal strategy for the agent following some contract offer. Equivalently, the contract (M, t) *implements* F. F can be *implemented efficiently* if it can be implemented at the first-best cost.

Any contract (M, t) offered to the agent produces a *value function*

$$W(\mathbf{x}) := \max_{m \in M} \mathbb{E}_{\mathbf{x}}[t(m, \theta)] - \kappa c(\mathbf{x}) \ ,$$

which is the highest payoff the agent can obtain–if he accepts the contract and does not walk away after acquiring information–when his posterior is $\mathbf{x}$. By construction, W is continuous and piecewise strictly concave. The agent chooses a distribution over posteriors to maximize his *ex ante* value.

Caplin et al. (2022) observe that the agent's optimal behavior corresponds to the hyperplane that is tangent to the concavified value function at the prior μ.[10] We denote this hyperplane by $\mathcal{H}$ and sometimes refer to it as the *concavifying hyperplane*. As is standard, we can identify this supporting hyperplane $\mathcal{H}$ with an

[9] The support of a distribution, denoted by $\text{supp}(\cdot)$, is the smallest closed set that has probability one.

[10] The concavified value function is the pointwise lowest concave function that majorizes the value function.

affine function $f_{\mathcal{H}}(\mathbf{x}) : \Delta(\Theta) \to \mathbb{R}$. This hyperplane is the central object in our contracting problem under study: as we will shortly discover, the principal's implementation problem is essentially one of choosing this hyperplane, which pins down the required transfers. The agent's optimal *ex ante* value is given by $f_{\mathcal{H}}(\mu)$.

Caplin et al. (2022) point out that the optimal posteriors are the points at which this hyperplane intersects W; we denote the set of such points by $P_{(M,t)}$:

$$P_{(M,t)} := \{\mathbf{x} \in \Delta(\Theta) : f_{\mathcal{H}}(\mathbf{x}) = W(\mathbf{x})\}.$$

By construction, at every $\mathbf{x} \in P_{(M,t)}$, it is optimal for the agent to report the realized posterior honestly. Therefore, a necessary condition for a distribution F to be implemented by a contract (M,t) is that $\mathrm{supp}(F) = P_{(M,t)}$.[11] This need not be sufficient for implementation: the contract must also prevent the agent from walking away at any point in the interaction. In particular, no matter what the realized posterior is, the agent cannot deviate profitably by taking his outside option without making a report; this requires

$$f_{\mathcal{H}}(\mathbf{x}) \geq v_0 - \kappa c(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \Delta(\Theta)\,. \tag{$IR - v_0$}$$

As this constraint imposes restrictions *ex interim*, we often call it the interim individual rationality constraint (or participation constraint).[12] It is stronger than the *ex ante* participation constraint $f_{\mathcal{H}}(\mu) \geq v_0$.

Thus,

Lemma 3.1. *A contract (M,t) implements distribution F if and only if*

(i) ***Incentive Compatibility****:* $\mathrm{supp}(F) = P_{(M,t)}$*; and*

(ii) ***Individual Rationality****: Constraint $IR - v_0$ holds; and*

(iii) ***(Ex Post) Limited Liability****: if there is limited liability, $t(m,\theta) \geq 0$ for all $\theta \in \Theta$ and $m \in M$.*

[11] Technically, this is incorrect: the necessary condition is that $\mathrm{supp}(F) \subseteq P_{(M,t)}$. The stated equality anticipates Proposition 4.2, in which we argue that WLOG F has affinely-independent support.

[12] The relevance of this constraint is illustrated in an example in Section 3.3.

The individual rationality constraint ($IR - v_0$) is similar to the (interim) limited liability constraint in Rappoport and Somma (2017) (p.11)–since Rappoport and Somma (2017) allow for contracting upon the realized posterior, their limited liability is of *ex interim* nature. Indeed, ($IR - v_0$) requires that the value of the agent at any posterior in the support of the desired distribution (i.e., at the interim stage) cannot be too low. Here, this is a consequence of preventing the agent from learning according to a different distribution and not making a report following some posterior (rather than the direct imposition of Rappoport and Somma (2017)). We stick to the term "individual rationality" because it also ensures that the agent accepts the contract *ex ante* and prefers not to send the null message *ex interim*. Limited liability (iii) in our work imposes restrictions *ex post*. This concern is absent from Rappoport and Somma (2017), as contracts there are not state-contingent. To streamline exposition, we frequently drop "ex post" and refer to this constraint merely as "limited liability."

To solve the principal's contracting problem, we adopt a two-step approach: first, for every implementable distribution F, we solve the principal's cost minimization problem:

$$\min_{(M,t)} \mathbb{E}_{F,\mathbf{x}}\left[v^{-1}\left(t\left(\mathbf{x},\theta\right)\right)\right], \tag{‡}$$

subject to (i), (ii), and (iii) in Lemma 3.1; denote its value by $\Gamma(F)$. Second, the principal chooses an implementable distribution F to maximize her payoff under agency, $\int V(\mathbf{x})\,dF(\mathbf{x}) - \Gamma(F)$. Like most papers studying moral hazard, we focus on the first step.

3.3 An Example

Let $\Theta = \{\theta_L, \theta_H\}$ and x denote the posterior probability that the state is θ_H. Suppose the principal intends to implement distribution $\{x_L, x_H\}$. Consider the contract $(\tilde{M}, \tilde{t})$ where $\tilde{M} = \{x_L, x_H, \varnothing\}$, $x_L = 1/4$ and $x_H = 3/4$, and $\tilde{t}$ is such that $\tilde{t}(x_L, \theta_L) = 1.0125$, $\tilde{t}(x_L, \theta_H) = 0$, $\tilde{t}(x_H, \theta_L) = 0.0125$, and $\tilde{t}(x_H, \theta_H) = 1$. The gross payoff to the

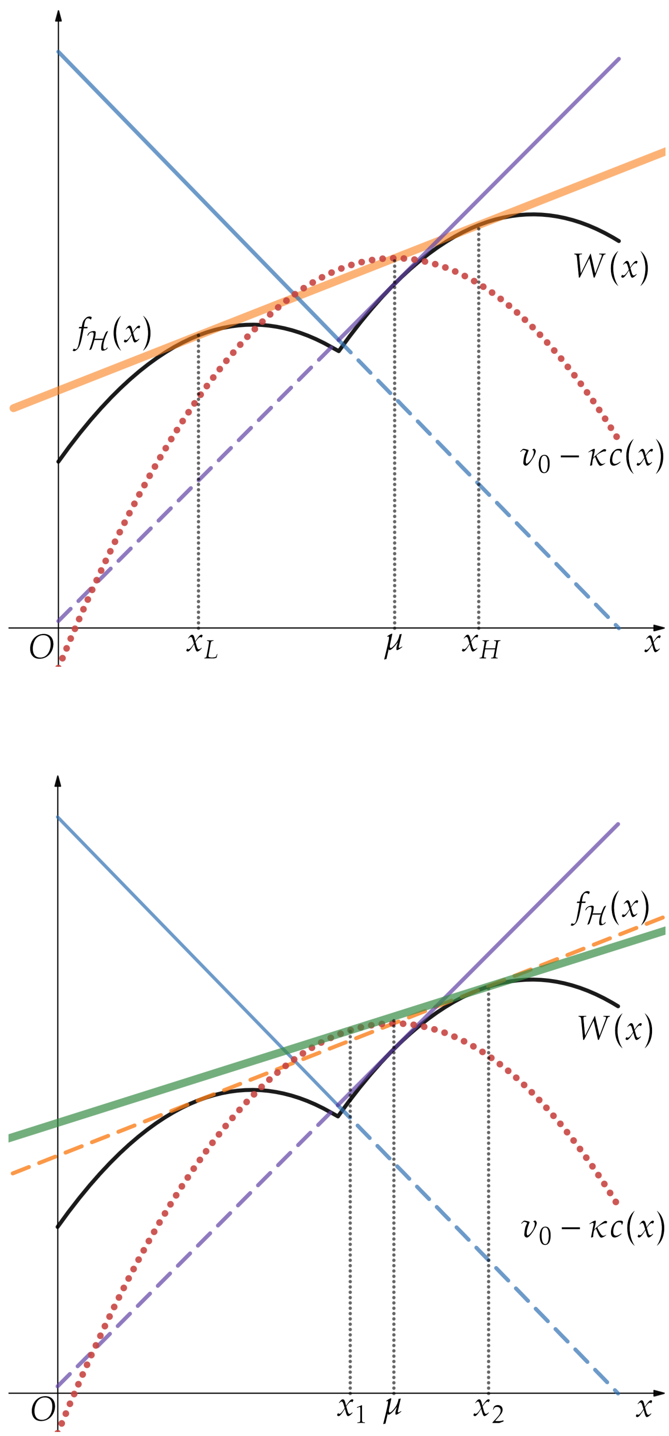


Figure 1: Contract $(\tilde{M}, \tilde{t})$ fails to implement $\{x_L, x_H\}$, where $x_L = 1/4$ and $x_H = 3/4$, when $\mu = 3/5$, $\kappa = 2$, and $v_0 = 0.65$, and with quadratic cost: $c(x) = (x - \mu)^2$. This contract satisfies the limited liability constraints, but the interim IR ($IR - v_0$) is violated.

agent from sending message $m \in \{x_L, x_H\}$, as a function of posterior x, is given by

$$\mathbb{E}_x[\tilde{t}(m,\theta)] = x\tilde{t}(m,\theta_H) + (1-x)\tilde{t}(m,\theta_L).$$

This example is illustrated in Figure 1. The blue and purple lines depicts the gross payoff to the agent from sending x_L and x_H, respectively. The maximum of these functions, net of the agent's learning cost, is the agent's induced value function, W, depicted in black. The concavifying line $f_{\mathcal{H}}$–which pins down the agent's optimal learning–is in orange. Finally, the agent's net payoff from taking the outside option v_0 is the dashed red curve.

If we ignore the individual rationality constraint, the contract $(\tilde{M}, \tilde{t})$ induces the agent to acquire $\{x_L, x_H\}$ and report truthfully because incentive compatibility holds: as shown in the upper panel of Figure 1, the optimal posteriors for the agent are the points at which $f_{\mathcal{H}}$ intersects W, which are indeed $\{1/4, 3/4\}$. However, the IR constraint is violated because $f_{\mathcal{H}}$ does not lie entirely above $v_0 - \kappa c$, which prevents the principal from implementing her desired distribution.[13] As shown in the lower panel of Figure 1, the agent can profitably deviate by acquiring $\{x_1, x_2\}$:[14] if x_1 realizes he opts out by sending message $\varnothing$ and takes his outside option, and he reports x_H if x_2 realizes. This way, his *ex ante* payoff is given by the line through $(x_1, v_0 - \kappa c(x_1))$ and $(x_2, W(x_2))$ (depicted in green) evaluated at the prior μ, which is strictly higher than $f_{\mathcal{H}}(\mu)$.

4 Preliminary Results

We begin by arguing that any distribution over posteriors with support on n or fewer points can be implemented by some contract.

[13]This example also illustrates that interim IR is stronger than *ex ante* IR: as shown in the upper panel, $f_{\mathcal{H}}(\mu) = v_0 - \kappa c(\mu) = v_0$, and hence *ex ante* IR holds.

[14]In this example, $x_1 = 0.522$ and $x_2 = 0.766$.

Lemma 4.1. *If F is a distribution over posteriors with $|\text{supp}(F)| \leq n$ and $\text{supp}(F) \subseteq \text{int}\,\Delta(\Theta)$, there exists a contract (M,t) that implements F, and the expected cost to the principal is finite.*

The proof of Lemma 4.1, and all other proofs omitted from the main text, are collected in Appendix A. For each F supported on n or fewer interior points of $\Delta(\Theta)$, because the cost function c is bounded and differentiable on $\text{int}\,\Delta(\Theta)$, Lemma 2 of Caplin et al. (2022) guarantees that there is a decision problem such that F is optimal. Therefore, we can construct a contract with bounded transfers such that the agent finds it optimal to first acquire F then report the realized posterior truthfully. Moreover, by adding a finite constant to the transfer, we can make Constraint $IR - v_0$ hold. Therefore, every such distribution can be implemented at finite cost.

Because the support of any extreme point of $\mathcal{F}(\mu)$ is on n or fewer affinely-independent points, any $F \in \mathcal{F}(\mu)$ can be obtained by randomizing over a set of contracts each of which implements a distribution with support on at most n affinely-independent points–consequently, any distribution whose support is on the interior of $\Delta(\Theta)$ can be induced at a finite expected cost. As it is cheaper for the principal to randomize first rather than implement F directly, it is without loss of generality for the principal to implement a distribution over posteriors with support on at most n affinely-independent points.

Proposition 4.2. *(i) Every $F \in \mathcal{F}(\mu)$ with $\text{supp}(F) \subseteq \text{int}\,\Delta(\Theta)$ can be implemented at a finite cost.*

(ii) Without loss of generality, the principal only implements distributions with support on at most n affinely-independent points.

By Proposition 4.2 (ii), we can restrict our attention to distributions over posteriors with support on $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$, where n is the number of states, and $s \leq n$. In our next result, we discover that incentive compatibility allows us to reduce transfers to a single variable for each state.

For each state $k = 1,\ldots,n$, define $\Omega^k(i,j) := t_i^k - t_j^k$ $(i,j = 1,\ldots,s)$, where $t_i^k := t(\mathbf{x}_i,\theta_k)$ is the promised payment to the agent from sending message i in state k. Accordingly, each $\Omega^k(i,j)$ specifies the difference between the payoff to the agent from sending any message i versus message j in state k. Importantly, because (on path) each message corresponds to a different posterior, the collection of differences $\left(\Omega^k(i,j)\right)_{k=1}^n$ captures the relative benefit to the agent from obtaining posterior j rather than posterior i.

Theorem 4.3. *The relative incentives $\left(\Omega^k(i,j)\right)_{i,j=1,\ldots,s;k=1,\ldots,n}$ are completely pinned down by incentive compatibility.*

Consequently, the principal's problem of optimally inducing a distribution over posteriors reduces to an n-variable optimization problem, where n is the number of states. For each state k, the principal fixes a benchmark message $j(k)$, then chooses $\left(t_{j(k)}^k\right)_{k=1}^n$; the payoff to the agent from sending message $j(k)$ in state k.

Theorem 4.3 is reminiscent of the standard result that truthtelling only identifies relative payments in adverse selection settings. Here; however, the relative incentives are pinned down jointly by the optimality of the desired distribution *ex ante* and truthful reporting *ex interim*. As part (i) of Lemma 3.1 states, incentive compatibility for the agent requires that the value function of the agent, W, intersects the concavifying hyperplane at the support points of the distribution over posteriors the contract aims to implement. Such a hyperplane pins down the transfers in each state for each posterior in the support of the agent's learning. Consequently, the principal's problem is equivalent to one of choosing a hyperplane, which is an n-variable optimization problem.[15]

That was a technical explanation, here is an economic one. Fix a desired distribution; if the agent wants to deviate by slightly increasing the probability of a message realization in a certain state, basic probability implies that there must be

[15]Framed in this manner, this theorem is closely related to Lemma 2 in Caplin et al. (2022), which states that when constructing a decision problem the tangent hyperplane is arbitrary.

a commensurate decrease in the probability of another message realization in that state. At the optimum, no such local deviation in the agent's information acquisition strategy can be profitable. Hence for any two posteriors in the support of the desired distribution, any deviation of the sort described above must generate a marginal value to the agent equal to the marginal cost. Because the marginal value of varying the probability of a message realization in a state is determined by the transfer for sending that message in that state, this "zero net marginal gain" observation generates an equality that connects the transfers for sending two distinct posteriors in the support of the desired distribution in the same state.

Recall that we specified early on that the principal suffers a disutility greater than v_0 should the agent take his outside option. This ensures that the principal does not want to replace one of the messages with the null message, i.e., have the agent exit the relationship, sending the null message with positive probability. In principle, if the principal is not hurt (severely) by the agent's exit, it could be optimal for the principal to write a contract in which the null message is sent with positive probability (inducing the desired posterior) thereby allowing the principal to save on paying the agent. By Theorem 4.3 the belief to which the null message corresponds pins down the other transfers. Thus, if the principal's penalty from an agent's exit is less than v_0, one must check at most s additional contracts (other than those in which the null message is never sent), in which the null message is sent after each belief, in turn.[16]

[16]In the supplementary appendix we discuss an example in which it is optimal for the agent to exit the relationship with positive probability.

5 Main Results I. No (*Ex Post*) Limited Liability

5.1 Risk-Neutral Agent

We first assume that the agent is risk neutral; without loss of generality, $v(\cdot) = \cdot$. Theorem 4.3 implies that choosing a contract is tantamount to choosing a concavifying hyperplane $\mathcal{H}$. Recalling that $f_{\mathcal{H}}$ is the function that identifies $\mathcal{H}$, and the agent's value from acquiring information for the principal (from an *ex ante* perspective) is $f_{\mathcal{H}}(\mu)$. To implement distribution F efficiently, the principal must only pay the first best cost, namely $v^{-1}(C(F)+v_0) = C(F)+v_0$. Hence, efficient implementation requires $f_{\mathcal{H}}(\mu) = v_0$. Furthermore, for ($IR-v_0$) to hold, the graph of $f_{\mathcal{H}}$ must lie entirely above the graph of $v_0 - \kappa c$. Then, because $f_{\mathcal{H}}$ is affine and c is strictly convex,

Observation 5.1. *When the agent is risk neutral, a principal can implement a distribution efficiently if and only if $f_{\mathcal{H}}$ is tangent to $v_0 - \kappa c$ at μ.*

Applying Observation 5.1, efficient implementation is equivalent to the following n conditions:

$$t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right) = -\kappa c_k(\mu) \quad \text{for all} \quad k = 1, \ldots, n-1, \qquad (IR-R)$$

and $f_{\mathcal{H}}(\mu) = v_0$. We are able to pick an arbitrary support point $\mathbf{x}_j$ of F, as implementability requires that at each $\mathbf{x} \in \operatorname{supp}(F)$, the agent's value function W induced by the contract intersects the same supporting hyperplane ($\mathcal{H}$).

By Theorem 4.3, the solution to this system, $\left(t_j^k\right)_{k=1}^n$, if it exists, identifies a contract. Accordingly, whether a distribution can be implemented efficiently boils down to whether the system of equations defined by the $n-1$ equations in Constraint $IR-R$ and $f_{\mathcal{H}}(\mu) = v_0$ has a solution. A solution always exists:

Proposition 5.2. *If the agent is risk neutral and not protected by limited liability, every* *(feasible) distribution F with* $\operatorname{supp}(F) \subseteq \operatorname{int}\Delta(\Theta)$ *can be implemented efficiently.*[17]

When there is no limited liability, the amount of incentive constraints is "just right" such that there exists a transfer scheme that delivers the right incentives and keeps the agent's surplus at his outside option. Figure 2 illustrates this construction: the principal can always find a contract such that the concavifying hyperplane $f_{\mathcal{H}}$ (depicted in orange) of the agent's value function W (in black) is a supporting hyperplane of the graph of $v_0 - \kappa c$ (the red curve), which is the agent's payoff from exiting the relationship. Therefore, the interim IR constraint ($IR - v_0$) is always satisfied. The following Interactive Link illustrates the optimal contract (for an arbitrary binary distribution with support $\{l, h\}$) when there are two states, the agent's information acquisition cost is entropy-reduction, and his outside option is 0.

It is instructive to compare Proposition 5.2 to Proposition 2 in Rappoport and Somma (2017), which states that when the realized posteriors are contractible (but the true state is not), efficient implementation is possible when the agent is risk neutral, even if he is protected by limited liability. As we discussed after Lemma 3.1, their limited liability is of *ex interim* nature, and is similar to our individual rationality constraint ($IR - v_0$). Consequently, one interesting way to interpret Proposition 5.2 is that so long as the realized state can be contracted upon, Rappoport and Somma (2017)'s Proposition 2 still holds when the principal must elicit the agent's belief instead of being able to contract on it.

However, if we require *ex post* limited liability, for some distributions and outside option values, efficient implementation cannot be achieved (irrespective of the interim IR constraint's presence). In our model, to induce the agent to gather information, the transfers must be "rewarding" when the agent "gets the state right" and "punishing" when he is wrong. This gap between the two outcomes must be

[17] The supplementary appendix reveals that this holds even when there are uncountably many states.

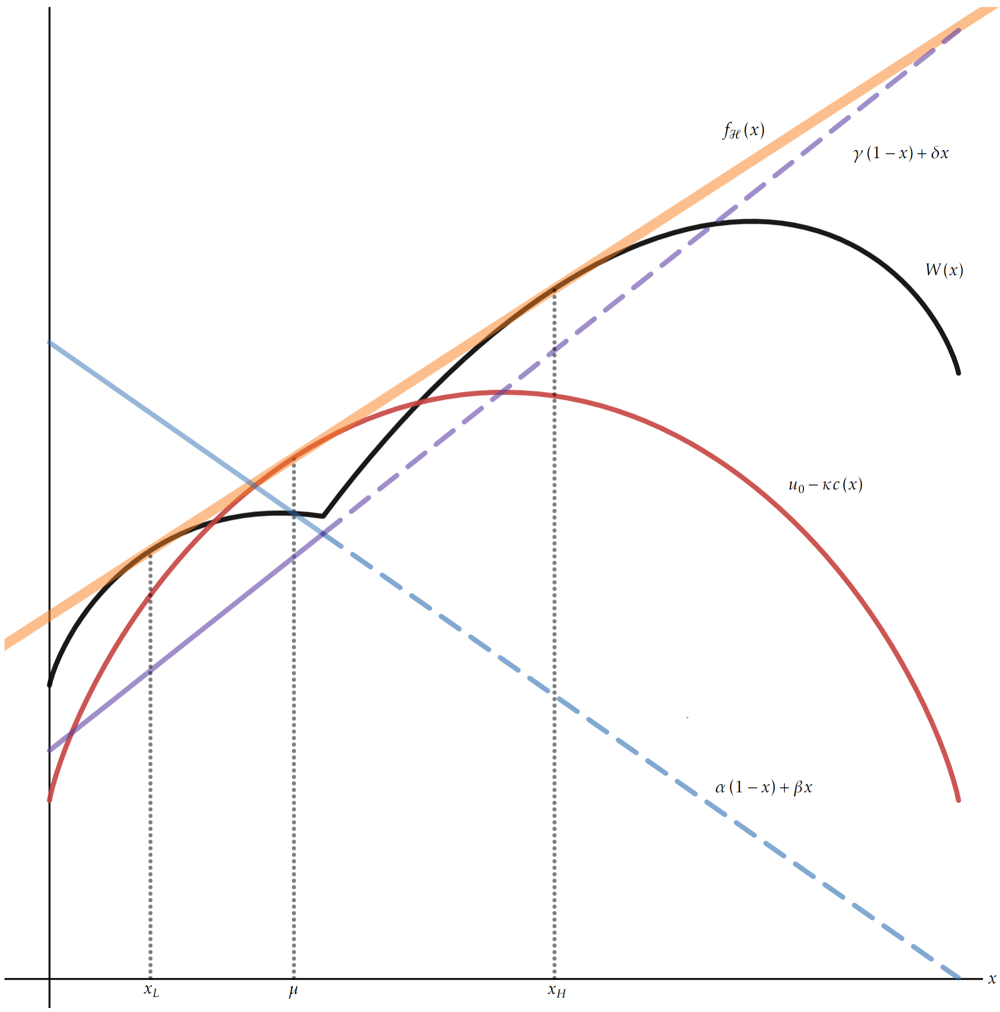


Figure 2: Efficient implementation of $\{x_L, x_H\}$, where $x_L = 1/9$ and $x_H = 5/9$, when $\mu = 1/(1+e)$, $\kappa = 1$, $v_0 = \log\{9/(1+e)\}$, and with entropy cost. This contract satisfies the limited liability constraints–as stated in Proposition 6.3, the specified ratio v_0/κ is the minimum such ratio such that efficient implementation is feasible under limited liability.

large enough to justify the cost of learning. Therefore, when v_0 is small enough, to achieve an expected transfer of $\Gamma(F) = C(F) + v_0$, some "punishing" transfer(s) must be negative.

5.1.1 Comparison to Classical Moral Hazard with an Interim Participation Constraint

A natural question is whether an analog of Proposition 5.2 holds if we add in an interim participation constraint to the canonical moral hazard problem. That is, is the ability of the principal to accommodate the interim participation constraint a special feature of our information acquisition problem or does it also hold in the classical environment?

As we show in the supplementary appendix, in the classical environment, the ability of the agent to exit the relationship after (privately) observing his output realization is tantamount to limited liability: clearly a contract cannot promise the agent a payoff less than his outside option for any (divulged) output. Moreover, because a contract must be incentive compatible, unless the principal implements the lowest possible effort, i.e., pays a constant wage, the agent must get strictly positive rents.

5.1.2 Selling the Project to the Agent?

One might also wonder whether Proposition 5.2 is really needed. In the standard moral hazard problem, when the agent is risk neutral and there are no limited liability constraints, the principal can attain the first best by "selling the project to the agent" (henceforth the **STP** contract). In our setting,[18] that corresponds to the principal writing the contract such that the agent's net utility as a function of his

[18] We are also assuming here that the set of actions in the principal's decision problem is finite. The construction when the principal has infinitely many actions is analogous but more ungainly, so we omit it.

posterior $\mathbf{x}$ is $V(\mathbf{x}) - \kappa c(\mathbf{x}) - \tau$, where $\tau = f_{\mathcal{H}}(\mu) - v_0$. That is, the principal writes a contract so that the agent's problem, gross of the cost, is precisely that faced by the principal, then lowers the transfers to the agent uniformly to leave his *ex ante* expected payoff equal to his outside option.

Already, the similarity between our interim IR constraint and (interim) limited liability suggests that this may not be possible (generically). Indeed, that is so in the standard moral hazard setting. This hunch is correct: when the agent can exit *ex interim*, the principal cannot implement a distribution over posteriors at the first-best cost generically by selling the project. Recall that the optimal contract must be robust to double deviations in which the agent learns differently then takes her outside option with positive probability: $f_{\mathcal{H}}$ must lie everywhere above $v_0 - \kappa c(\mathbf{x})$. Moreover, efficient implementation is even more demanding: $f_{\mathcal{H}}$ must be tangent to $v_0 - \kappa c(\mathbf{x})$ at μ. As we show in the supplementary appendix, this tangency property is a non-generic property of the principal's information acquisition problem.

5.2 Risk-Averse Agent

When the agent is risk averse, but unprotected by limited liability, characterizing the optimal contract is more involved. Fix an arbitrary benchmark message for all states, say j; the principal's payoff is strictly decreasing in each of the n control variables $\left(t_j^k\right)_{k=1}^n$ and so the principal wants to set each one as low as possible. Unencumbered by limited liability, the lone constraint is $IR{-}v_0$, which necessarily binds (since otherwise, the principal could reduce the control variables). Thus,

Observation 5.3. *When the agent is risk averse, there exists an* $\mathbf{x}^* \in \Delta(\Theta)$ *such that* $f_{\mathcal{H}}(\mathbf{x})$ *is tangent to* $v_0 - \kappa c(\mathbf{x})$ *at* $\mathbf{x}^*$.

Given this, solving for the optimal implementation of a distribution over posteriors F can be turned into an $n-1$ variable optimization problem by using the tangency conditions to substitute in for each t_j^k. This yields the principal an objective that is a function of $\mathbf{x}^*$.[19] Unless $\mathbf{x}^* = \mu$, which does not hold in general, the agent obtains positive rents. This finding is a consequence of the interim participation constraint, which requires that $f_{\mathcal{H}}$ lie above $v_0 - \kappa c$ everywhere. Otherwise–with only *ex ante* IR–the agent would not obtain rents. Indeed, without the interim IR constraint, the lone constraint is the *ex ante* participation constraint, which obviously binds; hence, $f_{\mathcal{H}}(\mu) = v_0$. Importantly, the agent's risk aversion engenders inefficiencies: the agent must be exposed to risk in order to acquire information and report honestly, which destroys surplus due to his risk aversion.

Proposition 5.4. *Suppose the agent is risk averse and not protected by limited liability.*

(i) For every distribution over posteriors F with support in $\operatorname{int}\Delta(\Theta)$*, an optimal contract exists, and the transfers can be found by choosing* $\mathbf{x}^* \in \Delta(\Theta)$*.*

(ii) If the agent can exit ex interim, he gets strictly positive rents unless $\mathbf{x}^* = \mu$*. If the agent cannot exit ex interim, he gets zero rents.*

(iii) Only the degenerate distribution of posteriors can be implemented efficiently.

In choosing $\mathbf{x}^*$, the principal optimally trades off between risk sharing and conceding rents: when a contract that makes the agent break even entails too much risk, moving $\mathbf{x}^*$ away from μ mitigates this risk. Then, although the agent receives strictly positive rents, implementing the new contract can be cheaper to the principal. This is reminiscent of the trade-off studied in Proposition 5 in Rappoport and Somma (2017) though the exact mechanisms are different:[20] in their work, the

[19]More precisely, for each $\mathbf{x}^* \in \Delta(\Theta)$, $\left(t_j^k\right)_{k=1}^n$ can be solved from the following n equations: $t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right) = -\kappa c_k(\mathbf{x}^*)$ for all $k = 1, \ldots, n-1$, and $f_{\mathcal{H}}(\mathbf{x}^*) = v_0 - \kappa c(\mathbf{x}^*)$. Moreover, the relative incentives identified in Theorem 4.3 allow us to obtain the other transfers. Plugging the transfers into (‡), the principal's objective can then be written as a function of $\mathbf{x}^*$.

[20]The resemblance stems from the fact that our interim IR constraint is similar to the (interim) limited liability constraint in Rappoport and Somma (2017), which leads to similar tangency conditions.

most cost-efficient way for compelling the agent to choose the right distribution is to have the hyperplane determined by the wage contract (which, in their setting, is a function of the *verifiable* posterior) to be tangent to the agent's value function. In our problem, averting double deviations to the outside option is what begets the tangency condition mentioned in Observation 5.3.

6 Main Results II. (*Ex Post*) Limited Liability

Throughout this section, we assume that the agent is protected by limited liability. In Subsection 6.1, we solve for the optimal incentives when the agent's value for his outside option is sufficiently small. In Subsection 6.2, we allow for an arbitrary outside option but impose that the agent is risk neutral.

6.1 Low Outside Option

For simplicity, we set $v_0 = 0$; it is not hard to see that all the results in this subsection go through for all sufficiently small $v_0 > 0$. By Theorem 4.3, for any desired distribution, the relative incentives are identified. Consequently, for each state we can pinpoint a benchmark message that determines the lowest payment.

Lemma 6.1. *For every state $k = 1, \ldots, n$, there exists $j^*(k)$ such that $t\left(\mathbf{x}_{j^*(k)}, \theta_k\right) \leq t(\mathbf{x}_i, \theta_k)$ for all $i = 1, \ldots, s$.*

Lemma 6.1 allows us to completely identify the optimal transfers when the agent's outside option is sufficiently low.

Proposition 6.2. *Suppose $v_0 = 0$, and the agent is protected by limited liability. Then for each state $k = 1, \ldots, n$, there exists $j^*(k)$ such that $t\left(\mathbf{x}_{j^*(k)}, \theta_k\right) = 0$, and all other transfers are nonnegative and determined by the relative incentives identified in Theorem 4.3.*

Proposition 6.2 is intuitive: for a sufficiently small outside option, Constraint $IR - v_0$ always holds, so the transfer scheme is pinned down by optimal learning and limited liability. Optimal learning leaves, for each state, one degree of freedom to the principal; and to satisfy limited liability, the best that the principal can do is to find the smallest transfer in each state and set it to zero. The following Interactive Link illustrates the optimal contract (for an arbitrary binary distribution with support $\{l, h\}$) when there are two states, the risk-neutral agent's information acquisition cost is entropy-reduction, and his outside option is 0.

6.2 Risk-Neutral Agent

Now, we dispense with the assumption that the outside option is small–v_0 can take any value. For expository ease, we start with the two state case and then argue that our results generalize when there are more than two states.

6.2.1 Two States

When there are just two states, $\Theta = \{\theta_1, \theta_2\}$. By Proposition 4.2 (ii), we can identify a distribution by its support $\{x_L, x_H\}$. Our first result characterizes the distributions over posteriors that a principal can implement efficiently; *viz.*, at the first-best cost. Defining

$$\eta(x_L, x_H) := \max\{-\mu c'(\mu) - c(x_H) + c'(x_H)x_H, (1-\mu)c'(\mu) - c(x_L) - (1-x_L)c'(x_L)\} ,$$

we have

Proposition 6.3. *The principal can implement $\{x_L, x_H\}$ efficiently if and only if $v_0/\kappa \geq \eta(x_L, x_H)$.*

Proposition 6.3 states that a given distribution can be implemented efficiently if and only if either the agent has a sufficiently high outside option, or he can acquire information sufficiently cheaply. Intuitively, efficient implementation under limited liability requires that (1) the average payment to the agent net of the cost of information acquisition is v_0 (and Proposition 5.2 shows that ($IR - v_0$) can be satisfied by "tilting the hyperplane"), and (2) the payments from sending the "wrong message" in both states cannot be negative. (1) and (2) together imply that the differential payments between the "right message" and the wrong one cannot be too large compared to v_0. The differential payments are exactly the relative incentives identified in Theorem 4.3, which are pinned down jointly by the desired distribution and model primitives (i.e., κ and c). This produces the inequality in Proposition 6.3.

Consequently, as v_0 gets larger, efficient implementation is easier. Furthermore, for a smaller κ or a (Blackwell) less informative distribution, the principal only needs smaller differential payments to incentivize information acquisition, which also makes efficient implementation easier to achieve. Therefore, the left-hand side of Proposition 6.3's necessary and sufficient condition is strictly decreasing in the information cost parameter κ. Moreover, the function η is decreasing in x_L and increasing in x_H. This suggests the following corollary:

Corollary 6.4. *(i) For any pair of posteriors $\{x_L, x_H\}$ with $0 < x_L \leq \mu \leq x_H < 1$, if v_0/κ is sufficiently large, $\{x_L, x_H\}$ can be implemented efficiently.*[21]

(ii) Efficient implementation is monotone with respect to the Blackwell order: if $\{x_L, x_H\}$ can be implemented efficiently, then any distribution that corresponds to a less informative experiment can be implemented efficiently.

(iii) If $v_0 > 0$ then any distribution that corresponds to a sufficiently uninformative experiment can be implemented efficiently.

In the canonical moral hazard problem with a risk-averse agent, no matter what outside option the agent has, only the lowest action can be implemented efficiently. Corollary 6.4 has a flavor of that classical result: efficient implementation is possible whenever the agent is not asked to learn too much. For distributions more

[21] If $c'(0)$ and $c'(1)$ are finite, this is true for any $0 \leq x_L \leq \mu \leq x_H \leq 1$.

spread out than some set of threshold distributions; however, positive rents must be provided to the agent. To implement such distributions efficiently it must be that the relative incentives are high enough for the agent to acquire that much information and so when v_0 is small limited liability is always violated.

When the first-best implementation of $\{x_L, x_H\}$ is infeasible, there are three other possibilities, listed in our next proposition. Denoting $\gamma := t_2^1$ the transfer from sending message x_L in state θ_2, and $\beta := t_1^2$ the transfer from sending message x_H in state θ_1, we have

Proposition 6.5. *One of the following must occur at the optimum. Either*

(i) $\{x_L, x_H\}$ *can be implemented efficiently (and Constraint* $IR - v_0$ *binds); or*

(ii) $\{x_L, x_H\}$ *cannot be implemented efficiently; and either*

  (a) Constraint $IR - v_0$ *binds and* $\beta = 0$*; or*

  (b) Constraint $IR - v_0$ *binds and* $\gamma = 0$*; or*

  (c) Constraint $IR - v_0$ *does not bind and* $\gamma = \beta = 0$*.*

When the cost function is the entropy cost, it is straightforward to characterize the four regions of $\{x_L, x_H\}$ pairs. They are depicted in Figure 3. Here is an Interactive Link, where one can adjust the sliders for $m \equiv \mu$ and $u \equiv \frac{v_0}{\kappa}$, to see how the optimal contract changes.

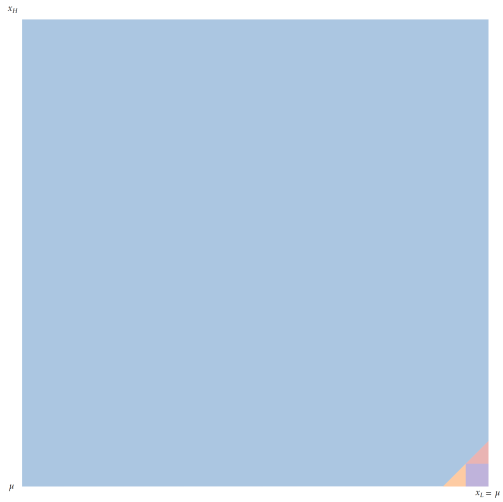
6.2.2 Two States and No Interim IR

When there is no interim IR constraint, we need only impose $f(\mu) \geq v_0$ to guarantee that the agent accepts the contract. Defining

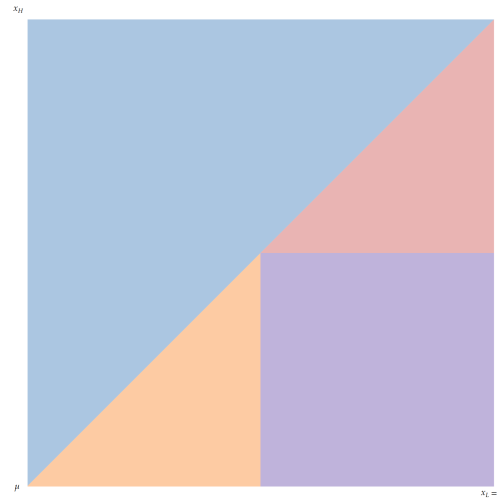
$$\zeta(x_L, x_H) := (1-\mu)(x_H c'(x_H) - c(x_H)) - \mu((1-x_L)c'(x_L) + c(x_L)) \ ,$$

we have

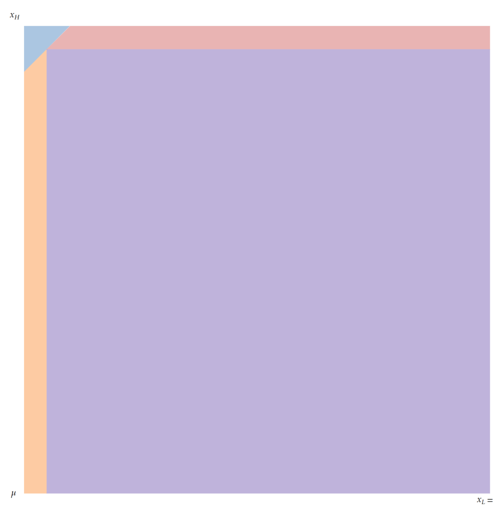
Proposition 6.6. *The principal can implement* $\{x_L, x_H\}$ *efficiently if and only if* $v_0/\kappa \geq \zeta(x_L, x_H)$*. Otherwise,* $\gamma = \beta = 0$*.*



(a) $v_0/\kappa = .05$.



(b) $v_0/\kappa = \log(1/(1-\mu))$.



(c) $v_0/\kappa = 3$.

Figure 3: **Implementation Regions for $\mu = 1/2$:** x_L is on the horizontal axis, ranging from 0 to $\mu = 1/2$; and x_H is on the horizontal axis, ranging from $\mu = 1/2$ to 1. Pairs (x_L, x_H) in the purple region can be implemented efficiently, (x_L, x_H) in the blue region are optimally implemented by $\gamma = \beta = 0$, (x_L, x_H) in the orange region are optimally implemented by $\beta = 0$ and some $\gamma \geq 0$; and (x_L, x_H) in the red region are optimally implemented by $\gamma = 0$ and some $\beta \geq 0$.

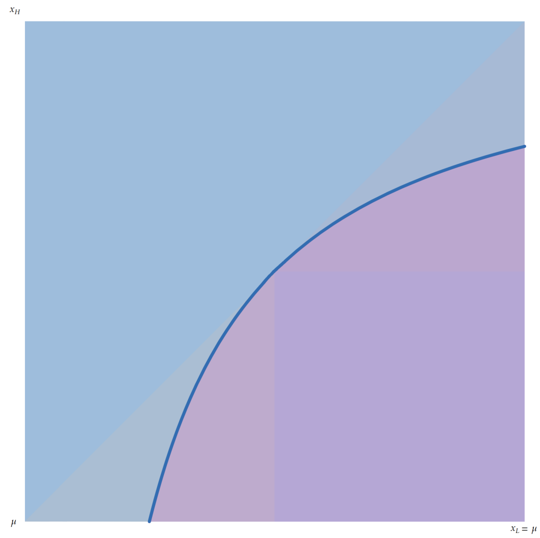


Figure 4: **Implementation Regions for $\mu = 1/2$ and $v_0/\kappa = \log(1/(1-\mu))$ (No Interim IR):** Pairs (x_L, x_H) in the purple region can be implemented efficiently and (x_L, x_H) in the blue region are optimally implemented by $\gamma = \beta = 0$.

It is obvious that an exact analog of Corollary 6.4 holds when there is no interim IR constraint. *Viz.*, any pair of posteriors can be implemented efficiently if the outside option is sufficiently large and the cost of acquiring information κ is sufficiently small. Moreover, the more information an agent is asked to acquire, the more difficult it is to implement the distribution efficiently.

When the cost function is the entropy cost, it is straightforward to characterize the two regions of $\{x_L, x_H\}$ pairs. They are depicted in Figure 4, superimposed over the four regions present when there is an interim participation constraint.

6.2.3 More Than Two States

By Lemma 6.1, for each state k we can find a message $j^*(k)$ that delivers the lowest payment; and by Theorem 4.3, to pin down the transfer scheme, it suffices to determine $t^k_{j^*(k)}$ for each state k. Thus, there are n unknowns. When we impose the interim IR constraint, there are n equations: efficient implementation is equivalent to $f_{\mathcal{H}}(\mu) = v_0$, and the other $n-1$ equations are given by Constraint $IR-R$:

$$t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right) = -\kappa c_k(\mu) \ ,$$

where $k = 1, \ldots, n-1$ indicates the state, and j, which indexes the posterior, is arbitrary. Then the distribution can be efficiently implemented if and only if $t_{j^*(k)}^k \geq 0$ for each k; consequently, Proposition 6.3, Corollary 6.4 (i), Proposition 6.5, and Proposition 6.6 naturally extend to more than two states.

7 Discussion

We conclude by discussing a couple of our assumptions.

Agent has no intrinsic preferences for learning. We make this assumption to zero in the incentive provision problem when the principal has to delegate information acquisition to an agent who cannot make verifiable reports. To allow the agent to have intrinsic motivation, we assume that the agent's intrinsic value from posterior $\mathbf{x}$ is $\phi(\mathbf{x})$, which is known to both parties. Then it is as if the agent's cost of arriving at posterior $\mathbf{x}$ is $\kappa c(\mathbf{x}) - \phi(\mathbf{x})$, and all of our results survive intact.

Restricted learning. Our agent is unconstrained in how she learns: she may choose any Bayes-plausible distribution. How might our results change if the agent instead could only choose from some subset thereof? In general, such restrictions must make implementation (of those available distributions) cheaper, as the agent has fewer possible deviations. A corollary of this observation (proved in the supplementary appendix), is that the principal can implement any (feasible) distribution efficiently if the agent is risk neutral and there are no (*ex post*) limited liability constraints.

References

Jacopo Bizzotto, Eduardo Perez-Richet, and Adrien Vigier. Information design with agency. *Mimeo*, 2020.

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715, 2022.

Gabriel Carroll. Robust incentives for information acquisition. *Journal of Economic Theory*, 181:382–420, 2019.

Hector Chade and Natalia Kovrijnykh. Delegated Information Acquisition with Moral Hazard. *Journal of Economic Theory*, 162:55–92, 2016.

Aubrey Clark and Giovanni Reggiani. Contracts for acquiring information. *arXiv: 2103.03911*, 2021.

Rahul Deb, Mallesh M. Pai, and Maher Said. Evaluating strategic forecasters. *American Economic Review*, 108(10):3057–3103, 2018.

Joel S Demski and David E M Sappington. Delegated expertise. *Journal of Accounting Research*, 25:68–89, 1987.

Tilmann Gneiting and Adrian E Raftery. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

Denis Gromb and David Martimort. Collusion and the organization of delegated expertise. *Journal of Economic Theory*, 137(1):271–299, 2007.

Sanford J Grossman and Oliver D Hart. An Analysis of the Principal-Agent Problem. *Econometrica*, 51(1):7, 1983.

Samuel Häfner and Curtis R Taylor. On young turks and yes men: Optimal contracting for advice. *The RAND Journal of Economics*, 53(1):63–94, 2022.

Benjamin Hébert and Michael Woodford. Neighborhood-based information costs. *American Economic Review*, 111(10):3225–3255, 2021.

Bengt Holmström. Moral hazard and observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

Bengt Holmström and Paul Milgrom. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, 55(2):303–328, 1987.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615, 2011.

Andreas Kleiner and Mark Whitmeyer. Whether to randomize with subjective evaluations. *Mimeo*, In Progress.

Nicolas S Lambert. Elicitation and Evaluation of Statistical Forecasts. *Mimeo*, 2019.

Richard A Lambert. Executive Effort and Selection of Risky Projects. *The RAND Journal of Economics*, 17(1):77–88, 1986.

Bartosz Maćkowiak, Filip Matějka, and Mirko Wiederholt. Rational inattention: A review. *Journal of Economic Literature*, Forthcoming.

Filip Matějka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1): 272–98, 2015.

J A Mirrlees. The theory of moral hazard and unobservable behaviour: Part i. *The Review of Economic Studies*, 66(1):3–21, 1999.

Michèle Müller-Itten, Roc Armenter, and Roc Stangebye. Rational inattention via ignorance equivalence. *Mimeo*, 2021.

Kent Osband. Optimal Forecasting Incentives. *Journal of Political Economy*, 97(5): 1091–1112, 1989.

Luciano Pomatto, Philipp Strack, and Omer Tamuz. The cost of information: The case of constant marginal costs. *American Economic Review*, 113(5):1360–1393, 2023.

Daniel Rappoport and Valentin Somma. Incentivizing Information Design. *Available at SSRN 3001416*, 2017.

Karl H Schlag, James Tremewan, and Joël J van der Weele. A Penny for Your Thoughts: a Survey of Methods for Eliciting Beliefs. *Experimental Economics*, 18:457–490, 2015.

Salil Sharma, Elias Tsakas, and Mark Voorneveld. Procuring unverifiable information. *Mimeo*, 2020.

Christopher A Sims. Stickiness. In *Carnegie-rochester conference series on public policy*, volume 49, pages 317–356, 1998.

Christopher A Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

Spyros Terovitis. Motivating Information Acquisition Under Delegation. *Available at SSRN 3264177*, 2018.

Cristian Martín Vidal. *La Masia. Developing People Beyond Sport*. Editorial Base, 2019.

Han Wang. Contracting with heterogeneous researchers. *Mimeo*, 2023.

Nathan Yoder. Designing incentives for heterogeneous researchers. *Journal of Political Economy*, 130(8):2018—-2054, 2022.

Luis Zermeño. A principal-expert model and the value of menus. *Mimeo*, 2011.

A Omitted Proofs

A.1 Lemma 4.1 Proof

Proof. Let $\text{supp}(F) = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$, where $s = |\text{supp}(F)| \leq n$. Consider a contract (M, t) where $M = \text{supp}(F)$, and for each $j = 1, \ldots, s$,

$$t\left(\mathbf{x}_j, \theta_k | \tau\right) = \kappa c\left(\mathbf{x}_j\right) - \sum_{i=1}^{n-1} x_j^i \kappa c_i\left(\mathbf{x}_j\right) + \kappa c_k\left(\mathbf{x}_j\right) + \tau \text{ for all } k = 1, \ldots, n-1 \ ,$$

$$t\left(\mathbf{x}_j, \theta_n | \tau\right) = \kappa c\left(\mathbf{x}_j\right) - \sum_{i=1}^{n-1} x_j^i \kappa c_i\left(\mathbf{x}_j\right) + \tau \ ,$$

where x_j^i is the i-th entry of $\mathbf{x}_j$, c_i is the partial derivative of c with respect to its i-th entry, and τ is a constant that scales the transfers.

For any $m \in M$, we define the agent's *net utility* $N(\mathbf{x}|m)$ as the expected utility of sending message m net of the cost of $\mathbf{x}$:

$$N(\mathbf{x}|m) = \mathbb{E}_{\mathbf{x}}\left[t(m, \theta)\right] - \kappa c(\mathbf{x}) \ .$$

Let G be a distribution over posteriors, and let $\sigma : \Delta(\Theta) \to \Delta(M)$ denote a reporting strategy. Then, the agent's *ex ante* value of choosing (G, σ) is given by

$$\Upsilon(G, \sigma) = \sum_{\mathbf{x} \in \text{supp}(G)} \sum_{m \in M} G(\mathbf{x}) \sigma(m|\mathbf{x}) N(\mathbf{x}|m) \ .$$

We claim that (F, σ^*), where $\sigma^*(\cdot|\mathbf{x}_j) = \delta_{\mathbf{x}_j}$ is an optimal strategy for the agent.[22] By Lemma 1 in Caplin et al. (2022), it suffices to show that, for every $\mathbf{x}_j$, $j = 1, \ldots, s$, there exists a $n-1$ dimensional vector $\lambda = (\lambda_1, \ldots, \lambda_s)$ such that

$$N(\mathbf{x}|m) - \sum_{i=1}^{n-1} \lambda_i x^i \leq N\left(\mathbf{x}_j \middle| \mathbf{x}_j\right) - \sum_{i=1}^{n-1} \lambda_i x_j^i \ ,$$

[22] $\delta_{\mathbf{x}_j}$ denotes the degenerate distribution at $\mathbf{x}_j$; that is, the agent truthfully reports the posterior that he obtained from learning.

for all $\mathbf{x} \in \Delta(\Theta)$ and $m \in M$. We set λ to be the zero vector, so the above inequality reduces to $N(\mathbf{x}|m) \leq N\left(\mathbf{x}_j \middle| \mathbf{x}_j\right)$. We first show that for any fixed $m \in M$, $N(\mathbf{x}|m) \leq N\left(\mathbf{x}_j \middle| m\right)$, and then we show that $N\left(\mathbf{x}_j \middle| m\right) \leq N\left(\mathbf{x}_j \middle| \mathbf{x}_j\right)$. To establish the first inequality, since $c(\mathbf{x})$ is strictly convex, the first-order conditions (FOCs) are sufficient; the FOCs are

$$t\left(m, \theta_i | \tau\right) - t\left(m, \theta_n | \tau\right) - \kappa c_i(\mathbf{x}) = \kappa\left(c_i\left(\mathbf{x}_j\right) - c_i(\mathbf{x})\right) = 0 \text{ for all } i = 1, \ldots, n-1 ,$$

clearly setting $\mathbf{x} = \mathbf{x}_j$ makes all of them hold. For the second inequality,

$$N\left(\mathbf{x}_j \middle| \mathbf{x}_j\right) - N\left(\mathbf{x}_j \middle| m\right) = \kappa\left(c\left(\mathbf{x}_j\right) - c(m) - \sum_{i=1}^{n-1}\left(x_j^i - m_i\right) c_i(m)\right) \geq 0 ,$$

where m_i is the i-th coordinate of m, and the inequality follows from the convexity of c. Therefore, (F, σ^*) is indeed optimal, and it is direct that the agent's payoff is $\Upsilon(F, \sigma^*) = \tau$. Moreover, there exists $\tau^* < \infty$ large enough, since c is bounded and differentiable on $\operatorname{int}\Delta(\Theta)$, such that Constraint $IR - v_0$ holds. Thus, contract (M, t) implements F. The principal's expected cost is finite since $t\left(\mathbf{x}_j, \theta_k | \tau^*\right)$ is finite for all j, k. ■

A.2 Proposition 4.2 Proof

Proof. Let $\operatorname{ext}\mathcal{F}(\mu)$ denote the set of extreme points of $\mathcal{F}(\mu)$. Because $\mathcal{F}(\mu)$ is convex and compact, by Choquet's theorem, for any $G \in \mathcal{F}(\mu)$ there exists a probability measure Λ_G that puts probability 1 on $\operatorname{ext}\mathcal{F}(\mu)$, and

$$G = \int_{\operatorname{ext}\mathcal{F}(\mu)} H \, \mathrm{d}\Lambda_G(H) . \tag{R}$$

Therefore, any distribution G with support on $\operatorname{int}\Delta(\Theta)$ can be obtained by randomizing over distributions supported on at most n affinely-independent points. Then by Lemma 4.1, G can be implemented at a finite cost by randomizing over contracts we constructed therein. This establishes part (i).

For part (ii), suppose there exists a contract (M,t) under which the agent chooses G, where $|\text{supp}(G)| > n$, and $(G,\hat{\sigma})$ is the induced optimal strategy of the agent. Without loss of generality, $M = \text{supp}(G)$ and $\hat{\sigma}(\cdot|\mathbf{x}) = \delta_{\mathbf{x}}$ for all $\mathbf{x} \in \text{supp}(G)$. Then for every posterior $\mathbf{x} \in \text{supp}(G)$ and every $m \in M$ with $\hat{\sigma}(m|\mathbf{x}) > 0$,

$$N(\mathbf{x}|m) + \sum_{i=1}^{n-1} \left(t(m,\theta_i) - t(m,\theta_n) - \kappa c_i(\mathbf{x})\right)(\tilde{x}_i - x_i) \geq N(\tilde{\mathbf{x}}|m') \tag{H}$$

for all $\tilde{\mathbf{x}} \in \Delta(\Theta)$ and $m' \in M$. By Equation R, for every $F \in \text{supp}(\Lambda_G)$, and every posterior $\mathbf{x}$, $\mathbf{x} \in \text{supp}(G)$. Hence, the strategy $\left(F, \hat{\sigma}\big|_{\text{supp}(F)}\right)$ is also optimal for the agent since Inequality H holds for every $\mathbf{x} \in \text{supp}(F)$ and every $m \in M$ with $\hat{\sigma}\big|_{\text{supp}(F)}(m|\mathbf{x}) > 0$. Now it is direct that each $F \in \text{supp}(\Lambda_G)$ can be implemented by the contract (M_F, t_F) where $M_F = \text{supp}(F)$, and t_F is the restriction of t to M_F; thus, G can be implemented at the same cost by randomizing over $\text{supp}(\Lambda_G)$.

Note that; however, for all $F \in \text{supp}(\Lambda_G)$, (M_F, t_F) need not be the least costly contract under which the agent chooses F: randomizing over $\text{supp}(\Lambda_G)$ and finding the least costly contract for each F is at least cheaper than (M,t). Therefore, without loss of generality, the principal only implements distributions with support on at most n affinely-independent points. This concludes the proof of part (ii). ■

A.3 Theorem 4.3 Proof

Proof. The principal wants to implement a distribution F using some contract (M,t). By part (i) of Lemma 3.1, a necessary condition for implementation is that $\text{supp}(F) = P_{(M,t)}$; this condition holds if and only if the contract is such that the following s expressions

$$\begin{array}{l}\sum_{k=1}^{n-1} x_1^k t_1^k + \left(1 - \sum_{k=1}^{n-1} x_1^k\right) t_1^n - \kappa c\left(\mathbf{x}_1\right) + \sum_{k=1}^{n-1}\left(t_1^k - t_1^n - \kappa c_k\left(\mathbf{x}_1\right)\right)\left(x_k - x_1^k\right)\\ \sum_{k=1}^{n-1} x_2^k t_2^k + \left(1 - \sum_{k=1}^{n-1} x_2^k\right) t_2^n - \kappa c\left(\mathbf{x}_2\right) + \sum_{k=1}^{n-1}\left(t_2^k - t_2^n - \kappa c_k\left(\mathbf{x}_2\right)\right)\left(x_k - x_2^k\right)\\ \vdots\\ \sum_{k=1}^{n-1} x_s^k t_s^k + \left(1 - \sum_{k=1}^{n-1} x_s^k\right) t_s^n - \kappa c\left(\mathbf{x}_s\right) + \sum_{k=1}^{n-1}\left(t_s^k - t_s^n - \kappa c_k\left(\mathbf{x}_s\right)\right)\left(x_k - x_s^k\right)\end{array},$$

define the same hyperplane, where $t_j^k := t\left(\mathbf{x}_j, \theta_k\right)$, x_j^i is the i-th entry of $\mathbf{x}_j$, and c_i is the partial derivative of c with respect to its i-th entry. Accordingly, for all $k = 1, \dots, n-1$ and $i, j = 1, \dots s$

$$t_i^k - t_i^n - \kappa c_k\left(\mathbf{x}_i\right) = t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right) \quad \text{and} \quad t_i^n = t_j^n + \Xi_{ij} ,$$

where Ξ_{ij} is some function of the primitives (but not directly of the ts). Combining these two equations, we obtain

$$\Omega^k(i,j) = \kappa c_k\left(\mathbf{x}_i\right) - \kappa c_k\left(\mathbf{x}_j\right) + \Xi_{ij}, \text{ for } k = 1, \dots, n-1 \text{ and } \Omega^n(i,j) = \Xi_{ij} .$$

Accordingly, for each state $k = 1, \dots, n$, once the principal chooses the transfer for one of the messages in state k, the transfers for all other messages are automatically pinned down. In other words, the principal has one degree of freedom for each of the states. In every state $k = 1, \dots, n$, and for every $i, j = 1, \dots, s$, we can write $t_i^k = t_j^k + \Omega^k(i,j)$. ■

A.4 Proposition 5.2 Proof

Proof. Let F be such that $\operatorname{supp}(F) = \{\mathbf{x}_1, \dots, \mathbf{x}_s\} \subseteq \operatorname{int}\Delta(\Theta)$, where $s \leq n$. As noted in the main text, there are $n-1$ equations given by Constraint $IR-R$: $t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right) = -\kappa c_k(\mu)$ for all $k = 1, \dots, n-1$, and efficient implementation requires $f_{\mathcal{H}}(\mu) = v_0$, which can be written as

$$\sum_{k=1}^{n-1}\left(t_j^k - t_j^n - \kappa c_k\left(\mathbf{x}_j\right)\right)\mu_k + t_j^n = Q \ ,$$

where μ_k is the k-th entry of μ, and Q does not depend on t's. To show that F can be efficiently implemented, it suffices to find a solution of this system of n equations. Using $IR\!-\!R$, the equality above can be reduced to $t_j^n = Q+\sum_{k=1}^{n-1}\kappa\mu_k c_k(\mu)$; plugging this into the other $n-1$ equations, we get $t_j^k = Q+\sum_{i=1}^{n-1}\kappa\mu_i c_i(\mu)+\kappa\left(c_k\left(\mathbf{x}_j\right)-c_k(\mu)\right)$ for each $k=1,\ldots,n$. We have thus found a solution. Because F is an arbitrary distribution over posteriors supported on at most n points, the principal can implement any distribution G with $\text{supp}(G)\subseteq \text{int}\,\Delta(\Theta)$ efficiently by randomizing *ex ante*. ■

A.5 Proposition 5.4 Proof

Proof. Suppose first that the agent can exit *ex interim*. Because c is strictly convex, $v_0 - c(\mathbf{x})$ is strictly concave. By Observation 5.3, $f_{\mathcal{H}}(\mathbf{x})$ is tangent to $v_0 - c(\mathbf{x})$ at $\mathbf{x}^*$. Thus, $\mathbf{x}^* \neq \mu$ implies that $f_{\mathcal{H}}(\mu) > v_0$, and hence the agent gets strictly positive rents. When the agent cannot exit *ex interim*, the fact that he gets zero rents is almost immediate: if $f_{\mathcal{H}}(\mu) > v_0$, because there is no limited liability, the transfer can be lowered by some small $\varepsilon > 0$.

Let F be a nondegenerate distribution with $\text{supp}(F) = \{\mathbf{x}_1,\ldots,\mathbf{x}_s\}$, where $\mathbf{x}_i \neq \mathbf{x}_j$ for all $i,j=1,\ldots,s$ with $i\neq j$. Suppose to the contrary that F can be efficiently implemented. To simplify notation, let $x_j^n := 1-\sum_{k=1}^{n-1}x_j^k$ for each $j=1,\ldots,s$. The principal wishes to minimize $\sum_{j=1}^{s}\sum_{k=1}^{n}p_j x_j^k v^{-1}\left(t_j^k\right)$. Because F can be efficiently implemented,

$$\sum_{j=1}^{s}\sum_{k=1}^{n}p_j x_j^k v^{-1}\left(t_j^k\right) = v^{-1}\left(C(F)+v_0\right) = v^{-1}\left(\sum_{j=1}^{s}\sum_{k=1}^{n}p_j x_j^k t_j^k\right).$$

Because v is strictly concave, v^{-1} is strictly convex; Jensen's inequality then implies $t_j^k = \tilde{t}$ for all $k=1,\ldots,n$ and $j=1,\ldots,s$. But then since c is strictly convex, learning according to the degenerate distribution is uniquely optimal to the agent, and hence the contract with constant transfer cannot implement F. A contradiction. ■

A.6 Lemma 6.1 Proof

Proof. Fix any state k and an arbitrary message, say s. Define $N(k) = \left\{i : \Omega^k(i,s) < 0\right\}$. If $N(k) = \varnothing$, let $j^*(k) = s$; then since $t_i^k = t_{j^*(k)}^k + \Omega^k(i,s)$, we have $t_{j^*(k)}^k \leq t_i^k$ for all $i = 1,\ldots,s$. Otherwise, let $j^*(k)$ be an arbitrary selection of $\arg\min_{j\in N(k)} \Omega^k(j,s)$. Optimal learning requires, for any i, $t_i^k = t_s^k + \Omega^k(i,s)$ and $t_{j^*(k)}^k = t_s^k + \Omega^k(j^*(k),s)$, which implies $t_{j^*(k)}^k - t_i^k = \Omega^k(j^*(k),s) - \Omega^k(i,s) \leq 0$. Again, $t_{j^*(k)}^k \leq t_i^k$ for all $i = 1,\ldots,s$. ■

A.7 Proposition 6.2 Proof

Proof. By Lemma 6.1, for every state $k = 1,\ldots,n$, there exists $j^*(k)$ such that $t_{j^*(k)}^k \leq t_i^k$ for all $i = 1,\ldots,s$. Then by setting $t_{j^*(k)}^k = 0$, the agent's honesty is not affected, and the limited liability constraints are satisfied. For every $i \neq j^*(k)$, we have

$$t_i^k = \Omega^k(i, j^*(k)) = \kappa c_k(\mathbf{x}_i) - \kappa c_k\left(\mathbf{x}_{j^*(k)}\right) + \Xi_{i j^*(k)}$$

for each $k = 1,\ldots,n-1$; and $t_i^n = \Xi_{i j^*(n)}$ for $i \neq j^*(n)$. ■

A.8 Proposition 6.3 Proof

Proof. Without loss of generality $\alpha := t_1^1 \geq t_2^1 =: \gamma$; and $\delta := t_2^2 \geq t_1^2 =: \beta$. In this case, it is convenient to write down the agent's value function:

$$W(x) = \begin{cases} \alpha(1-x) + \beta x - \kappa c(x), & \text{if} \quad 0 \leq x \leq \frac{\alpha-\gamma}{\alpha-\gamma+\delta-\beta} \\ \gamma(1-x) + \delta x - \kappa c(x), & \text{if} \quad \frac{\alpha-\gamma}{\alpha-\gamma+\delta-\beta} \leq x \leq 1 \end{cases}.$$

Consequently, the equations that pin down the agent's optimal learning simplify to

$$\kappa\left(c'(x_H) - c'(x_L)\right) = A + B \quad \text{and} \quad A + \kappa\left(c(x_H) - c(x_L)\right) = \kappa\left(c'(x_H)x_H - c'(x_L)x_L\right),$$

where $A := \alpha - \gamma \geq 0$, $B := \delta - \beta \geq 0$. Because c is strictly convex, both A and B are strictly positive if $x_L < \mu < x_H$, and zero if $x_L = x_H = \mu$. Furthermore, the concavifying line is

$$f(x) = (\beta - \gamma - A - \kappa c'(x_L))\, x + \gamma + A - \kappa (c(x_L) - x_L c'(x_L)) \ . \tag{$\star$}$$

The principal chooses γ and β in order to maximize

$$-\gamma(1-\mu) - \beta\mu - p x_H B - (1-p)(1-x_L) A \ ,$$

where $p = (\mu - x_L)/(x_H - x_L)$ is the (unconditional) probability that posterior x_H realizes, subject to limited liability: $\beta, \gamma \geq 0$, and

$$f(x) \geq v_0 - \kappa c(x) \quad \text{for all} \quad x \in [0,1] \ , \tag{IR-v_0}$$

where F is given in Equation $\star$. By construction, the agent cannot deviate profitably by learning differently *and* reporting to the principal. Constraint IR-v_0 ensures that the agent cannot deviate profitably by learning differently and taking his outside option.

Using the concavifying line ($\star$), $\{x_L, x_H\}$ can be implemented efficiently if and only if

(i) $(\beta - \gamma - A - \kappa c'(x_L))\,\mu + \gamma + A - \kappa(c(x_L) - x_L c'(x_L)) = v_0$; and

(ii) $\beta - \gamma - A - \kappa c'(x_L) = -\kappa c'(\mu)$; and

(iii) $\beta, \gamma \geq 0$.

From (i) and (ii),

$$\gamma = v_0 + \kappa c'(\mu)\mu - A - \kappa(c'(x_L)x_L - c(x_L)) = v_0 + \kappa c'(\mu)\mu - \kappa(c'(x_H)x_H - c(x_H)) \ ,$$

and

$$\beta = v_0 - \kappa(1-\mu)c'(\mu) + \kappa(1 - x_L)c'(x_L) + \kappa c(x_L) \ .$$

(iii) requires $v_0/\kappa \geq \eta(x_L, x_H)$, as stated in the result. ■

A.9 Proposition 6.5 Proof

Proof. (i) is a consequence of Proposition 6.3. Suppose that $v_0/\kappa < \eta(x_L, x_H)$ so that efficient implementation is infeasible. Recall that P wants to maximize $-\gamma(1-\mu)-\beta\mu$. Thus, if $\gamma = \beta = 0$ is implementable, they are obviously optimal. Substituting them into the concavifying line ($\star$) we get

$$h(x) = -(A + \kappa c'(x_L))x + A - \kappa(c(x_L) - x_L c'(x_L)) \ .$$

We need to check for which values of x_L and x_H h lies above $v_0 - \kappa c(x)$. To that end, we define function $g(x) := h(x) - v_0 + \kappa c(x)$. Then,

$$g'(x) = -(A + \kappa c'(x_L)) + \kappa c'(x) \ ,$$

and observe that g is strictly convex in x. Evidently, $g'(0) < 0$, so F is either minimized at $x^\circ = x^\circ(x_L, x_H)$, implicitly defined as $g'(x^\circ) = 0$ (if such an $x \leq 1$ exists) or $x = 1$. Define $x^\dagger := \min\{x^\circ, 1\}$. Thus, $\gamma = \beta = 0$ is optimal if and only if $g\left(x^\dagger\right) \geq 0$. Note that there is a knife-edge case where $v_0/\kappa = \eta(x_L, x_H)$, $x^\dagger = \mu$, and $\beta = \gamma = 0$ (and the first-best is attained). This is the only way for all three constraints to bind.

Can we have one of the non-negativity constraints bind, $\gamma = 0$, say; and the other constraints all be slack, i.e., $\beta > 0$ and $f(x) > v_0 - \kappa c(x)$ for all $x \in [0,1]$? No: otherwise the principal could decrease β by a sufficiently small $\varepsilon > 0$, strictly increasing her payoff and still leaving Constraint IR-v_0 satisfied. This yields (ii)a and (ii)b of the result. ■

A.10 Proposition 6.6 Proof

Proof. Simply rearrange the inequality $f(\mu) \geq v_0$ to get

$$\gamma \geq \frac{v_0 - \kappa((1-\mu)(x_H c'(x_H) - c(x_H)) - \mu((1-x_L)c'(x_H) + c(x_L)))}{1-\mu} - \frac{\mu}{1-\mu}\beta \ ,$$

then set $\beta = 0$ and solve for when the right-hand side of this inequality is positive.

■