MOTIVATING AGENTS TO ACQUIRE INFORMATION∗

Charles Angelucci†

First Version: January 2014
This Version: October 2016

Abstract

It is difficult to motivate advisors to acquire information through standard performance-contracts when outcomes are uncertain and/or only observed in the distant future. Decision-makers may however exploit advisors’ desire to influence the outcome. We build a model in which a decision-maker and two advisors care about the decision’s consequences, which depend on an unknown parameter. To reduce uncertainty, the advisors can produce information of low or high accuracy; this information becomes public but the decision-maker cannot assess its accuracy with certainty, and monetary transfers are unavailable. Moreover, the information the two advisors produce may be conflicting. We show the optimal decision-rule involves making the final decision exceedingly sensitive to the jointly provided information.

Keywords: Information acquisition, experts, communication, advice.

JEL Classification Numbers: D00, D20, D80, D73, P16.

∗I am deeply indebted to Philippe Aghion, Emmanuel Farhi, Oliver Hart, Patrick Rey, and Jean Tirole for their invaluable guidance and support. This paper has also benefited from discussions with Nava Ashraf, Julia Cagé, Wouter Dessein, Alfredo Di Tillio, Matthew Gentzkow, Renato Gomes, Ben Hebert, Christian Hellwig, Daisuke Hirata, Emir Kamenica, Nenad Kos, Thomas Le Barbanchon, Yanxi Li, Lucas Maestri, Eric Mengus, Simone Meraglia, Marco Ottaviani, Nicola Persico, Alexandra Roulet, Andrei Shleifer, Alp Simsek, Kathryn Spier, David Sraer, Eric Van den Steen, and audiences at Harvard, HBS, Yale SOM, Chicago Booth, Columbia, Kellogg, WUSTL, Toulouse, Queen’s, Bocconi, CUNEF, HEC, INSEAD, Rochester, Polytechnique, Oxford, Université Paris Dauphine, and the Petralia Workshop. Opinions and errors are mine.

†Columbia Business School. Email: ca2630@gsb.columbia.edu
1 Introduction

Information acquisition is central to decisions made under uncertainty, yet it often relies on intermediaries. For instance, regulatory agencies benefit from internal staff expertise; judges rely on prosecutors to collect evidence; and managers ask subordinates to estimate projects’ returns. However, delegating information acquisition responsibilities triggers a natural concern about the quality of the evidence upon which the decision will be made.

In many environments, performance-based contracts offer a powerful means ensuring the provision of accurate information from experts (e.g., portfolio managers). In other environments, performance is difficult to measure and so alternative levers must be found. This is particularly true when the appropriateness of decisions becomes apparent only in the distant future (e.g., regulatory actions), if ever (e.g., judicial sanctions).\(^1\) In addition, specifying the desired quality of advice in a contract—to reward the information itself rather than the outcomes—is often prohibitively costly or unenforceable in a court of law (e.g., because too subjective), so that not only is performance difficult to assess, but also transfers are a rather coarse instrument.\(^2\) However, in many cases, a decision-maker may be capable of motivating information-acquisition by exploiting her advisors’ desire to influence the outcome.\(^3\) This alternative channel of influence is the object of this paper.

---

\(^1\) The appropriateness of a past decision may also be hard to measure in absence of clear counterfactuals (e.g., pollution standards and merger decisions).

\(^2\) In many environments, monetary transfers contingent on performance are banned altogether (e.g., regulatory hearings, expert witnesses in courts of law, etc). See Szalay [2005] and Alonso and Matouschek [2008] for discussions of this matter.

\(^3\) An advisor’s desire to influence the outcome may be due to his reputational concerns, or simply because the decision affects him directly. See for instance Prendergast [2007] on the intrinsic motivation of bureaucrats.
In particular, we investigate the issue of how to optimally motivate multiple advisors. Decision-makers often rely on several advisors, and doing so presents advantages and drawbacks whose consequences for the decision-making process we analyze.

A decision-maker interacts with one or two agents before taking an action. All parties are interested in the decision, which depend on some unknown state of nature. For most of the analysis, players agree on which decision is optimal when they have access to the same information. Agents privately acquire one of two informationally-ranked signals, with the more accurate signal being also the costliest one. Agents have access to identical signals and signal realizations are conditionally independent. The information they produce becomes public, but its accuracy remains privately known. The decision-maker wishes to motivate agents to acquire the more accurate signal, but is unable to infer with certainty which signal was acquired by looking at its realization. Moreover, agents have insufficient incentives to acquire the more informative signal when the ex post optimal action is implemented. Monetary transfers are ruled out, and thus the choice of the final action is the only means by which the decision-maker can motivate her agents to each acquire the more informative signal. The decision-maker is capable of committing to a decision rule, which specifies the action to implement as a function of the information produced by the agents. This provides a good approximation of the settings in which decision-makers can restrain their behavior through pre-specified decision-making procedures (e.g., regulatory procedures, evidentiary rules in the judicial system, etc). Finally, the informational environment we consider is such that higher realizations of either signal lead to correspondingly higher desired actions.
We first look at the version of the model with a single agent. In this environment we show that, to induce the acquisition of the more accurate signal, it is optimal to implement a higher (resp. lower) action than the ex post optimal action when observing a realization such that the agent would desire a lower (resp. higher) action under the less accurate signal. Departing from the ex post optimal action in this manner hurts the agent under the more accurate signal, but relatively more so under the less accurate signal. This asymmetric effect occurs because players suffer increasingly more as the implemented action becomes more distant from their desired action.

We then restrict attention to environments in which signals induce schedules of desired actions that are rotations one of another around the status-quo (i.e., the action desired under the prior). In these environments, it is optimal to implement a higher action than desired when desiring an action higher than the status-quo, and a lower action than desired when desiring an action lower than the status-quo. Distorting the final decision away from the status-quo motivates the acquisition of the more informative signal because it punishes the agent whose posteriors are induced by the less informative signal.

We extend the baseline model by incorporating a second agent. Relying on a second agent is potentially valuable because of the additional information he provides, but also because it may allow the decision-maker to better control her agents by comparing their information. However, introducing a second agent also creates difficulties: the information the agents provide may be conflicting. Specifically, the two signal realizations may be such that one agent desires a lower action than the ex post optimal action when deviating and acquiring the less informative signal, whereas the other agent desires a higher
action than the ex post optimal action when deviating and acquiring the less informative signal. As a consequence, departing from the ex post optimal action necessarily dampens one agent’s incentive to acquire the more informative signal. This phenomenon occurs because each agent, upon unilaterally deviating and acquiring the less informative signal, puts more weight on the other agent’s signal realization than on his own. We show it is optimal for the decision-maker to introduce a distortion that targets the agent whose signal realization is the most extreme. This, in turn, implies it is again optimal to implement higher actions than desired when desiring actions higher than the status-quo, and lower actions than desired when desiring actions lower than the status-quo.

The main insight of our model is thus that the decision-maker should make her final decision exceedingly sensitive to the provided information. This, we show, is the optimal way of structuring the final decision in a way that makes it privately interesting for her agents to acquire costly but accurate information.

It is common for organizations to commit to ex post inefficient rules to motivate information acquisition by its members (for instance, through the delegation of decision-rights, the choice of default-actions or standards of proof, and the non-admissibility of certain decisions or relevant information). In the course of the analysis, we discuss mechanisms through which organizations can implement (or approximate) the decision rule characterized in this paper. All involve making the status-quo (and actions near it) relatively hard to implement and the final decision exceedingly sensitive to the provided information.

The rest of the paper is structured as follows. A review of the literature concludes the introduction. Section 2 presents the setup of the model. Section
3 characterizes the optimal decision rule for the baseline model with one agent (Section 3.1); it then turns to the analysis with two agents (Section 3.2). Section 4 provides several extensions and discussions, and Section 5 concludes.

**Related Literature**

This paper contributes to a literature that incorporates information acquisition in principal-agent models without monetary transfers or, more generally, with limited or costly transfers (see, e.g., Prendergast [1993], Aghion and Tirole [1997], Dewatripont and Tirole [1999], Li [2001], Szalay [2005], Gerardi and Yariv [2008b], and Krishna and Morgan [2008]). To the best of our knowledge, our paper is the first to investigate the issue of designing rules that motivate information acquisition by multiple agents in environments of non-transferable utility and principal-commitment. Our paper is closely related to Li [2001] and Szalay [2005], who consider single-agent settings. In both papers, the second-best decision rule is such that the status-quo (and actions near it) are implemented less often than ex post optimal. In Szalay [2005], the decision-maker implements a higher action than recommended when the recommended action is above the status-quo, and vice versa. This “exaggeration” property manifests itself in various settings, including the present paper. It, for instance, also arises in procurement models with information acquisition (see, e.g., Cremer et al. [1998], and Szalay [2009]). Our analysis suggests it

---

4See also Austen-Smith [1994], Che and Kartik [2009], and Argenziano et al. [2016] for models with information acquisition but lack of commitment. See Demski and Sappington [1987] on performance-contracts.

5Multiple experts have been considered in other environments (often without costly information acquisition). See, e.g., Krishna and Morgan [2001], Battaglini [2004], Gentzkow and Kamenica [2016], Gul and Pesendorfer [2012], and Bhattacharya and Mukherjee [2013].
also holds in advisee/multiple advisors settings.

This paper also relates to the literature on experts (see Scharfstein and Stein [1990], Dewatripont et al. [1999], and Holmstrom [1999]), and that on committees (see Gilligan and Krehbiel [1987], Persico [2004], Gerardi and Yariv [2008a], Cai [2009], and Gershkov and Szentes [2009]). Although delegation is not optimal in our setting, our paper is related to the literature on optimal delegation (see, e.g., Melumad and Shibano [1991] and Alonso and Matouschek [2008]). Aspects of our analysis are reminiscent of the early analyses of moral hazard (see Holmstrom [1979], Holmstrom [1982], and Mirrlees [1999]). Finally, our way of modeling information acquisition is related to work by Persico [2000], Athey and Levin [2001], Johnson and Myatt [2006], and Shi [2012].

2 The Setting

A decision-maker, DM, relies on two agents, $A_1$ and $A_2$. DM has payoff $u(a, \omega)$ and $A_i$ has payoff $u(a, \omega) - C(e_i)$, for $i = 1, 2$. Payoffs depend on DM’s action $a \in A \subset \mathbb{R}$ (where $A$ is compact) and the unobservable state of the world $\omega \in \Omega \subseteq \mathbb{R}$, where $\omega$ is the realization of a random variable $W$ with c.d.f. $F_W(\omega)$ (the common prior) on full support $\Omega$. The function $C(e_i)$ represents $A_i$’s cost of exerting effort $e_i$. The decision-maker and the agents have identical gross preferences, but the decision-maker does not internalize the agents’ cost $C(\cdot)$ of acquiring information. We assume homogeneous gross preferences to isolate the consequences of DM’s failure to internalize the agents’ costs.\(^6\) In Section

\(^6\)We could have assumed DM’s payoff is the sum of the agents’ payoffs, i.e., $2u(a, \omega) - \sum_{i=1}^{2} C(e_i)$. This model could capture, for instance, a committee in which the solution to
4, we provide an extension with heterogeneous gross preferences.

We suppose DM is able to commit to a decision rule but unable to make monetary transfers to her agents. For simplicity, we also ignore participation constraints at the agent level. We comment on both assumptions in Section 4.

2.1 Payoff Functions

We restrict attention to payoff functions \( u(a, \omega) \) of the form

\[
u(a, \omega) = l(\omega) + a \cdot (\eta \omega + \kappa) - \psi a^2, \tag{1}\]

where \( l(\omega) : \Omega \to \mathbb{R}, \kappa \in \mathbb{R}, \) and \( \eta, \psi > 0. \) For instance, \( u(a, \omega) = -(\eta \omega + \kappa - a)^2 \) or \( u(a, \omega) = \omega a - \frac{1}{2} a^2. \) Let \( \hat{a}(\omega) \equiv \arg\max_{a \in \mathbb{R}} u(a, \omega) = \frac{n \omega + \kappa}{2 \psi} \) denote the players’ desired action when \( W = \omega. \) Because \( \eta, \psi > 0, \) \( \hat{a}(\omega) \) is strictly increasing in \( \omega. \) We set \( \eta = \psi = 1 \) and \( \kappa = 0 \) because they play no role in the analysis.

First, \( u(a, \omega) \) is strictly concave in \( a \) and symmetric around \( \hat{a}(\omega), \) that is, \( u(\hat{a}(\omega) + \epsilon, \omega) = u(\hat{a}(\omega) - \epsilon, \omega), \forall \epsilon. \) Second, the players’ desired action when there exists uncertainty about \( W \) depends on the expected value of \( W \) only. Finally, a player’s expected payoff under a given belief about \( W \) is a translation of this player’s expected payoff under any other belief about \( W. \)

2.2 Effort and Information Acquisition

\( A_i \) unobservably chooses effort \( e_i \in E \subset \mathbb{R}, \) which determines the acquisition of signal \( X_{e_i} \) from a set of signals \( \{X_e\}_{e \in E} \) identical for both agents. We

\( DM’s \) problem is the optimal aggregation rule. This approach is qualitatively identical to ours because it shares the essential element whereby each agent’s net benefit from acquiring information is lower than \( DM’s, \) and the key insights our model generates therefore apply.
denote by $G_{X_e|W}(x|\omega)$ signal $X_e$’s conditional c.d.f., where $x$ denotes a realization. All signals have full, connected, and common support $\mathcal{X} \subseteq \mathbb{R}$. Agents’ signal realizations are conditionally independent, and $x_i$ denotes $A_i$’s signal realization. Also, $G_{X_e}(x) = \mathbb{E}_W[G_{X_e|W}(x|\omega)] = \int_{\Omega} G_{X_e|W}(x|\omega) dF_W(\omega)$ represents the marginal distribution of $X_e$, with associated density function $g_{X_e}(x)$.

Let $x \equiv [x_1, x_2]$ and $G_{X_{e_1},X_{e_2}}(x) \equiv G_{X_{e_1}}(x_1) G_{X_{e_2}}(x_2)$.

We assume effort levels to be unobservable, but signal realizations $x_1$ and $x_2$ to be publicly observable. Each agent produces public information of privately known and unverifiable accuracy.\(^7\) This setting captures, for instance, a situation in which both $DM$ and the agents are experts, but $DM$ is unable or unwilling to verify every aspect of the agents’ information gathering activities.\(^8\) If effort levels/signals were observable, the players’ desired action when observing $X_{e_1} = x_1$ and $X_{e_2} = x_2$ would be:

$$\hat{a}_{X_{e_1},X_{e_2}}(x) \equiv \arg\max_{a \in \mathbb{R}} \int_{\Omega} u(a, \omega) dF_{W|X_{e_1},X_{e_2}}(\omega | x) = \mathbb{E}_{W|X_{e_1},X_{e_2}}(x),$$

where Bayes’ rule applies. We assume signals are such that $\hat{a}_{X_{e_1},X_{e_2}}(x_1, x_2)$ is nondecreasing in each signal realization, $\forall X_{e_1}, X_{e_2}$.\(^9\) For any signals, players wish to match higher signal realizations with higher actions. We refer to $\hat{a}_{X_{e_1},X_{e_2}}(x)$ as players’ schedule of desired actions under signals $X_{e_1}$ and $X_{e_2}$.

---

\(^7\)DM cannot infer which signal was acquired from observing its realization due to the common support assumption. Suppose an agent is in charge of data collection and analysis. Although DM can replicate the analysis if given the data, she may lack the resources to verify the representativeness of the sample, and cannot infer it from observing the output.

\(^8\)Assuming observable realizations but unobservable signals allows us to model partially verifiable information, which provides a reasonable approximation of settings in which agents cannot fully manipulate information because, for instance, subject to procedures (or because the decision-maker is also an expert).

\(^9\)To ensure this property, it is enough, for instance, to restrict attention to signals that induce conditional densities for $W$ that satisfy MLRP.
For most of the analysis, we set \( E = \{ e, \overline{e} \} \) and let \( \underline{X} \equiv X_e \) and \( \overline{X} \equiv X_{\overline{e}} \). Agents choose one signal from two available signals. In Section 4, we provide an extension with continuous effort. We use the subscript \( X_i \) to represent a given signal acquired by \( A_i \), and the subscript \( \underline{X}_i \) (resp. \( \overline{X}_i \)) to mean that \( A_i \) acquired signal \( \underline{X}_i \) (resp. \( \overline{X}_i \)). For instance, \( \hat{a}_{\underline{X}_i, \overline{X}_j} (x) \) denotes the players’ schedule of desired actions when \( A_i \) acquires signal \( \underline{X}_i \) and \( A_j \) acquires signal \( \overline{X}_j \). When indexes are omitted, the pair \( \{ X, X' \} \) always denotes \( \{ X_1, X_2 \} \).

We assume signal \( \underline{X} \) is more informative than signal \( \overline{X} \), in the sense that, for any \( X_j \), DM’s highest possible expected payoff when \( A_i \) acquires \( \underline{X} \),

\[
\int_{X \times X} \int_{\Omega} u \left( \hat{a}_{\underline{X}_i, X_j} (x), \omega \right) dF_{W \mid \underline{X}_i, X_j} (\omega \mid x) dG_{\underline{X}_i, X_j} (x),
\]

is higher than DM’s highest possible expected payoff when \( A_i \) acquires \( \overline{X} \),

\[
\int_{X \times X} \int_{\Omega} u \left( \hat{a}_{\overline{X}_i, X_j} (x), \omega \right) dF_{W \mid \overline{X}_i, X_j} (\omega \mid x) dG_{\overline{X}_i, X_j} (x),
\]

where \( i, j = 1, 2 \) and \( i \neq j \). We set \( C (e) = 0 < C (\overline{e}) = c \). Acquiring the more informative signal is more costly to the agents.

### 2.2.1 Rotations

We sometimes focus on environments that generate schedules of desired actions that are rotations one of another, where the schedule induced by a more informative pair of signals is a steeper rotation of the schedule induced by a less informative pair of signals around some “rotation-point.” To fix ideas, suppose there exists one agent only, and let \( \hat{a}_X (x) \) denote the players’ desired action under signal \( X \): the schedule of desired actions \( \hat{a}_X (x) \) is a steeper rotation of
the schedule of desired actions $\hat{a}_X(x)$ around some rotation-point $\tilde{x}$ if

$$(\hat{a}_X(x) - \hat{a}_X(x)) (x - \tilde{x}) > 0, \quad \forall x \neq \tilde{x}. \quad (5)$$

Figure 1 provides an illustrative example (where $\tilde{x} = 0$). As discussed below, many commonly assumed informational environments generate rotations.

We refer to the realizations higher (resp. lower) than the rotation-point $\tilde{x}$ as high (resp. low) realizations. The desired action when observing a high (resp. low) realization is higher (resp. lower) under signal $X$ than signal $\hat{X}$.

Insert Figure 1 about here.

In what follows, we sometimes focus on the more stringent case in which the conditional expectation of $W$ is a convex combination of the prior and the signal realization $x$, where the weights are independent of $x$, that is, $\hat{a}_X(x) = \mathbb{E}_{W|X}[\omega | x] = \beta x + (1 - \beta) \omega_0$, and $\hat{a}_X(x) = \mathbb{E}_{W|X}[\omega | x] = \beta x + (1 - \beta) \omega_0$, with $0 \leq \beta < \bar{\beta} \leq 1$. One verifies $\hat{a}_X(x)$ is a steeper rotation of $\hat{a}_X(x)$ around $\tilde{x} \equiv \omega_0$, where $\omega_0 = \mathbb{E}_W[\omega] = \int_{\Omega} \omega dF_W(\omega)$. Signal realizations higher (resp. lower) than $\omega_0$ are high (resp. low). Moreover, under either signal, the desired action when observing $x > \omega_0$ (resp. $x < \omega_0$) is higher (resp. lower) than the action desired under the prior: the status-quo (denoted $a_0$, where $a_0 = \omega_0$).

Two agents. For tractability, in the two-agent case, we focus exclusively on environments that generate schedules that are rotations one of another. Consider a given signal $X_j$ and realization $x_j$. Schedule $\hat{a}_{X_i,X_j}(x_i, x_j)$ is a

---

10 This terminology is borrowed from Athey and Levin [2001], who offer a notion of informativeness that builds upon the distinction between low and high realizations.
steeper rotation of schedule \( \hat{a}_{\overline{x}, x_j} (x_i, x_j) \) around rotation-point \( \tilde{x}_i (x_j, \omega_0) \) if
\[
\left( \hat{a}_{\overline{x}, x_j} (x_i, x_j) - \hat{a}_{\overline{x}, x_j} (x_i, x_j) \right) (x_i - \tilde{x}_i (x_j, \omega_0)) > 0, \quad \forall x_i \neq \tilde{x}_i (x_j, \omega_0). \quad (6)
\]
Specifically, with two agents, we always require the prior belief and the signals to generate expressions for the conditional expectation of \( W \) that are a convex combination of the prior and the signal realizations, where the weights are independent of the signal realizations, that is
\[
\begin{align*}
E_{W \mid X_i, X_j} [\omega \mid x_i, x_j] &= \alpha (x_i + x_j) + (1 - 2\alpha) \omega_0, \quad (7) \\
E_{W \mid \overline{X}, X_j} [\omega \mid x_i, x_j] &= \alpha x_i + \alpha x_j + (1 - \alpha - \alpha) \omega_0, \quad (8)
\end{align*}
\]
where \( 0 \leq \underline{\alpha} < \alpha < \overline{\alpha} \leq 1 \). Equal weight \( \alpha \) is put on signals of equal informativeness, more weight is put on \( \overline{X} \) than \( X \) when both signals are acquired (i.e., \( \overline{\alpha} > \alpha \)), more weight is put on signal \( \overline{X} \) when the other acquired signal is \( X \) rather than \( \overline{X} \) (i.e., \( \overline{\alpha} > \alpha \)), and the weight put on signal \( \overline{X} \) is lower than the weight put on signal \( \overline{X} \) when the other acquired signal is \( \overline{X} \) (i.e., \( \underline{\alpha} < \alpha \)).

There exist many informational environments that generates conditional means that can be expressed as (7)-(8). Suppose, for example, that \( W \sim N \left( 0, \frac{1}{h} \right) \), \( \overline{X} = W + \overline{\epsilon} \), and \( X = W + \epsilon \), where \( \overline{\epsilon} \) and \( \epsilon \) are independently distributed and such that \( \overline{\epsilon} \sim N \left( 0, \frac{1}{\overline{h}} \right) \) and \( \epsilon \sim N \left( 0, \frac{1}{h} \right) \). Suppose also \( \frac{1}{h} < \frac{1}{\overline{h}} \). Then,
\[
\begin{align*}
E_{W \mid \overline{X}, X_j} [\omega \mid x_i, x_j] &= \frac{h x_i + \overline{\epsilon} x_j}{h + 2\overline{h}} \quad \text{and} \quad E_{W \mid \overline{X}, \overline{X}_j} [\omega \mid x_i, x_j] = \frac{h x_i + \overline{\epsilon} x_j}{h + 2\overline{h}}.
\end{align*}
\]
Signals whose associated expressions for the conditional expectation of \( W \) are given by (7)-(8) generate schedules of desired actions that are rotations one of another. In particular, if \( A_j \) acquires signal \( \overline{X} \), then \( \hat{a}_{\overline{X}, \overline{X}_j} (x_i, x_j) \) is a
steeper rotation of $\hat{a}_{X_i, X_j} (x_i, x_j)$ around rotation point

$$\bar{x}_i (x_j, \omega_0) = \left( \frac{\alpha - \alpha}{\alpha - \alpha} \right) x_j + \left( \frac{2\alpha - \alpha - \alpha}{\alpha - \alpha} \right) \omega_0.$$

Building on the Normal distribution example, $\hat{a}_{X_i, X_j} (x_i, x_j) = \frac{\bar{\pi}(x_i + x_j)}{h+2h}$ is a steeper rotation of $\hat{a}_{X_i, X_j} (x_i, x_j) = \frac{h x_i + \pi x_j}{h + h + h}$ around $\bar{x}_i (x_j, \omega_0) = \frac{\pi}{h+h} x_j$.

Exactly as with a single agent, we refer to the realizations $x_i$ higher (resp. lower) than $\bar{x}_i (x_j, \omega_0)$ as high (resp. low) realizations, where again the desired action when observing a high (resp. low) realization is higher (resp. lower) under signal $\tilde{X}$ than under signal $X$. However, observe $\bar{x}_i (x_j, \omega_0)$ is a function of $x_j$. With two agents, whether $x_i$ is high/low depends on the value $x_j$ takes. In particular, $x_i$ is high if and only if $x_j$ is sufficiently low. Figure 2 depicts $A_2$’s schedules when $x_2$ is high (b) and low (c), assuming $X_1 = \tilde{X}$.

Insert Figure 2 about here.

**LR Condition.** In the two-agent case, in which conditional means are a convex combination of the produced information and the prior, we suppose the likelihood ratio $LR (x) \equiv \frac{g_X(x)}{g_\tilde{X}(x)}$ is weakly increasing in $|x - \omega_0|$ and symmetric (i.e., $LR (\omega_0 + \epsilon) = LR (\omega_0 - \epsilon)$, $\forall \epsilon$). Notice the action $A_i$ desires when observing $X_i = \omega_0$ coincides, $\forall X_i$, with the action he desires under the prior, that is, $E_{W|\tilde{X}} [\omega | \omega_0] = E_{W|X} [\omega | \omega_0] = \omega_0$. Assuming $LR (x)$ is increasing in $|x - \omega_0|$ amounts to assuming that realizations that are increasingly distant from $\omega_0$—that is, realizations that are increasingly “extreme”—become increasingly more likely under $X$ than $\tilde{X}$.

$^{11}$This assumption imposes enough “monotonicity” for us to characterize a simple decision rule in the two-agent case. In some sense, it is the analogue of the MLRP assumption for the
Heuristically, more noisy information is produced when the agents shirk.\textsuperscript{12}

Many environments (i) satisfy the LR condition and (ii) generate schedules that are convex in the prior and the realizations. For instance, these conditions are generally satisfied if the joint distribution of $W$, $X$, and $\bar{X}$ belongs to the class of Elliptical distributions, which includes the Normal distribution.\textsuperscript{13} In the course of the analysis, we sometimes provide explicit solutions assuming that $W \sim N(0, 1)$, $X = \rho W + \sqrt{1 - \rho^2} \xi$, and $\bar{X} = \overline{\rho} W + \sqrt{1 - \overline{\rho}^2} \tau$, with $0 \leq \rho < \overline{\rho} \leq 1$ and where $\tau$ and $\xi$ are i.i.d. and such that $\xi, \tau \sim N(0, 1)$.\textsuperscript{14}

### 2.3 Decision Rules and Timing

We characterize the optimal deterministic decision rule $a^*(x_1, x_2) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that ensures the existence of an equilibrium in which $A_1$ and $A_2$ acquire signal $\bar{X}$. The timing is as follows. First, $DM$ announces and commits to $a(x_1, x_2)$. Second, each agent $A_i$ privately chooses his signal $X_i \in \{X, \bar{X}\}$. Third, the realizations $(x_1, x_2)$ are publicly observed and, fourth, $DM$ implements $a(x_1, x_2)$.\textsuperscript{15}

\textsuperscript{12}In our model, agents behave “opportunistically” by shirking, which, for most of the analysis will mean providing less accurate information. However, one could imagine other forms of opportunism. For instance, agents may be tempted to choose signals that fool $DM$ into not revising its prior (see Prendergast [1993] on a related issue). Our approach would naturally extend to these alternative scenarios but, for reasons of space, we leave these considerations for future research.

\textsuperscript{13}See Deimen and Szalay [2016] and references therein.

\textsuperscript{14}Joint distributions that do not belong to the Elliptical class can also satisfy our conditions. This is the case, for instance, if $W \sim U[0, 1]$, $X = W + \xi$, and $\bar{X} = W + \tau$, with $\xi$ and $\tau$ independently distributed and such that $\xi \sim N(0, \frac{1}{2})$ and $\tau \sim N(0, \frac{1}{2})$, with $\overline{h} > h$ and $\mathcal{X}(\omega) = [\omega - \nu, \omega + \nu]$. As an additional example, consider also $W \sim N\left(\frac{1}{2}, \frac{1}{3}\right)$ on $\Omega = [0, 1]$, with $\bar{X} \mid W = W$ and $X \mid W \sim U[0, 1]$.

\textsuperscript{15}Analyzing alternative timings (in which, for instance, the agents gather their signals sequentially) is beyond the scope of this paper. On this issue, see Gerardi and Yariv [2008b].
3 The Model

3.1 The Single-Agent Case

3.1.1 The General Case

Suppose $DM$ relies on a single agent $A$, and let $\pi(x) \equiv \hat{a}_\pi(x)$ and $a(x) \equiv \hat{a}_X(x)$. To rule out trivial cases, we assume $c$ is sufficiently large that $A$ has insufficient incentives to acquire signal $X$ if $DM$ implements the first-best decision rule $\pi(x)$. Formally, we assume that, under the decision rule $\pi(x)$, $A$’s expected payoff when gathering signal $X$,

$$
\Pi(\pi) - c \equiv \int_X \int_\Omega u(\pi(x), \omega) dF_{W|X}(\omega | x) dG_X(x) - c,
$$

is lower than his expected payoff when gathering signal $X$,

$$
\Pi(a) \equiv \int_X \int_\Omega u(a(x), \omega) dF_{W|X}(\omega | x) dG_X(x).
$$

When $c > \Pi(\pi) - \Pi(a)$, $DM$ must depart from the first-best to motivate $A$ to acquire $X$. By contrast, inducing the acquisition of $X$ can straightforwardly be achieved by implementing $a(x)$. As a result, inducing the acquisition of $X$ is desirable only if the informational gain outweighs the payoff losses due to the inefficiencies necessary to motivate $A$. We now characterize the nature of these inefficiencies, that is, we characterize the “second-best” decision rule.

Suppose $DM$ wishes $A$ to acquire signal $X$. She then solves

$$
\max_{a(x) \in Q(x)} \int_X \int_\Omega u(a(x), \omega) dF_{W|X}(\omega | x) dG_X(x),
$$

when $\ Pi(\pi) - \Pi(a)$.
subject to inducing $A$ to acquire signal $X$ rather than signal $\bar{X}$:

$$
\int_{\mathcal{X}} \int_{\Omega} u(a(x),\omega) \, dF_{W|\bar{X}}(\omega \mid x) \, dG_{\bar{X}}(x) - c \geq 0
$$

$$
\int_{\mathcal{X}} \int_{\Omega} u(a(x),\omega) \, dF_{W|X}(\omega \mid x) \, dG_{X}(x).
$$

To ensure the existence of a solution, we restrict $a(x)$ to belong to $Q(x) = [a_l(x), a_h(x)], \forall x$.\(^{16}\) We also assume (i) $a_l(x) = \bar{a}(x) - q(x)$ and $a_h(x) = \bar{a}(x) + q(x)$, with $q(x) > 0, \forall x$, and (ii) $\underline{a}(x) \in Q(x), \forall x$.\(^{17}\) The optimal decision rules under $\bar{X}$ and $\bar{X}$ are feasible. Because $c > \Pi(\bar{a}) - \Pi(\bar{a})$, we anticipate (13) binds at the optimum and set up a Lagrangian, where $\lambda$ is the multiplier associated with (13). Also, we assign the multipliers $\gamma_l(x)$ and $\gamma_h(x)$ to $a(x) \geq a_l(x)$ and $a(x) \leq a_h(x)$, respectively. $DM$ chooses $a(x)$ to maximize pointwise:

$$
\int_{\mathcal{X}} L(x,a) \, dx = \int_{\mathcal{X}} \left( \int_{\Omega} u(a(x),\omega) \, dF_{W|\bar{X}}(\omega \mid x) \, g_{\bar{X}}(x) \right)
$$

\(^{16}\)Let $TV^b_{\bar{a}}(a)$ be the total variation of function $a$ in $[b, b']$. We assume that for some sufficiently large interval $[\underline{a}, \bar{a}] \subseteq \mathcal{X}$, there exists at least one function $a_0(x)$ such that $a_0(x) \in [a_l(x), a_h(x)], \forall x \in \mathcal{X}$, with $TV^b_{\underline{a}}(a_0) < \infty$, and such that

$$
\int_{[\underline{a}, \bar{a}]} \int_{\Omega} u(a_0(x),\omega) \, dF_{W|\bar{X}}(\omega \mid x) \, dG_{\bar{X}}(x) - c - \int_{[\underline{a}, \bar{a}]} \int_{\Omega} u(a_0(x),\omega) \, dF_{W|X}(\omega \mid x) \, dG_{X}(x) >
$$

$$
\int_{[\underline{a}, \bar{a}]} \int_{\Omega} u(a(x),\omega) \, dF_{W|\bar{X}}(\omega \mid x) \, dG_{\bar{X}}(x) - \int_{[\underline{a}, \bar{a}]} \int_{\Omega} u(a(x),\omega) \, dF_{W|X}(\omega \mid x) \, dG_{X}(x), \forall a(x).
$$

One can prove the existence of a solution by restricting attention to functions $a$ (i) such that $a \in L^1(\mathcal{X})$ and (ii) such that, for any $[b, b'] \subseteq \mathcal{X}$, and some parameter $H > \frac{TV^b_{\bar{a}}(a_0)}{(\bar{\theta} - \underline{\theta})}$, $TV^b_{\bar{a}}(a) \leq H (\bar{\theta} - \underline{\theta})$ (i.e., we restrict attention to functions with a uniformly bounded total variation). We assume $H$ to be large enough for our class of functions to include all functions of economic interest that satisfy (i) and (ii).

\(^{17}\)We assume the bounds $a_l(x)$ and $a_h(x)$ depend on $x$ to avoid corner solutions in some examples provided below. The assumption is inessential for our main results, and may be replaced with $a(x) \in [a_l, a_h]$. Also, we suppose the bounds are equidistant from $\bar{a}(x)$ to shorten some proofs. However, this assumption can be relaxed without affecting our results.
\[
\gamma g_X(x) \int_\Omega u(a(x),\omega) dF_{W\mid X}(\omega \mid x) - c
\]

\[
- g_X(x) \int_\Omega u(a(x),\omega) dF_{W\mid X}(\omega \mid x)
\]

\[
+ g_X(x) \left( \gamma_l(x)(a(x) - a_l(x)) + \gamma_h(x)(a_h(x) - a(x)) \right) dx.
\]

Let \( K \equiv \{ x \in \mathcal{X} : \bar{a}(x) = a(x) \} \) be the set of all realizations such that the desired actions under signals \( \mathcal{X} \) and \( \bar{X} \) coincide. Similarly, let \( K^+ \equiv \{ x \in \mathcal{X} : \bar{a}(x) > a(x) \} \) be the set of high realizations and \( K^- \equiv \{ x \in \mathcal{X} : \bar{a}(x) < a(x) \} \) be the set of low realizations. Also, let \( \tilde{\lambda} \equiv \frac{1 + \lambda}{\lambda} \). We now state the relationship between the first- and second-best decision rules.

The proof of Proposition 1, and most other proofs, are in the Appendix.

**Proposition 1** Suppose DM wishes A to acquire signal \( \mathcal{X} \). Then, it is optimal to implement, for almost every \( x \), a decision rule \( a^*(x) \) such that:

1. \( a^*(x) \in (\bar{a}(x), a_h(x)) \) if \( x \in K^- \),

2. \( a^*(x) \in [a_l(x), \bar{a}(x)) \) if \( x \in K^+ \),

3. \( a^*(x) = \bar{a}(x) \) if \( x \in K \) and \( \tilde{\lambda} \geq LR(x) \), and either \( a^*(x) = a_l(x) \) or \( a^*(x) = a_h(x) \) if \( x \in K \) and \( \tilde{\lambda} < LR(x) \).

Although DM does not observe which signal A acquires, she knows, for any realization \( x \), which action A desires both in case he has acquired signal \( \mathcal{X} \) and in case he has acquired signal \( \bar{X} \). As a result, for all realizations such that the desired action differs depending on which signal was acquired (i.e., \( \forall x \notin K \)), DM implements an action that departs from the first-best \( \bar{a}(x) \) in the direction opposite to the action \( a(x) \) desired under signal \( \bar{X} \). DM thus implements
higher actions than desired when observing \textit{high realizations}, and lower actions than desired when observing \textit{low realizations}. These distortions induce the acquisition of $\bar{X}$ because, even though they hurt both players in equilibrium, they would hurt $A$ relatively more so if the latter was to deviate and acquire $X$. This asymmetric effect occurs because players suffer increasingly more as the implemented action becomes more distant from their desired action.\footnote{Specifically, departing from $\bar{\pi}(x)$ in the opposite direction than $a(x)$ involves a second-order loss to $DM$ and $A$ in equilibrium, but a first-order loss to $A$ in case he deviates.}

When the realization $x$ is such that $A$ desires the same action regardless of the acquired signal’s accuracy (i.e., when $x \in K$), the logic highlighted above breaks downs: distortions are ex post equally damaging to $A$ under both signals. However, $DM$ chooses an action other than the first-best $\bar{a}(x)$ in case $g_X(x)$ is sufficiently larger than $g_{\bar{X}}(x)$, that is, in case the likelihood ratio $LR(x)$ is sufficiently large. Heuristically, in these instances, departing from the first-best hurts—\textit{in expectation}—much more severely $A$ in case he deviates and acquires signal $X$ than $DM$ and $A$ under signal $\bar{X}$, and $DM$ thus introduces as large a distortion as possible (by implementing either $a_l(x)$ or $a_h(x)$).

\textbf{Corollary 1} When $x \notin K$, all else equal, the size of the distortion $|a^*(x) - \bar{a}(x)|$ is weakly increasing in both $|\bar{\pi}(x) - a(x)|$ and $LR(x)$.

Because players suffer increasingly more as the chosen action becomes more distant from their desired action, the effectiveness of a departure from the first-best in punishing $A$ in case he acquires $X$ increases with the distance between $\bar{a}(x)$ and $a(x)$. Further, the likelihood ratio $LR(x) = \frac{g_X(x)}{g_{\bar{X}}(x)}$ determines how effective distortions are \textit{in expected terms}, so that $DM$ introduces larger distortions for the realizations with a higher likelihood ratio.
3.1.2 Rotations

We suppose that, under either signal, the conditional expectation of $W$ can be expressed as a convex combination of the prior and the realization $x$. Recall $\pi(x)$ is then a steeper rotation of $a(x)$ around $\tilde{x} = \omega_0$. Recall also $a_0$ denotes the action desired under the prior, that is, $a_0$ denotes the status-quo. In this environment, $\pi(\omega_0) = a(\omega_0) = a_0$, and $\pi(x), a(x) > a_0$ if $x > \omega_0$ (and conversely).

**Corollary 2** Suppose DM wishes $A$ to acquire signal $X$. Then, it is optimal to implement, for almost every $x$, a decision rule $a^*(x)$ which is a rotation of the schedule $\pi(x)$ around $\omega_0$, that is,

$$
(a^*(x) - \pi(x))(x - \omega_0) > 0, \quad \forall x \neq \omega_0.
$$

The final decision is made exceedingly sensitive to the provided information, so that implementing an action close to $a_0$ is less likely than under the first-best.

**Proof.** By assumption, $(\pi(x) - a(x))(x - \omega_0) = (\beta - \beta)(x - \omega_0)^2 > 0, \forall x \neq \omega_0$. It follows from Proposition 1 that $(a^*(x) - \pi(x))(x - \omega_0) > 0, \forall x \neq \omega_0$. 

The decision rule is such that, compared to the first-best, it is less likely that an action close to $a_0$ is implemented, which hurts $A$ in case he acquires signal $X$, because the actions he desires are are then closer to the status-quo (i.e., $\bar{\beta} > \beta$). Notice DM does not make it less likely to implement an action close to the status-quo by introducing random noise in the final decision. Instead, she makes her final decision exceedingly sensitive to the provided information: DM implements higher actions than desired when the first-best action is above

---

\textsuperscript{19} Formally, $\Pr[(a^*(x) - a_0) > \epsilon] > \Pr[|\pi(x) - a_0| > \epsilon], \forall \epsilon > 0$. 

19
\(a_0\), and conversely when the first-best action is below \(a_0\). In other words, \(DM\) overreacts to the provided information in order to motivate her agent.

### 3.1.3 An Explicit Solution

Suppose \(u(a, \omega) = - (\omega - a)^2\), \(W \sim N(0, 1)\), \(\mathbf{X} = \rho W + \sqrt{1 - \rho^2} \epsilon\), and \(\overline{X} = \overline{\rho W + \sqrt{1 - \overline{\rho}^2} \epsilon}\), where (i) \(0 \leq \rho < \overline{\rho} \leq 1\) and (ii) \(\epsilon \sim N(0, 1)\). Here, \(W | X = x \sim N(\rho x, 1 - \rho^2)\), so that \(\overline{a}(x) = \overline{\rho} x\) and \(q(x) = \rho x\). To limit the number of cases, suppose \(\overline{\rho} > \sqrt{\rho}\). Also, \(\forall x, q(x)\) is sufficiently large that \(a^*(x) \in \text{int} \, \mathbb{Q}(x)\).

**Proposition 2** Suppose \(DM\) wishes \(A\) to acquire signal \(\overline{X}\). If \(c \leq 2 \overline{\rho} (\overline{\rho} - \rho)\), \(DM\) implements, for almost every \(x\), the first-best decision rule \(a^*(x) = \overline{\rho} x\). Otherwise, for almost every \(x\), \(DM\) implements:

\[
a^* (x) = \left( \frac{c}{2 (\overline{\rho} - \rho)} \right) x, \quad \text{where} \quad \frac{c}{2 (\overline{\rho} - \rho)} > \overline{\rho}.
\]

When the cost to \(A\) of acquiring signal \(\overline{X}\) is large, \(DM\) must depart from the first-best decision rule and adopt a decision rule which is a steeper rotation of the first-best rule \(\overline{a}(x)\) around 0. Specifically, \(DM\) puts excessive weight on the provided information (compared to the first-best), thereby making it less likely that an action close to the status-quo is implemented. Also, notice the weight put on \(A\)'s information is increasing in the cost \(c\) and decreasing in the informational superiority of signal \(\overline{X}\), as measured by \(\overline{\rho} - \rho\).\(^{20}\)

\(^{20}\)Recall that—when \(c > \Pi (\overline{\pi}) - \Pi (\pi)\)—inducing the acquisition of \(\overline{X}\) may not be optimal because of the distortions. In this example, one computes it is optimal to induce the acquisition of \(\overline{X}\) when \(c \leq 2 (\overline{\rho} - \rho) \left( \overline{\rho} + \sqrt{\overline{\rho}^2 - \rho} \right)\).
Corollary 3 The variance $\mathbb{V}_X[a^*(x)]$ of the optimal decision rule $a^*(x)$ is larger than the variance $\mathbb{V}_X[\overline{a}(x)]$ of the first-best decision rule $\overline{a}(x) = \rho x$.

Proof. $\mathbb{V}_X[a^*(x)] = \int_X (a^*(x) - \mathbb{E}_X[a^*(x)])^2 dG_X(x) > \mathbb{V}_X[\overline{a}(x)] = \int_X (\overline{a}(x) - \mathbb{E}_X[\overline{a}(x)])^2 dG_X(x)$ because (i) $\mathbb{E}_X[a^*(x)] = \mathbb{E}_X[\overline{a}(x)] = 0$ and (ii) $(a^*(x) - \overline{a}(x))(x - \tilde{x}) > 0, \forall x \neq \tilde{x}$. $lacksquare$

In our setting, signal $X$ is more informative than signal $X'$ if and only if $\mathbb{V}_X[\hat{a}_X(x)] > \mathbb{V}_X'[\hat{a}_X'(x)]$, that is, if and only if the spread of the desired actions under signal $X$ is higher than that under signal $X'$. Because the second-best decision rule $a^*(x)$ is a rotation of the first-best decision rule $\overline{a}(x)$, and because the expected value of both decision rules is the same, $\mathbb{V}_X[a^*(x)] > \mathbb{V}_X[\overline{a}(x)]$. In order to motivate her agent, $DM$ behaves as if the provided information’s underlying accuracy is higher than what it actually is. In other words, $DM$ behaves as if overconfident about the information she is provided with.

3.2 The Two-Agent Model

Suppose $DM$ now relies on two agents, $A_1$ and $A_2$, where each may either acquire signal $X$, at zero cost, or signal $\overline{X}$, at cost $c$. Recall our focus on environments in which conditional expectations of $W$ are a convex combination of the signal realizations $(x_1,x_2)$ and the prior $\omega_0$. Relying on a second agent is potentially valuable because of the additional information, but also because it may allow $DM$ to better control agents by comparing their information.

Our objective is to characterize the optimal decision-rule that ensures the existence of an equilibrium in with both agents acquire signal $\overline{X}$.

\footnote{As in any multiple-agent setting, a concern exists about equilibrium multiplicity. How-}
ticular, we wish to investigate whether and how the main intuition of the single-agent case—i.e., that $DM$ makes her final decision excessively sensitive to the provided information so as to make it harder to implement actions close to $a_0$—survives with two agents. We do not seek conditions under which motivating both agents to acquire signal $\bar{X}$ is optimal (as opposed to, for instance, relying on one agent, or relying on two agents but requesting them to acquire different signals). Computing closed-form conditions would require first specifying many of our primitives.\textsuperscript{22} Intuitively, requesting both agents to acquire $X$ will be optimal whenever the informational gain outweighs the necessary distortions. We provide conditions that ensure motivating both agents to gather $\bar{X}$ is optimal when solving for specific examples.

Suppose $DM$ wishes both agents to acquire signal $\bar{X}$. She then solves

$$\max_{a(x_1,x_2)\in\mathcal{Q}(x_1,x_2)} \int_{X\times X} \int_{\Omega} u(a(x),\omega) dF_{W|\bar{X},\bar{X}}(\omega|X) dG_{\bar{X},\bar{X}}(x),$$

subject to inducing $A_i$ to acquire signal $\bar{X}$ rather than signal $\bar{X}$:

$$\Pi_{\bar{X},\bar{X}}(a) - c := \int_{X\times X} \int_{\Omega} u(a(x),\omega) dF_{W|\bar{X},\bar{X}}(\omega|X) dG_{\bar{X},\bar{X}}(x) - c \geq (17)$$

$$\Pi_{\bar{X},\bar{X}}(a) := \int_{X\times X} \int_{\Omega} u(a(x),\omega) dF_{W|\bar{X},\bar{X}}(\omega|X) dG_{\bar{X},\bar{X}}(x),$$

for $i = 1, 2$ and $j \neq i$. To ensure the existence of a solution, we restrict

\textsuperscript{22}Note that our analysis is a necessary step in $DM$’s larger optimization problem. To analyze whether, for instance, relying on two agents dominates relying on a single agent, one must necessarily compute the two associated optimal decision rules first.

\textsuperscript{22}Note that our analysis is a necessary step in $DM$’s larger optimization problem. To analyze whether, for instance, relying on two agents dominates relying on a single agent, one must necessarily compute the two associated optimal decision rules first.
(\text{TV} (a_1(X), a_2(X)) = \text{TV} (a_1(\hat{x}), a_2(\hat{x})) + \text{TV} (a_1(\bar{x}), a_2(\bar{x})) - 2 \text{TV} (a_1(\hat{x}), a_2(\bar{x}))).

\text{(TV) is the total variation of function } a \text{ in rectangle } [a_1, a_2] \times [\hat{x}, \bar{x}].

We assume that for some sufficiently large rectangle \([a_1, a_2] \times [\hat{x}, \bar{x}]\), there exists at least one function \(a_0(x)\) such that \(a_0(x) \in \langle a_l(x), a_h(x) \rangle, \forall x \in X \times \hat{x} \times \bar{x}\), with TV \((a_0, I_\theta^2) < \infty\), and such that \(\Pi_{X_i, X_j} (a_0) - \Pi_{\bar{x}, \hat{x}} (a_0) \geq c\) for \(i = 1, 2\). One can prove the existence of a solution by restricting attention to functions \(a\) such that \(a \in L^1(X^2)\) and \(a \in L^1(\hat{x}, \bar{x})\), and some parameter \(H \geq TV (a_0, I_\theta^2), H \geq TV (a, I_\theta^2)\).

\text{Again, we assume the bounds of } Q(x) \text{ depend on } x \text{ to avoid corner solutions in some of the examples provided below. One may assume } a \in Q \text{ without qualitatively affecting our main results. Similarly, we assume } a_l(x) \text{ and } a_h(x) \text{ are equidistant from } a \text{ to shorten some of the proofs. The assumption can be relaxed without affecting our main results.}
unilateral deviation, and conversely for the other agent.

The existence of such polarized out-of-equilibrium desired actions makes it unclear whether and how the intuition of the single-agent case—i.e., that DM departs from the first-best in the direction opposite to what the (single) agent desires under signal $\bar{X}$—survives with two agents. Figure 2 depicts both agents’ schedules of desired actions under signals $\bar{X}$ and $\bar{X}$, holding fixed the other agent’s signal realization (and assuming it acquired $\bar{X}$). Figure 2 ((a)-(b)) considers a situation with congruent information in which $x_1$ and $x_2$ are both high. Figure 2 ((c)-(d)) considers a situation with conflicting information in which $x_1$ is high whereas $x_2$ is low.

### 3.2.1 Costless Public Signal

To distinguish between the modifications to the second-best decision rule (compared to the single-agent case) that occur because of the presence of a second signal from those that stem from having a second agent to motivate, take it as given, for the moment, that $A_j$ acquires signal $\bar{X}$. Equivalently, suppose there exists a costless public signal $\bar{X}_j$.\(^{25}\) We maintain all other assumptions. To rule out trivial cases, we suppose DM wishes $A_i$ ($i \neq j$) to acquire signal $\bar{X}$ and assume $c$ is large enough that $A_i$ has insufficient private incentives to gather $\bar{X}$ in case the first-best $\hat{a}_{\bar{X}_i,\bar{X}_j}(x_i, x_j)$ is implemented.\(^{26}\)

We denote by $a^*_{i,\bar{X}_j}(x)$ the solution to DM’s problem when $A_i$’s effort constraint binds and signal $\bar{X}_j$ is observed. Let $K(x_j) \equiv \{x_i \in \mathcal{X} : a_{\bar{X}_i,\bar{X}_j}(x_i, x_j) = a^*_{X_i,\bar{X}_j}(x_i, x_j)\}$, that is, let $K(x_j)$ be the set of re-

\(^{25}\)We set $X_j = \bar{X}$ for notational convenience only. Our results are identical if $X_j = \bar{X}$.

\(^{26}\)Formally, we suppose $c > \Pi_{\bar{X}_i,\bar{X}_j}(\hat{a}_{\bar{X}_i,\bar{X}_j}) - \Pi_{\bar{X}_i,\bar{X}_j}(\check{a}_{\bar{X}_i,\bar{X}_j})$. 

24
alizations $x_i$ such that $A_i$ desires the same action under $X$ and $\bar{X}$, for a given $\bar{X}_j = x_j$. Also, $K^+ (x_j) \equiv \{ x_i \in X : a^i_{\bar{X}_i, \bar{X}_j} (x_i, x_j) > a^i_{X_i, \bar{X}_j} (x_i, x_j) \}$, $K^- (x_j) \equiv \{ x \in X : a^i_{\bar{X}_i, \bar{X}_j} (x_i, x_j) < a^i_{X_i, \bar{X}_j} (x_i, x_j) \}$, and $\lambda_i = \frac{1 + \lambda_j}{\lambda_i}$. $K^+ (x_j)$ (resp. $K^- (x_j)$) is the set of high (resp. low) realizations, for a given $\bar{X}_j = x_j$. From (9), we know $x_i \in K^+ (x_j)$ if $x_i > \hat{x}_j (x_i, \omega_0)$ and $x_i \in K^- (x_j)$ if $x_i < \hat{x}_j (x_i, \omega_0)$.

**Proposition 3** Suppose DM wishes $A_i$ to acquire signal $\bar{X}$. Then, it is optimal to implement, for almost every $(x_i, x_j)$, a decision rule that specifies an action $a^*_i, \bar{X}_j (x)$ such that:

1. $a^*_i, \bar{X}_j (x) \in \left( \hat{a}^i_{\bar{X}_i, \bar{X}_j} (x), a_h (x) \right)$ if $x_i \in K^+ (x_j)$,

2. $a^*_i, \bar{X}_j (x) \in \left[ a_l (x), \hat{a}^i_{\bar{X}_i, \bar{X}_j} (x) \right)$ if $x_i \in K^- (x_j)$,

3. $a^*_i, \bar{X}_j (x) = \hat{a}^i_{\bar{X}_i, \bar{X}_j} (x)$ if $x_i \in K (x_j)$ and $\lambda_i \geq LR (x_i)$, and either $a^*_i, \bar{X}_j (x) = a_l (x)$ or $a^*_i, \bar{X}_j (x) = a_h (x)$ if $x_i \in K (x_j)$ and $\lambda_i < LR (x_i)$.

For all joint realizations $(x_i, x_j)$, DM chooses an action that departs from the first-best (i.e., $\hat{a}^i_{\bar{X}_i, \bar{X}_j} (x)$) in the direction opposite to the action $A_i$ desires when deviating and acquiring signal $X$ (i.e., $\hat{a}^i_{X_i, \bar{X}_j} (x)$). As in the single-agent case, this policy motivates information acquisition because it ensures the decision rule is a sufficiently better match with the posterior beliefs jointly induced $\{ \bar{X}_i, \bar{X}_j \}$ than with those jointly induced by $\{ X_i, \bar{X}_j \}$.

In the single-agent case without additional signal, DM compared $x$ to the prior $\omega_0$ to determine the sign of the distortion. For instance, DM would introduce an upward distortion when $x > \omega_0$. Here, instead, DM compares $x_i$ to the rotation-point $\hat{x}_i (x_j, \omega_0)$ induced by $X_j$. Thus, $a^*_i, \bar{X}_j (x) > \hat{a}^i_{\bar{X}_i, \bar{X}_j} (x)$ if
$x_i > \tilde{x}_i (x_j, \omega_0)$ and $a^*_{i, \overline{X}} (x) < \hat{a}_{\overline{X}, \overline{X}} (x)$ if $x_i < \tilde{x}_i (x_j, \omega_0)$.

Therefore, $DM$ exploits signal $\overline{X}_j$ not only because of its informational content (notice $a^*_{i, \overline{X}} (x)$ is a function of $\hat{a}_{\overline{X}, \overline{X}} (x)$), but also to better control $A_i$.

Further, unlike the single-agent case without additional signal, $DM$ does not necessarily introduce an upward distortion when the first-best is above the status-quo (i.e., when $\hat{a}_{\overline{X}, \overline{X}} (x) > a_0$), or a downward distortion when the first-best is below the status-quo (i.e., when $\hat{a}_{\overline{X}, \overline{X}} (x) < a_0$). For some realizations, $DM$ may instead implement an action that lies closer to $a_0$ than the first-best. To see this, suppose, for instance, that $\frac{1}{2} (x_i + x_j) > \omega_0$, which in our setting implies $\hat{a}_{\overline{X}, \overline{X}} (x) > a_0$ (see (7)-(8)). In words, suppose $(x_i, x_j)$ is such that, in equilibrium, the desired action is above the status-quo. If the realizations are also such that $x_i < \tilde{x}_i (x_j, \omega_0)$, that is, if $x_i$ is low given $x_j$, the final decision $a^*_{i, \overline{X}}$ is lower than $\hat{a}_{\overline{X}, \overline{X}} (x)$ and thus possibly closer to $a_0$. More generally, joint realizations $(x_1, x_2)$ such that either “$\frac{1}{2} (x_i + x_j) > \omega_0$ and $x_i < \tilde{x}_i (x_j, \omega_0)$” or “$\frac{1}{2} (x_i + x_j) < \omega_0$ and $x_i > \tilde{x}_i (x_j, \omega_0)$” occur with positive probability and may generate a decision-rule that does not make it less likely to implement actions close to the status-quo compared to the first-best.

**Explicit Solution.** Suppose $u (a, \omega) = - (\omega - a)^2$, $W \sim N (0, 1)$, and let $\overline{X} = \epsilon$ and $\overline{X} = \overline{p}W + \sqrt{1 - \overline{p}}\epsilon$, where $0 < \overline{p} < 1$ and $\epsilon \sim N (0, 1)$. To avoid corner solutions, we assume that, $\forall (x_i, x_j)$, $q (x)$ is large enough that $a^* (x) \in \text{int } Q (x)$. $A_i$’s schedule of desired actions under $\{\overline{X}_i, \overline{X}_j\}$, $\hat{a}_{\overline{X}, \overline{X}} (x) = \frac{p}{1 + p} (x_i + x_j)$, is a

---

27 Formally, $\forall x_j$, $a^*_{i, \overline{X}} (x_i, x_j)$ is a steeper rotation of $\hat{a}_{\overline{X}, \overline{X}} (x_i, x_j)$ around rotation-point $\tilde{x}_i (x_j, \omega_0)$. Notice $\tilde{x}_i (x_j, \omega_0)$ is random from an ex ante perspective.

28 Similarly to Corollary 1, one can also show that, $\forall x_j$, if $x_i \notin K (x_j)$, the size of the distortion $|a^*_{i, \overline{X}} (x) - \hat{a}_{\overline{X}, \overline{X}} (x)|$ is weakly increasing in $|x_i - \tilde{x}_j (x_i, \omega_0)|$ and LR ($x$).
steeper rotation of \( \hat{a}_{\bar{X}_i, X_j} = \bar{p} x_j \) around rotation-point \( \tilde{x}_i (x_j, \omega_0) = \bar{p}^2 x_j \). Notice the first-best decision rule can be expressed as a function of the sum of the signal realizations, that is, \( (x_i + x_j) \) is a sufficient statistic for \( (x_i, x_j) \).

Suppose DM wishes \( A_i \) to acquire signal \( \bar{X} \) and suppose also \( c > \Pi_{\bar{X}_i, X_j} (\hat{a}_{\bar{X}_i, X_j}) - \Pi_{\bar{X}_i, X_j} (\hat{a}_{\bar{X}_i, X_j}) = \frac{2p^2}{(1+p^2)^2} \). Solving DM’s problem yields

\[
a^{*}_{i, \bar{X}_j} (x) = \hat{a}_{\bar{X}_i, X_j} (x) + \lambda^* (c) \left( \frac{\bar{p}}{1 + \bar{p}^2} \right) (x_i - \bar{p}^2 x_j),
\]

(18)

where \( \lambda^* (c) = \frac{1}{\sigma_{X_j}^2 - 3} \left( \frac{1}{\bar{p}} \sqrt{c (1 + \bar{p}^2)^2 (5\bar{p}^4 - 3) + \bar{p}^2 (7 - 8\bar{p}^4 + \bar{p}^6) - (1 + \bar{p}^4)} \right) \).

The second-best decision rule \( a^{*}_{i, \bar{X}_j} (x) \) is a steeper rotation of \( \hat{a}_{\bar{X}_i, X_j} (x) \) around \( \tilde{x}_i (x_j, \omega_0) = \bar{p}^2 x_j, \forall x_j \). Also, \( (x_i - \bar{p}^2 x_j) \) determines both the sign and size of the distortion. Finally, notice the optimal decision rule (18) cannot generally be rewritten as a function of \( (x_1 + x_2) \). This occurs because DM uses \( X_j \) not only for its informational content, but also to better motivate \( A_i \).

### 3.2.2 Two Agents to Motivate

Suppose now DM must motivate both agents to acquire signal \( \bar{X} \). Let \( \lambda_i \) denote the multiplier associated to (17), for \( i = 1, 2 \). To focus on the interesting case, assume \( c \) is sufficiently large that—when \( \hat{a}_{\bar{X}, X_j} (x_1, x_2) \) is implemented—no agent has private incentives to acquire signal \( \bar{X} \) (when anticipating the other agent is acquiring \( \bar{X} \)). Similarly, we also assume \( c \) is sufficiently large that \( A_j \) (\( j = 1, 2 \)) acquires signal \( \bar{X} \) if \( a^{*}_{i, \bar{X}_j} (x) \) is implemented.\(^{29}\) Finally, we assign the multipliers \( \gamma_l (x) \) and \( \gamma_h (x) \) to \( a(x) \geq a_l (x) \) and \( a(x) \leq a_h (x) \), respectively.

\(^{29}\)Formally, we set \( c > \max \left[ \Pi_{\bar{X}_i, X_j} (\hat{a}_{\bar{X}_i, X_j}) - \Pi_{\bar{X}_i, X_j} (\hat{a}_{\bar{X}_i, X_j}), \Pi_{\bar{X}_i, X_j} (a^{*}_{i, \bar{X}_j}) - \Pi_{\bar{X}_i, X_j} (a^{*}_{i, \bar{X}_j}), \Pi_{\bar{X}_i, X_j} (a^{*}_{i, \bar{X}_j}) - \Pi_{\bar{X}_i, X_j} (a^{*}_{i, \bar{X}_j}) \right] \). When this inequality holds, (17) binds (i.e., \( \lambda_i > 0 \)) for \( i = 1, 2 \).
DM chooses \( a(x) \in \mathbb{Q}(x) \) to maximize \( \int_{\mathcal{X} \times \mathcal{X}} L(x, a) \, dx \) pointwise, where

\[
L(x, a) = \left( g_{\mathcal{X}, \mathcal{X}}(x) \int_{\Omega} u(a(x), \omega) \, dF_{W|\mathcal{X}, \mathcal{X}}(\omega | x) \right) \\
+ \sum_{i=1}^{2} \lambda_{i} \left( g_{\mathcal{X}}(x_{i})g_{\mathcal{X}}(x_{2}) \int_{\Omega} u(a(x), \omega) \, dF_{W|\mathcal{X}, \mathcal{X}}(\omega | x) - c \\
- g_{\mathcal{X}}(x_{i})g_{\mathcal{X}}(x_{j}) \int_{\Omega} u(a(x), \omega) \, dF_{W|\mathcal{X}, \mathcal{X}}(\omega | x) \right) \\
+ g_{\mathcal{X}, \mathcal{X}}(x) \left( \gamma_{l}(x)(a(x) - a_{l}(x)) + \gamma_{h}(x)(a_{h}(x) - a(x)) \right) \right).
\]

(19)

The next proposition compares the first- and second-best decision rules. In what follows, let \( \tilde{\lambda} = \frac{1+2\lambda}{X} \).

**Proposition 4** Suppose DM wishes both agents to acquire signal \( \mathcal{X} \). Then, \( \lambda_{i}^* = \lambda^* \), for \( i = 1, 2 \), and it is optimal to implement, for almost every \((x_{1}, x_{2})\), a decision rule \( a^*(x) \) such that:

1. \( a^*(x) \in \left( \hat{a}_{\mathcal{X}, \mathcal{X}}(x), a_{h}(x) \right] \) if \( \hat{a}_{\mathcal{X}, \mathcal{X}}(x) > a_{0} \),

2. \( a^*(x) \in \left[ a_{l}(x), \hat{a}_{\mathcal{X}, \mathcal{X}}(x) \right) \) if \( \hat{a}_{\mathcal{X}, \mathcal{X}}(x) < a_{0} \), and

3. \( a^*(x) = \hat{a}_{\mathcal{X}, \mathcal{X}}(x) \) if \( \hat{a}_{\mathcal{X}, \mathcal{X}}(x) = a_{0} \) and \( \tilde{\lambda} \geq \frac{2}{\sum_{i=1}^{2} LR(x_{i})} \), and either \( a^*(x) = a_{l}(x) \) or \( a^*(x) = a_{h}(x) \) if \( \hat{a}_{\mathcal{X}, \mathcal{X}}(x) = a_{0} \) and \( \tilde{\lambda} < \frac{2}{\sum_{i=1}^{2} LR(x_{i})} \).

For all joint realizations \((x_{1}, x_{2})\), DM implements an action that departs from the first-best \( \hat{a}_{\mathcal{X}, \mathcal{X}}(x_{1}, x_{2}) \) in the direction opposite to the status quo \( a_{0} \). Therefore, DM implements higher actions than desired when the first-best action is above \( a_{0} \), and lower actions than desired when the first-best
action is below \( a_0 \). Exactly as in the single-agent case, \( DM \) makes it less likely to implement actions close to \( a_0 \). This feature arises despite the occurrence of polarized out-of-equilibrium desired actions, and despite \( DM \)'s desire to compare the provided information so as to better control her agents.

When information is congruent—for instance, when both realizations are high—the analysis is identical to the single-agent case. Both agents would desire an action lower than the first-best upon unilaterally deviating, so that \( DM \) introduces an upward distortion. The distortion hurts all players in equilibrium, but it would hurt the agents relatively more so if they were to deviate.

However, when information is conflicting, introducing, say, an upward distortion will be useful in terms of incentive provision for the agent whose realization is high, since his desired action when deviating is lower than the first-best. However, for the other agent, whose realization is low, introducing an upward distortion will dampen his incentives, because it means implementing an action closer to his desired action when deviating. When faced with this trade-off, \( DM \) introduces a distortion that targets the agent whose signal realization is the most “extreme” (i.e., the farthest away from \( \omega_0 \)). The reason for this is twofold. First, given the focus on an environment in which schedules of desired actions are linear in the produced information and rotations one of another, the agent whose signal realization is the most extreme is the agent’s whose desired action is the most sensitive to the accuracy of his own signal. Specifically, if \(|x_i - \omega_0| > |x_j - \omega_0|\), then

\[
\left| \hat{a}_{X_i,X_j}(x) - \hat{a}_{X_i,X_j}(x) \right| > \left| \hat{a}_{X_i,X_j}(x) - \hat{a}_{X_i,X_j}(x) \right|. \tag{30}
\]

Because agents suffer increasingly more as the implemented action is farther from their ideal, and ignoring likelihood ratios for now, it pays the most to

\[\text{This inequality is derived using (7)-(8).}\]
introduce a distortion targeting the agent whose signal realization is the most extreme. In addition, because more extreme realizations become increasingly more likely under signal $X$ than $\bar{X}$—that is, because $|x_i - \omega_0| > |x_j - \omega_0|$ implies $LR(x_i) \geq LR(x_j)$, distortions become increasingly effective at providing incentives (in “expected terms”), which reinforces the first effect.\footnote{Removing the assumption whereby (i) $LR(x)$ is increasing in $|x - \omega_0|$ and (ii) $LR(\omega_0 + \epsilon) = LR(\omega_0 - \epsilon)$, $\forall \epsilon$, would mean that, in case of conflicting information, $DM$ targets the most extreme realization if and only if its associated likelihood ratio is sufficiently high. She otherwise targets the less extreme realization. As a result, whether $a^* > \hat{a}_{\bar{X},X}(x)$ (resp. $a^* < \hat{a}_{\bar{X},X}(x)$) when $\hat{a}_{\bar{X},X}(x) > a_0$ (resp. $\hat{a}_{\bar{X},X}(x) < a_0$) would be ambiguous.}

Unlike the single-agent case with an additional signal, here $DM$ always implements an action that lies farther away from the status quo than the first-best. As a result, actions close to $a_0$ are unambiguously less likely under the second-best decision rule. In our setting, when realizations are congruent and, say, high, then it is necessarily the case that $\hat{a}_{\bar{X},X}(x) > a_0$, so that $DM$ straightforwardly implements $a^*(x) > \hat{a}_{\bar{X},X}(x)$.\footnote{Formally, using (7)-(8), one shows $x_1 \geq \tilde{x}_1(x_2)$ and $x_2 \geq \tilde{x}_2(x_1)$ together imply $\frac{1}{2} (x_1 + x_2) > \omega_0$, and thus $\hat{a}_{\bar{X},X}(x) \geq a_0$.} When information is conflicting, say $x_1$ is high (i.e., $x_1 \geq \tilde{x}_1(x_2)$) and $x_2$ is low (i.e., $x_2 \leq \tilde{x}_2(x_1)$), the optimal decision rule targets the most extreme realization. Thus, in our example, if $|x_1 - \omega_0| > |x_2 - \omega_0|$, $DM$ implements an upward distortion. However, the most extreme realization determines whether $\frac{1}{2} (x_1 + x_2) \leq \omega_0$, that is, the most extreme realization determines whether $\hat{a}_{\bar{X},X}(x) \leq a_0$.\footnote{Because (i) $|x_1 - \omega_0| > |x_2 - \omega_0|$ and (ii) $x_1$ is high whereas $x_2$ is low, it must be the case that either $x_1 > x_2 > \omega_0$ or $x_1 > \omega_0 > x_2$.} Thus, again in our example, $|x_1 - \omega_0| > |x_2 - \omega_0|$ implies that $\frac{1}{2} (x_1 + x_2) > \omega_0$, and the sign of $(a^*(x) - \hat{a}_{\bar{X},X}(x))$ is necessarily the same as the sign of $(\hat{a}_{\bar{X},X}(x) - a_0)$.

$DM$ is limited in its ability to control her agents by comparing their information because of the occurrence of conflicting information.
is conflicting, the agents’ out-of-equilibrium desired actions are polarized and DM’s distortion can only target one agent at most.

**Explicit Solution.** Suppose \( u(a, \omega) = - (a - \omega)^2 \), \( W \sim N(0,1) \), \( X = \rho W + \sqrt{1 - \rho^2} \epsilon \), and \( \bar{X} = \epsilon \), where \( \epsilon \sim N(0,1) \). To shorten expressions, suppose also \( \rho = \frac{2}{3} \) and \( c \leq \frac{1}{2} \).\(^{34}\) One shows implementing the first-best \( \hat{a}_{X,Y}(x_1, x_2) = \frac{6}{13} (x_1 + x_2) \) fails to motivate the agents to gather signal \( X \) if \( c > \frac{72}{109} \).\(^{35}\) When \( c \in \left[ \frac{72}{109}, \frac{1}{2} \right] \), the second-best decision rule is given by

\[
 a^*(x) = \left( \frac{6}{13} + \lambda^*(x) \frac{56}{117} \right) (x_1 + x_2),
\]

where \( \lambda^*(c) = \frac{27}{224} \left( 5 - 13\sqrt{1 - 2c} \right) \). DM’s associated expected payoff is \( \frac{1}{4} \left( -9 + 5\sqrt{1 - 2c} + 13c \right) \). Notice the decision rule (20) is a function of the sufficient statistic \( (x_1 + x_2) \), unlike expression (18). This occurs because DM is unable to control her agents by comparing the information they provide. Notice also \( V[a^*(x)] > V[\bar{a}(x)] \) again holds: DM behaves as if the jointly provided information were more accurate than what it actually is.

Recall that it may not always be optimal to induce both agents to gather signal \( \bar{X} \). We now provide conditions specific to this example which ensure it is. When \( c \leq \frac{72}{109} \), the first-best is implementable so that relying on two agents is clearly optimal (DM’s expected payoff is then \( -\frac{5}{13} \)). When \( c > \frac{72}{109} \), relying on two agents guarantees a higher payoff than relying on a single-agent if and only if \( c \leq \frac{4}{1521} \left( 142 + 15\sqrt{10} \right) \).\(^{36}\) When \( c \in \left[ \frac{4}{1521} \left( 142 + 15\sqrt{10} \right), \frac{1}{2} \right] \), DM

\[^{34}\] We also assume that, \( V(x_i, x_j), q(x) \) is large enough that \( a^*(x) \in \text{int} Q(x) \).

\[^{35}\] Implementing \( a^*_{i,X_j}(x_1, x_2) \) also fails to motivate \( A_j \) to gather signal \( X \) when \( c > \frac{72}{109} \).

One computes \( a^*_{i,X_j}(x_1, x_2) = \frac{6}{13} (x_i + x_j) + \frac{81}{163} \left( \frac{97}{81} - \frac{3}{2} \sqrt{\frac{143260}{59049} - \frac{27547}{6561}} \right) \frac{6}{13} (x_i - \frac{4}{9} x_j) \).

\[^{36}\] Because in this example \( \rho = 0 \), the single-agent case coincides with the two-agent case.
optimally relies on a one agent and her expected payoff is $-\frac{5}{9}$.\textsuperscript{37}

4 Extensions

We extend the single-agent case to allow for both continuous effort and heterogeneous preferences. We also comment on two of our modeling assumptions, and discuss ways for organizations to implement our decision-rule.

4.1 The Continuous Effort Case

Let $u(a, \omega) = -(a - \omega)^2$. Suppose $A$ exerts effort $e \in E \equiv [0, \overline{e}]$, and incurs a cost $c(e)$ ($c_e(e) > 0$ and $c_{ee}(e) > 0$). To avoid corner solutions, let $c_e(0) = 0$ and $c_e(\overline{e}) = +\infty$. Also, suppose $A$’s effort does not affect the marginal distribution $G(x)$ but only the posterior beliefs $F_W(\omega | x, e)$, where $F_W(\omega | x, e)$ is twice differentiable with respect to $e$. Moreover, for any $e' > e$, we assume:

$$
(\hat{a}(x, e') - \hat{a}(x, e))(x - \bar{x}) > 0, \quad \forall x \neq \bar{x}.
$$

(21)

Finally, for simplicity, suppose (i) $\omega_0 = 0$ and (ii) $\hat{a}(\bar{x}, e) = 0$, $\forall e$. DM chooses $a(x)$ and $e \in [0, \overline{e}]$ to maximize her expected payoff:

$$
-\int_X \int_\Omega (\omega - a(x))^2 f(\omega | x, e) \, d\omega dG(x),
$$

(22)

subject to:

$$
-\int_X \int_\Omega (\omega - a(x))^2 f(\omega | x, e) \, d\omega dG(x) - C(e) \geq 0
$$

(23)

in which agents gather signals of different accuracy.

\textsuperscript{37}In the single-agent case, the effort constraint binds if and only if $c > \frac{8}{9}$. 32
\[- \int_X \int_\Omega (\omega - a(x))^2 f(\omega | x, e') \, d\omega dG(x) - C(e') ,\]

for $\forall e' \in [0, \pi]$. We replace the global constraint (23) by the local constraint:\textsuperscript{38}

\[- \int_X \int_\Omega (\omega - a(x))^2 f_e(\omega | x, e) \, d\omega dG(x) = C_e(e) , \tag{24}\]

which is a necessary condition for an interior effort $e$ to constitute an optimum.

We first characterize the solution to the relaxed problem, and then derive sufficient conditions for this solution to also satisfy (23). Let the multiplier of (24) be $\lambda$. Pointwise maximization with respect to $a(x)$ yields:

\[ a(x) = \int_\Omega \omega f(\omega | x, e) \, d\omega + \lambda \int_\Omega \omega f_e(\omega | x, e) \, d\omega , \quad \forall x , \tag{25}\]

and the first-order condition associated to $e$ is:

\[- \int_X \int_\Omega (\omega - a(x))^2 f_e(\omega | x, e) \, d\omega dG(x) \tag{26}\]

\[ + \lambda \left( - \int_X \int_\Omega (\omega - a(x))^2 f_{ee}(\omega | x, e) \, d\omega dG(x) - C_{ee}(e) \right) = 0 . \]

From (25), because $\lambda \geq 0$, $a(x)$ is positive if $x > \tilde{x}$, equal to zero if $x = \tilde{x}$, and negative if $x < \tilde{x}$. Provided $\lambda > 0$, $a(x)$ is a steeper rotation of $\hat{a}(x, e)$ around $\tilde{x}$. We now show $\lambda > 0$. $A$’s expected payoff under $a(x)$ and effort $e$ is:

\[ - \text{Var}_W[\omega] - \int_X a(x)^2 dG(x) + 2 \int_X a(x) \int_\Omega \omega f_W(\omega | x, e) \, d\omega dG(x) - C(e) . \tag{27}\]

Differentiating (27) twice with respect to $e$ yields:

\[ 2 \int_X a(x) \mathbb{E}_{ee} [\omega | x, e] dG(x) - C_{ee}(e) . \tag{28}\]

It is sufficient that (i) $E_{ee} [\omega | x, e] < 0$ if $x > \bar{x}$ and (ii) $E_{ee} [\omega | x, e] > 0$ if $x < \bar{x}$ for (28) to be negative; that is, for $A$’s expected payoff to be concave. Suppose this condition holds. Finally, note the first term in equation (26) is positive, by virtue of condition (24) and the fact that $c_e (e) > 0$. Because the expression within brackets in equation (26) is negative, $\lambda > 0$ necessarily.

Because $A$’s expected payoff is concave for the solution of the relaxed problem, (24) is sufficient to characterize the incentive constraints. The first-order conditions of the original problem are the same as those of the relaxed problem, and $a^*(x)$ and $e^*$ are characterized by equations (24), (25), and (26).

Proposition 5 The solution to the relaxed problem is the solution to the general problem if $\int_\Omega \omega f_{ee} (\omega | x, e) d\omega < 0 \ (> 0) \ \forall x > \bar{x} \ (\forall x < \bar{x})$, in which case the optimal decision rule, for a given $e^*$, is such that:

1. $a^*(x) < \hat{a}(x, e^*)$ if $\hat{a}(x, e^*) < a_0$,

2. $a^*(x) = \hat{a}(x, e^*)$ if $\hat{a}(x, e^*) = a_0$, and

3. $a^*(x) > \hat{a}(x, e^*)$ if $\hat{a}(x, e^*) > a_0$.

When effort is continuous, the decision-maker finds it optimal to make the final decision excessively sensitive to the information provided by the agent.

4.2 Heterogeneous Preferences

Take the framework of Section 3.1.3 and let $DM$’s payoff be $u (a, \omega, \eta, \kappa) = - (\eta \omega + \kappa - a)^2$, with $\kappa > 0$. Not only has $DM$ a different sensitivity to the underlying state of nature relative to $A$, she is also upward biased.
**Proposition 6** Suppose \( c \leq 2\eta\bar{p}(\bar{p} - \rho) \). Implementing \( a^*(x) = \eta\bar{p}x + \kappa \) is optimal. Suppose instead \( c > 2\eta\bar{p}(\bar{p} - \rho) \). The optimal decision rule is then:

\[
a^*(x) = \left(\frac{c}{2(\bar{p} - \rho)}\right) x + \kappa,
\]

where \( \frac{c}{2(\bar{p} - \rho)} > \eta\bar{p} \).

Compared to the solution in Section 3.1.3, the second-best decision rule when \( c \) is sufficiently high is shifted upwards by an amount equal to \( \kappa \).

### 4.3 Monetary Transfers

Our environment was one of non-transferable utility. This begs the question of whether the main insights are preserved if transfers contingent on the provided information are feasible. Observe that if the agents’ payoffs do not depend on the action \( a \), and if transfers contingent on \( x \) are possible, our analysis becomes a standard moral hazard problem. Unlike the one-dimensional decision \( a \) in our setting, in a such a hypothetical environment, \( DM \) could offer individual schedules of transfers. \( DM \) would tend to make high payments when observing realizations with a low likelihood ratio \( LR(x) \), and low payments when observing realizations with a high likelihood ratio. Also, \( DM \) would make each agent’s transfers contingent on both agents’ signal realizations to exploit the underlying common uncertainty (see Holmstrom [1982]).

However, if monetary transfers are possible and the agents are also affected by \( a \) (and protected by limited liability), \( DM \) should use both channels (money and distortions to the decision-rule) to induce the acquisition of accurate information. Making her final decision exceedingly sensitive to the provided
information would allow $DM$ to reduce expected payments.$^{39}$

4.4 Participation Constraints

Our analysis assumed the agents' participation was guaranteed. This is a natural assumption when agents cannot escape the consequences of the decision. However, extending the model to allow $DM$ to make flat payments would, on the one hand, not contradict the view that it is difficult in practice to design a contract specifying transfers as a function of the provided information and, on the other, allow the former to ensure the agents' participation. Because a flat payment does not impact the agents' incentives to acquire information, the main features of our decision-rules would remain unchanged.$^{40}$

4.5 Organizational Decision-Making

It is common for organizations to commit to ex post inefficient rules to foster information acquisition by its members (e.g., through the delegation of decision rights or the choice of admissible actions, default-actions, and standards of proof).$^{41}$ Organizations can implement (or approximate) the decision-rule characterized in this paper in several ways. One approach is for organizations to put excessive weight on the information produced by their own staff to force the latter to internalize the importance of producing accurate information. Choosing a weight to put on certain types of information is routinely


$^{40}$Our results are qualitatively unaffected when payments are infeasible and the decision-rule must also guarantee agents' participation (as long a solution continues to exist). A proof of this robustness check for the single-agent case is available upon request.

done by administrations which must commit to transparent and predictable
decision-making procedures (e.g., US Sentencing Guidelines). Organizations
can also affect the pros and cons of the possible decisions they face
by manipulating enactment costs. In our setting, it is optimal to raise the
enactment costs associated to the status-quo, and to lower those associated
to actions different from the status-quo. An alternative approach consists
of dismissing or garbling already available information (e.g., reputation, past
history, etc). By restricting the role played by the prior, the newly produced
information becomes more pivotal. For instance, one rationale often put forth
to justify the non-admissibility of certain relevant pieces of evidence in the
inquisitorial system of law (e.g., criminal records) is that it gives incentives to
the prosecutor to find better information. More generally, it has been docu-
mented by organizational sociologists that decision-makers within firms ignore
relevant information or have poor “organizational memory”, (see for instance
Feldman and March [1981] and Walsh and Ungson [1991]). Our paper provides
one possible efficiency rationale: by disregarding available information, firms

---

42 The Sentencing Guidelines are rules that establish a uniform sentencing policy for de-
defendants (individuals or organizations) convicted in the federal court system. The guidelines
operate as a formula that maps facts (nature of conduct, harm, etc) into a fine or jail sen-
tence. For antitrust violations, for instance, the harm to consumers (often estimated by the
staff of the Department of Justice) is taken into account. Conceivably, the weight put on
such estimates of harm affects the motivation of the staff in charge of its production.

43 This logic can rationalize the often-criticized tendency by administrations (or, more
generally, by decision-makers) to over-rely on (or be over-confident about) their expertise.

44 See Stephenson [2011] on enactment costs.

45 See Lester et al. [2012] on a related matter.

46 Similarly, an administration such as the US Food and Drug Administration can com-
mmit to a policy of ignoring other jurisdictions’ publicly observable administrative decisions
(say, from Europe or Canada) when reaching their own decisions. Among many other con-
sequences, such a policy mechanically makes the information provided by the FDA’s staff
more pivotal. A similar reasoning holds for the FDA’s policy of not taking into account
Pharmaceutical companies’ track records when determining whether to market a new drug.
motivate employees by making their (better) information pivotal. Finally, organizations can delegate decision-rights to less experienced managers—with a more diffused prior—who over-react to the information provided by their more experienced subordinates.\textsuperscript{47}

5 Concluding Remarks

This paper has investigated a situation in which a decision-maker exploits her agents’ desire to influence the outcome to motivate information acquisition. The main insight that emerges from the single-agent analysis is that the decision-maker should make her final decision excessively sensitive to the provided information. This policy motivates information acquisition because the implemented action is then a very poor match with what the agent desires when he produces low-quality information. A direct consequence of this strategy is that actions near the status-quo become less likely to be implemented. Whether this insight generalizes to more complex organizations that comprise two agents was unclear. With multiple agents, what each agent desires when unilaterally producing low-quality information may be drastically different, so that distorting the final decision away from the status-quo may actually reward one agent for shirking. Despite this difficulty, we showed it is again optimal to make the final decision exceedingly sensitive to the jointly provided information, and to make it hard to implement actions near the status-quo.

\textsuperscript{47}Detailed computations regarding these statements are available upon request.
References


Appendix

Proof of Proposition 1 and Corollary 1

Let $U(a,x) \equiv \int_{\Omega} u(a,\omega) dF_{W|x}(\omega \mid x)$ and $U(a,x) \equiv \int_{\Omega} u(a,\omega) dF_{W|x}(\omega \mid x)$. We ignore the constraints $a(x) \geq a_l(x)$ and $a(x) \leq a_h(x)$, and identify ex post when they hold. From (14), it follows

$$\frac{\partial}{\partial a} L(x,a) = (1 + \lambda) g(x) \frac{\partial}{\partial a} U(a,x) - \lambda g(x) \frac{\partial}{\partial a} U(a,x).$$

Because $\frac{\partial^2}{\partial a^2} u(a,\omega) = -1$, $\text{Sign} \left[ \frac{\partial^2}{\partial a^2} L(x,a) \right] = \text{Sign} \left[ \lambda g(x) - (1 + \lambda) g(x) \right]$. Therefore, $L(x,a)$ is concave if $\tilde{\lambda} \equiv \frac{1 + \lambda}{\lambda} > LR(x) \equiv g(x)$, and otherwise convex.

Case 1: $x \in K^+$. We first show $a^* \geq a$. Suppose $a_l \leq a^* < a$. This cannot be optimal because, for instance, $L(2a - a^*,\omega) > L(a^*,\omega)$, due to (i) $U(2a - a^*,x) \geq U(a^*,x)$ and (ii) $U(2a - a^*,x) = U(a^*,x)$. Suppose now $a^* \in [a,\bar{a}]$. The terms $I$ and $-II$ are strictly positive for $\forall a \in [a,\bar{a}]$, so that $a^* \geq \bar{a}$. To see that $a^* > \bar{a}$, note that, on the one hand, $I(\bar{a}) = 0$, while, on the other, $-II(\bar{a}) > 0$, implying $\frac{\partial}{\partial a} L(x,a) > 0$ when $a = \bar{a}$. Moreover, if $\tilde{\lambda} > LR(x)$, $L(x,a)$ is strictly concave, so that $a^*$ is equal to the minimum between $L(x,a)$’s global maximizer on $\mathbb{R}$ and $a_h$. If $\tilde{\lambda} \leq LR(x)$, $L(x,a)$ is convex and $a^* = a_h$.

Case 2: $x \in K^-$. The proof proceeds through identical steps and is left out.

Case 3: $x \in K$. Because $\bar{a} = a$, $I(\bar{a}) = II(\bar{a}) = 0$. If $\tilde{\lambda} > LR(x)$, $L(x,a)$ is strictly concave and thus $a^* = \bar{a} = a$. If $\tilde{\lambda} < LR(x)$, $L(x,a)$ is strictly convex,
\( \bar{a} = a \) is \( L(x,a) \)'s global minimizer, and \( a^* \) is equal to either \( a_l \) or \( a_h \). Finally, if \( \tilde{\lambda} = LR(x) \), \( L(x,a) \) is flat and \( a^* \in [a_l,a_h] \).

**Corollary 1.** Suppose \( x \in K^+ \). If \( \tilde{\lambda} > LR(x) \), \( L(x,a) \) is concave and \( a^* \) is equal to the minimum between the value of \( a \) that solves \( I - II = 0 \) and \( a_h \). Suppose \( a^* \) is given by \( -\tilde{I}(a^*) = -\tilde{II}(a^*) \). Divide both sides by \( g(x) \), and relabel \( I(a) \) and \( II(a) \) as \( \tilde{I}(a) \) and \( \tilde{II}(a) \). Interpret \( -\tilde{I}(a) \) as the marginal cost of increasing \( a \) starting at \( a = \bar{a} \), and \( -\tilde{II}(a) \) as the marginal benefit. \( -\tilde{I}(a) \) and \( -\tilde{II}(a) \) are increasing in \( a \) for \( a \in [\bar{a},a_h] \). Moreover, (i) \( \tilde{I}(a) \) is independent of \( LR(x) \) and \( |a - \bar{a}| \) on \( [\bar{a},a_h] \) whereas (ii) \( -\tilde{II}(a) \) is higher on \( [\bar{a},a_h] \) the higher \( LR(x) \) and \( |a - \bar{a}| \) are. The greater \( LR(x) \) is, the higher the function \( -\tilde{II} \) is in \( [\bar{a},a_h] \), and thus the weakly higher \( a^* \) is. Similarly, the greater \( |a - \bar{a}| \) is, the higher the function \( -\tilde{II} \) is in \( [\bar{a},a_h] \), so that the weakly higher \( a^* \) is. Similarly, if \( a^* = a_h \), then, all else equal, increases in \( LR(x) \) and \( |a - \bar{a}| \) do not affect \( a^* \). Finally, if \( \tilde{\lambda} \leq LR(x) \), \( L(x,a) \) is (weakly) convex and thus \( a^* = a_h \) independently of \( LR(x) \) and \( |a - \bar{a}| \). The proof for \( x \in K^- \) is symmetric and thus left out.

**Proofs of Propositions 2 and 6**

Suppose \( DM \) has payoff function \( -(\eta \omega + \kappa - a)^2 \) and \( A \) has gross payoff function \( -(\omega - a)^2 \). \( DM \) chooses \( a(x) \) to maximize pointwise

\[
\int_{\mathcal{X}} L(x,a) \, dx = \int_{-\infty}^{\infty} \left( - \int_{-\infty}^{\infty} (\eta \omega + \kappa - a(x))^2 \, dF_{\omega | X}(\omega | x) \right) g_X(x) + \lambda \left( - \int_{-\infty}^{\infty} (\omega - a(x))^2 \, dF_{\omega | X}(\omega | x) g_X(x) - c \right)
\]
which yields, \( \forall x \), first-order condition \( a(x) = \eta p x + \kappa + \lambda (\bar{p} - \rho) x \). Substituting \( a(x) \) in \( A \)'s effort constraint and rearranging yields \( \lambda = \frac{c - 2\eta(\bar{p} - \rho)}{2(\sigma - \rho)^2} \). \( A \)'s effort constraint binds if and only if \( c > 2\eta(\bar{p} - \rho) \). Substituting \( \lambda \) in \( a(x) \) and rearranging concludes (with \( \eta = 1 \) and \( \kappa = 0 \) for Proposition 2).

Proof of Proposition 3

Let \( \bar{U}(a, x) \equiv \int_{\Omega} u(a, \omega) dF_{\bar{X}, \bar{X}}(\omega | x) \) and \( \underline{U}(a, x) \equiv \int_{\Omega} u(a, \omega) dF_{\underline{X}, \underline{X}}(\omega | x) \). DM chooses \( a(x) \) to maximize pointwise

\[
\int_{\chi^2} L(x, a) \, dx = \int_{\chi^2} \left( g_{\bar{X}_i}(x_i) g_{\bar{X}_j}(x_j) \bar{U}(a, x) \right) + g_{\bar{X}_i}(x_i) g_{\bar{X}_j}(x_j) \left( \gamma_l(x) (a(x) - a_l(x)) + \gamma_h(x) (a_h(x) - a(x)) \right) \, dx.
\]

For any given \( x_j \), the proof is identical to the proof of Proposition 1, replacing \( K \) with \( K(x_j) \), \( K^+ \) with \( K^+(x_j) \), and \( K^- \) with \( K^-(x_j) \).

Proof of Proposition 4

DM chooses \( a(x) \in [a_l(x), a_h(x)] \) to maximize \( L(x, a) \), \( \forall x \). \( ^{48} \)

\[
\frac{\partial}{\partial a} L(x, a) = \frac{g_{\bar{X}_1}(x_1) g_{\bar{X}_2}(x_2) (1 + \lambda_1 + \lambda_2)}{\bar{I}(a)} \int_{\Omega} \frac{\partial}{\partial a} u(a, \omega) dF_{W|\bar{X}, \bar{X}}(\omega | x)
\]

\(^{48}\)For the sake of brevity, we ignore \( a_l(x) \leq a(x) \) and \( a(x) \leq a_h(x) \) and identify ex post when these constraints are susceptible to hold.
\[-g_X(x_1)g_X(x_2)\lambda_1 \int_{\Omega} \frac{\partial}{\partial a} u(a, \omega) \ dF_{W|X,X}(\omega \mid x) \]

\[-g_X(x_1)g_X(x_2)\lambda_2 \int_{\Omega} \frac{\partial}{\partial a} u(a, \omega) \ dF_{W|X,X}(\omega \mid x) . \]

Since $\frac{\partial^2}{\partial a^2} u(a, \omega) = -1$, \( \text{Sign} \left[ \frac{\partial^2}{\partial a^2} L(x, a) \right] = \text{Sign} \left[ \sum_{i=1}^2 \text{LR}(x_i) \lambda_i - \left( 1 + \sum_{i=1}^2 \lambda_i \right) \right] . \)

\( L(\cdot) \) is concave if $\lambda_1 \text{LR}(x_1) + \lambda_2 \text{LR}(x_2) < 1 + \lambda_1 + \lambda_2$, and otherwise convex. Therefore, the value of $a$ that solves $\frac{\partial}{\partial a} L(x, a) = 0$ is $L(\cdot)$’s maximizer if (i) it belongs to $[a_l(x), a_h(x)]$ and (ii) if $L(\cdot)$ is concave. Otherwise, $L(\cdot)$’s maximizer is equal to either $a_l(x)$ or $a_h(x)$. In what follows, we exploit (i) whether $L(\cdot)$ is concave/convex, (ii) $\text{Sign} \left[ I \left( \hat{a}_{X,X}(x) \right) - \Pi \left( \hat{a}_{X,X}(x) \right) - \text{III} \left( \hat{a}_{X,X}(x) \right) \right]$, and (iii) the properties of $g_X(x_1)g_X(x_2)$.

\[ U(a, x) := g_X(x_1)g_X(x_2)\lambda_1 \int_{\Omega} u(a, \omega) \ dF_{W|X,X}(\omega \mid x) \]

\[ + g_X(x_1)g_X(x_2)\lambda_2 \int_{\Omega} u(a, \omega) \ dF_{W|X,X}(\omega \mid x) , \]

where $\frac{\partial}{\partial a} U(\cdot) = \Pi(a) + \text{III}(a)$. Let $\hat{a}(x) := \arg\max_{a \in \mathbb{R}} U(a, x) . \]

**Lemma A.1**

1. Function $U(a, x)$ is strictly concave, such that $U(\hat{a} + \epsilon, x) = U(\hat{a} - \epsilon, x)$, $\forall \epsilon$, and such that (i) $\hat{a} \in \left[ \min \left\{ \frac{1}{2} \left( \hat{a}_{X,X} + \hat{a}_{X,X}, \hat{a}_{X,X} \right), \hat{a}_{X,X} \right\} \right] ,$

\[ \max \left[ \frac{1}{2} \left( \hat{a}_{X,X} + \hat{a}_{X,X}, \hat{a}_{X,X} \right), \hat{a}_{X,X} \right] \] if $\text{LR}(x_i) \lambda_i > \text{LR}(x_j) \lambda_j$, and (ii) $\hat{a} = \frac{1}{2} \left( \hat{a}_{X,X} + \hat{a}_{X,X} \right)$ if $\text{LR}(x_i) \lambda_i = \text{LR}(x_j) \lambda_j .$

2. Moreover, $\frac{\partial}{\partial \lambda_i} \left| \hat{a} - \hat{a}_{X,X} \right| < 0$ and $\frac{\partial}{\partial \text{LR}(x_i)} \left| \hat{a} - \hat{a}_{X,X} \right| < 0 . \]
Proof. The concavity and symmetry of $U(a,x)$ follows from (1). Also, given that $\int_\Omega u(a,\omega) dF_{\mathbb{X}^i,\mathbb{X}^j}(\omega \mid x)$ is a translation of $\int_\Omega u(a,\omega) dF_{\mathbb{X}^i,\mathbb{X}^j}(\omega \mid x)$, $\tilde{a} = \frac{1}{2} \left( \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} + \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right)$ if LR $(x_i) \lambda_i = LR(x_j) \lambda_j$. The rest of Part i and Part ii follow from the observation that $\lambda_i$ and LR $(x_i)$ are weights put on $\int_\Omega u(a,\omega) dF_{\mathbb{X}^i,\mathbb{X}^j}(\omega \mid x_i,x_j)$ in $U(a,x)$, for $i = 1, 2$. ■

Let $i, j = 1, 2$ and $i \neq j$. We consider all 4 possible rankings of $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$, $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$, and $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$ in turn.

Case 1: $x \in K^{++} = \left\{ x \in \mathcal{X} \times \mathcal{X} \mid \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \geq \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \& \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} > \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right\}$

Here, $\tilde{a} \leq \hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$. Setting $a_l \leq a < \hat{a}$ is not optimal because $L(2\hat{a} - a, x) > L(a, x)$.\textsuperscript{49} Similarly, setting $a \in \left[ \hat{a}, \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right]$ is not optimal because $\frac{\partial}{\partial a} L(x,a) > 0$ throughout the interval. Finally, $a^* > \hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$ is strictly optimal because $I \left( \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right) - II \left( \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right) - III \left( \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right) > 0$. Also, $x \in K^{++} \Rightarrow \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} > a_0$ because $x_i \geq \bar{x}_i(x_i)$ and $x_j \geq \bar{x}_j(x_i)$ imply $x_i + x_j > 2a_0$, which implies $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j} > a_0$.

Case 2: $x \in K^{--} = \left\{ x \in \mathcal{X} \times \mathcal{X} \mid \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \leq \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \& \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} < \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right\}$

The proof of Case 2 is symmetric to the proof of Case 1, and thus omitted. When $x \in K^{--}$, $a^* < \hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$. Also $x \in K^{--} \Rightarrow \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} < a_0$.

Case 3: $x \in K = \left\{ x \in \mathcal{X} \times \mathcal{X} \mid \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \& \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} \right\}$

Here, $\tilde{a} = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$ and $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$ is $L(\cdot)$’s global maximizer if $\lambda_i LR(x_i) + \lambda_j LR(x_j) < 1 + \lambda_i + \lambda_j$. Otherwise, $L(\cdot)$ is convex and $a^*$ is equal to either $a_l$ or $a_h$. Moreover, $\hat{a}_{\mathbb{X}^i,\mathbb{X}^j} = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j}$ if and only if $x_i = x_j = \omega_0$, implying $x \in K \Rightarrow \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} = a_0$ (see (7)-(8)).\textsuperscript{49}

\textsuperscript{49} Moreover, $2\hat{a} - a < a_h$ because $a_l = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} - q \leq a < \hat{a} \leq \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} < a_h = \hat{a}_{\mathbb{X}^i,\mathbb{X}^j} + q$. 46
Case 4: $\mathbf{x} \in K^+=\{\mathbf{x} \in \mathcal{X} \times \mathcal{X} | \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\}$

a) $|x_j - \omega_0| < |x_i - \omega_0|$

We show $a_0 < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} < a^*$. First, we establish $|x_i - \tilde{x}| > |x_j - \tilde{x}|$ implies $\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$. Using (7)-(8), one computes that the necessary and sufficient condition for inequality $\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$ to hold is $x_i + x_j > 2\omega_0$. This strict inequality always holds because either (i) $x_i > \omega_0 > x_j$ (with $|x_i - \omega_0| = x_i - \omega_0 > |x_j - \omega_0| = \omega_0 - x_j$) or (ii) $x_i > x_j > \omega_0$. Indeed, any other ranking of $(x_i, x_j, \omega_0)$ violates either $\mathbf{x} \in K^+$ and/or $|x_i - \omega_0| > |x_j - \omega_0|$.

To continue, suppose $\lambda_i \geq \lambda_j$. Also, recall that $|x_i - \omega_0| > |x_j - \omega_0|$ implies $LR(x_i) \geq LR(x_j)$. Then, from Lemma A.1, we know $\tilde{a} \in \left[\hat{a}_{\mathbf{x}_i,\mathbf{x}_j}, \frac{1}{2} \left(\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} + \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right)\right]$. Further, $\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$ implies $\frac{1}{2} \left(\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} + \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right) < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$. It follows $\tilde{a} < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$. Because (i) $U(a, \mathbf{x})$ is symmetric and (ii) $\tilde{a} < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$, $a_t \leq a < \tilde{a}$ cannot be optimal. Indeed, $L(x, a) < L(x, 2\tilde{a} - a)$.

Moreover, $a \in \left[\tilde{a}, \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right]$ cannot be optimal either, because $\frac{\partial}{\partial a} \left(\hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right) = 0$ and $\Pi_{\mathbf{x},\mathbf{x}_j} \left(\hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right) + \Pi_{\mathbf{x},\mathbf{x}_j} \left(\hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right) < 0$, so that $\frac{\partial}{\partial a} L(x, \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}) > 0$. Therefore, when $\lambda_i \geq \lambda_j$, $a^* > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$. Finally, note that $\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > a_0$ if and only if $x_i + x_j > 2\omega_0$, which we have shown to hold. Thus, when $\lambda_i \geq \lambda_j$, $a^* > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} > a_0$.

Suppose now $\lambda_i < \lambda_j$, and recall $\frac{\partial}{\partial a} \left|\hat{a} - \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}\right| < 0$. Because $\hat{a}_{\mathbf{x}_i,\mathbf{x}_j} < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$, there exists a threshold, denoted $\tilde{\lambda}_j(x_i, x_j)$, such that $\tilde{a} > \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$ if and only if $\lambda_j > \tilde{\lambda}_j(x_i, x_j)$.

If $\lambda_i < \lambda_j < \tilde{\lambda}_j(x_i, x_j)$, $\tilde{a} < \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$ and the proof is identical to that when $\lambda_i \geq \lambda_j$. If $\lambda_j = \tilde{\lambda}_j(x_i, x_j)$, then $\tilde{a} = \hat{a}_{\mathbf{x}_i,\mathbf{x}_j}$ and $a^*$ is

---

50 Again, $2\tilde{a} - a < a_0$ because $a_t = \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} - q \leq a < \tilde{a} \leq \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} < a_h = \hat{a}_{\mathbf{x}_i,\mathbf{x}_j} + q$.

51 Notice $\lambda_j(x_i, x_j) > \lambda_i$ because $|x_i - \omega_0| > |x_j - \omega_0|$.
equal to either \( \hat{a}_{\bar{x}, \bar{x}_j} \) if \( \lambda_i \text{LR}(x_i) + \lambda_j \text{LR}(x_j) < 1 + \lambda_i + \lambda_j \), or otherwise either \( a_l \) or \( a_h \). Finally, if \( \lambda_j > \tilde{\lambda}_j (x_i, x_j) \), then \( \tilde{a} > \hat{a}_{\bar{x}, \bar{x}_j} \) and the proof follows the same steps as those in the previous paragraph, with however \( a^* < \hat{a}_{\bar{x}, \bar{x}_j} \).

b) \(|x_i - \omega_0| = |x_j - \omega_0|\)

Here, \( x_i > \tilde{x} > x_j \). Any other ranking violates \( x \in K^+ \) or \(|x_i - \omega_0| = |x_j - \omega_0|\). Thus, \( x_i + x_j = 2 \omega_0 \) and \( \hat{a}_{\bar{x}, \bar{x}_j} - \hat{a}_{\bar{x}, \bar{x}_j} = \hat{a}_{\bar{x}, \bar{x}_j} - \hat{a}_{\bar{x}, \bar{x}_j} \), thereby implying \( \hat{a}_{\bar{x}, \bar{x}_j} = \frac{1}{2} (\hat{a}_{\bar{x}, \bar{x}_j} + \hat{a}_{\bar{x}, \bar{x}_j}) \). Also, \(|x_i - \omega_0| = |x_j - \omega_0| \) implies \( \text{LR}(x_i) = \text{LR}(x_j) \). Therefore, \( \tilde{a} \in [\min [\hat{a}_{\bar{x}, \bar{x}_j}, \hat{a}_{\bar{x}, \bar{x}_j}], \max [\hat{a}_{\bar{x}, \bar{x}_j}, \hat{a}_{\bar{x}, \bar{x}_j}]] \) if \( \lambda_i > \lambda_j \), and \( \tilde{a} = \hat{a}_{\bar{x}, \bar{x}_j} \) if \( \lambda_i = \lambda_j \). Here, \( \hat{a}_{\bar{x}, \bar{x}_j} = a_0 \) because \( x_i + x_j = 2 \omega_0 \). Thus, \( a^* > a_{\bar{x}, \bar{x}_j} = a_0 \) if \( \lambda_i > \lambda_j \) and \( a^* < a_{\bar{x}, \bar{x}_j} = a_0 \) if \( \lambda_j > \lambda_i \). If \( \lambda_i = \lambda_j \), then \( a^* = a_0 \) if \( \lambda_i \text{LR}(x_i) + \lambda_j \text{LR}(x_j) < 1 + \lambda_i + \lambda_j \), and otherwise \( a^* \) is either \( a_l \) or \( a_h \).

c) \(|x_i - \omega_0| < |x_j - \omega_0|\)

If \( \lambda_j \geq \lambda_i \), then \( a^* < \hat{a}_{\bar{x}, \bar{x}_j} < a_0 \). Also, there exists a threshold \( \tilde{\lambda}_i (x) > \lambda_j \) such that (i) \( a^* \leq \hat{a}_{\bar{x}, \bar{x}_j} \) if \( \lambda_j < \lambda_i \leq \tilde{\lambda}_i (x) \) and (ii) \( a^* > \hat{a}_{\bar{x}, \bar{x}_j} \) if \( \lambda_i > \tilde{\lambda}_i (x) \).

**Proof that** \( \lambda_1 = \lambda_2 \). Because agents have identical payoffs and available signals, and both effort constraints are binding, \( a^*(x) \) must be such that \( \lambda_1^* = \lambda_2^* > 0 \). By contrast, suppose, for instance, that \( \lambda_1 > \lambda_2 \geq 0 \). Then, the LHS of both effort constraints is identical. However, the RHSs will not generally be equal, which contradicts the fact that both constraints bind at the optimum.

Because \( \lambda_1^* = \lambda_2^* > 0 \), it follows from Cases 1-4 that \( a^* > \hat{a}_{\bar{x}, \bar{x}} \) if \( \hat{a}_{\bar{x}, \bar{x}} > a_0 \), \( a^* < \hat{a}_{\bar{x}, \bar{x}} \) if \( \hat{a}_{\bar{x}, \bar{x}} < a_0 \), and \( a^* \in \{a_l, \hat{a}_{\bar{x}, \bar{x}}, a_h\} \) if \( \hat{a}_{\bar{x}, \bar{x}} = a_0 \).
Figure 1: The Single-Agent Case
Figure 2: Congruent and Conflicting Information