The Design of Information Acquisition and Sharing*

Dimitri Migrow† and Francesco Squintani‡

November 23, 2022

Abstract

We investigate how to design remunerations in order to incentivize acquisition and sharing of information in multi-divisional organizations. If the information between the divisions is sufficiently correlated, and is not too costly to acquire, the headquarters manager optimally chooses a remuneration scheme that depends on the performance of the entire organization. Otherwise, the best incentive to encourage information acquisition is to make each manager's remuneration depend only on the performance of her own division.

Keywords: organizational design, cheap talk, information acquisition
JEL codes: D24, D82, D83, M31

*For helpful comments and discussions we thank Li Hao, Alexander Jakobsen, Wei Li, Shuo Liu, Niko Matouschek, Heikki Rantakari, Sergei Severinov, Colin Stewart, and Chris Wallace, as well as seminar participants at Calgary, 2018 Canadian Economic Theory Conference, 2018 North American Summer Meeting of the Econometric Society and 2018 Royal Economic Society Conference.

†University of Edinburgh, School of Economics, dimitri.migrow@ed.ac.uk

‡University of Warwick, Department of Economics, f.squintani@warwick.ac.uk

1 Introduction

Efficient decisions in organizations rely on high-quality information. Often, information relevant for decision-making is not readily available. It needs to be acquired, and shared between the organizational participants. For instance, consider multi-divisional organizations where division managers operate in local markets and can obtain local information. If a multi-divisional organization supplies similar products in multiple markets, locally obtained information can be useful for other divisions. We ask, how should such an organization design transfers in order to incentivize acquisition and sharing of information? In particular, we study transfer schemes that allow for rewarding individual division performance, team performance (joint bonuses and penalties) and relative performance (tournaments).[1]

Remunerating division managers on the basis of their relative performance fosters information acquisition but can harm incentives to share information. Remunerating different managers as a "team" facilitates communication but harms individual incentives to acquire information in the first place. Alternatively, the managers could be remunerated only on the basis of their own division performance. However, in the context of correlated information this might not be the best way to cut back the managers' rents.

To study the optimal incentives scheme, we develop a model that comprises a principal (she) and two agents (he). The former can be thought as the company headquarters manager and the latter as local division managers. There are two unobserved local states, drawn from two continuous and correlated distributions. Each agent takes a decision in his division. The company's profit is separable across the divisions' choices and increasing in how closely each agent's decision matches his division's local state. Prior to making his decision, each division manager can obtain a costly private signal about his local state. The agents can inform each other about their signals using cheap talk communication. Then the agents simultaneously make their decisions. The profit determined within each division is verifiable in a court of law, but the managers' decisions and the local states are not.

The principal offers and commits to a linear transfer scheme which is separable in the performances of both agents: it may remunerate each agent for his own performance and/or for the other agent's performance. The contracts can feature a competitive, tournament element if better performance by the other local manager leads to a worse payoff for the agent, or a cooperative, team element if better performance by the other agent increases the agent's own payoff. The agents are protected by limited liability and negative transfers are ruled out. We ask, what is the optimal pattern of communication and signal acquisition from the principal's perspective and what are the cheapest incentives to achieve it?

We model the information exchange between the agents as cheap talk which reflects the

[1]Multiple studies in organizational economics capture organizational design by incentive schemes that facilitate information acquisition and sharing. However, existing theories typically disregard transfers and instead focus on organizing communication and assignments of decision rights. See, e.g., Aghion and Tirole (1997); Alonso, Dessein, and Matouschek (2008, 2015); Dessein (2002); Rantakari (2008).

observation that in many organizations managers are unable to commit to choices based on agents' reports. At the same time, we allow for contracting based on the performances of the local divisions that are often observable by the headquarters management and therefore verifiable in the court of law. Indeed, often the headquarters manager lack time or expertise to monitor actions of local managers and therefore has to design an incentive scheme based on divisional outcomes. Our principal therefore has an intermediate commitment power between incomplete contracts approach where transfers are ruled out[2] and the mechanism approach with full commitment power.

To discuss our main results and intuitions, consider an environment where a division manager i expects the other manager j to acquire and supply information about j's own market, and the market characteristics such as consumer preferences are sufficiently correlated, in which case the supplied information has a large impact on the manager's decision. If the other manager j shirks, does not acquire information, and therefore sends a report that is not based on any signal informative about j's local state, a distortion in the manager i's decision can be substantial. In particular, under large correlation in market characteristics this distortion is larger than a distortion for the division, whose manager did not acquire information. Therefore, to ensure that local information is acquired and shared, the headquarters manager optimally links the pay of each manager to the performance of the other manager.

In contrast, if the market in which the organization operates are sufficiently heterogeneous and therefore the signal of each agent i has a limited value for the other agent j's action, the headquarters manager optimally remunerates each manager only for his own performance.

To understand the first result, suppose momentarily that the local states are perfectly correlated, so that each agent's researched information is equally valuable for both agents' decisions. Suppose that agent i shirks and does not exert effort in collecting information. In the *optimal* equilibrium, agent j exerts effort researching information, reports this information to i, and expects i to do the same. To fulfill this expectation, agent i has to put together something to report to j. Because it is not based on research, what i reports is just noise. Agent j takes i's report seriously and uses it to make his decisions, together with his own costly and valuable research. Agent j would make a better decision if he simply ignored agent i's report, and based his decision only on his own research. This is exactly what agent i does: He bases his decision on j's report only. The shirking agent i does more damage to his peer j's performance than to his own: on top of not providing useful information, i also biases j's decision. As a result, the most potent incentive to prevent an agent i from shirking information acquisition is making i's payment sensitive not to his own performance, but to the performance of his peer j.

This finding holds as long as the states are sufficiently correlated. When they are not, the value of the non-shirking agent j's information for the shirking agent i's decision is too low, and

[2]Various models with weak commitment by organizational decision-makers are discussed in Gibbons, Matouschek, and Roberts (2012).

the bias induced on j's decision by the noise in agent i's report is too low, so that if an agent i shirks information acquisition, he damages his own performance more than the other agent j's performance. As a result, the most potent incentive to prevent shirking is then making each agent i's payment depend on i's own performance.

As our second result, we show that the above characterization is not limited to the case of low information acquisition costs, and it is valid as long as the research cost is not prohibitively high to prevent any information acquisition by the division managers. The only difference is that, for such intermediate research costs, it is optimal for the organization that only one division manager acquires and shares information. But, again, if the states are sufficiently correlated, this is optimally achieved by making that division manager's remuneration depend mainly on the other agent's performance.

Interestingly, for intermediate research costs and low correlation among the states, the optimal contract is such that both division managers acquire information without sharing it. The reason for this result is that, in this parameter range, it is too costly (relative to the expected profit increase) to motivate a division manager i to acquire information to improve the performance of his own division i if he already expects valuable information to be reported from the other division manager j. Company profits turn out to be maximized by asking division managers to research information without sharing it, in exchange for lower expected transfers.

The results described so far refer to the setup in which division managers would be unwilling to reveal that they shirked their assignments, were they to deviate and not acquire information. However, we also study a setup in which managers would tell each other that they shirked if they did not acquire information. Our third set of results shows that in this case the optimal contract always links each manager's remuneration only to the performance of his own division. As long as the states are imperfectly correlated, a shirking manager who reveals himself to be uninformed does more damage to his own division's performance than to the performance of the other division. Therefore the optimal way to prevent shirking information acquisition is to link the manager's remuneration to his own division's performance.

Importantly, we show that the profits of the organization as a whole are higher in case the designer is able to prevent reporting shirking behavior compared to the case in which the organizational design makes managers willing to reveal when they are not informed. Impeding collusive off-path behavior of shirking managers is an important ingredient to foster on-path cooperation in information acquisition and sharing, to the benefit of the whole organization.

We complete the paper by generalizing our main result with respect to the statistical model. Further, we allow for the possibility that the principal can be biased towards one of the divisions. In this case, if the correlation between the local states is not too hight, the principal can prefer an allocation with two acquired signal but only one sided information where an agent of her least preferred division communicates his signal to the other agent. We also study mixed extension of the baseline game and allow for partial revelation of information. In both cases the principal's preferred equilibrium is achieved within the baseline model. Finally, we study decentralized sequential communication where one of the agents decides whether to acquire a signal *after* observing the message from the other agent. We show that sequential communication protocol is dominated by simultaneous communication.

In terms of the empirical/testable implications of our analysis, we emphasize that incentivizing information acquisition and sharing by multiple division managers displays a feature of joint performance remuneration. Information sharing is possibly the most evident form of cooperation in teams of division managers. And there is empirical evidence that incentive schemes based in part on joint performance evaluation improve division manager's productivity and profit relative to fixed wage structures and individual performance evaluations. As we detail in the literature review, Kruse (1993) documents a productivity increase of about 4.5%-5.5% in companies that adopt profit sharing plans in the forms of cash transfers. Further, Kandel and Lazear (1992), Che and Yoo (2001) and Alonso et al. (2008) discuss case studies of major corporations that experienced productivity increases following the restructuring of managerial incentives to include elements of joint performance evaluation.

Related literature: This paper is related to several literatures. It relates to the literature on organizational design in the presence of strategic communication (Alonso et al. (2008), Rantakari (2008)). The optimal design of incentives for a team of agents is an important topic in the economics of organizations. As mentioned in the introduction, much of the literature assumes that the agents attach more weight to the profits of their own division relative to the other division(s). We endogenize those weights and show that this need not be optimal – in fact, a profit-maximizing principal might optimally link agent's incentives to the entire organizational profits and not just to the profit of the agent's own division.

Our paper also relates to the literature on contract design with multiple agents. Earlier theoretical contributions have not always precisely identified the form of individual efforts to be remunerated, nor the form of the uncooperative behavior that makes wage spread undesirable. We explicitly posit that individual effort is placed in costly information acquisition, and that the uncooperative behavior consists of not sharing the information acquired, which is regarded as an important setting when studying cooperation in teams.[3]

A contract can foster either a competitive or cooperative behavior (or both). The competition element is usually associated with a tournament-based contract structure. In a seminal contribution, Lazear and Rosen (1981) show that with risk-neutral agents the tournaments are optimal and result in the same outcomes as piece rates. Green and Stokey (1983) study risk-averse agents and show that tournaments can outperform piece rates when agents' performances are influenced by a common shock unobservable to the principal. There is no common shock in the agents' performances in our model, and as a result there is no role for tournaments in the optimal contracts.

[3]In personnel economics, Lazear (1989) motivated his study precisely with the practical example where workers exert effort to obtain information and cooperation is captured by information sharing. Che and Yoo (2001) also discuss how information sharing captures cooperation in teams.

The cooperation element in contracts is studied in Holmström and Milgrom (1990), Lazear (1989), Kandel and Lazear (1992), and Itoh (1991), among others. These papers study reduced-form models of multi-agent production and point out that production externalities lead to the optimality of team based agents' remuneration. In those cases, rewards which are only contingent on individual performances or on relative performances can harm cooperation and so the principal's objectives. Itoh (1993) shows that optimal contracts include team bonuses when the agents can monitor each other and stipulate self-enforceable side-contracts, and Che and Yoo (2001) study the merits of relative versus joint performance evaluation in a repeated game in which such self-enforcing contracts arise as an equilibrium phenomenon. By considering a model of information acquisition and sharing in teams of agents, we identify a precise case in which joint performance evaluations improve workers' productivity.

Empirical studies find a strong positive relationship between the adoption of profit sharing schemes and a productivity increase.[4] For instance, Kruse (1993) presents evidence based on a survey of 500 U.S. companies with publicly traded stock. He documents a productivity increase of about 4.5%-5.5% in companies that adopt profit sharing plans in form of cash transfers. This effect is more pronounced in smaller firms, and under a larger profit sharing. (The mechanism is not determined: it may be because of increased monitoring in small teams, but also because of workers' cooperation.) In the analysis of an apparel factory by Hamilton, Nickerson, and Owan (2003), a move from individual piece rates to team production and team-based incentive pay raised productivity substantially. Interestingly, the high-productivity workers were the first to voluntarily join the newly forming teams. There are several case studies that compare the effectiveness of different incentive schemes and show the value of joint bonuses. For example, Alonso et al. (2008) discuss the case of the management restructuring of BPX, the oil and gas exploration division of British Petroleum, in the early '90s by the then head of BPX (and future CEO of BP) John Browne.[5]

Our setting relates to a large literature studying environments with cheap talk communication (Crawford and Sobel (1982)) and endogenous information acquisition. In a setting with a single sender and a single receiver Di Pei (2015) shows that when a sender can choose a partitional information structure at a cost, she reveals all her acquired information to the receiver; Argenziano, Severinov, and Squintani (2016) and Deimen and Szalay (2019) study the implications of information acquisition on organizational design. These papers show that communication results in better incentives to acquire signals, compared to delegation. Organi-

[4]The empirical importance of joint performance evaluation is discussed, e.g., in Ichniowski and Shaw (2003)

[5]"Browne decentralized BPX in the early 1990s, creating almost 50 semi-autonomous business units. Initially, since "business unit leaders were personally accountable for their units' performance, they focused primarily on the success of their own businesses rather than on the success of BPX as a whole." (Hansen and Von Oetinger (2001)) To encourage coordination between the business units, BPX established changes in the implicit and explicit incentives of business unit leaders to reward and promote them, not just based on the success of their own division, but also for contributing to the successes of other business units. As a result, "lone stars who deliver outstanding business unit performance but engage in little cross-unit collaboration can survive within BP, but their careers typically plateau." (Hansen and Von Oetinger (2001))." Alonso et al. (2008), page 164-165.

zational design in a multi-agent setting is studied in the context of coordinated adaptation by Alonso et al. (2008) and Rantakari (2008) who focus on whether a multi-divisional organization should implement decentralization with horizontal communication, or a centralized architecture. In a two-agent setting where the headquarters can either choose prices or quantities, Alonso, Matouschek, and Dessein (2010) show how different headquarters' choices affect quality of communicated information. In these papers decision-relevant information is exogenous. In contrast, Angelucci (2017) studies a model with two agents and endogenous acquisition of costly information. Different to our setting, in the above papers contractual transfers based on information are absent and so the principal has to rely on different instruments than monetary transfers to incentivize her agents.

We also relate to Callander and Harstad (2015) who show that it can be optimal for the principal to constrain agents' actions in order to maximize informational sharing. In our setup in contrast, the principal solves the incentive problem of information acquisition and sharing while letting the agents choose their respective actions.

2 Model

An organization consists of a headquarter manager and two division managers. For simplicity we refer to the division managers as agents 1 and 2 (he), and to the headquarter manager simply as a principal (she). Each agent i is assigned to a division in which he chooses an action $y_i \in [0,1]$. There are two unobserved local states, $\theta_1 \in [0,1]$ and $\theta_2 \in [0,1]$. The profit of a division i is $\pi_i(y_i,\theta_i) = \bar{\pi} - \ell_i(y_i,\theta_i)$ where $\ell_i(y_i,\theta_i) = (y_i - \theta_i)^2$ for each agent i. Each agent i's decision y_i gives a higher profit $\pi_i(y_i,\theta_i)$ the more precisely it matches the local state θ_i. The loss $\ell_i(y_i,\theta_i)$ is expressed in a simple quadratic form that is standard in organization design since Dessein (2002). The principal wants to maximize the profit function π which separable in the agent's performances and takes the form $\pi = \pi_1(y_1,\theta_1) + \pi_2(y_2,\theta_2)$.[6] Since for each $i = 1,2$, $\ell_i \in [0,1]$, we simply set $\bar{\pi} = 1$.

The states θ_1, θ_2 are correlated: with probability r they are identical and randomly drawn from the uniform distribution $U[0,1]$. With probability $1-r$ each state is drawn independently from $U[0,1]$. Prior to choosing y_i, each agent i can exert a costly effort $c > 0$ which enables him to observe a private signal $s_i \in \{0,1\}$ about the local state θ_i, such that $\Pr(s_i = 1|\theta_i) = \theta_i$. If an agent exerts no effort he cannot obtain an informative signal. After the signals are received and before the decisions (y_1, y_2) are taken the agents can simultaneously communicate with each other using cheap talk messages. We assume that each agent i has an arbitrary large set of messages M_i with a typical element $m_i \in M_i$.

[6] Our results generalize qualitatively to the case in which the profits dependence on performances is asymmetric: $\pi = \lambda_1\pi_1(y_i,\theta_i) + \lambda_2\pi_2(y_i,\theta_i)$. Such asymmetries can arise, for example, due to different sizes of the organizational divisions (Rantakari (2008)), the leniency bias of the principal (Bol (2011); Breuer, Nieken, and Sliwka (2013)) or the different prospects of markets where the corresponding divisions operate (Liu and Migrow (2018)).

The principal does not observe local states and the agents information acquisition and communication choices. She observes the divisional performances. Thus, contracting is based on the performances π_1 and π_2. Specifically, at the beginning of the game the principal offers (and commits to) a linear contract of the form

$$t_i(\ell_i, \ell_j) := w_i - a_i\ell_i - b_i\ell_j = w_i - a_i(y_i - \theta_i)^2 - b_i(y_j - \theta_j)^2, \tag{1}$$

where $w_i \in \mathbb{R}$, $a_i \in \mathbb{R}$, $b_i \in \mathbb{R}$, and $j \neq i$ denotes the other agent.

As standard in the literature, agents are protected by limited liability, and cannot be paid negative transfers: $t_i(\ell_i, \ell_j) \geq 0$ for all possible performances $\pi_i \in [0, 1]$ and $\pi_j \in [0, 1]$. Normalizing the value of their outside option to zero, it must also be the case that each agent i is willing to accept the contract t_i ex-ante, i.e., before deciding whether to acquire signal s_i, and before sending signal m_i and making choice y_i.

Linear contracts are a standard tool of much of organizational design literature. We note that the above linear specification allows for multiple contracts found in organizations, as explained below. For instance, because $\ell_i = (y_i - \theta_i)^2$ is the loss determined by agent i's imprecise matching of y_i with θ_i, the contract t_i is a piece-wise linear contract, composed of a fixed wage $w_i = \bar{w}_i - a_i - b_i$, a bonus payment $a_i(1-\ell_i)$ that depends on agent i's performance, and a payment $b_i(1 - \ell_j)$ based on the other agent j's performance.[7] The contract t_i can also be interpreted as a mixture of relative performance evaluation and joint performance evaluation based transfers, by letting $\overline{a}_i = (a_i - b_i)/2$ and $\overline{b}_i = (a_i + b_i)/2$ and obtaining: $t_i(\ell_i, \ell_j) = w_i - \overline{a}_i(\ell_i - \ell_j) - \overline{b}_i(\ell_i + \ell_j)$. The parameter $\overline{a}_i$ is a weighting factor for agent i's relative performance and $\overline{b}_i$ is a weighting factor for the team performance: agent i's payment $t_i(\ell_i, \ell_j)$ is more sensitive to the relative loss $(\ell_i - \ell_j)$ the higher is $\bar{a}_i$ and to the aggregate loss $(\ell_i + \ell_j)$ the higher is $\bar{b}_i$.

We also note that while our communication protocol is the same as communication under decentralized decision making in Alonso et al. (2008) and Rantakari (2008), in Section 5 we study decentralized sequential communication where signal acquisition decision of one of the agent happens after he receives signal-contingent information from the other agent.

The timing of our game proceeds as follows: First, nature (privately) chooses (θ_1, θ_2) and the principal offers, and commits to, contracts (t_1, t_2). Second, the agents decide whether to acquire signals (s_1, s_2). Then they send simultaneous cheap talk messages m_i to each other. Finally, each agent i chooses y_i, performances π_1 and π_2 are publicly observed and the transfers are paid as specified in the contracts. The structure of the model is common knowledge.

Finally, we note that given the contracts (t_1, t_2), multiple equilibria may exist in the agents' game, for example there is always an equilibrium in which agents do not communicate with

[7] We will later show that in the optimal linear contracts, it is the case that $a_i < 1$, $b_i < 1$, and that the limited liability constraint binds, so that $t_i(\ell_i, \ell_j) = 0$ for $\ell_i = 1$ and $\ell_j = 1$, and hence $w_i - a_i - b_i = 0$. Then, the contracts $t_i(\ell_i, \ell_j)$ can be decomposed as $t_i(\ell_i, \ell_j) = w_i + a_i(1 - \ell_i) + b_i(1 - \ell_j)$.

each other. As customary, we only consider equilibria with meaningful flow of information.

3 Conditional optimal transfers

We first study subgames that follow every possible principal's choice of agents' strategies, where each agent's strategy specifies the information acquisition and communication decision, as well as the optimal action y_i. For the moment, we operate under the assumption that, given the principal's choices of contracts t_1 and t_2, the agents coordinate on the equilibrium preferred by the principal. This analysis allows us to characterize the cheapest way for the principal to implement a targeted allocation. At the end of the section, we discuss other equilibria and their plausibility.

Consider the first case where the principal wants to implement full signal acquisition and complete sharing. The principal's optimization problem goes as follows:

$$\max_{t_1,t_2} E[\pi_1(\ell_1(y_1(s_1,s_2),\theta_1)) + \pi_2(\ell_2((y_2(s_2,s_1),\theta_2)) - \sum_{i=1}^{2} t_i(\ell_1(y_1(s_1,s_2),\theta_1),\ell_2(y_2(s_2,s_1),\theta_2))]$$

where the expectation is taken with respect to the tuple $(\theta_1, \theta_2, s_1, s_2)$, subject to the following constraints. First, the contract space is defined by (1). Second, given information $s_i \in \{0,1\}$ and $s_j \in \{0,1\}$, and the common belief that signals are exchanged truthfully, each agent i chooses $y_i\left(s_i, s_j\right)$ to solve:

$$y_i\left(s_i, s_j\right) = \arg\max_{y_i} E[t_i(\ell_i(y_i(s_i,s_j),\theta_i),\ell_j(y_j\left(s_j,s_i\right),\theta_j)|s_i,s_j],$$

where the expectation is taken with respect to θ_i, θ_j (since agent i observes his own signal is observed, and s_j is believed to be known after the message m_j has been received by i). Third, the communication constraint specifies that i communicates $s_i \in \{0,1\}$ truthfully to j:

$$E[t_i(\ell_i(y_i(s_i,s_j),\theta_i),\ell_j(y_j\left(s_j,s_i\right),\theta_j)|s_i] \geq E[t_i(\ell_i(y_i(s_i,s_j),\theta_i),\ell_j(y_j\left(s_j,1-s_i\right),\theta_j)|s_i],$$

where the expectation is taken with respect to θ_i, θ_j and s_j. Forth, the information acquisition constraint is:

$$E_{s_i,s_j}[E_{\theta_i,\theta_j}[t_i(\ell_i(y_i(s_i,s_j),\theta_i),\ell_j(y_j(s_j,s_i),\theta_j))|s_i,s_j]] - c \geq$$
$$E_{s_j}[E_{\theta_i,\theta_j}[t_i(\ell_i(y_i(s_j),\theta_i),\ell_j(y_j(s_j,\hat{s}_i),\theta_j))|s_j]]$$

where $\hat{s}_i$ is arbitrary, and can take either value 0 or 1. Fifth, the exogenous limited liability constraint specifies that for all $\ell_i \in [0,1]$, $\ell_j \in [0,1]$,

$$t_i\left(\ell_i, \ell_j\right) \geq 0.$$

Finally, the joint feasibility constraint is:

$$t_1(\ell_1, \ell_2) + t_2(\ell_2, \ell_1) \leq 2 - \ell_1 - \ell_2.$$

We proceed by noting that each agent $i = 1, 2$ is motivated to choose decision y_i so as to minimize the loss $\ell_i = (y_i - \theta_i)^2$ for any $a_i \geq 0$: we observe here that it is a dominant strategy by the principal to choose $a_i \geq 0$.[8] At the decision stage, he matches y_i to $E_i(\theta_i|s_i, m_j)$, the posterior expectation of θ given his signal s_i and the message m_j that is *presumed* by i to be truthful. We observe here that the expected loss $E(\ell_i|s_i, m_j) = E[(E(\theta_i|s_i, m_j) - \theta_i)^2|s_i, m_j]$ for a given (s_i, m_j) is the residual variance of the state θ_i given the estimator $y_i(s_i, m_j) = E(\theta_i|s_i, m_j)$.

Proceeding backwards, we next consider the incentives to motivate sharing acquired signals. The following Lemma formalizes the result that, given that $a_i \geq 0$ for both $i = 1, 2$, each agent i is motivated to truthfully report $m_i = s_i$ by setting $b_i \geq 0$.

Lemma 1. *A necessary condition for the existence of an equilibrium in which each agent $i = 1, 2$ acquires signal s_i, truthfully communicates $m_i = s_i$ to the other agent j and chooses $y_i = E(\theta_i|s_i, s_j)$ is that $b_i \geq 0$. This condition ensures that agent i does not deviate from truthtelling.*

The result is intuitive: to incentivize truthful communication by agent i, the principal has to make i's payoff contingent on j's performance. Because $a_j \geq 0$, agent j chooses y_j to match his expectation of θ_j given her signal s_i and the equilibrium belief that $m_i = s_i$. Since communication is costless, an arbitrary small b_i ensures that an informed agent i sends a truthful message to agent j.

We next turn to incentives to acquire signals. Denote by u_i the expected equilibrium payoff of agent i prior to observing s_i, if he chooses to acquire a signal. In this case the expected on-path payoff has a simple form

$$u_i(s_i) = \bar{w}_i - (a_i + b_i)E[E(\ell_i|s_i, s_j)|s_i] - c \tag{2}$$

where in $E[E(\ell_i|s_i, s_j)|s_i]$ given his signal s_i the expectation is taken with respect to the other agent's signal s_j, unknown to i at the time when she chooses whether to acquire signal s_i.

We show in the Appendix B that because of the model's symmetry across agents and signal realizations, $E[E(\ell_i|s_i, s_j)|s_i]$ is the same regardless of whether $s_i = 0$ or $s_i = 1$, and specifically, $E[E(\ell_i|s_i, s_j)|s_i] = \frac{3-r^2}{6(9-r^2)}$. Hence, it is also the case that the *unconditional* expected loss is

[8] An arbitrary small but positive a_i ensures that i chooses y_i to minimize ℓ_i. Instead, a_i needs to be strictly bounded above zero if agent i's effort has a direct effect on the profit π_i, instead of just an indirect effect through information acquisition. In order to motivate agent i to exert such productive effort, the transfer $t_i(\ell_i, \ell_j)$ needs to be significantly sensitive to the loss ℓ_i.

$E[E(\ell_i|s_i,s_j)] = \frac{3-r^2}{6(9-r^2)}$: here, the expectation is taken with respect to s_i and s_j. Hence, we can write the expected on-path payoff of agent i before observing s_i as:

$$u_i = w_i - (a_i+b_i)E[E(\ell_i|s_i,s_j)] - c = \bar{w}_i - (a_i+b_i)\frac{3-r^2}{6(9-r^2)} - c \tag{3}$$

When the states θ_1 and θ_2 become less correlated, s_j becomes less informative about θ_i, and so the decision $y_i(s_i,s_j) = E(\theta_i|s_i,s_j)$ becomes less precise about θ_i. This results in a larger expected loss $E(\ell_i|s_i,s_j)$ for both $s_j = 0,1$.

Now, suppose that agent i deviates at the signal acquisition stage, shirks and thus obtains no signal about θ_i. This deviation in unobserved by j: indeed, j continues to believe that i acquired s_i and hence interprets any possible message realization $m_i \in M_i$ as either meaning that $s_i = 0$ or that $s_i = 1$. Like in Argenziano et al. (2016), the equilibrium language is fixed by the on-path communication strategy.

Now, agent i's choice is only based on j's truthful message $m_j = s_j$. His optimal decision is therefore $y_i(s_j) = E(\theta_i|s_j)$, which yields the expected loss $E[E(\ell_i|s_j)]$. Agent j however makes the biased decision $y_j(s_j,m_i) = E(\theta_j|s_j,m_i)$ because he mistakenly believes that agent i acquired a signal $s_i \in \{0,1\}$ and truthfully reported $m_i = s_i$. The misled decision y_j results in a larger expected loss $E[(E(\theta_j|s_j,m_i)-\theta_j)^2]$ than if agent j knew that agent i did not acquire s_i.

We show in Appendix B that agent i's expected off-path payoff when shirking is:

$$u_i^D = w_i - a_iE[E(\ell_i|s_j)] - b_iE[(E(\theta_j|s_j,m_i)-\theta_j)^2] = w_i - a_i\frac{3-r^2}{36} - b_i\frac{27+r^4}{6(9-r^2)}. \tag{4}$$

Whether agent i's shirking at the information acquisition stage induces a larger or a smaller expected loss for division i than for division j, $E[E(\ell_i|s_j)] >$ (or $<$)$E[(E(\theta_j|s_j,m_i)-\theta_j)^2]$, depends on the correlation r across the states θ_i and θ_j i.e., it depends on the informativeness of the truthful message $m_i = s_i$ for the optimal choice of y_i, and so on the distortion that agent i incurs on j's performance when shirking and not acquiring a signal. In the Appendix B we show that there exists a threshold $r_1 \approx 0.803$ (determined precisely in the Appendix B) such that the expected loss is smaller for division i than for division j if and only if $r < r_1$. That is, for highly correlated states θ_1 and θ_2, the dominant effect determined by agent i not acquiring his signal s_i consists in misleading the choice y_j of agent j, and not in choosing y_i with less information.

We now turn to the principal's cost minimization program, $\min E[t_1(\pi_1,\pi_2)+t_2(\pi_2,\pi_1)]$, subject to the constraints we have outlined. To recap, each agent $i = 1,2$ picks the optimal decision $y_i(s_i,s_j) = E[\theta_i|s_i,s_j]$ only if $a_i \geq 0$, communicates truthfully $m_i = s_i$ only if $b_i \geq 0$, and acquires his signal s_i only if $u_i \geq u_i^D$.

We restrict attention to a symmetric pair of linear contracts $t_1 = t_2$. This is without loss of generality, because the principal's cost minimization program is linear.[9] Imposing symmetry across agents, we show in the Appendix B that the principal's cost-minimization program is reduced to the following program, for either $i = 1, 2$, and $j \neq i$,

$$\min_{\substack{a_i \geq b_i \geq 0 \\ w_i \geq a_i + b_i}} \left\{ w_i - (a_i + b_i) E[E(\ell_i | s_i, s_j)] = w_i - (a_i + b_i) \frac{3 - r^2}{6(9 - r^2)} \right\} \qquad \text{s.t. } u_i \geq u_i^D, \tag{5}$$

where $w_i \geq a_i + b_i$ comes from the limited liability constraint, and where we use the previously obtained result that $E[E(\ell_i | s_i, s_j)] = \frac{3-r^2}{6(9-r^2)}$. Proposition 1, proved in the Appendix B, shows that the solution of this program is such that if the agents' states θ_1 and θ_2 are sufficiently uncorrelated (specifically, r is smaller than the threshold r_1), then each agent i is only rewarded for his own performance π_i. For $r > r_1$ each agent i's pay is optimally based on the performance of the other agent, π_j. When the states θ_1 and θ_2 are sufficiently correlated, each agent i's signal s_i is so informative on the other agent's state θ_j that it is optimal to base agent i's remuneration mainly on the performance of the other agent.

Proposition 1. *The optimal contract t_1, t_2 to make each agent $i = 1, 2$ acquire his signal s_i, truthfully transmit $m_i = s_i$ to the other agent j and optimally choose $y_i = E(\theta_i | s_i, m_j)$ in the principal-preferred equilibrium is as follows:*[10]

1. For $r < r_1$, every optimal contract t_i is such that $\bar{w}_i = a_i > 0$ and $b_i = 0$, each agent i's remuneration is based only on his own performance π_i.

2. For $r > r_1$, the optimal contract t_i is such that $a_i = 0$ and $\bar{w} = b_i > 0$, each agent i's remuneration is based only on the performance of the other agent, π_j.

We posited information acquisition in a standard framework in which an agent who receives a signal, forms a noisy estimate of the underlying state. The specific beta-binomial setup we adopt is widely used in the literature on information acquisition due to the tractability of the summary statistics. In Section 5 we show that Proposition 1 holds under fairly weak conditions on the general statistical model. In a nutshell, in a setting where the local states are sufficiently correlated, the requirement is that each signal realization entails some information about the underlying state.

An alternative way to model information acquisition, in the tradition of Aghion and Tirole (1997), allows for the possibility that agents observe entirely uninformative signals with some probability. It is easy to show that our main result, Proposition 1, holds in a variety of models

[9] Each agent i's constraints are linear in the maximization arguments w_i, a_i and b_i. Thus, the constraint set is convex. Hence, suppose that an asymmetric pair of linear contracts $t_1 \neq t_2$ minimized the sum of expected transfers to the agents. Because the model is symmetric, the antisymmetric pair of contracts $t_1' = t_2$, $t_2' = t_1$ is also optimal. But then, the constraint set being convex, it contains also the symmetric pair of contracts obtained by averaging these two pairs. As the objective is linear, this symmetric pair of contracts is also optimal.

[10] The precise value of the solutions a_1 and b_1, of the threshold r_1, and of the analogous solutions and thresholds in the following results are expressed in Appendix B.

that allow agents to receive uninformative signal realizations, as long as the probability of being uninformed is sufficiently small. To illustrate, consider a framework similar to ours, but where with some probability the agent fails to observe the outcome of the trial. The principal's preferred equilibrium is semi-pooling.[11]

Having determined the optimal contracts to make both agents i acquire signal s_i and share it with the other agent j, we calculate the optimal contracts for the other two cases in which both agents acquire information. In the first one, both agents acquire information and only one of them shares with the other agent. In the second case, both agents acquire information and neither of them shares it with the other agent. The principal's optimization is similar to the first case, and is stated in the proof of the following Proposition 2.

The optimal contract t_i for an agent i who is not intended to share signal s_i with the other agent j is such that $w_i = a_i > 0$ and $b_i = 0$ for all values of c and r. The optimal contract is based on i's performance alone as the principal does not want to make i share his information with j. Again, the information acquisition constraints lead to a trade-off within the optimal contract for an agent i who is incentivized to share signal s_i. For low values r of correlation among the states θ_1 and θ_2, optimality requires that $w_i = a_i > 0$ and $b_i = 0$. When θ_1 and θ_2 are highly correlated, the optimal contract is such that $a_i = 0$ and $w_i = b_i > 0$.

Proposition 2. *The optimal contract t_1, t_2 to induce each agent $i = 1, 2$ to acquire signal s_i, only one of them, say agent 1, to transmit $m_1 = s_1$ to the other agent 2 and both agents to choose y_i optimally in the principal's preferred equilibrium is such that $w_2 = a_2 > 0$, $b_2 = 0$ for all r, and that $w_1 = a_1 > 0$, $b_1 = 0$ for $r < r_2$, and $a_1 = 0$, $w_1 = b_1 > 0$ for $r > r_2$. The optimal contract t_1, t_2 to induce both agents $i = 1, 2$ to acquire s_i, not to transmit it to the other agent j, and choose y_i optimally, is such that $w_i = a_i > 0$, $b_i = 0$.*

This result shows that whenever the principal wants both agents to acquire information without sharing it, the optimal contract links each agent's performance to his own output.

The next result characterizes the optimal contracts to induce a single agent i to acquire information, and then either share it with agent j or not. Agent j is not asked to acquire information. The characterization in Proposition 3 mirrors the case for two agents (Propositions 1 and 2), but the optimal contract for agent j is irresponsive of the profits π_i and π_j, i.e., it has $w_j = a_j = b_j = 0$.

Proposition 3. *The optimal contract t_1, t_2 to make one agent, say agent 1, acquire signal s_1 and transmit $m_1 = s_1$ to the other agent 2, agent 2 not acquire information, and both agents i to choose y_i optimally in the principal's preferred equilibrium is such that $w_2 = a_2 = b_2 = 0$ for all r, and that $w_1 = a_1 > 0$ and $b_1 = 0$ for $r < r_3$ and $a_1 = 0$, $w_1 = b_1 > 0$ for $r > r_3$. The*

[11]Each agent i either reports observing one of the outcomes of the trial (say, $s_i = 0$), or he reports not observing that outcome. In the latter case, the agent pools the outcome $s_i = 1$ with the outcome of not observing s_i. In practice, when not observing s_i, agent i pretends that $s_i = 1$.

optimal contract t_1, t_2 to induce only one agent, say agent 1, to acquire signal s_1, but not to transmit $m_1 = s_1$ to agent 2, and both agents i to choose y_i optimally is such that $w_1 = a_1 > 0$, $b_1 = 0$ and $w_2 = a_2 = b_2 = 0$.

Finally, the optimal contracts t_1, t_2 in the case that both agents $i = 1, 2$ are not supposed to acquire information are such that $w_i = a_i = b_i = 0$. The only role played by contracts is to ensure that each agent i matches y_i with the state θ_i to the best of the shared knowledge that θ_i is uniformly distributed on $[0, 1]$. Because neither agent i derives any (dis-)utility from the decision y_i and state θ_i, this objective can be achieved with any $w_i = a_i \geq 0$ and $b_i = 0$.

Equilibrium selection

The analysis so far has presumed that the agents coordinate on the principal's preferred equilibrium, given the principal's choices of contracts t_1 and t_2. Under this specification, were an agent i to shirk and not acquire his signal s_i, he would not reveal to the other agent j that he shirked and has no useful information. Even in case such admission of cheating is not contractually sanctioned, hiding it is part of an equilibrium. Because agent j expects i to acquire the signal $s_i \in \{0, 1\}$, he interprets any possible message m_i as either meaning that $s_i = 0$ or that $s_i = 1$. This is the equilibrium play preferred by the principal, given the offered contracts t_1 and t_2, because it makes the information acquisition constraint $u_i \geq u_i^D$ less demanding.

We now consider the alternative 'agent collusive' equilibrium in which agents who deviate from the equilibrium path and do not acquire their signals, are willing to reveal each other that they shirked. The next result shows that in the consequent optimal contract, the principal links the remuneration of any agent i who is expected to acquire s_i, and possibly share it with the other agent j, to the agent i's own division profit.

Proposition 4. *Suppose that the agents $i = 1, 2$ play the collusive equilibrium, given the contracts t_i. Then, the optimal contract to make any agent i acquire his signal s_i, possibly transmit $m_i = s_i$, and then play y_i optimally is such that $w_i = a_i > 0$ and $b_i = 0$. If an agent i is not supposed to acquire s_i, then $a_i = b_i = w_i = 0$. Each agent i's remuneration is only linked to his own performance π_i.*

To understand the result, note that here an agent i who shirks and reveals he did acquire his signal s_i does more damage to his own division i's profit than to the profit of the other division j. This is because agent j is alerted that i has no useful information, and optimally chooses y_j based on his signal s_j only. Even if j is supposed to share s_j with i, the expected loss $E[E(y_j|s_j)]$ of j's division is smaller than the expected loss $E[E(y_i|s_j)]$ of i's division. This is because the states θ_1 and θ_2 are only imperfectly correlated, and hence signal s_j is more informative about θ_j than about θ_i. (Of course, the argument holds a fortiori in case j is not supposed to transmit s_j to i.) Because i's shirking information acquisition damages his performance π_i more than the other agent j's, the optimal contract t_i links agent i's remuneration to his own performance.

The plausibility of the principal preferred equilibrium relative to the agent collusive equilibrium ultimately depends on the intended application. Whenever the principal is in the position of monitoring information transmission across divisions, it is plainly the case that agents would not volunteer any information that they shirked. Because the messages are contractible, the principal will sanction a sender who reveals that he has shirked. The agent collusive equilibrium is ruled out by savvy engineering of communication channels in the organization.

If the agent' communications take forms more difficult or illegal to monitor, it may sound appealing that an agent i "colludes" by alerting j that he does not have information useful for j's decision y_j. However, basic career concern considerations make it implausible that manager i would admit to j that he shirked his assignments. Reporting lack of research findings bears substantial costs for the sender's career prospects both within internal and external markets. The latter can place a high value on the player's expertise where a history of information acquisition acts as a proxy for the agents' informativeness (or abilities), and thus can have a first-order effect on the agent's wages.[12] Competition in the internal markets for better jobs (or wages) likely leads agent i to never admit being uninformed. This can happen even if the principal cannot contract future job opportunities on today's performance, because many organizations have implicit contracts where future wages and promotions are influenced by the current actions of the workers (Holmström (1999)).

Even leaving aside career concern considerations, there is a further reason why agent i might be reluctant to admit to the other agent j that he shirked and did not acquire information. The hard-working agent j may feel "cheated" by the agent i's lack of cooperation, especially because i's failure to acquire information directly reduces j's expected remuneration. Indeed, it is an insight since at least Lazear (1989), that such cheating behavior may lead to retaliation among coworkers.

4 Optimal contracts

For different values of the information acquisition cost c and state correlation r parameters, the principal's expected profit may be maximized by different information acquisition and communication choices, together with the associated optimal contracts. We begin the analysis by showing that some of the agents' choices are dominated for *all* parameter values of c and r. We consider the optimal contracts to make only one agent i acquire and share his signal s_i with the other agent j, and j not acquire his signal s_j. We show that these contracts yield a higher profit than the optimal contracts that make i acquire his signal s_i without sharing it with j, and j not acquire s_j. For brevity, we henceforth omit reference to the agents $i = 1, 2$

[12]There is a large literature studying players' market value in "experts markets" who try to influence markets' beliefs, for example, through cheap talk communication, signaling or selective disclosure of biographies: see, e.g., Ottaviani and Sørensen (2006), Gow, Wahid, and Yu (2018), and, for a more recent discussion of the literature, Meloso, Nunnari, and Ottaviani (2018).

choosing y_i optimally given their information, when describing optimal contracts.

Lemma 2. *For all cost c and correlation values r, the expected profit $E\pi_{11}$ of the optimal contracts t_1, t_2 inducing only one agent i to acquire s_i and share it with the other agent j is strictly larger than $E\pi_{10}$ for almost all values of c and r, where $E\pi_{10}$ is the optimal profit obtained when only one agent i acquires s_i and does not share it with j.*

This result is intuitive. Given that the optimal contracts t_1, t_2 induce only one agent i to acquire his signal s_i, there is no reason not to make him also share s_i with the other agent j. The only consequence of transmitting $m_i = s_i$ is that the precision of agent j's decision y_j improves, and this increases division j's expected profit. Further, it is very cheap to make agent i share s_i, as this can be achieved with any $b_i \geq 0$.

However, this simple logic does not extend to the optimal contracts that make both agents acquire information. For some c and r, it is not true that the expected profit $E\pi_{22}$ of the optimal contracts that makes both agents i acquire s_i and share it with the other agent j is larger than $E\pi_{21}$, the expected profit of optimally inducing both agents i to acquire s_i but only one of them to share it. Nor it is true that $E\pi_{22}$ is larger than $E\pi_{20}$, the expected profit of optimally making both agents i acquire s_i without sharing it.

The reason for this result is as follows. Suppose that both agents $i = 1, 2$ are asked to acquire their signals s_i by the principal. Consider an agent j, and suppose that he expects to receive signal s_i from agent i in equilibrium. Then, the informational benefit of acquiring signal s_j is smaller than when he does not expect to receive s_i. As a result, the contractual transfer needed to make agent j acquire s_j has to reward the precision of agent j's action y_j more than when j does not receive s_i. When the cost of information c is sufficiently high, it becomes so expensive to simultaneously make agent i send s_i to agent j and agent j acquire s_j, that the principal is better off not asking agent i to share s_i with agent j.

Note that this intuition does not apply to the comparison between $E\pi_{11}$ and $E\pi_{10}$, because in this case agent j is not asked to acquire s_j by the principal. Further, this intuition does not entirely invalidate the possibility of comparing expected profit in the three cases in which both agents are asked to acquire their signals by the principal. It turns out that for every information cost value c and every correlation value r, the choice of asking both agents i to acquire s_i and only one of them to share s_i with the other agent j is either dominated by asking both agents i to acquire and also share their signals s_i, or by asking both agents i to acquire s_i without sharing it.

Lemma 3. *For all cost c and correlation values r, the expected profit $E\pi_{21}$ of the optimal contracts t_1, t_2 inducing both agents i to acquire s_i and only one of them to share it with the other agent j is (generically strictly) smaller than either $E\pi_{22}$, the optimal profit obtained when both agents $i = 1, 2$ acquire and share s_i, or than $E\pi_{20}$, the optimal profit obtained when both agents $i = 1, 2$ acquire s_i without sharing it with the other agent j.*

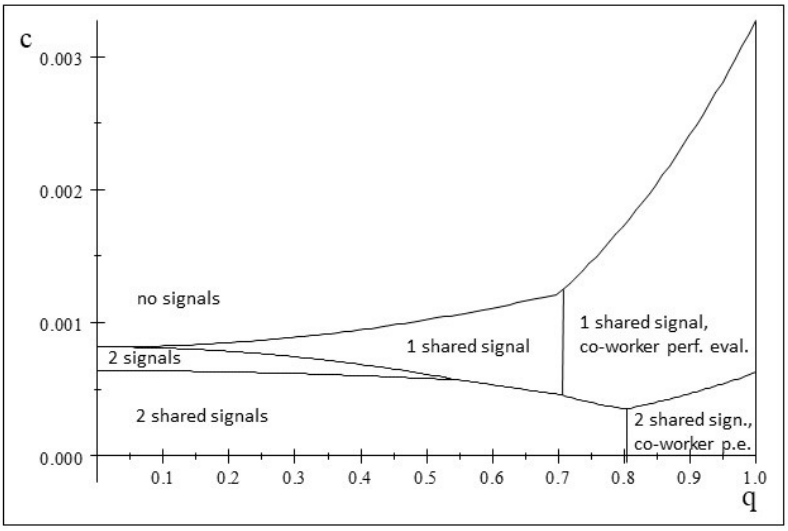

Figure 1: Optimal Linear Contracts

Given that both agents are asked to acquire their signals, it is either more advantageous to also ask each agent i to share s_i so as to improve j's decision precision, or not to incentivize any sharing. Because of symmetry across the players, it cannot be the case that it is optimal to ask one agent to share his signal and the other one not to.

The optimal contracts inducing the remaining four possible course of actions (both agents $i = 1, 2$ acquiring and sharing signals s_i, both agents $i = 1, 2$ acquiring s_i without sharing it, only one agent i acquiring and sharing s_i, and neither agent $i = 1, 2$ acquiring s_i) turn out to maximize the organization's profit in different areas of the parameter space defined by the information acquisition cost c and state correlation r.

The complete characterization is summarized in the proposition that follows, and depicted in Figure 1, which also identifies the areas in which the optimal contracts t_1, t_2 reward information acquisition and sharing by an agent i by making i's remuneration depend mainly on the performance of the other agent.

Proposition 5. *The profit maximizing agents' actions achieved through the optimal linear contracts are as follows.*
1. For research cost $c < c_{22-20}(r)$ and correlation $r < \tilde{r}$, and for $r > \tilde{r}$ and $c < c_{22-11}(r)$, both agents $i = 1, 2$ collect signal s_i and share it with the other agent j.
2. When $r < \tilde{r}$ and $c_{22-20}(r) < r < c_{20-11}(r)$, both agents $i = 1, 2$ collect signal s_i but do not share it with j.
3. When $c_{22-11}(r) < c < c_{11-00}(r)$, for all r, only one agent i collects signal s_i and shares it with j.
4. When $c > c_{11-00}(r)$, for all r, neither agent i collects signal s_i.

We conclude this section by combining Proposition 1-3 and Proposition 5 to present the main result of our analysis. We describe the optimal contracts t_1, t_2 that induce agent(s) i to acquire and share information by making i's payment depend on the performance of the other agent.

Corollary 1. *If the states are sufficiently correlated ($r > r' \approx 0.803$) and signal acquisition cheap enough ($c < c(r)_{22-11}$), then $a_i = 0$ and $\bar{w}_i = b_i > 0$ for both $i = 1, 2$. Each agent $i = 1, 2$ is induced to collect signal s_i and share it with the other agent j with a reward based on the other agent j's performance.*
For sufficient state correlation ($r > \sqrt{\frac{33}{67}} \approx 0.701$) and intermediate signal costs ($c(r)_{22-11} < c < c(r)_{11-00}$), only one agent i is induced to collect s_i and share with j with a reward based j's performance. (The other agent receives a flat payment.)
For all other values of r and c, each agent i is induced to collect s_i (and possibly share s_i with j) only with rewards based on agent i's own performance.

This section has determined the company profit maximizing contracts, and has uncovered an important role for joint performance evaluations. Making one division manager's remuneration depend on the performance of the other division manager may be a very potent incentive for information acquisition and sharing. In the extreme case in which each manager's information is equally useful for both divisions, we have shown that such an incentive is always more potent than remunerating the manager for his own performance.

5 Extensions

Decentralized sequential communication

In the baseline scenario at the beginning of the game the principal designs, and commits to implement decentralized simultaneous communication. Such one-round simultaneous cheap talk communication is the natural benchmark in studying incentives within multi-divisional organizations (see, e.g., Gibbons et al. (2012)). However, it is natural to ask whether the principal could improve her payoff by implementing sequential communication. The mechanism of sequential communication allows the principal to make one of the agents better informed (both about the states and about the signal of the other agent) before that agent decides on whether to acquire and share his signal.

In particular, suppose that one of the agents, say, agent 1 acquires a signal s_1 and then sends a message to agent 2. Only then agent 2 decides whether to acquire a signal and which information to communicate to agent 1. As a result, in the sequential setup agent 2 possesses additional information about agent 1's signal before acquiring and sharing his information.

The next proposition shows that sequential communication does not benefit the principal, compared to the simultaneous communication.

Proposition 6. *For all pairs (r, c) decentralized simultaneous communication generates at least the same expected payoff for the principal as simultaneous sequential communication, and for some pairs (r, c) the simultaneous decentralized communication strictly outperforms decentralized sequential communication.*

When agent 2 listens to agent 1 (and agent 1 is believed to be truthful about his own signal), agent 2 improves his information about local states (and, of course, he also believes to know the signal that agent 1 holds). As a result, however, it can become more expensive for the principal to incentivize agent 2 to acquire a signal. This is because, sometimes, a distortion that a shirking agent 2 expects to impose on agent 1's output need not be as substantial as in the simultaneous case. Therefore, for some parameter pairs (r, c) the principal finds it too costly to motivate complete acquisition and sharing, and therefore prefers to only motivate information acquisition without sharing. In contrast, under the same parameters the allocation under simultaneous communication also involves sharing, which is better for the principal: this is because in the latter case she always has the option to implement an allocation without sharing, but chooses not to do so.

General statistical model

In the following we derive sufficient conditions for the statistical model that generalize our main results. It will become clear that such assumptions are satisfied by most well-behaved statistical models. Suppose that the states (θ_1, θ_2) are distributed on $\mathbb{R}^2$ according to a joint cumulative distribution F. For tractability, consider symmetric distributions i.e., $F(\theta_1, \theta_2) = F(\theta_2, \theta_1)$ for all θ_1, θ_2.

As before, at a cost $c > 0$ each agent i can collect signal $s_i \in \mathbb{R}$ informative of θ_i. The signals $s_i|\theta_i$ are i.i.d. We focus on sufficiently low costs of signal acquisition such that the principal wants to implement an allocation with complete information acquisition and sharing. The principal offers a contract $t_i = w_i - a_i(y_i - \theta_i)^2 - b_i(y_j - \theta_j)^2$ for each agent i. Because the principal wants each agent to set $y_i = E_i[\theta_i|\Omega_i]$ given an agent's information set Ω_i: therefore she *optimally* chooses $a_i \geq 0$ (the argument is exactly the same as the one derived earlier).

Firstly, we make the following minimal assumptions on the distributions of θ_i and $s_i|\theta_i$,

$$\begin{aligned} E[\ell_i(y_i(s_i, s_j))|s_i, s_j] &= E[(E[\theta_i|s_i, s_j] - \theta_i)^2 |s_i, s_j] = EVar(\theta_i|s_i, s_j) \\ &< EVar(\theta_i|s_i) = E[(E[\theta_i|s_i] - \theta_i)^2 |s_i] = E[\ell_i(y_i(s_i))|s_i] \\ &< EVar(\theta_j|s_j) = E[(E[\theta_i|s_j] - \theta_i)^2 |s_j] = E[\ell_i(y_i(s_j))|s_j] \\ &< Var(\theta_j) = E[(E[\theta_i] - \theta_i)^2] = E[\ell_i(y_i(\varnothing))], \end{aligned} \tag{6}$$

where $EVar(\theta_i|s_i, s_j)$ denotes the expectation of the variance $Var(\theta_i|s_i, s_j)$ which is as a function of the random signals s_i and s_j, and $EVar(\theta_i|s_i)$ and $EVar(\theta_j|s_j)$ follow analogous definitions.[13]

Essentially, these conditions require that (1) s_i and s_j are informative of θ_i, and (2) in expectation s_i is more informative than s_j about θ_i – without these conditions, the problem would not be meaningful.

Second, as the final requirement we require the following inequality to hold:

$$EVar[\theta_i|s_j] - EVar[\theta_j|s_j] < E[(E[\theta_j|s_j, s_i] - E[\theta_j|s_j])^2], \text{ for all } s_i. \tag{7}$$

To understand the above requirement, consider the case of perfect correlation, $r = 1$. In this case, the conditional expected variances of the two states θ_i and θ_j, given the signal s_j, are equal: $EVar[\theta_i|s_j] = EVar[\theta_j|s_j]$. At the same time, the RHS of (7) requires that a signal s_i is informative about θ_j, on top of the signal s_j. We note that this must hold for *every* signal realization s_i: i.e., there must not exist a signal realization which is completely uninformative about the underlying state.

The reason for this requirement is that if an agent deviates at the signal acquisition stage and communicates an *uninformative* signal to the other agent, he is unable to impose any harm on the other agent's output. Such distortion, however, is central for our main result that stipulates the optimality of linking each agent's remuneration to the performance of the other agent. In other words, in the case of perfect correlation, if every possible signal realization conveys *some* information about the unobserved state, our main result holds.

Proposition 7. *Suppose that information acquisition costs are sufficiently low, under conditions (6) and (7), the optimal non-linear contract t_1, t_2 to make each agent $i \in \{1, 2\}$ collect information and transmit it to the other agent j in the most informative equilibrium features $a_i = 0$ and $w_i = b_i > 0$. Each agent i's remuneration depends on the other agent j's performance.*[14]

Suppose that the principal wants to incentivize each agent $i \in \{1, 2\}$ to acquire and share his signal. In this case the contract t_1, t_2 must be designed such that agent i does not benefit from deviating and not collecting information. If a deviation results in $E\ell_j$ being higher compared to $E\ell_i$ then – exactly as in Proposition 1 – the principal optimally remunerates agent i based on the other agent j's performance.

We conclude by showing that conditions (6) and (7) are not specific to the statistical model presented in Section 2. For example, suppose that $\theta_1 = \theta_2$ are normally distributed with zero mean and variance $\sigma^2 = 1/\alpha$, i.e., $\theta_1 = \theta_2 \sim N(0, \sigma^2)$. For $i = 1, 2$, the signal $s_i = \theta_i + \varepsilon_i$

[13] In econometrics terms, $EVar(\theta_i|s_i, s_j)$ denotes the variance of the residual of θ_i after it has been regressed on s_i and s_j.

[14] This result follows from simple algebraic permutations showing that (7) implies that $E\ell_i = E[(E(\theta_i|s_j) - \theta_i)^2|s_j] < E\ell_j = E[E(\theta_j|s_j, m_i) - \theta_j)^2|s_j]$, for every possible m_i. The former denotes the expected loss induced by agent i's decision $y_i(s_j) = E(\theta_i|s_j)$ if agent i does not collect a signal s_i but receives a signal s_j from the other agent j. The latter is agent i's expectation of the loss generated by j's decision $y_i(s_j, m_i) = E(\theta_j|s_j, m_i)$ in case j collects signal s_j and mistakenly believes that i collected signal $s_i = m_i$.

is such that $\varepsilon_i \sim N(0, \tau^2)$, $\tau^2 = 1/\beta$, and ε_i is independent of θ_i and ε_j. Then, (6) obviously holds because $Var(\theta_i) = E[\theta_i^2] = \frac{1}{\alpha}$, $EVar(\theta_i|s_i) = E[\theta_i^2|s_i] = \frac{1}{\alpha+\beta} = EVar(\theta_i|s_j) = E[\theta_i^2|s_j]$ and $EVar(\theta_i|s_i, s_j) = E[\theta_i^2|s_i, s_j] = \frac{1}{\alpha+2\beta}$. At the same time, because $\theta_1 = \theta_2$, (7) simplifies to $E[(E[\theta_j|s_j, s_i] - E[\theta_j|s_j])^2] > 0$ for all $s_i \in \mathbb{R}$, and simple algebraic manipulations reported in Appendix B show that this condition always holds.

More general contracts

Linear contracts are not uncommon in organizations. At the same time, it is natural to inquire whether our results are robust to richer contracts. In particular, we extend our baseline model in the simplest possible way and allow for a direct interaction between the agents' losses.

While the extension is simple, we believe that a broader lesson can be learnt from this exercise: the non-linearity of the agents' objectives leads to distorted action choices by the agents *given* their signals (as the agents are not setting their actions equal to the expectations of the respective states). We show that the distortion can harm the principal who then wants to adhere to linear contracts.

Specifically, suppose that for each agent i the principal offers a contract:

$$t_i(\ell_1, \ell_2) = w_i - a_i\ell_i - b_i\ell_2 - g_i\ell_1\ell_2. \tag{8}$$

We continue to assume limited liability, and allow for any realization of the triple $(a_i, b_i, g_i) \in \mathbb{R}^3$ for each agent i. The interaction component provides principal with additional flexibility to incentive the agent to acquire a signal and share it with the other agent.

We focus on the most interesting case of high correlation where our main result states that under the optimal linear scheme the principal should optimally link each agent's pay to the performance of the other agent.

The following proposition shows that if the states are sufficiently correlated then the principal does not benefit from deviating from the linear transfer specification in the baseline model to the non-linear specification of the form (8) provided that the cost of information acquisition is sufficiently low.

Proposition 8. *There exists $\bar{r} < 1$ and $\bar{c} > 0$ such that for sufficiently high correlation, $r > \bar{r}$, and a sufficiently low information acquisition cost, $0 < c < \bar{c}$, the principal cannot increase her payoff under the optimal linear contract case by using the transfer scheme of the form (8).*

As we show in the proof, if the principal offers a contract with $g_i \neq 0$ for each agent i, she may induce a substantial distortion of the agents' actions given the agents' information sets. In particular, if the agents receive signals (and fully share information) that either indicate that in expectation the state is very low, or that it is very high, the optimal profile of actions (y_1, y_2) features a significant distortion relative to the agents' actions under the optimal linear contract. As a result, the upper bound of the principal's expected payoff when using the non-linear contract scheme of the form (8) is below the principal's payoff under optimal linear contracts.

Asymmetric weights on agents.

Suppose that the principal's objective is $\pi = \lambda_1\pi_1 + \lambda_2\pi_2$ where, recall, $\pi_i = 1-\ell_i = 1-(y_i-\theta_i)^2$ for each agent i. The transfer for each agent is the same as before:

$$t_i(\ell_i, \ell_{-i}) = w_i - a_i(y_i-\theta_i)^2 - b_i(y_i-\theta_i)^2,\ i = 1, 2,$$

which means that the conditional optimal transfers remain the same. However, as we show below, for $\lambda_1 \neq \lambda_2$ an asymmetric contract with two signals and one-sided communication is optimal even if the costs of information acquisition are very low.

Proposition 9. *An allocation in which both agents acquire information and only agent j informs agent i dominates all other allocations if for $\lambda_j > \lambda_i$, it is either the case that $r < \bar{r}_1$ and $c \in (\lambda_i \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}, \lambda_j \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)})$, or that $\bar{r}_1 < r < \tilde{r}_2$ and*

$$c \in (\lambda_i \frac{2r^4(3-r^2)^2}{3\,(9-r^2)\,(5r^8-330r^6+2484r^4-7290r^2+4131)}, \lambda_j \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}).$$

First, notice that, different to the case of an unbiased principal ($\lambda_1 = \lambda_2$), whenever the correlation coefficient r is sufficiently low and $\lambda_j > \lambda_i$, there is a range of costs where the principal wants to implement an asymmetric allocation with two signals and one-sided information. The principal wants both agents to obtain a signal and agent i to communicate his signal to the principal's most preferred agent j. The larger the weight that the principal assigns to her favorite agent, relative to the other one, the larger the range of cost which rationalizes the above asymmetric allocation as the most preferred for the principal. For the intermediate values of correlation, the value of communicating a signal increases. Thus, the principal only prefers an asymmetric allocation with two signals and i communicating to j if the gap $\lambda_j - \lambda_i$ is sufficiently large.

Noisy communication

Our baseline model studies all-or-nothing communication: if the principal wants to incentivize information sharing then the agent always completely reveals his signal. One way inquire whether partial or noisy communication could improve principal's payoffs. In this section we study two extensions of the baseline communication game. The first extension allows for both fully revealing messages and a pooling message that allows the agent to withhold his signal, and we allow for mixing. The second (related) extension studies mixed strategy equilibria in the two-message scenario. In both extensions we study symmetric strategies. In both cases we show that adding noise to communication is not beneficial for the principal.

Consider the following extension of the baseline communication game: each agent's message set M contains three messages, $0, 1$ and m_p.[15] Consider the following strategy profile where we assume that the agents' strategies are symmetric.

$$\text{for } s_i = 0: \ \Pr(m_i = m_p) = \mu_1^i \in [0,1], \ \Pr(m_i = 0) = 1 - \mu_1^i,$$

$$\text{for } s_i = 1: \ \Pr(m_i = m_p) = \mu_2^i \in [0,1], \ \Pr(m_i = 1) = 1 - \mu_2^i.$$

As standard in cheap talk games, messages have equilibrium meaning and the players' beliefs upon receiving messages are derived from Bayes rule and agents' strategies. Given the above strategies we observe that the messages 0 and 1 fully reveal the signal. Further, player j's belief about s_i given (m_i, s_j) is obtained by Bayes Rule: $Pr(s_i = 0|m_i, s_j) = \frac{\Pr(s_j)\Pr(s_i=0|s_j)\mu_1^i}{\Pr(s_j)[\Pr(s_i=0|s_j)\mu_1^i + \Pr(s_i=1|s_j)\mu_2^i]}$.

In the Appendix B we show that also under this specification of the communication game, an optimal contract still requires $b_i > 0$ for each player i who has acquired a signal, and is expected to communicate it according to the above strategies. However, under $b_i \geq 0$ the equilibrium strategies can only feature $\mu_1^i = \mu_2^i = 0$ or $\mu_1^i = \mu_2^i = 1$. The first case is, of course, the revealing equilibrium in the baseline model. The second case results in agents always choosing the pooling message. This, however, leads to informational losses for the principal: indeed, the resulting allocation is akin to the allocation in which both agents are expected to acquire information and not share it. Therefore, the modified communication game does not lead to higher expected payoffs for the principal compared to our baseline game.

As an alternative, we consider a simple mixed strategy extension of the baseline communication game as follows. Suppose that each agent i has access to two messages, $\{m_i^1, m_i^2\}$.[16] Each agent randomizes between two messages, m_i^1 and m_i^2, where each signal realization $s_i \in \{0,1\}$ results in a distribution over the message space $\{m_i^1, m_i^2\}$. Since the messages are binary, the signal-contingent distribution over the message space can be summarized by two probabilities: $\mu_1^i = \text{Prob}(m_i^1|s_i = 0)$ and $\mu_2^i = \text{Prob}(m_i^2|s_i = 1)$. We focus on symmetric strategies such that $\mu_1^1 = \mu_1^2$ and $\mu_2^1 = \mu_2^2$.

In the Appendix B we show that the only symmetric equilibrium entails $\mu_1^i = 1 - \mu_2^i$ for each agent i. It is easy to see that the informational losses make the allocation equivalent to an allocation where each agent acquires a signal and does not share it. We obtain the following result.

[15] The same analysis can be conducted with an arbitrarily large message space where *in equilibrium* each message has one of the three meanings.

[16] Again, the same argument can be constructed with a richer message set that is partitioned in two subsets, and where the agent, first, chooses which subset to use, and then uniformly randomizes between all messages within the chosen subset.

Proposition 10. *Partial revelation as specified above can never result in better payoffs for the principal compared to communication that involves full revelation.*

6 Conclusion

This paper sheds light on how to design transfers in order to incentivize acquisition and sharing of information. This question is relevant, for instance, in the context of multi-divisional organizations where division managers operate in local markets and can obtain local information, while the headquarters manager often lacks time and expertise to monitor the actions of local managers. We show that if a division manager expects another manager to acquire and supply information about their own market, and the market characteristics such as consumer preferences are sufficiently correlated, the supplied information has a large impact on the manager's decisions. In such an environment with sufficiently similar characteristics of local markets, the principal wants to incentivize signal acquisition and sharing by optimally choosing a linear contract that links each manager's pay mainly to the performance of the other division.

To our knowledge, this paper is the first to propose endogeneizing the weights that division managers should attach to the different organizational divisions. We provocatively suggest that optimal managerial incentives may involve arrangements entirely different from what is usually assumed. It may be optimal for a company to condition a division manager's transfers mainly on the performance of other divisions, provided that the local states associated with these divisions are sufficiently correlated.

Our analysis is also relevant within the context of personnel economics. One novel conclusion is that the optimal remuneration may require actively discouraging agents from "collusion," that takes the form of revealing each other when they shirk information acquisition assignments.

The model we solved is kept simple with the purpose of presenting our results in the cleanest manner. We believe that the current framework can be extended in multiple ways. One venue is to introduce diversity in skills. Empirical studies show that introduction of team-based compensation in heterogeneous teams leads to a significant improvement of organizational outcomes.[17] The exact mechanism, however, is still an open question and could be addressed through the lens of the optimal contract design in the environment with costless information transmission.

References

AGHION, P. AND J. TIROLE (1997): "Formal and real authority in organizations," *Journal of Political Economy*, 105, 1–29.

[17] See, e.g., Hamilton et al. (2003) and Boning, Ichniowski, and Shaw (2007).

Alonso, R., W. Dessein, and N. Matouschek (2008): "When does coordination require centralization?" *American Economic Review*, 98, 145–79.

——— (2015): "Organizing to adapt and compete," *American Economic Journal: Microeconomics*, 7, 158–87.

Alonso, R., N. Matouschek, and W. Dessein (2010): "Strategic communication: prices versus quantities," *Journal of the European Economic Association*, 8, 365–376.

Angelucci, C. (2017): "Motivating agents to acquire information," *Available at SSRN 2905367*.

Argenziano, R., S. Severinov, and F. Squintani (2016): "Strategic information acquisition and transmission," *American Economic Journal: Microeconomics*, 8, 119–55.

Bol, J. C. (2011): "The determinants and performance effects of managers' performance evaluation biases," *The Accounting Review*, 86, 1549–1575.

Boning, B., C. Ichniowski, and K. Shaw (2007): "Opportunity counts: Teams and the effectiveness of production incentives," *Journal of Labor Economics*, 25, 613–650.

Breuer, K., P. Nieken, and D. Sliwka (2013): "Social ties and subjective performance evaluations: an empirical investigation," *Review of Managerial Science*, 7, 141–157.

Callander, S. and B. Harstad (2015): "Experimentation in federal systems," *The Quarterly Journal of Economics*, 130, 951–1002.

Che, Y.-K. and S.-W. Yoo (2001): "Optimal incentives for teams," *American Economic Review*, 91, 525–541.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 1431–1451.

Deimen, I. and D. Szalay (2019): "Delegated expertise, authority, and communication," *American Economic Review*, 109, 1349–74.

Dessein, W. (2002): "Authority and communication in organizations," *The Review of Economic Studies*, 69, 811–838.

Di Pei, H. (2015): "Communication with endogenous information acquisition," *Journal of Economic Theory*, 160, 132–149.

Gibbons, R., N. Matouschek, and J. Roberts (2012): "Decisions in organizations," in *The Handbook of Organizational Economics*, ed. by R. Gibbons and J. Roberts, Princeton, NJ: Princeton University Press, 373–431.

Gow, I. D., A. S. Wahid, and G. Yu (2018): "Managing reputation: Evidence from biographies of corporate directors," *Journal of Accounting and Economics*, 66, 448–469.

Green, J. R. and N. L. Stokey (1983): "A comparison of tournaments and contracts," *Journal of Political Economy*, 91, 349–364.

Hamilton, B. H., J. A. Nickerson, and H. Owan (2003): "Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation," *Journal of Political Economy*, 111, 465–497.

Hansen, M. T. and B. Von Oetinger (2001): "Introducing T-shaped managers: Knowledge management's next generation," *Harvard Business Review*, 79, 106–117.

Holmström, B. (1999): "Managerial incentive problems: A dynamic perspective," *The Review of Economic Studies*, 66, 169–182.

Holmström, B. and P. Milgrom (1990): "Regulating trade among agents," *Journal of Institutional and Theoretical Economics*, 85–105.

Ichniowski, C. and K. Shaw (2003): "Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices," *Journal of Economic Perspectives*, 17, 155–180.

Itoh, H. (1991): "Incentives to help in multi-agent situations," *Econometrica*, 611–636.

——— (1993): "Coalitions, incentives, and risk sharing," *Journal of Economic Theory*, 60, 410–427.

Kandel, E. and E. P. Lazear (1992): "Peer pressure and partnerships," *Journal of Political Economy*, 100, 801–817.

Kruse, D. L. (1993): "Does profit sharing affect productivity?" Tech. rep., National Bureau of Economic Research.

Lazear, E. P. (1989): "Pay equality and industrial politics," *Journal of Political economy*, 97, 561–580.

Lazear, E. P. and S. Rosen (1981): "Rank-order tournaments as optimum labor contracts," *Journal of Political Economy*, 89, 841–864.

Liu, S. and D. Migrow (2018): "Designing Organizations in Volatile Markets," *Available at SSRN 3275742*.

Meloso, D., S. Nunnari, and M. Ottaviani (2018): "Looking into crystal balls: a laboratory experiment on reputational cheap talk," .

Ottaviani, M. and P. N. Sørensen (2006): "Reputational cheap talk," *The RAND Journal of Economics*, 37, 155–175.

Rantakari, H. (2008): "Governing adaptation," *The Review of Economic Studies*, 75, 1257–1285.

Appendix A: Beliefs Updating

It is useful to see how the players update their beliefs based on obtained signals and received messages. Suppose, first, that only agent 1 obtains a signal. The posterior density of θ_1 given s_1 is obtained via Bayes rule:

$$f(\theta_1|s_1) = \frac{f(\theta_1)f(s_1|\theta_1)}{\int_0^1 f(s_1|\theta_1)d\theta_1}, \quad f(s_1|\theta_1) = \theta^{s_1}(1-\theta_1)^{1-s_1}.$$

Thus, for $s_1 = 0$ the density is $f(\theta_1|s_1) = 2(1-\theta_1)$ with the expected value $E[\theta_1|s_1] = \frac{1}{3}$ and for $s_1 = 1$ the density is $f(\theta_1|s_1) = 2\theta_1$ with the expected value $E[\theta_1|s_1] = \frac{2}{3}$.

Next, suppose that only agent 2 obtains a signal and truthfully communicates it to agent 1. The posterior density of agent 1 is

$$f(\theta_1|s_2) = \frac{f(\theta_1)f(s_2|\theta_1)}{\int_0^1 f(s_2|\theta_1)d\theta_1}, \quad f(s_2|\theta_1) = r\underbrace{\theta_1^{s_2}(1-\theta_1)^{1-s_2}}_{Pr(s_2|\theta_1)|\theta_1=\theta_2} + (1-r)\underbrace{(1/2)}_{Pr(s_2|\theta_1)|\theta_1\neq\theta_2}.$$

The densities and the expected values of θ_1 depending on the realization of $s_2 \in \{0,1\}$ are

$$f(\theta_1|s_2=0) = 1 + r(1-2\theta_1), \quad E(\theta_1|s_2=0) = \frac{3-r}{6},$$
$$f(\theta_1|s_2=1) = 1 - r(1-2\theta_1), \quad E(\theta_1|s_2=1) = \frac{3+r}{6}.$$

Naturally, if $r = 0$ then the posterior $f(\theta_1|s_2)$ is equal to the prior. For $r > 0$ and $s_2 = 0$ ($s_2 = 1$) the posterior puts a larger mass to the left (right) of $\frac{1}{2}$. As r increases, the expected value converges to $\frac{1}{3}$ ($\frac{2}{3}$).

The conditional distributions that agent 1 assigns to the signal realization of agent 2 are

$$\begin{aligned}
\Pr(s_2=1|s_1=1) &= r\Pr(s_2=1|s_1=1,\theta_1=\theta_2) + (1-r)\Pr(s_2=1|s_1=1,\theta_1\neq\theta_2) \\
&= r\frac{2}{3} + (1-r)\frac{1}{2} = \frac{3+r}{6}, \\
\Pr(s_2=0|s_1=1) &= r\frac{1}{3} + (1-r)\frac{1}{2} = \frac{3-r}{6}.
\end{aligned}$$

Suppose that both agents acquire and truthfully communicate their signals. We consider θ_1, the case for θ_2 is symmetric. The density of θ_1 after obtaining s_1 and receiving $m_2 = s_2$ is

$$f(\theta_1|s_1,s_2) = \frac{f(\theta_1,s_1,s_2)}{f(s_1,s_2)} = \frac{f(s_1,s_2|\theta_1)f(\theta_1)}{\int_0^1 f(s_1,s_2|\theta_1)f(\theta_1)d\theta_1}.$$

To derive $f(s_1,s_2|\theta_1)$ notice that the following. First, the ex ante probability of $s_1+s_2 = 0$ and $s_1+s_2 = 1$ (it means when $s_1 = s_2$) is $\frac{1}{3}$, whereas the ex ante probability of both signals being different is $\frac{1}{6}$. To see this notice that $Pr(l|n=2) = \int_0^1 Pr(l|\theta_1,n=2)d\theta_1 = \frac{1}{n+1}$ and that *conditional* on a particular l all sequences of signals which result in the same sum of signals l are equiprobable.

Second, if both states are correlated which happens with probability r, the probability of $l = s_1 + s_2$ is $\frac{n!}{l!(n-l)!}\theta_1^l(1-\theta_1)^{n-l}$. With the converse probability $1-r$ the probability of s_1 is $\theta_1^{s_1}(1-\theta_1)^{1-s_1}$ and the realization of s_2 is independent of θ_1 (and so of s_1) and is equal to $\frac{1}{2}$.

Therefore, for $s_1 + s_2 = l \in \{0, 2\}$ we have

$$f(s_1, s_2|\theta_1) = r \underbrace{[\theta_1^l(1-\theta_1)^{2-l}]}_{Pr(s_1,s_2|(\theta_1,\theta_1=\theta_2))} + (1-r)\underbrace{[\theta_1^{s_1}(1-\theta_1)^{1-s_1}]\frac{1}{2}}_{Pr(s_1,s_2|(\theta_1,\theta_1\neq\theta_2))}$$

and for $s_1 + s_2 = 1$ we have

$$f(s_1, s_2|\theta_1) = r \underbrace{\frac{1}{2}[2\theta_1(1-\theta_1)]}_{Pr(s_1,s_2|(\theta_1,\theta_1=\theta_2))} + (1-r)\underbrace{[\theta_1^{s_1}(1-\theta_1)^{1-s_1}]\frac{1}{2}}_{Pr(s_1,s_2|(\theta_1,\theta_1\neq\theta_2))}$$

The corresponding densities of the posterior are, for $s_1 + s_2 = l \in \{0, 2\}$

$$\begin{aligned}
f(\theta_1|s_1, s_2) &= \frac{r\theta_1^l(1-\theta_1)^{2-l} + (1-r)\theta_1^{s_1}(1-\theta_1)^{1-s_1}\frac{1}{2}}{\int_0^1 r\theta_1^l(1-\theta_1)^{2-l} + (1-r)\theta_1^{s_1}(1-\theta_1)^{1-s_1}\frac{1}{2}d\theta_1} \\
f(\theta_1|s_1, s_2) &= \frac{r\frac{1}{2}2\theta_1(1-\theta_1) + (1-r)\theta_1^{s_1}(1-\theta_1)^{1-s_1}\frac{1}{2}}{\int_0^1[r\frac{1}{2}2\theta_1(1-\theta_1) + (1-r)\theta_1^{s_1}(1-\theta_1)^{1-s_1}\frac{1}{2}]d\theta_1}.
\end{aligned}$$

The calculations for θ_2 are symmetric.

Assume two efforts and truthful communication. The corresponding posteriors, and the expected values for agent 1 (the analysis for agent 2 is analogous) are:

$$\begin{aligned}
f(\theta_1|s_1 = s_2 = 0) &= \frac{r[\frac{2!}{0!(2-0)!}\theta_1^0(1-\theta_1)^{2-0}] + (1-r)[\frac{1!}{0!(1-0)!}\theta_1^0(1-\theta_1)^{1-0}]\frac{1}{2}}{\frac{3+r}{12}} \\
&= \frac{6(1-\theta_1)(1+r-2r\theta_1)}{3+r}, \\
E(\theta_1|s_1 = s_2 = 0) &= \int_0^1 \theta_1 \frac{6(1-\theta_1)(1+r-r2\theta_1)}{3+r}d\theta_1 = \frac{1}{3+r}.
\end{aligned}$$

Further,

$$\begin{aligned}
f(\theta_1|s_1 = 0, s_2 = 1) &= \frac{6(1-\theta_1)(1-r+r2\theta_1)}{3-r}, \quad E(\theta|s_1 = 0, s_2 = 1) = \frac{1}{3-r}. \\
f(\theta_1|s_1 = 1, s_2 = 0) &= \frac{6\theta_1(1+r-2r\theta_1)}{3-r}, \quad E(\theta_1|s_1 = 1, s_2 = 0) = \frac{2-r}{3-r}. \\
f(\theta_1|s_1 = s_2 = 1) &= \frac{6\theta_1(1-r+2r\theta_1)}{3+r}, \quad E(\theta_1|s_1 = s_2 = 1) = \frac{2+r}{3+r}.
\end{aligned}$$

Appendix B: Proofs

Proof of Lemma 1 and Derivation of Equation 3 Recall that under $a_1 \geq 0, a_2 \geq 0$, the optimal actions are $y_i(s_i, s_j) = E(\theta_i|s_i, s_j)$. Given the optimal choices (y_1, y_2) consider the incentives of agent $i = 1, 2$ at the communication stage if he holds a signal s_i. For a common belief that both agents are truthful each agent i chooses an action that matches his posterior of θ_i, given his own private signal s_i and the message from agent j, m_j. The expected payoff is then

$$
\begin{aligned}
u_i(s_i) = w_i - a_i \sum_{s_j=0,1} \Pr(s_j|s_i) E[(E(\theta_i|s_i, s_j) - \theta_i)^2|s_i, s_j] \\
-b_i \sum_{s_j=0,1} \Pr(s_j|s_i) E[(E(\theta_j|s_j, s_i) - \theta_j)^2|s_i, s_j] - c,
\end{aligned}
$$

where the expected losses are conditioned on truthful communication. Because of the symmetry across the agents, $E(\ell_j|s_i, s_j) = E[(E(\theta_j|s_i, s_j) - \theta_j)^2|s_i, s_j] = E[(E(\theta_i|s_i, s_j) - \theta_i)^2|s_i, s_j]$ when $s_i = s_j$. Adding symmetry across signal realizations, it is also the case that $E[(E(\theta_j|s_j, s_i) - \theta_j)^2|s_j, s_i] = E[(E(\theta_i|s_i, s_j) - \theta_i)^2|s_i, s_j]$ when $s_i \neq s_j$. By letting

$$
E[E(\ell_i|s_i, s_j)|s_i] = \sum_{s_j=0,1} \Pr(s_j|s_i) E(\ell_i|s_i, s_j)
$$

we obtain Equation 3.

We now calculate $E[E(\ell_i|s_i, s_j)|s_i]$, taking the case $s_i = 0$. The case $s_i = 1$ is symmetric. We begin by calculating $E(\ell_i|s_i, s_j)$, for $s_j = 0, 1$. As shown in Appendix A,

$$
\begin{aligned}
E(\theta_i|s_i = 0, s_j = 0) &= \frac{1}{3+r}, \quad f(\theta_i|s_i = 0, s_j = 0) = \frac{6(1-\theta_i)(1+r-2r\theta_i)}{3+r}, \\
E(\theta_i|s_i = 0, s_j = 1) &= \frac{1}{3-r}, \quad f(\theta_i|s_i = 0, s_j = 1) = \frac{6(1-\theta_i)(1+r-2r\theta_i)}{3-r}.
\end{aligned}
$$

Substituting in the expected losses definitions and simplifying, we obtain,

$$
\begin{aligned}
E(\ell_i|s_i = 0, s_j = 0) &= \int_0^1 (E(\theta_i|s_i = 0, s_j = 0) - \theta_i)^2 f(\theta_i|s_i = 0, s_j = 0) d\theta_i = \frac{5+2r-r^2}{10(3+r)^2} \\
E(\ell_i|s_i = 0, s_j = 1) &= \frac{5-2r-r^2}{10(3-r)^2}.
\end{aligned}
$$

Substituting the expected losses into the agent's payoff function, together with the conditional posteriors assigned to signal s_j (see Appendix A),

$$
\Pr(s_j = 1|s_i = 1) = \frac{3+r}{6}, \ \Pr(s_j = 0|s_i = 1) = \frac{3-r}{6},
$$

and simplifying, we obtain

$$E[E(\ell_i|s_i,s_j)|s_i] = \sum_{s_j=0,1} \Pr(s_j|s_i)E(\ell_i|s_i,s_j) = \frac{3-r^2}{6(9-r^2)}.$$

If agent i deviates at the communication stage and informs agent j that his signal is $1-s_i$ instead of the true signal s_i he expects the payoff

$$\begin{aligned} u_i^L(s_i) = w_i - a_i \sum_{s_j=0,1} \Pr(s_j|s_i)E[(y_i(s_i,s_j)-\theta_i)^2|s_i,s_j] \\ -b_i \sum_{s_j=0,1} \Pr(s_j|s_i)E[(y_j(s_j,1-s_i)-\theta_j)^2|s_j,s_i]. \end{aligned}$$

Of course, the expression for $\sum_{s_j=0,1} \Pr(s_j|s_i)E[(y_i(s_i,s_j)-\theta_i)^2|s_i,s_j]$ is unchanged. We calculate $\sum_{s_j=0,1} \Pr(s_j|s_i)E[(y_j(s_j,1-s_i)-\theta_j)^2|s_j,s_i]$ assuming that $s_i = 0$ as the case $s_i = 1$ is symmetric. The agents' decisions and densities for $s_j = 1$, and $s_j = 0$, are, respectively:

$$\begin{aligned} y_j(s_j=1,m_i=1) = E(\theta_j|s_j=1,s_i=1) = \frac{2+r}{3+r}, \quad f(\theta_j|s_j=1,s_j=0) = \frac{6\theta_j(1+r-2r\theta_j)}{3-r}, \\ y_j(s_j=0,m_i=1) = \frac{1}{3-r}, \quad f(\theta_j|s_j=s_i=0) = \frac{6(1-\theta_j)(1+r-2r\theta_j)}{3+r}. \end{aligned}$$

Hence, we obtain:

$$\begin{aligned} E[(y_j(s_j=1,1-s_i)-\theta_j)^2|s_j=1,s_i] &= \int_0^1 E(\theta_j|s_j=1,s_i=1)f(\theta_j|s_j=1,s_j=0)d\theta_j \\ &= \frac{15+9r+11r^2+r^3}{10(3-r)(3+r)^2}. \\ E[(y_j(s_j=0,1-s_i)-\theta_j)^2|s_i,s_j] &= \frac{15-9r+11r^2-r^3}{10(3-r)^2(3+r)} \end{aligned}$$

Wrapping up:

$$\begin{aligned} \sum_{s_j=0,1} \Pr(s_j|s_i)E[(y_j(s_j,1-s_i)-\theta_j)^2|s_j,s_i] &= \frac{3-r}{6} \cdot \frac{15+9r+11r^2+r^3}{10(3-r)(3+r)^2} + \frac{3+r}{6} \cdot \frac{15-9r+11r^2-r^3}{10(3-r)^2(3+r)} \\ & \frac{(9+r^2)(3+r^2)}{6(9-r^2)^2}. \end{aligned}$$

The expected deviation payoff can be written as

$$u_i^L(s_i) = w_i - a_i\frac{3-r^2}{6(9-r^2)} - b_i\frac{(9+r^2)(3+r^2)}{6(9-r^2)^2}$$

so that agent 1 does not deviate at the communication stage if

$$w_i - (a_i+b_i)\frac{3-r^2}{6(9-r^2)} \geq w_i - a_i\frac{3-r^2}{6(9-r^2)} - b_i\frac{(9+r^2)(3+r^2)}{6(9-r^2)^2}$$

which implies $b_i \frac{4r^2}{(9-r^2)^2} \geq 0$, or $b_i \geq 0$.
Q.E.D.

Derivation of Equation 4. First, we calculate the expected loss $E(\ell_i|s_j) = E[(E(\theta_i|s_j) - \theta_i)^2|s_j]$, supposing $s_j = 0$ (the case $s_j = 1$ is symmetric). From Appendix A,

$$f(\theta_i|s_j = 0) = 1 + r(1 - 2\theta_i), \quad E(\theta_i|s_j = 0) = \frac{3 - r}{6}.$$

Substituting in the expected loss definition and simplifying, we obtain:

$$E[(E(\theta_i|s_j) - \theta_i)^2|s_j] = \int_0^1 (E(\theta_i|s_j) - \theta_i)^2 f(\theta_i|s_j) d\theta_i = \frac{3 - r^2}{36}.$$

Because, $E(\ell_i|s_j) = \frac{3-r^2}{36}$ is the same regardless of whether $s_j = 0$ or $s_j = 1$, we also obtain that $E[E(\ell_i|s_j)] = \frac{3-r^2}{36}$.

To calculate the expected loss $E[(E(\theta_j|s_j, m_i) - \theta_j)^2|s_j]$, we assume w.l.o.g. that $m_i = 0$. From Appendix A, the conditional densities and expected values are:

$$\begin{aligned} f(\theta_j|s_j = 0) &= 2(1 - \theta_j) \quad E(\theta_j|s_j = 0, m_i = 0) = \frac{1}{3 + r}, \\ f(\theta_j|s_j = 1) &= 2\theta_j \quad E(\theta_j|s_j = 1, m_i = 0) = \frac{2 - r}{3 - r}, \end{aligned}$$

the consequent expected losses are:

$$\begin{aligned} E[(E(\theta_j|s_j = 0, m_i = 0) - \theta_j)^2|s_j = 0] &= \int_0^1 (E(\theta_j|s_j = 0, m_i = 0) - \theta_j)^2 f(\theta_j|s_j = 0) d\theta_j \\ &= \frac{3 + 2r + r^2}{6(3 + r)^2}. \\ E[(E(\theta_j|s_j = 1, m_i = 0) - \theta_j)^2|s_j = 1] &= \frac{3 - 2r + r^2}{6(3 - r)^2}. \end{aligned}$$

Plugging the expected loss formulas into the unconditional loss formula, we obtain:

$$\sum_{s_j=0,1} \Pr(s_j) E[(E(\theta_j|s_j = 0, m_i = 0) - \theta_j)^2|s_j = 0] = \frac{27 + r^4}{6(9 - r^2)^2}.$$

Appropriate rearranging yields Equation 4.
Q.E.D.

Proof of Proposition 1. Program 5 is obtained through the same process of simplification used to obtain equation 3, and based on the model's symmetry across agents and signal realizations. By linearity of the objective function and the information acquisition constraint $u_i \geq u_i^D$, the solution to the program 5 involves either $a_i > 0$, $b_i = 0$, or $a_i = 0, b_i > 0$. In each case $w_i = a_i + b_i$ and the constraint $u_i \geq u_i^D$ binds.

If $w_i = a_i > 0$, and $b_i = 0$, then the constraint $u_i = u_i^D$ becomes $a_i = \frac{36c(9-r^2)}{(3-r^2)^2}$ and the expected transfer to agent i results in

$$w_i - (a_i + b_i)E[E(\ell_i|s_i, s_j)] = \frac{36c(9-r^2)}{(3-r^2)^2}\left(1 - \frac{3-r^2}{6(9-r^2)}\right) = \frac{6c(51-5r^2)}{(3-r^2)^2},$$

and we verify the ex-ante agent i's participation constraint, because: $w_i - (a_i + b_i)E[E(\ell_i|s_i, s_j)] - c = c\frac{(9-r^2)(r^2+33)}{(3-r^2)^2} \geq 0$ for all $c \geq 0$.

If $a_i = 0, w_i = b_i > 0$, then the constraint $u_i = u_i^D$ becomes $b_i\frac{2r^2}{(9-r^2)^2} = c$. The expected transfer to agent i becomes:

$$w_i - (a_i + b_i)E[E(\ell_i|s_i, s_j)] = \frac{(9-r^2)^2 c}{2r^2}(1 - \frac{3-r^2}{6(9-r^2)}) = \frac{c(9-r^2)(51-5r^2)}{12r^2},$$

and again, we verify $w_i - (a_i + b_i)E[E(\ell_i|s_i, s_j)] - c \geq 0$ for all $c \geq 0$.

By comparing the two cases, we have $\frac{36c(9-r^2)}{(3-r^2)^2} < (>)\frac{(9-r^2)^2 c}{2r^2}$ for $r < (>) r_1$ where

$$r_1 \equiv \sqrt{5 - 10\sqrt[3]{\frac{2}{9\sqrt{29}-43}} + 2^{2/3}\sqrt[3]{9\sqrt{29}-43}} \approx 0.803.$$

As a result, in the case of $r > r_1$ the expected principal's payoff is:

$$\begin{aligned} E\pi_{22} &= E[2\,(1 - E(\ell_i|s_i, s_j))] - 2\bar{w}_i - 2(a_i + b_i)E[E(\ell_i|s_i, s_j)] \\ &= 2 - \frac{3-r^2}{3(9-r^2)} - \frac{(51-5r^2)(9-r^2)}{6r^2}c. \end{aligned}$$

When $r < r_1$, the expected principal's payoff is:

$$E\pi_{22} = 2 - \frac{3-r^2}{3(9-r^2)} - \frac{12(51-5r^2)}{(3-r^2)^2}c.$$

Q.E.D.

Proof of Proposition 2. We distinguish 2 cases.

Case '21': Two-sided acquisition and one-sided sharing.

We calculate the optimal linear contracts t_1, t_2 to induce both agents $i = 1, 2$ to acquire information and only agent, say 1, to transmit it to the other agent. First note that, again, each agent $i = 1, 2$ is motivated to choose decision y_i so as to minimize the loss $\ell_i = (y_i - \theta_i)^2$ by setting $a_i \geq 0$. Likewise, $b_1 \geq 0$ is needed so that 1 reports s_1 truthfully to 2.

The principal's cost minimization problem is:

$$\min_{\substack{a_2\geq 0,\\ b_1\geq 0,\ w_i\geq a_i+b_i}} \quad w_1-(a_1+b_1)E[E(\ell_1|s_1)]+w_2-(a_2+b_2)E[E(\ell_2|s_2,s_1)], \text{ s.t. } u_i\geq u_i^D. \tag{9}$$

Let us consider agent 2's information acquisition stage constraint $u_2 \geq u_2^D$. The equilibrium payoff of agent 2 is, using $m_1 = s_1$,

$$\begin{aligned} u_2 &= w_2 - a_2E[(y_2(s_2,s_1)-\theta_2)^2] - b_2E[(y_1(s_1)-\theta_1)^2] - c \\ &= w_2 - a_2\frac{3-r^2}{6(9-r^2)} - b_2\frac{1}{18} - c. \end{aligned}$$

If agent 2 deviates at the information acquisition stage, her payoff is

$$\begin{aligned} u_2^D &= w_2 - a_2E[(y_2(s_1)-\theta_2)^2] - b_2E[(y_1(s_1)-\theta_1)^2] \\ &= w_2 - a_2\frac{3-r^2}{36} - b_2\frac{1}{18}. \end{aligned}$$

using $$\begin{aligned} E[(y_2(s_1)-\theta_2)^2] &= \sum_{s_1=0,1}\frac{1}{2}\int_0^1 (E[\theta_2|s_1]-\theta_2)^2 f(\theta_2|s_1)d\theta_2 \\ &= \int_0^1 (E[\theta_2|s_1=0]-\theta_2)^2 f(\theta_2|s_1=0)d\theta_2 = \frac{3-r^2}{36}, \end{aligned}$$

because $f(\theta_2|s_1=0) = 1 + r(1-2\theta_2)$ and $E(\theta_2|s_1=0) = \frac{3-r}{6}$ (see Appendix A).

The constraint $u_2 \geq u_2^D$ is thus: $a_2\frac{1}{36}\frac{(3-r^2)^2}{9-r^2} \geq c$. This yields the optimal contract for agent 2: $w_2 = a_2 = 36\frac{9-r^2}{(3-r^2)^2}c$ and $b_2 = 0$. The agent's ex-ante participation constraint is satisfied as an equality.

Then, we consider the optimal contract of agent 1. Again, we note that he does not deviate from truthtelling if and only if $b_1 \geq 0$. Turning to the information acquisition constraint, we note that the equilibrium payoff of agent 1 is:

$$\begin{aligned} u_1 &= w_1 - a_1E[(y_1(s_1)-\theta_1)^2] - b_1E[(y_2(s_2,m_1)-\theta_2)^2] - c \\ &= w_1 - a_1\frac{1}{18} - b_1\frac{3-r^2}{6(9-r^2)} - c, \end{aligned}$$

using $$\begin{aligned} E[(y_1(s_1)-\theta_1)^2] &= \sum_{s_1=0,1}\frac{1}{2}\int_0^1 (E(\theta_1|s_1)-\theta_1)^2 f(\theta_1|s_1)d\theta_1 \\ &= \int_0^1 (E(\theta_1|s_1=0)-\theta_1)^2 f(\theta_1|s_1=0)d\theta_1 = \frac{1}{18}. \end{aligned}$$

If agent 1 deviates and does not acquire information, then his payoff is:

$$u_1^D = w_1 - a_1E[(y_1-\theta_1)^2] - b_1E[E[(y_2(s_2,m_1)-\theta_2)^2|s_2]]$$

$$w_1 - a_1\frac{1}{12} - b_1\frac{27+r^4}{6(9-r^2)^2},$$

using $E[(y_1-\theta_1)^2] = \int_0^1 (E(\theta_1)-\theta_1)^2 f(\theta_1)d\theta_1 = \frac{1}{12}$.

This yields the information acquisition constraint: $a_1\frac{1}{36} + b_1\frac{2r^2}{(9-r^2)^2} \geq c$.

Because the principal objective function is linear in w_1, a_1 and b_1, and such that $w_1 \geq 0, a_1 \geq 0, b_1 \geq 0$, there are two possibilities: either $w_1 = a_1 > 0$ and $b_1 = 0$, or $a_1 = 0, w_1 = b_1 > 0$.

In the first case, $w_1 = a_1 > 0$ and $b_1 = 0$, the constraint $u_1 \geq u_1^D$ becomes $a_1 \geq 36c$. Using $w_1 = a_1 = 36c$ the expected transfer to agent 1 becomes:

$$a_1\left[1 - E[(y_1(s_1)-\theta_1)^2]\right] = 34c.$$

In the second case, $a_1 = 0, w_1 = b_1 > 0$, and the constraint $u_1 \geq u_1^D$ becomes $b_1 \geq \frac{(9-r^2)^2 c}{2r^2}$. Using $w_1 = b_1 = \frac{(9-r^2)^2 c}{2r^2}$ and $a_1 = 0$, the expected transfer to agent 1 becomes:

$$b_1\left(1 - E[(y_2(s_2,s_1)-\theta_2)^2]\right) = \frac{(9-r^2)(51-5r^2)c}{12r^2}.$$

In either case, the ex-ante participation constraint is satisfied.

Now, we have

$$34 < (>)\frac{(9-r^2)(51-5r^2)c}{12r^2} \text{ if and only if } r < (>) r_2 = \sqrt{\frac{3}{5}(84-\sqrt{6801})}$$

As a result, we conclude that for $r < r_2$ the principal optimally chooses $w_1 = a_1 > 0$ and $b_1 = 0$. Otherwise, for $r \geq r_2$ the principal optimally chooses $a_1 = 0, w_1 = b_1 > 0$. The principal's payoff is

$$E\pi_{21} = 2 - \frac{9-2r^2}{9(9-r^2)} - \min\left\{36\left(1-\frac{1}{18}\right), \frac{(9-r^2)^2}{2r^2}\left(1-\frac{3-r^2}{6(9-r^2)}\right)\right\}c - \frac{6(51-5r^2)}{(3-r^2)^2}c.$$

Case '20': Two-sided acquisition and no sharing.

We calculate the optimal contract when the principal wants to incentivize both agents $i = 1,2$ to acquire and not to share their signal s_i. The principal's cost minimization problem is:

$$\min_{\substack{a_i \geq 0,\\ w_i \geq a_i+b_i}} \bar{w}_i - (a_i+b_i)E[(y_i(s_1)-\theta_i)^2] = \min_{\substack{a_i \geq 0,\\ \bar{w}_i \geq a_i+b_i}} \bar{w}_i - (a_i+b_i)\frac{1}{18}, \text{ s.t. } u_i \geq u_i^D.$$

For each agent i the information acquisition stage constraint $u_i \geq u_i^D$ becomes

$$w_i - a_i\frac{1}{18} - b_i\frac{1}{18} - c \geq w_i - a_i\frac{1}{12} - b_i\frac{1}{18}$$

resulting in $w_i = a_i \geq 36c$. The transfer paid to each agent i is $36c(1-\frac{1}{18})$. The agents' ex-ante participation constraint is satisfied as an equality. The principal's payoff is:

$$E\pi_{20} = 2 - \frac{1}{9} - 2a_i(1 - \frac{1}{18}) = 2 - \frac{1}{9} - 2 \cdot 36c(1 - \frac{1}{18}).$$

Q.E.D.

Proof of Proposition 3 and case '00': We distinguish 3 cases.

Case '11': One-sided acquisition and sharing.

We calculate the optimal linear contracts t_1, t_2 to induce agent 1 to acquire signal s_1 and share with the other agent 2, and agent 2 to not acquire information. The optimal contract of agent 2 is, trivially, $t_2(\ell_1, \ell_2) = 0$ for all ℓ_1 and ℓ_2 (the optimal linear contract is such that $w_2 = a_2 = b_2 = 0$). The principal's cost minimization problem for agent 1 is:

$$\min_{\substack{a_1 \geq 0, \geq b_1 \geq 0, \\ w_1 \geq a_1 + b_1}} w_1 - a_1 E[(y_1(s_1) - \theta_1)^2] - b_1 E[(y_2(s_1) - \theta_2)^2], \text{ s.t. } u_1 \geq u_1^D.$$

The equilibrium payoff of agent 1 is:

$$\begin{aligned} u_1 &= w_1 - a_1 E[(y_1(s_1) - \theta_1)^2] - b_1 E[(y_2(s_1) - \theta_2)^2] - c \\ &= w_1 - a_1 \frac{1}{18} - b_1 \frac{3 - r^2}{36} - c \end{aligned}$$

If agent 1 does not acquire information, he still sends a message m_1 to agent 2 who mistakenly believes that $m_1 = s_1$. The equilibrium payoff of agent 1 is:

$$\begin{aligned} u_1^D &= w_1 - a_1 E[(y_1 - \theta_1)^2] - b_1 E[(y_2(m_1) - \theta_2)^2] \\ &= w_1 - a_1 \frac{1}{12} - b_1 \frac{3 + r^2}{36}, \end{aligned}$$

using $E[(y_2(m_1) - \theta_2)^2] = \int_0^1 [(E[\theta_2 | m = 0] - \theta_2)^2] f(\theta_2) d\theta_2 = \frac{3 + r^2}{36}$.

Here, the information acquisitions constraint is:

$$-a_1 \frac{1}{18} - b_1 \frac{3 - r^2}{36} - c \geq -a_1 \frac{1}{12} - b_1 \frac{3 + r^2}{36}$$

So, the optimal contract is either $w_1 = a_1 = 36c$ and $b_1 = 0$, or $w_1 = b_1 = \frac{18c}{r^2}$ and $a_1 = 0$.

In case $w_1 = a_1 = 36c$ and $b_1 = 0$, the expected transfer to agent 1 is

$$a_1 \left[1 - E[(y_1(s_1) - \theta_1)^2]\right] = 34c,$$

and the participation constraint is met with equality.

In case, $a_1 = 0, w_1 = b_1 = \frac{18c}{r^2}$, the expected transfer to agent 1 is

$$(a_1 + b_1) \left[1 - E[(y_1(s_1) - \theta_1)^2]\right] = \tfrac{18c}{r^2} \left(1 - \tfrac{3 - r^2}{36}\right).$$

As a result, the principal's payoff is

$$E\pi_{11} = 2 - \frac{1}{18} - \frac{1}{36}(3 - r^2) - \min\left\{34, \frac{18}{r^2}(1 - \frac{1}{36}(3 - r^2))\right\} c.$$

Equating $34 = \frac{18}{r^2}(1 - \frac{1}{36}(3 - r^2))$, we obtain the admissible solution $r_3 := \sqrt{\frac{33}{67}}$.

Case '10': One-sided acquisition and no sharing.
We calculate the optimal linear contract to induce agent 1 to acquire signal s_1 and not to share it, and agent 2 not to acquire information. The optimal contract for agent 2 is $t_2(\ell_1, \ell_2) = 0$. The principal's cost minimization problem for agent 1 is:

$$\min_{\substack{a_1 \geq 0 \\ w_1 \geq a_1 + b_1}} w_1 - a_1 E[(y_1(s_1) - \theta_1)^2] - b_1 E[(y_2 - \theta_2)^2], \quad \text{s.t. } u_1 \geq u_1^D.$$

The equilibrium payoff of agent 1 is:

$$u_1 = w_1 - a_1 E[(y_1(s_1) - \theta_1)^2] - b_1 E[(y_2 - \theta_2)^2] = w_1 - a_1 \frac{1}{18} - b_1 \frac{1}{12} - c,$$

his deviation payoff at the information acquisition stage is:

$$u_1^D = w_1 - a_1 E[(y_1 - \theta_1)^2] - b_1 E[(y_2 - \theta_2)^2] = w_1 - a_1 \frac{1}{12} - b_1 \frac{1}{12},$$

so that the incentive compatibility constraint is: $a_1 \geq 36c$, and the optimal linear contract is $w_1 = a_1 = 36c$ and $b_1 = 0$. The principal's expected profit is:

$$E\pi_{10} = 2 - \frac{1}{18} - \frac{1}{12} - 36(1 - \frac{1}{18})c.$$

Case '00': No acquisition.
The optimal linear contracts t_1, t_2 in the case that both agents $i = 1, 2$ are not supposed to acquire information are such that $w_i = a_i = b_i = 0$. This leads to expected principal's profit:

$$E\pi_{00} = 2 - 2\frac{1}{12}.$$

Q.E.D.

Proof of Proposition 4. By the same logic as in Propositions 1 - 3, an agent i who is not expected to acquire a signal s_i, gets zero transfer: $w_i = 0, a_i = 0, b_i = 0$. An agent i expected to acquire a signal s_i, but not to share it, has a transfer linked only to his own performance: $w_i = a_i > 0, b_i = 0$. In the following we study the optimal transfer to any agent i who is expected to acquire and share his signal s_i.

Case '22': Two-sided acquisition and sharing.

The principal's cost minimization program is still (5): it needs to be that $a_i \geq 0$ for either agent to optimal choose $y_i(s_i, m_j)$ and that $b_i \geq 0$ to ensure truthful communication. The expected payoff of agent i if not acquiring signal s_i is:

$$u_i^D = w_i - a_i E[E(\ell_i|s_j)] - b_i E(\ell_i) = w_i - a_i \frac{3-r^2}{36} - b_i \frac{1}{18},$$

because if i does not acquire signal s_i, then she reveals that to j. The agent's equilibrium expected payoff u_i is unchanged, and hence the information acquisition constraint $u_i \geq u_i^D$ takes the following form: $a_i \frac{(3-r^2)^2}{36(9-r^2)} + b_i \frac{r^2}{81-9r^2} \geq c$. The principal's program has a linear objective function and a linear constraint. Because the coefficient $\frac{(3-r^2)^2}{36(9-r^2)}$ of a_i is larger than the coefficient $\frac{r^2}{81-9r^2}$ of b_i in the constraint $u_i \geq u_i^D$, and the two choice variables have the same coefficient in the objective function (5), the optimal contract is such that $b_i = 0$ and $a_i > 0$. Solving out $u_i = u_i^D$, we obtain $a_i = \frac{36c(9-r^2)}{(3-r^2)^2}$.

Case '21:' Two-sided acquisition and one-sided sharing.

The principal's cost minimization problem is still (9). The incentive constraint that prevents agent 1 to deviate at the information acquisition stage and report the lack of the signal to agent 2, is

$$w_1 - a_1 \frac{1}{18} - b_1 \frac{3-r^2}{6(9-r^2)} - c \geq w_1 - a_1 \frac{1}{12} - b_1 \frac{1}{18}$$

that can be rewritten as $a_1 \frac{1}{36} + b_1 \frac{r^2}{81-9r^2} \geq c$. The principal's program has a linear objective function and a linear constraint. The ratio between the coefficients of a_1 and b_1 in the (binding) constraint $u_1 = u_1^D$ is $\rho_{ab} = \frac{81-9r^2}{36r^2}$, whereas the same ratio in the objective function is $\bar{\rho}_{ab} = \frac{6(9-r^2)}{18(3-r^2)}$. Because $\bar{\rho}_{ab} < \rho_{ab}$, the optimal contract is such that $b_i = 0$ and $a_i > 0$. Solving out $u_1 = u_1^D$, we obtain: $a_1 = 36c$.

Case '10:' One-sided acquisition and sharing.

The principal's cost minimization problem is still (9). The incentive constraint $u_1 = u_1^D$ that prevents agent 1 to deviate at the information acquisition stage and report the lack of the signal to agent 2, is

$$\bar{w}_1 - a_1 \frac{1}{18} - b_1 \frac{3-r^2}{36} - c \geq \bar{w}_1 - a_1 \frac{1}{12} - b_1 \frac{1}{12}$$

that can be expressed as $a_1 \frac{1}{36} + b_1 \frac{r^2}{36} \geq c$. The ratio between the coefficients of a_1 and b_1 in the (binding) constraint $u_1 = u_1^D$ is $\rho_{ab} = \frac{1}{r^2}$, whereas the same ratio in the objective function is $\bar{\rho}_{ab} = \frac{2}{3-r^2}$. Because $\bar{\rho}_{ab} < \rho_{ab}$, the optimal contract is such that $b_i = 0$ and $a_i > 0$. Solving out $u_1 = u_1^D$, we obtain: $a_1 = 36c$.

Q.E.D.

Proof of Lemma 2: Subtracting the formulas of $E\pi_{11}(r,c)$ and $E\pi_{10}(r,c)$ and rearranging, we obtain

$$E\pi_{11} - E\pi_{10} = \frac{5}{36} - \frac{1}{36}(5 - r^2) + 34c - \min\left\{34, \frac{1}{2}\frac{r^2+33}{r^2}\right\} c,$$

which is obviously strictly positive.

Q.E.D.

Proof of Lemma 3: We first compare $E\pi_{22}(r,c)$ and $E\pi_{21}(r,c)$ and consider

$$\begin{aligned} E\pi_{22}(r,c) - E\pi_{21}(r,c) &= \frac{9-2r^2}{9(9-r^2)} - \frac{3-r^2}{3(9-r^2)} - D_{22-21}(q)c \\ &= \frac{q^2}{9(9-r^2)} - D_{22-21}(r)c \geq -D_{22-21}(r)c. \end{aligned}$$

where

$$D_{22-21}(r) = \min\left\{\frac{12}{(3-r^2)^2}, \frac{(9-r^2)}{6q^2}\right\}(51-5r^2) - \min\left\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\right\} - 6\frac{51-5r^2}{(3-r^2)^2}.$$

Calculations omitted for brevity show that $D_{22-21}(r) > 0$ for $0 \leq r < r_1$ and $D_{22-21}(r) < 0$ for $r_1 < r \leq 1$. We obtain that for $0 \leq r < r_1$, whether $E\pi_{22}(r,c)$ is larger or smaller than $E\pi_{21}(r,c)$ depends on whether c is smaller or larger than a strictly positive threshold $c_{22-21}(r)$ implicitly defined by the equation $E\pi_{22}(r,c) = E\pi_{21}(r,c)$, whereas for $r_1 \leq r \leq 1$ it is the case that $E\pi_{22}(r,c) > E\pi_{21}(r,c)$ for all c.

To complete the proof we show that, for almost all c and $0 \leq r \leq r_1$ it is either the case that $E\pi_{22}(r,c) > E\pi_{21}(r,c)$ or that $E\pi_{20}(r,c) > E\pi_{21}(r,c)$. We begin by noting that the functions $E\pi_{22}(r,c)$, $E\pi_{21}(r,c)$ and $E\pi_{20}(r,c)$ are all linear in c, and that $E\pi_{22}(r,c) > E\pi_{21}(r,c) > E\pi_{20}(r,c)$ for $c = 0$. As a result, we can proceed by comparing the threshold functions

$$c_{22-21}(r) = \frac{\frac{1}{18} - \frac{3-r^2}{3(9-r^2)} + \frac{3-r^2}{6(9-r^2)}}{\min\left\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6r^2}\right\}(51-5r^2) - \min\left\{34, \frac{9-r^2}{12r^2}(51-5r^2)\right\} - 6\frac{51-5r^2}{(r^2-3)^2}}$$

$$c_{21-20}(r) = \frac{\frac{1}{18} - \frac{3-r^2}{6(9-r^2)}}{\min\left\{34, \frac{9-r^2}{12r^2}(51-5r^2)\right\} + 6\frac{51-5r^2}{(r^2-3)^2} - 68}$$

implicitly defined by the equations $E\pi_{22}(r,c) = E\pi_{21}(r,c)$ and $E\pi_{21}(r,c) = E\pi_{20}(r,c)$, respectively. In fact, for any (r,c) such that $c < c_{22-21}(r)$, it is the case that $E\pi_{22}(r,c) > E\pi_{21}(r,c)$, and for any (r,c) such that $c > c_{21-20}(r)$, it is the case that $E\pi_{21}(r,c) < E\pi_{20}(r,c)$.

Calculations omitted for brevity show that $c_{22-21}(r) \geq c_{21-20}(r)$ for all $0 \leq r \leq r_1$. This completes the proof of the Lemma, because it implies that for almost all c and $0 \leq r \leq r_1$, it is either the case that $E\pi_{22}(r,c) > E\pi_{21}(r,c)$ or that $E\pi_{20}(r,c) > E\pi_{21}(r,c)$.

Q.E.D.

Proof of Proposition 5: We need to compare the profit functions $E\pi_{22}(r,c)$, $E\pi_{20}(r,c)$, $E\pi_{11}(r,c)$ and $E\pi_{00}(r,c)$. To determine the area in which $E\pi_{22}(r,c)$ is the largest, we note that all the profit functions are linear in c, and that $E\pi_{22}(r,c) > E\pi_{20}(r,c) > E\pi_{11}(r,c) > E\pi_{00}(r,c)$ for $c = 0$ and all r. As a result, we can proceed by comparing the threshold functions

$$c_{22-20}(r) = \frac{\frac{1}{9} - \frac{3-r^2}{3(9-r^2)}}{\min\left\{\frac{12}{(r^2-3)^2}, \frac{9-r^2}{6q^2}\right\}(51-5r^2) - 68}$$

$$c_{22-11}(r) = \frac{\frac{5-r^2}{36} - \frac{3-r^2}{3(9-r^2)}}{\min\left\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6r^2}\right\}(51-5r^2) - \min\left\{34, \frac{1}{2}\frac{r^2+33}{r^2}\right\}}$$

$$c_{22-00}(r) = \frac{\frac{1}{6} - \frac{3-r^2}{3(9-r^2)}}{\min\left\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6r^2}\right\}(51-5r^2)}$$

implicitly defined by the equations $E\pi_{22}(r,c) = E\pi_{20}(r,c)$, $E\pi_{22}(r,c) = E\pi_{11}(r,c)$ and $E\pi_{22}(r,c) = E\pi_{00}(r,c)$. For any such a threshold function $c_{22-(\cdot)}(r)$, and any value $r \in [0,1]$ for which $c_{22-(\cdot)}(r)$ is positive, it is the case that $E\pi_{22}(r,c) > E\pi_{(\cdot)}(r,c)$ if and only if $c < c_{22-(\cdot)}(r)$. Instead, for all r such that $c_{22-(\cdot)}(r) < 0$, it is the case that $E\pi_{22}(r,c) > E\pi_{(\cdot)}(r,c)$ for all c.

Calculations omitted for brevity prove that $c_{22-20}(r) > 0$ if and only if $r < \sqrt{\frac{252}{5} - \frac{3\sqrt{3}}{5}\sqrt{2267}}$, and that $c_{22-11}(r) > 0$ and $c_{22-00}(r) > 0$ for all $r \in [0,1]$. Further, comparing $c_{22-11}(r)$ and $c_{22-00}(r)$, omitted calculations show that $c_{22-11}(r) < c_{22-00}(r)$ for all $r \in [0,1]$, and that $c_{22-20}(r) < c_{22-11}(r)$ if and only if $r < \tilde{r} \approx 0.553$ on the relevant range $r \in [0, \sqrt{\frac{(28\sqrt{3}-\sqrt{2267})3\sqrt{3}}{5}}]$. The implication is that $E\pi_{22}(r,c) > \max\{E\pi_{20}(r,c),\ E\pi_{11}(r,c),\ E\pi_{00}(r,c)\}$ for every $c < c_{22-20}(r)$ for $r < \tilde{r}$ and for every $c < c_{22-11}(r)$ for $r > \tilde{r}$.

Likewise, to determine the area in which $E\pi_{00}(r,c)$ is larger than $E\pi_{22}(r,c)$, $E\pi_{20}(r,c)$ and $E\pi_{11}(r,c)$, we note that $E\pi_{00}(r,c) > \max\{E\pi_{22}(r,c), E\pi_{20}(r,c), E\pi_{11}(r,c)\}$ for $c \to \infty$ and all r. As a result, we can proceed by comparing the threshold function $c_{22-00}(r)$ reported above with the threshold functions

$$c_{11-00}(r) = \frac{r^2+1}{36\min\left\{34, \frac{1}{2r^2}(r^2+33)\right\}} \text{ and } c_{20-00}(r) = \frac{1}{1224},$$

implicitly defined by the equations $E\pi_{20}(r,c) = E\pi_{00}(r,c)$ and $E\pi_{11}(r,c) = E\pi_{00}(r,c)$. Omitted calculations show that, for all $r \in [0,1]$ all the functions $c_{22-20}(r)$, $c_{22-11}(r)$ and $c_{22-00}(r)$ are strictly positive. Hence, for every $r \in [0,1]$, it is the case that $E\pi_{00}(r,c) > \max\{E\pi_{22}(r,c),\ E\pi_{20}(r,c),\ E\pi_{11}(r,c)\}$ for every $c > \max\{c_{22-00}(r), c_{20-00}(r), c_{11-00}(r)\}$. Comparing $c_{22-00}(r)$, $c_{20-00}(r)$ and $c_{11-00}(r)$, omitted calculations show that $c_{11-00}(r) > c_{22-00}(r)$, and $c_{11-00}(r) > c_{20-00}(r)$ for all $r \in [0,1]$. The implication is that $E\pi_{00}(r,c) > \max\{E\pi_{22}(r,c),\ E\pi_{20}(r,c),\ E\pi_{11}(r,c)\}$ for every r and

$c > c_{11-00}(r)$.

For any r and cost c values that are below $c_{11-00}(r)$ and either above $c_{22-20}(r)$, for $r < \tilde{r}$, or above $c_{22-11}(r)$, for $r > \tilde{r}$, it is either the case that $E\pi_{20}(r,c)$ or $E\pi_{11}(r,c)$ is the highest profit function. Because $E\pi_{20}(r,c) > E\pi_{11}(r,c)$ for $c = 0$ and all r, this is once more determined by considering a threshold function:

$$c_{20-11}(r) = \frac{1 - r^2}{2448 - 36 \min\left\{34, \frac{r^2+33}{2r^2}\right\}},$$

implicitly defined by the equation $E\pi_{20}(r,c) = E\pi_{11}(r,c)$. Because $2448 - 36 \cdot 34 = 1224$, the threshold function $c_{20-11}(r)$ is strictly positive for all $r \in [0,1]$. Hence, for all r it is the case that $E\pi_{20}(r,c) > E\pi_{11}(r,c)$ if and only if $c < c_{20-11}(r)$.

Comparing $c_{20-11}(r)$ with $c_{22-20}(r)$, $c_{22-11}(r)$ and $c_{11-00}(r)$, omitted calculations show that $c_{20-11}(r) = c_{11-00}(r)$ for $r = 0$, that $c_{20-11}(r) < c_{11-00}(r)$ for all $r > 0$, that $c_{20-11}(r) > c_{22-20}(r)$ for $0 \leq r < \tilde{r}$, that $c_{20-11}(r) = c_{22-20}(r) = c_{22-11}(r)$ for $r = \tilde{r}$ and that $c_{20-11}(r) < c_{22-11}(r)$ for $\tilde{r} < r \leq 1$.

This concludes the proof of the Proposition. We have derived the result depicted in Figure 1: For $0 < r < \tilde{r}$ and $c < c(r)_{22-20}$, and for $\tilde{r} < r \leq 1$ and $r < c(r)_{20-11}$, it is the case that $E\pi_{22}(r,c) > \max\{E\pi_{20}(r,c),\ E\pi_{11}(r,c),\ E\pi_{00}(r,c)\}$. For $0 < r < \tilde{r}$ and $c(r)_{22-20} < c < c(r)_{20-11}$, $E\pi_{20}(r,c) > \max\{E\pi_{22}(r,c),\ E\pi_{11}(r,c),\ E\pi_{00(r,c)}\}$. For $c(r)_{22-11} < c < c(r)_{11-00}$, $E\pi_{11}(r,c) > \max\{E\pi_{22}(r,c),\ E\pi_{20}(r,c),\ E\pi_{00}(r,c)\}$. For $c > c(r)_{11-00}$, $E\pi_{00}(r,c) > \max\{E\pi_{22}(r,c),\ E\pi_{20}(r,c),\ E\pi_{11}(r,c)\}$.
Q.E.D.

Proof of Proposition 6. Let agent 1 be the first agent in a sequence who obtains and then shares information.

We start with the observation that $a_i \geq 0$ and $b_i \geq 0$ are the cheapest ways for the principal to ensure that (1) each agent $i = 1,2$ chooses an optimal action given his information set I_i, $y_i = E[\theta_i|I_i]$, and (2) that if i obtains a signal, $s_i \neq \emptyset$, then the signal is truthfully shared, $m_i = s_i$. Therefore, for the rest of the proof we focus on the incentives to acquire information. Clearly, the contracts are the same as in the baseline model when the agents are *not* expected to share signals. Therefore, in the following we focus on strategy profiles where, if i acquires a signal, he is expected to share it: that is, if $s_i \neq \emptyset$ then $m_i = s_i$.

Suppose the strategy profile that the principal wants to implement only involves agent 1 obtaining and sharing a signal: $s_1 \neq \emptyset, s_2 = \emptyset, m_1 = s_1$. First, since 2 does not acquire a signal, agent 1's optimal contract is the same as in the baseline model: $a_1 = w_1 > 0, b_1 = 0$ for $r < r_3$, and $a_1 = 0, b_1 = w_1 > 0$ otherwise, where r_3 is defined in Proposition 3. Second, as 2 is not expected to obtain a signal, $w_2 = a_2 = b_2 = 0$. The optimal contract is, of course, symmetric if the principal wants to implement one-sided signal acquisition and sharing by agent 2.

In the following we study two remaining strategy profiles. In case 1 (see below), both agents acquire and share signals. In case 2, 1 acquires and shares s_1, while 2 only acquires s_2 without sharing it.

Case 1. Two-sided acquisition and sharing. We start with the observation that given the common belief that 2 adheres to the strategy profile, and given that 1 does not observe m_2 when deciding over signal acquisition, the optimal contractual arrangement for 1 is the same as in the baseline model: $a_1 = w_1 > 0, b_1 = 0$ for $r < r_1$, and $a_1 = 0, b_1 = w_2 > 0$ otherwise.

Now, consider 2's incentives. First, suppose that $m_1 = 0$.

On-path payoffs conditional on $m_1 = 0$. We first observe that agent 2 now believes $\Pr(s_2 = 0|s_1 = 0) = \frac{3+r}{6}$. The conditional expected losses of 2, provided that $s_2 \neq \emptyset$ (i.e., 2 acquires s_2) and $m_2 = s_2$ are as follows. For $(s_2 = 0, m_1 = 0)$ we calculate

$$1 - E_{\theta_2}[\ell_2(s_2, m_1)|s_2 = m_1 = 0] = \frac{85 + r(58 + 11r)}{10(3 + r)^2},$$

and similarly,

$$1 - E_{\theta_1}[\ell_1(s_1, m_2 = s_2)|s_1 = m_2 = 0] = \frac{85 + r(58 + 11r)}{10(3 + r)^2}.$$

Similarly, for $(s_2 = 1, m_1 = 0)$ we calculate

$$1 - E[\ell_2(s_2, m_1)|s_2 = 1, m_1 = 0] = 1 - E[\ell_1(s_1, m_2)|s_1 = 0, m_2 = 1] = 1 + \frac{r^2 + 2r - 5}{10(3 - r)^2}.$$

Off-path payoffs conditional on $m_1 = 0$. Suppose that 2 deviates and does not obtain a signal. In this case

$$1 - E_{\theta_2}[\ell_2(m_1)|m_1 = 0, s_2 = \emptyset] = 1 - \frac{(33 + r^2)}{36}.$$

Now, the expected loss of the agent 1 depends on m_2. Recall that $m_2 \in \{0, 1\}$. For $m_2 = 0$ we calculate

$$1 - E_{\theta_1}[\ell_1(s_1, m_2)|s_1 = 0, m_2 = 0] = 1 - \frac{r^2 + 2r + 3}{6(r - 3)^2} \tag{10}$$

Similarly, for $m_2 = 1$ we calculate

$$1 - E[\ell_1(s_1, m_2)|s_1 = 0, m_2 = 1, s_2 = \emptyset] = 1 - \frac{r^2 - 2r + 3}{6(r - 3)^2}. \tag{11}$$

Using the above calculations, we can obtain the incentive constraint for 2 upon receiving $m_1 = 0$:

$$\Pr(s_2 = 0|s_1 = 0)\left(a_2(1 - E_{\theta_2}[\ell_2(s_2, m_1)|s_2 = 0, m_1 = 0]) + b_2(1 - E_{\theta_1}[\ell_1(s_1, m_2)|s_1 = 0, m_2 = 0])\right) +$$

$$\Pr(s_2 = 1|s_1 = 0)\left(a_2(1 - E_{\theta_2}[\ell_2(s_2, m_1)|s_2 = 1, m_1 = 0]) + b_2(1 - E_{\theta_1}[\ell_1(s_1, m_2)|s_1 = 0, m_2 = 1])\right) - c \geq$$

$$a_2(1 - E_{\theta_2}[\ell_2(m_1)|m_1 = 0]) + b_2(1 - E_{\theta_1}[\ell_1(s_1, m_2')|s_1 = 0, m_2'])$$

where depending on $m_2' = 0$ or $m_2' = 1$ we use either the losses (10) or (11).

In the linear case, the optimal contract either entails $a_2 > 0, b_2 = 0$ or $a_2 = 0, b_2 > 0$.

1. Suppose $s_2 = \emptyset$ and $m_2 = 0$. Then, depending on the choice of (a_2, b_2) we have one of the two cases:

$$a_2 > 0, b_2 = 0 \;\Rightarrow\; a_2 \geq \frac{36c(9-r^2)}{(3-r^2)^2},$$

$$a_2 = 0, b_2 > 0 \;\Rightarrow\; b_2 \geq b_2^0 := \frac{3c(3-r)(3+r)^2}{2r^2}.$$

For $a_2 = w_2 > 0, b_2 = 0$ the expected transfer to agent 2 is therefore $\frac{6c(51-5r^2)}{(3-r^2)^2}$, and for $a_2 = 0, b_2 = w_2 > 0$ it is $\frac{c(3+r)(51-5r^2)}{4r^2}$.

2. Suppose $s_2 = \emptyset$ and $m_2 = 1$. Then, depending on the choice of (a_2, b_2) we have one of the two cases:

$$a_2 > 0, b_2 = 0 \;\Rightarrow\; a_2 \geq \frac{36c(9-r^2)}{(3-r^2)^2},$$

$$a_2 = 0, b_2 > 0 \;\Rightarrow\; b_2 \geq b_2^1 := \frac{3c(3+r)(3-r)^2}{2r^2}.$$

For $a_2 = w_2 > 0, b_2 = 0$ the expected transfer to agent 2 is the same as above, while for $b_2 = w_2 > 0, a_2 = 0$ it is $\frac{c(3-r)(51-5r^2)}{4r^2}$. Observe that for all $(r > 0, c > 0)$, $b_2^0 > b_2^2$, and so

$$\frac{c(3+r)(51-5r^2)}{4r^2} > \frac{c(3-r)(51-5r^2)}{4r^2}.$$

The critical correlation coefficients are as follows:

$$\frac{6c(51-5r^2)}{(3-r^2)^2} \lessgtr \frac{c(r+3)(51-5r^2)}{4r^2} \iff r \lessgtr \hat{r}_1 (\approx 0.889),$$

$$\frac{6c(51-5r^2)}{(3-r^2)^2} \lessgtr \frac{c(3-r)(51-5r^2)}{4r^2} \iff r \lessgtr \hat{r}_2 \text{ with } \hat{r}_2 < \hat{r}_1.$$

We now turn to the other case with respect to the signal realization of agent 1, and assume $m_1 = 1$. Then, on-path losses for the pairs $(s_2 = 0, m_1 = 1)$ and $(s_2 = 1, m_1 = 1)$ are as follows. First, similar to Case 1:

$$1 - E_{\theta_2}[\ell_2(s_2, m_1)|s_2 = m_1 = 1] = 1 - E_{\theta_1}[\ell_1(s_1, m_2)|s_1 = m_2 = 1].$$

Further we calculate for $(s_2 = 0, m_1 = 1)$:

$$1 - E[\ell_2(s_2, m_1)|s_2 = 0, m_1 = 1] = 1 - E[\ell_1(s_1, m_2)|s_1 = 1, m_1 = 0] = 1 + \frac{2r + r^2 - 5}{10(3-r)^2},$$

and for $(s_2 = 1, m_1 = 1)$:

$$1 - E[\ell_2(s_2, m_1)|s_2 = 1, m_1 = 1] = 1 - E[\ell_1(s_1, m_2)|s_1 = 1, m_1 = 1] = \frac{11r^2 + 58r + 85}{10(r+3)^2}.$$

By symmetry, $\Pr(s_2 = 1|s_1 = 1) = \Pr(s_2 = 0|s_1 = 0)$. Then, it must be the case that the conditional expected losses are the same as in the case of $m_1 = 0$.

Off-path i.e. when $s_2 = \emptyset$ we obtain the following expected losses depending on $m_2 \in \{0, 1\}$.

$$1 - E_{\theta_1}[\ell_1(s_1, m_2')|s_1 = 1, m_2 = 1, s_2 = \emptyset] = 1 - E_{\theta_1}[\ell_1(s_1, m_2')|s_1 = 0, m_2 = 0, s_2 = \emptyset],$$

$$1 - E_{\theta_1}[\ell_1(s_1, m_2')|s_1 = 1, m_2 = 0, s_2 = \emptyset] = 1 - E_{\theta_1}[\ell_1(s_1, m_2')|s_1 = 0, m_2 = 1, s_2 = \emptyset].$$

Moreover, the expected loss of agent 2 himself who bases y_2 only on m_1 is such that

$$1 - E_{\theta_2}[\ell_2(m_1)|m_1 = 1] = 1 - E_{\theta_2}[\ell_2(m_1)|m_1 = 0].$$

Then, to satisfy the incentive constraint at the signal acquisition stage the principal optimally chooses the following contract for agent 2: for $r < \hat{r}_1$, set $a_2 = w_2 > 0, b_2 = 0$, and otherwise set $a_2 = 0, b_2 = b_2^0 = w_2 > 0$.

But then, the optimal contractual arrangement for agent 2 is exactly the same as in the subgame following the message $m_1 = 0$.

Case 2. 1 acquires and shares a signal, while 2 only acquires a signal, without sharing it. We observe that the optimal contract for 1 is exactly the same as in the baseline model. For agent 2, we require $a_2 = w_2 > 0, b_2 = 0$.

Given the analysis in the Case 1 above, it is easy to see that the on path payoff for agent 2 is independent of the realized s_1 and is always $\frac{a_2(51-5r^2)}{6(9-r^2)} - c$ whereas the off path payoff is $\frac{1}{36}a_2\left(r^2 + 33\right)$. Therefore, using $a_2 = \frac{36c(9-r^2)}{(3-r^2)^2}$, we obtain the expected transfer to agent 2: $\frac{6c(51-5r^2)}{(3-r^2)^2}$ for all $r \in [0, 1]$.

Optimal allocations. To show the optimality of different strategy profiles for various combinations of (r, c) we note that it is straightforward to show that $E\pi_{21}$ is domnated by either $E\pi_{20}$ or $E\pi_{22}$. Moreover, observe that Lemma 2 still applies. Therefore, the only difference to the optimal arrangements in the baseline model is between $E\pi_{22}$ and $E\pi_{20}$, provided that $r > r_1$ (where r_1 is defined in Proposition 1) (observe that for $r < r_1$ the optimal contract and the equilibrium allocations are the same because $r_1 < \hat{r}_1$).

We first consider $r \in (r_1, \hat{r}_1)$ (recall that this is because $E\pi_{22}$ changes at $\hat{r}_1$).

$$E\pi_{22} > E\pi_{20}$$

$$2 - \frac{3-r^2}{3(9-r^2)} - \left(\frac{c(51-5r^2)(9-r^2)}{12r^2} + \frac{6c(51-5r^2)}{(3-r^2)^2}\right) > 2 - \frac{1}{18} - \frac{3-r^2}{36} - \frac{18c}{r^2}\left(1 - \frac{3-r^2}{36}\right)$$

$$\tilde{c}_1(r) < -\frac{r^2\left(r^2-3\right)^2\left(r^4-2r^2+9\right)}{3\left(r^2-9\right)\left(5r^8-132r^6+558r^4+1188r^2+2349\right)}.$$

For $r \geq \hat{r}_1$, similarly, we calculate

$$E\pi_{22} > E\pi_{20}$$

$$\tilde{c}_2(r) < -\frac{r^2\left(r^4-2r^2+9\right)}{3\left(r^2-9\right)\left(r(r(5(r-3)r-147)+153)+720\right)}.$$

Therefore, the optimal allocation is the same as in the baseline model, apart from the case when $c \in [\max\{\tilde{c}_1(r), \tilde{c}_2(r)\}, c_{22-11}(r))$ in which case the principal implements one-sided information acquisition and sharing and earns $E\pi_{20}$.

Comparison to the baseline payoffs. Recall that in the baseline model, for all $r > r_1$ $E\pi_{22} > E\pi_{20}$ for $c < \frac{r^2(r^4-2r^2+9)}{6(9-r^2)(5r^4-99r^2+360)}$.

Therefore, the principal ends up with a worse allocation for $c \in [\tilde{c}_1(r), c_{22-11}(r))$ for $r < \hat{r}_1$ and for $c \in [\tilde{c}_2(r), c_{22-11}(r))$ for $r \geq \hat{r}_1$. This proves the Proposition.

Proof of Proposition 7. Suppose that the principal adopts linear contracts t_1, t_2 to induce both agents i to collect signal s_i and transmit it to the other agent j. Then again the principal optimally chooses $a_i \geq 0$ for both i and using symmetry across agents, each agent i ex-ante expected equilibrium payoff can be written

$$u_i = w_i - (a_i + b_i)E[(E(\theta_i|s_i, s_j) - \theta_i)^2|s_i, s_j] - c,$$

where the outer expectation is taken first with respect to the random variables s_i, s_j and then θ_i.

If agent i deviates and does not collect signal s_i, he will still communicate a message m_i to the other agent j, who will act in the mistaken belief that $m_i = s_i$. Such a deviation yields agent i the expected payoff:

$$u_i^0 = w_i - a_iE[(E(\theta_i|s_j) - \theta_i)^2|s_j] - b_iE[(E(\theta_i|m_i, s_j) - \theta_i)^2|s_j].$$

It then follows that, if

$$E[(E(\theta_i|s_j) - \theta_i)^2|s_j] < E[(E(\theta_i|m_i, s_j) - \theta_i)^2|s_j] \text{ for all } m_i, \tag{12}$$

then the principal's minimization of agents' equilibrium transfers $w_i-(a_i+b_i)E[(E(\theta_i|s_i, s_j)-\theta_i)^2|s_i, s_j]$ subject to the constraint $\bar{u}_i \geq u_i^0$ entails setting $b_i > 0$ and $a_i = 0$. That (12) is equivalent to condition (7) follows from the following chain of equivalences:

$$E_{\theta_i,s_j}([E(\theta_i|s_j) - \theta_i]^2|s_j) < E_{\theta_j,s_j}([E(\theta_j|s_j, m_i) - \theta_j]^2|s_j),$$

$$E_{s_j}Var(\theta_i|s_j) = E_{\theta_i,s_j}([E(\theta_i|s_j) - \theta_i]^2|s_j) < E_{\theta_j,s_j}([E(\theta_j|s_j, m_i) - E(\theta_j|s_j) + E(\theta_j|s_j) - \theta_j]^2|s_j),$$

$$E_{s_j}Var(\theta_i|s_j) < E_{\theta_j,s_j}([E(\theta_j|s_j, m_i) - E(\theta_j|s_j)]^2|s_j) + E_{\theta_j,s_j}([E(\theta_j|s_j) - \theta_j]^2|s_j)$$

because $E_{\theta_j,s_j}(E(\theta_j|s_j) - \theta_j|s_j) = 0$, so that:

$$E_{s_j}Var(\theta_i|s_j) - E_{s_j}Var(\theta_i|s_j) < E_{\theta_j,s_j}([E(\theta_j|s_j, m_i) - E(\theta_j|s_j)]^2).$$

Q.E.D.

Proof that (7) holds in the case $\theta_1 = \theta_2 \sim N(0, 1/\alpha)$ and $s_i = \theta_i + \varepsilon_i$ with $\varepsilon_i \sim N(0, 1/\beta)$. Because $EVar[\theta_i|s_j] = EVar[\theta_j|s_j]$, we only need to show that $E[(E[\theta_j|s_j, s_i] - E[\theta_j|s_j])^2] > 0$ for all $s_i \in \mathbb{R}$. Plugging in the expressions for $E[\theta_j|s_j, s_i]$ and $E[\theta_j|s_j]$, we obtain:

$$E_{s_j}[(E[\theta_j|s_i, s_j] - E[\theta_j|s_j])^2] = E_{s_j}\left[\left(\frac{\beta s_j}{\alpha+\beta} - \frac{\beta(s_i+s_j)}{\alpha+2\beta}\right)^2\right] = E_{s_j}\left[\left(\beta\frac{\beta s_j - (\alpha+\beta)s_i}{(\alpha+\beta)(\alpha+2\beta)}\right)^2\right]$$
$$= \frac{\beta^2 E_{s_j}\left[\beta^2 s_j^2 + (\alpha+\beta)^2 s_i^2 - 2\beta s_j(\alpha+\beta)s_i\right]}{(\alpha+\beta)^2(\alpha+2\beta)^2}$$
$$= \beta^2\frac{\beta^2(1/\alpha+1/\beta) + (\alpha+\beta)^2 s_i^2}{(\alpha+\beta)^2(\alpha+2\beta)^2} \geq \beta^2\frac{\beta^2(1/\alpha+1/\beta)}{(\alpha+\beta)^2(\alpha+2\beta)^2} > 0.$$

Q.E.D.

Proof of Proposition 8. Consider fully correlated states, $\theta_1 = \theta_2$ (i.e., $r = 1$). Applying backward reasoning, we first consider agents' actions conditional on received information, under the common belief that the signals are truthfully exchanged. Consider the signal profile $(s_1 = 0, s_2 = 1)$. The expected losses depending on the actions (y_1, y_2) are as follows

$$E[\ell_1|s_1 = 1; s_2 = 0] = \frac{5 + r - 20y_1 + 10y_1^2(3-r)}{120},$$

$$E[\ell_2|s_2 = 0, s_1 = 1] = \frac{1}{120}\left(r\left(-10y_2^2 + 20y_2 - 9\right) + 5\left(6y_2^2 - 8y_2 + 3\right)\right).$$

The interaction component is as follows

$$E[\ell_1|s_1 = 1; s_2 = 0]E[\ell_2|s_2 = 0, s_1 = 1] = \frac{1}{5040}\Big(35\left(6y_1^2 - 4y_1 + 1\right)\left(6y_2^2 - 8y_2 + 3\right) +$$
$$r\left(-42y_1^2(10(y_2-2)y_2 + 9) + 28y_1(3 - 4y_2) + 14y_2(3y_2 - 4) + 15\right)\Big).$$

First, consider the linear case where the principal sets $g_1 = 0$ in which case each agent i chooses $y_i = E[\theta|s_1 = 0, s_2 = 1]$. Recall that $y_1 = y_2 = \frac{1}{2}$.

If we allow for $g_i \neq 0$, then the first-order approach can be used to calculate the agents' optimal actions

$$\frac{t_1(\ell_1, \ell_2)}{\partial y_1} = 0, \ \frac{t_2(\ell_1, \ell_2)}{\partial y_2} = 0,$$

where under symmetric contracts we obtain

$$y_1^* = \frac{1}{2}, y_2^* = \frac{1}{2}.$$

Therefore, the optimal actions are the same as in the linear case.

Next, consider the signal profile $(s_1 = s_2 = 0)$ We first calculate

$$E[\ell_1|s_1 = s_2 = 0] = \int_0^1 (y_1 - \theta)^2 (1-\theta)^2 d\theta,$$

$$E[\ell_2|s_2 = s_1 = 0] = \int_0^1 (y_2 - \theta)^2 (1-\theta)^2 d\theta,$$

$$E[\ell_1|s_1 = s_2 = 0]E[\ell_2|s_2 = s_1 = 0, r = 1] = \int_0^1 (y_1 - \theta)^2 (y_2 - \theta)^2 (1-\theta)^2 d\theta.$$

Agent 1's payoff under limited liability is

$$a_1 (1 - E[\ell_1|s_1 = s_2 = 0, r = 1]) + b_1 (1 - E[\ell_2|s_2 = s_1 = 0, r = 1]) +$$

$$g_1 (1 - E[\ell_1|s_1 = s_2 = 0, r = 1]E[\ell_2|s_2 = s_1 = 0, r = 1]).$$

We note here that the agent's payoff need not be concave in y_1 for arbitrary values of (a_1, g_1). However, observe that conditionally on the signal profile $(0,0)$ a convex transfer function yields largest possible distortions (where agents' actions are corner solutions): this cannot be optimal for the principal.

Assume momentarily that the symmetric profile $(a_1 = a_2 = a, g_1 = g_2 = g)$ yields a unique interior maximum of each agent i's transfer in y_i. In this case, the optimal actions of the agents can then be obtained using the first order approach:

$$\tilde{y}_1(0,0) = \tilde{y}_2(0,0) =$$

$$\frac{30g - \frac{6\ 10^{2/3} d(80a+9g)}{\sqrt[3]{12\sqrt{15}\sqrt{g^3((80a+9g)^3+135g^3)}+540g^3}} + \sqrt[3]{120\sqrt{15}\sqrt{g^3\left((80a+9g)^3+135g^3\right)}+5400g^3}}{120g}.$$

In the following step, we analyze the upper bound on the principal's expected payoff conditioning it on the two signal realizations: $(s_1 = s_2 = 0)$ and $(s_1 = 0, s_2 = 1)$ – the complement of this case is completely symmetric and yields the same conditional upper bound for the principal.

To obtain the upper bound, given the signal-contingent actions $(\tilde{y}_1, \tilde{y}_2)$ (and assuming momentarily that the first-order approach holds when using the signal profile $(s_1 = s_2 = 0)$), assume that the principal has to transfer c to each agent to reimburse each agent for obtaining a signal. We then have

$$\tilde{u}_0(c) \equiv \frac{1}{3}\left[1 - \int_0^1 (\tilde{y}_1(0,0) - \theta)^2 (1-\theta)^2 d\theta - c\right] +$$

$$\frac{1}{6}\left[1 - \int_0^1 \left(\frac{1}{2} - \theta\right)^2 (1-\theta)\theta d\theta - c\right]$$

where recall that from the ex ante perspective $\Pr(s_s = s_2 = 0) = \frac{1}{3}$, $\Pr(s_1 = 0, s_2 = 1) = \frac{1}{6}$.

Now, if the principal chooses $g = 0$ and therefore offers a (symmetric) linear contract (so that

$y_i = E[\theta|s_1, s_2]$), using the result from Proposition (1) the conditional expected payoff by the principal is proportional to

$$u_1(c) \equiv \frac{1}{3}\left[1 - \int_0^1 \left(\frac{1}{4} - \theta\right)^2 (1-\theta)^2 d\theta\right] + \frac{1}{6}\left[1 - \int_0^1 \left(\frac{1}{2} - \theta\right)^2 (1-\theta)\theta d\theta\right] -$$

$$32c\left[\frac{1}{3}\left(1 - \int_0^1 \left(\frac{1}{4} - \theta\right)^2 (1-\theta)^2 d\theta\right) + \frac{1}{6}\left(1 - \int_0^1 \left(\frac{1}{2} - \theta\right)^2 (1-\theta)\theta d\theta\right)\right].$$

It requires tedious but straightforward calculations to show that for *any* profile (a, g), we have

$$\lim_{c \to 0} u_1(c) \geq \lim_{c \to 0} \tilde{u}_0(c),$$

where the strict inequality holds almost everywhere.

But then, the principal cannot profit from offering a contract of the form (8), compared to the optimal general contract from Proposition 1– provided that we can use the first-order approach.

Since the above expression holds with equality only in the limit, for any $c > 0$ there must exist $\bar{r} < 1$ such that the principal adheres to the optimal linear contract for all $r \in (\bar{r}, 1]$ and sufficiently small costs.

Finally, suppose that (a_i, b_i, g_i) for each agent i is such that the transfers are non-concave, resulting in $\tilde{y}(0,0) \in \{0,1\}$. Suppose that $\tilde{y}(0,0) = 0$, which is the preferred choice for the principal. Then,

$$\hat{u}(c) \equiv \frac{1}{3}\left[1 - \int_0^1 (0-\theta)^2 (1-\theta)^2 d\theta - c\right] + \frac{1}{6}\left[1 - \int_0^1 \left(\frac{1}{2} - \theta\right)^2 (1-\theta)\theta d\theta - c\right].$$

But then, we have:

$$u_1(c) - \hat{u}(c) = \frac{5 - 11032c}{720},$$

so that there exists a range of strictly positive $c > 0$ that makes the above expression strictly positive. In other words, the distortion is so large that the principal never benefits from deviating from linear contracts for c small enough, provided that $r \in (\bar{r}, 1]$.
Q.E.D.

Proof of Proposition 9: In the first step we show when an allocation with two signals and one-sided communication dominates the two other allocations with two signals. First, consider the following expected payoffs:

$$E\pi_{22}^{\lambda}(r,c) = (\lambda_1 + \lambda_2)(1 - \frac{3 - r^2}{6(9 - r^2)}) - (51 - 5r^2)\min\{\frac{12}{(3 - q^2)^2}, \frac{9 - r^2}{6r^2}\}c,$$

$$E\pi_{21}^{\lambda}(r,c) = (\lambda_1 + \lambda_2) - \lambda_1\frac{1}{18} - \lambda_2\frac{3 - r^2}{6(9 - r^2)} - \min\{34, (51 - 5r^2)\frac{(9 - r^2)}{12r^2}\}c - 6\frac{51 - 5r^2}{(3 - r^2)^2}c,$$

$$E\pi^{\lambda}_{12}(r,c) = (\lambda_1+\lambda_2) - \lambda_2\frac{1}{18} - \lambda_1\frac{3-r^2}{6(9-r^2)} - \min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\}c - 6\frac{51-5r^2}{(3-r^2)^2}c.$$

such that the corresponding differences are:

$$E\pi^{\lambda}_{22}(r,c) - E\pi^{\lambda}_{21}(r,c) = \lambda_1(\frac{1}{18} + \frac{3-r^2}{6(9-r^2)}) - D^{22}_{21}(r)c,$$
$$E\pi^{\lambda}_{22}(r,c) - E\pi^{\lambda}_{12}(r,c) = \lambda_2(\frac{1}{18} + \frac{3-r^2}{6(9-r^2)}) - D^{22}_{12}(r)c.$$

where

$$D^{22}_{21}(r)c = D^{22}_{12}(r) = (51-5r^2)\min\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6r^2}\} - \min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\} - 6\frac{51-5r^2}{(3-r^2)^2}.$$

As in the baseline model, $D^{22}_{21}(r) < 0$ for $r > \tilde{r}_2$ which implies that in this case $E\pi^{\lambda}_{22}(r,c) > E\pi^{\lambda}_{21}(r,c)$. For $r < \tilde{r}_2$, $D^{22}_{21}(r) > 0$, so it could be that $E\pi^{\lambda}_{22}(r,c) < E\pi^{\lambda}_{21}(r,c)$. Similar to the approach in the proof of Lemma 4, define the following cost thresholds:

$$c^{\lambda}_{22-21}(r) = \frac{\lambda_1(\frac{1}{18} - \frac{3-r^2}{6(9-r^2)})}{(51-5r^2)\min\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6r^2}\} - \min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\} - 6\frac{51-5r^2}{(3-r^2)^2}},$$
$$c^{\lambda}_{22-12}(r) = \frac{\lambda_2(\frac{1}{18} - \frac{3-r^2}{6(9-r^2)})}{(51-5r^2)\min\{\frac{12}{(3-r^2)^2}, \frac{9-r^2}{6q^2}\} - \min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\} - 6\frac{51-5r^2}{(3-r^2)^2}},$$
$$c^{\lambda}_{21-20}(r) = \frac{\lambda_2(\frac{1}{18} - \frac{3-r^2}{6(9-r^2)})}{\min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\} + 6\frac{51-5r^2}{(3-r^2)^2} - 68},$$
$$c^{\lambda}_{12-20}(r) = \frac{\lambda_1(\frac{1}{18} - \frac{3-r^2}{6(9-r^2)})}{\min\{34, (51-5r^2)\frac{(9-r^2)}{12r^2}\} + 6\frac{51-5r^2}{(3-r^2)^2} - 68}.$$

such that $E\pi^{\lambda}_{22}(r,c) > E\pi^{\lambda}_{21}(r,c)$ for all $c < c^{\lambda}_{22-21}(r)$ and $E\pi^{\lambda}_{20}(r,c) > E\pi^{\lambda}_{21}(r,c)$ for all r and $c > c^{\lambda}_{22-21}(r)$.

The analysis distinguishes different cases depending on r.

Suppose that $r \leq \bar{r}_1 < \tilde{r}_2$. To see when an asymmetric allocation can be optimal, consider the case $\lambda_1 > \lambda_2$ such that $c^{\lambda}_{22-21} > c^{\lambda}_{22-12}$. In case $c < c^{\lambda}_{22-12}(r)$, omitted calculations show that $E\pi^{\lambda}_{22} > \max\{E\pi^{\lambda}_{12}, E\pi^{\lambda}_{21}\}$. If instead $c \in (c^{\lambda}_{21-20}(r), c^{\lambda}_{12-20}(r)]$, then omitted calculations show that $E\pi^{\lambda}_{22} > E\pi^{\lambda}_{21}$ and $E\pi^{\lambda}_{20} > E\pi^{\lambda}_{21}$ so $E\pi^{\lambda}_{21}$ cannot be optimal. However, $E\pi^{\lambda}_{12} > E\pi^{\lambda}_{22}$ and at the same time $E\pi^{\lambda}_{12} > E\pi^{\lambda}_{20}$. Since the case for $\lambda_2 < \lambda_1$ is symmetric, we conclude that when $\lambda_1 > \lambda_2$, $E\pi^{\lambda}_{12} > \max\{E\pi^{\lambda}_{20}, E\pi^{\lambda}_{22}\}$ for $c \in (c^{\lambda}_{21-20}(r), c^{\lambda}_{12-20}(r)]$, whereas if $\lambda_2 > \lambda_1$, then, for $c \in (c^{\lambda}_{12-20}(r), c^{\lambda}_{21-20}(r)]$, $E\pi^{\lambda}_{21} > \max\{E\pi^{\lambda}_{20}, E\pi^{\lambda}_{22}\}$.

Second, suppose that $r \in (\bar{r}_1, \tilde{r}_2)$. Omitted calculations show that for $\lambda_1 > 0$, $\lambda_2 > 0$, it is the case that $c^{\lambda}_{22-21}(r) > c^{\lambda}_{12-20}(r)$ and $c^{\lambda}_{22-12}(r) > c^{\lambda}_{21-20}(r)$. To see when an asymmetric allocation can be optimal, consider the case $\lambda_1 > \lambda_2$ and therefore $c^{\lambda}_{22-21}(r) > c^{\lambda}_{22-12}(q)$. Hence, for $c < c^{\lambda}_{22-12}(r)$, it is the case that $E\pi^{\lambda}_{22} > \max\{E\pi^{\lambda}_{12}, E\pi^{\lambda}_{21}\}$. Instead, if $c^{\lambda}_{12-20}(r) > c^{\lambda}_{22-12}(r)$, then for $c \in (c^{\lambda}_{12-20}(r), c^{\lambda}_{22-21}(r)]$, $E\pi^{\lambda}_{20} > \max\{E\pi^{\lambda}_{21}, E\pi^{\lambda}_{12}\}$. However, for $c \in [c^{\lambda}_{22-12}(r), c^{\lambda}_{12-20}(r))$, $E\pi^{\lambda}_{21} < E\pi^{\lambda}_{20} < E\pi^{\lambda}_{12}$. Thus, in this region $E\pi^{\lambda}_{12} > \max\{E\pi^{\lambda}_{20}, E\pi^{\lambda}_{22}\}$. Since the case for $\lambda_2 < \lambda_1$ is symmetric, we conclude that when $\lambda_1 > \lambda_2$, it is the case that $E\pi^{\lambda}_{12} > \max\{E\pi^{\lambda}_{20}, E\pi^{\lambda}_{22}\}$ for $c \in [c^{\lambda}_{22-12}(r), c^{\lambda}_{12-20}(r))$. If instead $\lambda_2 > \lambda_1$, then $E\pi^{\lambda}_{21} > \max\{E\pi^{\lambda}_{20}, E\pi^{\lambda}_{22}\}$ for $c \in [c^{\lambda}_{22-21}(r), c^{\lambda}_{21-20}(r))$.

In the next step we compare an allocation with two signals and one-sided communication to all other allocation with at most one acquired signal. First, notice that the non–informative allocation is dominated by an allocation with two acquired and not shared signals for

$$-(\lambda_1 + \lambda_2)\frac{1}{18} - 68c \geq -(\lambda_1 + \lambda_2)\frac{1}{12},$$

which implies $c \leq \frac{\lambda_1+\lambda_2}{2488}$. Further, the non-informative allocation is dominated by an allocation with one acquired and communicated signal (by agent i) for

$$-\lambda_i\frac{1}{18} - \lambda_j\frac{3-r^2}{36} - \min\{34, \frac{r^2+33}{2r^2}\}c > -(\lambda_i + \lambda_j)\frac{1}{12}.$$

Since the left-hand side of the above inequality increases in r, the inequality is satisfied if the following is true

$$-\lambda_i\frac{1}{18} - 34c > -\lambda_i\frac{1}{12},\ i = 1, 2,$$

which is true for $c \leq \frac{\lambda_i}{1224}$. In the following we assume that the above constraints are true.

In the following we focus on the case that the optimal allocation with two signals and one-sided communication entails communication by agent 1, as the case in which it is 2 who communicates is symmetric and can be addressed in the same way. This requires that $E\pi^{\lambda}_{21} > E\pi^{\lambda}_{12}$, that is,

$$\begin{aligned}
E\pi^{\lambda}_{21} &= \lambda_1 + \lambda_2 - \lambda_1\frac{1}{18} - \lambda_2\frac{3-r^2}{6(9-r^2)} - 34c - 6\frac{51-5r^2}{(3-r^2)^2}c \\
&> E\pi^{\lambda}_{11} = (\lambda_1 + \lambda_2) - \lambda_1\frac{1}{18} - \lambda_2\frac{3-r^2}{36} - \min\{34, \frac{r^2+33}{2r^2}\}c,
\end{aligned}$$

which implies:

$$\lambda_2\left(\frac{3-r^2}{36} - \frac{3-r^2}{6(9-r^2)}\right) + \min\{34, \frac{r^2+33}{2r^2}\}c - 34c - 6\frac{51-5r^2}{(3-r^2)^2}c > 0. \tag{13}$$

We distinguish between two cases. When $r < \bar{r}_3$, the inequality (13) becomes

$$\lambda_2\left(\frac{3-r^2}{36} - \frac{3-r^2}{6(9-r^2)}\right) - 6\frac{51-5r^2}{(3-r^2)^2}c > 0$$

which implies

$$c < \frac{\lambda_2(3-r^2)^4}{216(9-r^2)(51-5r^2)}.$$

When $r \in [\bar{r}_3, \tilde{r}_2)$, the inequality (13) becomes

$$\lambda_2(\frac{3-r^2}{36} - \frac{3-r^2}{6(9-r^2)}) + \frac{r^2+33}{2r^2}c - 34c - 6\frac{51-5r^2}{(3-r^2)^2}c > 0.$$

which implies

$$c \leq \frac{\lambda_2 r^2 \left(3-r^2\right)^4}{18\left(9-r^2\right)\left(67r^6-495r^4+1413r^2-297\right)}$$

which is strictly positive for the above range of parameters.

Now, we consolidate the obtained conditions. Given the discussion at the start, we restrict attention to costs such that $c \leq \frac{3-(\lambda_1+\lambda_2)}{1244}$ and $c \leq \frac{6-(2\lambda_1+3\lambda_2)}{1224}$.

Suppose first that $r < \bar{r}_3$. It turns out that $\frac{(3-r^2)^4}{216(9-r^2)(51-5r^2)} > \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}$ for $r < \tilde{r}_1$, so that for $r < \tilde{r}_1$ the asymmetric contract leading to $E\pi_{21}^{\lambda}$ is a global optimum for $c < \frac{\lambda_2(3-r^2)^4}{216(9-r^2)(51-5r^2)}$. If $\lambda_2 > \lambda_1$, then $E\pi_{21}^{\lambda}$ is a global optimum for

$$c \in (\lambda_1 \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}, \lambda_2 \frac{(3-r^2)^4}{18(9-r^2)(87-17r^2)}).$$

Then, suppose $\bar{r}_3 \leq r < \bar{r}_1$. It turns out that for the above range of parameters

$$\frac{r^2\left(3-r^2\right)^4}{18\left(9-r^2\right)\left(67r^6-495r^4+1413r^2-297\right)} > \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}$$

and so $E\pi_{21}^{\lambda}$ is a global optimum for

$$c \in (\lambda_1 \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}, \lambda_2 \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}).$$

Finally, say that $\bar{r}_1 < r < \tilde{r}_2$. Given the previous point, we conclude that the asymmetric contract yielding $E\pi_{21}^{\lambda}$ is a global optimum for

$$c \in (\lambda_1 \frac{2r^4\left(r^2-3\right)^2}{3\left(9-r^2\right)\left(5r^8-330r^6+2484r^4-7290r^2+4131\right)}, \lambda_2 \frac{(3-r^2)^2}{18(9-r^2)(87-17r^2)}).$$

Q.E.D.

Mixed extension I: introducing a single pooling message. First, consider the Bayesian updated depending on the pair (s_i, m_p) where recall that m_p denotes the pooling message. Without loss of generality, we consider agent 1 (since the strategies are symmetric across agents).

$$f(\theta_1|s_1 = 0, m_2 = m_p) = \frac{6(1-\theta_1)(\mu_1(-2r\theta_1 + r + 1) + \mu_2 r(2\theta_1 - 1) + \mu_2)}{\mu_1(3+r) + \mu_2(3-r)}$$

with $E[\theta_1|s_1=0, m_2=m_p] = \frac{\mu_1+\mu_2}{\mu_1(3+r)+\mu_2(3-r)}$, and similarly

$$f(\theta_1|s_1=1, m_2=m_p) = \frac{6\theta_1(\mu_1(-2r\theta_1+r+1)+\mu_2 r(2\theta_1-1)+\mu_2)}{\mu_1(3-r)+\mu_2(3+r)}$$

with $E[\theta_1|s_1=1, m_2=m_p] = 1-\frac{\mu_1+\mu_2}{\mu_1(3-r)+\mu_2(3+r)}$. The expected loss $E\ell_2$ if sending the message $m_1=0$ given the obtained signal $s_1=0$ is

$$\frac{2}{3}\left(1-\int_0^1 (\frac{1}{3+r}-\theta_2)^2\frac{6(1-\theta_2)(-2r\theta_1+r+1)}{r+3}\right)d\theta_2+$$

$$\frac{1}{3}\left(1-\int_0^1\left(1-\frac{1}{3-r}-\theta_2\right)^2\frac{6\theta_2(-2r\theta_2+r+1)}{3-r}\right)d\theta_2,$$

whereas if agent 1 sends the message $m_1=m_p$, the expected loss $E\ell_2$ is

$$\frac{2}{3}\left(1-\int_0^1(\frac{\mu_1+\mu_2}{\mu_1(r+3)-\mu_2(r-3)}-\theta_2)^2\frac{6(\theta_2-1)(\mu_1(-2r\theta_2+r+1)+\mu_2 r(2\theta_2-1)+\mu_2)}{\mu_2(r-3)-\mu_1(r+3)}\right)d\theta_2+$$

$$\frac{1}{3}\left(1\int_0^1\left(\frac{\mu_1+\mu_2}{\mu_1(r-3)-\mu_2(r+3)}+1-\theta_2\right)^2\left(-\frac{6\theta_2(\mu_1(-2r\theta_2+r+1)+\mu_2 r(2\theta_2-1)+\mu_2)}{\mu_1(r-3)-\mu_2(r+3)}\right)\right)d\theta_2.$$

Solving the indifference condition, we obtain $\mu_2=0$. In a similar way, the indifference condition for agent 1 to send either $m_1=1$ or the pooling message m_p requires $\mu_1=0$. As a result, all information is pooled in a single message: if the principal decides to link agent 1's pay to 2's performance. Thus, under $b_i\geq 0$ randomization results in a loss of any information coming from the other agent. This is inefficient for the principal because she cannot implement some allocations that are optimal in the baseline scenario where information is fully shared. If, instead, $b_i<0$, then each agent aims to create a largest possible distortion of ℓ_j, which is dominated for the principal compared to $b_i\geq 0$.

Mixed extension II: two-message scenario. As the strategies are symmetric across agents, it is without loss of generality to focus on agent 1.

$$f(\theta_1|s_1=0, m_2=0)=$$

$$\frac{r[(1-\theta_1)^2\mu_1+\theta_1(1-\theta_1)(1-\mu_2)]+(1-r)[(1-\theta_1)((1/2)\mu_1+(1/2)(1-\mu_2))]}{\int_0^1 r[(1-\theta_1)^2\mu_1+\theta_1(1-\theta_1)(1-\mu_2)]+(1-r)[(1-\theta_1)((1/2)\mu_1+(1/2)(1-\mu_2))d\theta_1},$$

where we obtain $E[\theta_1|s_1=0, m_2=0]=\frac{1+\mu_1-\mu_2}{\mu_1(3+r)+(3-r)(1-\mu_2)}$. In a similar fashion, we find the following expected values depending on the other profiles $(s_1, m_2)\in\{(0,1),(1,0),(11)\}$.

$$E[\theta_1|s_1=0, m_2=1]=\frac{1+\mu_2-\mu_1}{3+\mu_2(3-r)+r-\mu_1(3+r)},$$

$$E[\theta_1|s_1=1, m_2=0]=1-\frac{1+\mu_1-\mu_2}{\mu_1(3-r)+(1-\mu_2)(3+r)},$$

$$E[\theta_1|s_1=1,m_2=1]=1-\frac{1+\mu_2-\mu_1}{(1-\mu_1)(3-r)+\mu_2(3+r)}.$$

Communication strategy affects the losses of the other agent 2. The optimal contract for the principal who wants information to be shared, must specify $b_1 \geq 0$. The following two indifference conditions at the communication stage, which depend on 1's signal, have to be satisfied:

$$\Pr(s_2=0|s_1=0)(1-E\ell_2(s_2=0,m_1=0))+\Pr(s_2=1|s_1=0)(1-E\ell_2(s_2=1,m_1=0))=$$

$$\Pr(s_2=0|s_1=0)(1-E\ell_2(s_2=0,m_1=1))+\Pr(s_2=1|s_1=0)(1-E\ell_2(s_2=1,m_1=1)),$$

$$\Pr(s_2=0|s_1=1)(1-E\ell_2(s_2=0,m_1=1))+\Pr(s_2=1|s_1=1)(1-E\ell_2(s_2=1,m_1=1))=$$

$$\Pr(s_2=0|s_1=1)(1-E\ell_2(s_2=0,m_1=0))+\Pr(s_2=1|s_1=1)(1-E\ell_2(s_2=1,m_1=0)).$$

We now calculate the respective expected losses:

$$E\ell_2(s_2=0,m_1=0)=\frac{-r^2(\mu_1+\mu_2-1)^2+2r(\mu_1+\mu_2-1)(\mu_1-\mu_2+1)+5(\mu_1-\mu_2+1)^2}{10(\mu_1(r+3)+(\mu_2-1)(r-3))^2},$$

$$E\ell_2(s_2=0,m_1=1)=\frac{-r^2(\mu_1+\mu_2-1)^2+2r(\mu_1-\mu_2-1)(\mu_1+\mu_2-1)+5(-\mu_1+\mu_2+1)^2}{10(-r(\mu_1+\mu_2)-3\mu_1+3\mu_2+r+3)^2},$$

$$E\ell_2(s_2=1,m_1=0)=-\frac{r^2(\mu_1+\mu_2-1)^2+2r(\mu_1+\mu_2-1)(\mu_1-\mu_2+1)-5(\mu_1-\mu_2+1)^2}{10(\mu_1(r-3)+(\mu_2-1)(r+3))^2},$$

$$E\ell_2(s_1=1,m_2=1)=-\frac{r^2(\mu_1+\mu_2-1)^2+2r(\mu_1-\mu_2-1)(\mu_1+\mu_2-1)-5(-\mu_1+\mu_2+1)^2}{10(\mu_1(r-3)+\mu_2(r+3)-r+3)^2}.$$

Using the above expected losses and the fact that $\Pr(s_i=0|s_j=0)=\Pr(s_i=1|s_j=1)=2/3$, we find a simple solution to the two indifference conditions:

$$\mu_1=1-\mu_2.$$

The implication is as follows. Consider the communication strategy of agent 2 with $\mu_1=1-\mu_2$. Then, the signal-contingent expectations of agent 1 are:

$$Pr(s_2=0|s_1=0,m_2=0)=\frac{(2/3)(1-\mu_2)}{(2/3)(1-\mu_2)+(1/3)(1-\mu_2)}=\frac{2}{3}=\Pr(s_2=0|s_1=0),$$

$$Pr(s_2=1|s_1=0,m_2=1)=\frac{(1/3)\mu_2}{(1/3)\mu_2+(2/3)\mu_2}=\frac{1}{3}=\Pr(s_2=1|s_1=0).$$

Therefore, under the above communication strategies the allocation is the same as under no communication. We know, however, that the principal can implement a no-sharing allocation also in the baseline model. Thus, the mixed extension does not generate an improvement for the principal, compared to the baseline scenario. In fact, in the range of (c,r) in which the principal wants to implement full acquisition and sharing, the randomization lowers principal's payoffs.

It is easy to see that in general non-revealing information is weakly dominated for the principal.

The proof of Proposition 10 follows directly from the analysis of the two mixed extensions of the baseline communication game above.

Online Appendix

Proof of Lemma 2: Subtracting the formulas of $E\pi_{11}(r,c)$ and $E\pi_{10}(r,c)$ and rearranging, we obtain

$$E\pi_{11} - E\pi_{10} = \frac{5}{36} - \frac{1}{36}(5-r^2) + 34c - \min\left\{34, \frac{67+r^2}{1+2r^2}\right\} c,$$

which is strictly positive.

Q.E.D.

Proof of Lemma 3: In the first step we show that for all $r \leq r_1 \approx 0.803$, either $E\pi_{21} \leq E\pi_{22}$ or $E\pi_{21} \leq E\pi_{20}$. First, note that $E\pi_{21} \leq E\pi_{22}$ if

$$2 - \frac{3-r^2}{6(9-r^2)} - \frac{1}{18} - 36c(1-\frac{1}{18}) - 36c\frac{9-r^2}{(3-r^2)^2}\Big(1-\frac{3-r^2}{6(9-r^2)}\Big) \leq$$

$$2 - \frac{2(3-r^2)}{6(9-r^2)} - \frac{72c(9-r^2)}{(3-r^2)^2}\Big(1-\frac{3-r^2}{6(9-r^2)}\Big)$$

that implies

$$c \leq \frac{(r^2-3)^2}{18(r^2-9)(17r^2-87)}.$$

Further, $E\pi_{21} \leq E\pi_{20}$ implies that

$$2 - \frac{3-r^2}{6(9-r^2)} - \frac{1}{18} - 36c(1-\frac{1}{18}) - 36c\frac{9-r^2}{(3-r^2)^2}\Big(1-\frac{3-r^2}{6(9-r^2)}\Big) \leq$$

$$2 - \frac{1}{9} - 68c$$

that implies

$$c \geq \frac{(r^2-3)^2}{18(r^2-9)(17r^2-87)}.$$

Thus, for $r \leq r_1$, either $E\pi_{21} \leq E\pi_{22}$ or $E\pi_{21} \leq E\pi_{20}$.

Next, consider $r_1 \leq r \leq r_2$. The inequality $E\pi_{21} \leq E\pi_{22}$ is

$$2 - \frac{3-r^2}{6(9-r^2)} - \frac{1}{18} - 36c(1-\frac{1}{18}) - 36c\frac{9-r^2}{(3-r^2)^2}\Big(1-\frac{3-r^2}{6(9-r^2)}\Big) \leq$$

$$2 - \frac{2(3-r^2)}{6(9-r^2)} - \frac{72c(9-r^2)^2}{81+9r^2+15r^4-r^6}\Big(2-\frac{2(3-r^2)}{6(9-r^2)}\Big)$$

that implies

$$c \leq \hat{c}_1(r) := \frac{r^2(r^2-3)^2(r^6-15r^4-9r^2-81)}{18(r-3)(r+3)(17r^{10}-312r^8+396r^6+8046r^4-36693r^2+24786)}.$$

Further, $E\pi_{20} \leq E\pi_{22}$ is true for

$$c \leq \hat{c}_2(r) := \frac{r^2(r^6 - 15r^4 - 9r^2 - 81)}{18(r-3)(r+3)(17r^6 - 225r^4 - 729r^2 + 1377)}.$$

Since $\hat{c}_2(r) \leq \hat{c}_1(r)$ we conclude that $E\pi_{21}$ cannot be optimal.

Finally, consider $r \geq r_2$. Here, $E\pi_{21} \leq E\pi_{22}$ implies

$$2 - \frac{3-r^2}{6(9-r^2)} - \frac{1}{18} - 36c\frac{9-r^2}{81+54r^2+r^4}\Big(2 - \frac{1}{18} - \frac{3-r^2}{6(9-r^2)}\Big) - 36c\frac{9-r^2}{(3-r^2)^2}\Big(1 - \frac{3-r^2}{6(9-r^2)}\Big) \leq$$

$$2 - \frac{2(3-r^2)}{6(9-r^2)} - \frac{72c(9-r^2)^2}{81+9r^2+15r^4-r^6}\Big(2 - \frac{2(3-r^2)}{6(9-r^2)}\Big)$$

that implies

$$2c\Big(\frac{15}{3-r^2} + \frac{108}{(3-r^2)^2} + \frac{54(3-43r^2)}{r^4+54r^2+81} + \frac{12(5r^4-96r^2+459)}{r^6-15r^4-9r^2-81} + 32\Big) + \frac{r^2}{81-9r^2}$$

which is positive for the given range of r.

Q.E.D.

Generalized symmetric Beta prior

We study the case where the principal wants to incentivize full information acquisition and sharing.

A generalized beta prior has the form $\theta^{\alpha-1}(1-\theta)^{\beta-1}\frac{1}{B(\alpha,\beta)}$, where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and $\alpha > 0, \beta > 0$.

We study symmetric beta priors so that we set $\beta = \alpha$. Then, $f(\theta_i|\alpha) = \frac{(\theta(1-\theta))^{\alpha-1}\Gamma(2\alpha)}{(\Gamma(\alpha))^2}$. The priors from this family are centered around $\theta = \frac{1}{2}$.

Proposition 11. *Under a symmetric beta prior there exists a unique $r^*(\alpha) \in (0,1)$ such that the optimal linear contract incentivizing two-sided information acquisition and sharing specifies $a_i > 0, b_i = 0$ for all $r \leq r^*(\alpha)$ and $a_i = 0, b_i > 0$ otherwise.*

Proof of Proposition 11. First we note that, similar to the baseline model with the uniform prior, an arbitrarily small coefficient $b_i \geq 0$ for each agent i ensures truthful communication.

We only need to focus on incentives to acquire information.

Suppose that agent i deviates at the information acquisition stage. Then, if i sends $m_i \in \{0,1\}$ to agent j (who believes that i is truthful), then agent i expects the following loss by j's division:

$$E_{s_j,\theta_j}\ell_j(s_j, m_i = 0, \theta_j) = E_{s_j,\theta_j}\ell_j(s_j, m_i = 1, \theta_j) = \frac{\alpha\left(2(\alpha-1)(2\alpha+1)r^2 + (2\alpha+1)^3 + r^4\right)}{2(2\alpha+1)\left((2\alpha+1)^2 - r^2\right)^2},$$

while the loss that i incurs on their own division is

$$E_{s_j,\theta_i}\ell_i(s_j,\theta_i) = \frac{4\alpha^2(r-1) - 2\alpha(r+2) + 3r^2 - 2r - 1}{12(2\alpha+1)^2}.$$

Both losses decrease in r. The function $E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j)$ is convex on $r \in [0,1]$ for any $m_i \in \{0,1\}$: $\frac{\partial E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j)}{\partial r} = \frac{2\alpha^2 r((2\alpha+1)^2+3r^2)}{((2\alpha+1)^2-r^2)^3} > 0$ and $\frac{\partial^2 E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j)}{\partial r^2} = \frac{2\alpha^2(14(2\alpha r+r)^2+(2\alpha+1)^4+9r^4)}{((2\alpha+1)^2-r^2)^4} > 0$, since $\alpha > 0$ and $r \geq 0$. Furthermore, $E_{s_j,\theta_i}\ell_i(s_j,\theta_i)$ is concave on $r \in [0,1]$ (the first derivative with respect to r can be shown to be positive for $1+\alpha-2\alpha^2-3r > 0$, and otherwise negative, and moreover $\frac{\partial^2 E_{s_j,\theta_i}\ell_i(s_j,\theta_i)}{\partial r^2} = -\frac{1}{2(1+2\alpha)^2} < 0$).

Further, at $r = 0$, $E_{s_j,\theta_i}\ell_i(s_j,\theta_i) = \frac{1}{12} > \frac{\alpha}{2(1+2\alpha)^2} = E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j)$ for all $\alpha > 0$. At the same time, for $r = 1$, $E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j) - E_{s_j,\theta_i}\ell_i(s_j,\theta_i) = \frac{2\alpha(\alpha+1)+1}{8(\alpha+1)^2(2\alpha+1)^2} > 0$. But then, on the space $r \in [0,1]$ there exists a unique intersection between $E_{s_j,\theta_i}\ell_i(s_j,\theta_i)$ and $E_{s_j,\theta_j}\ell_j(s_j,m_i,\theta_j)$ at some strictly interior $r^*(\alpha) \in (0,1)$ such that the optimal linear contract for i yields $a_i > 0, b_i = 0$ for $r \leq r^*(\alpha)$ and $a_i = 0, b_i > 0$ for $r > r^*(\alpha)$.

Q.E.D.