Designing Organizations in Volatile Markets

Shuo Liu and Dimitri Migrow*

December 2018

Please click here for the latest version

Abstract

Multinational and multiproduct firms often experience uncertainty in the relative return of conducting activities in different markets due to, for example, exchange rate volatility or the changing prospects of different products. We study how a multi-divisional organization should optimally allocate decision-making authority to its managerial members when operating in such volatile markets. To be able to adapt its decisions to local conditions, the organization has to rely on self-interested division managers to collect and disseminate the relevant information. We show that if communication takes the form of verifiable disclosure, then centralized decision-making does not suffer from information asymmetry and it allows the headquarter of the organization to better cope with the inter-market uncertainty. However, a downside of centralization is that it can discourage information acquisition, and this negative effect is amplified by the need for coordinating the activities of different divisions. As a result, the optimality of decentralized decision-making can actually be driven by a large coordination motive.

Keywords: centralization, decentralization, volatile markets, coordinated adaptation, information acquisition, verifiable disclosure, costly exaggeration

JEL Classification: D82, M52

*Liu: Department of Economics, University of Zurich. Migrow: Department of Economics, University of Calgary. Emails: shuo.liu@econ.uzh.ch and dimitri.migrow@ucalgary.ca. We thank Dan Bernhardt, Julia Grünseis, Yingni Guo, Alexander Jakobsen, Navin Kartik, Alexey Kushnir, Igor Letina, Nick Netzer, Christian Oertel, Rob Oxoby, Harry Pei, Michael Powell, Marek Pycia, Heikki Rantakari, Armin Schmutzler, Alastair Smith, Jakub Steiner, Kremena Valkanova, Alex Whalley, and seminar participants at Bern, Calgary, Warwick and Zurich for many useful comments and suggestions. Shuo Liu would like to acknowledge the financial support by the Forschungskredit of the University of Zurich (grant no. FK-17-018).

1 Introduction

Many modern organizations operate in multiple markets. The most immediate example, perhaps, is that of multiproduct firms: Apple offers both smartphones and watches, BMW sells both cars and motorcycles, and Google's business is not restricted to running a search engine. In the era of globalization, multinational firms provide another important case in point. According to a recent study by Lincoln and McCallum (2018), the median number of destination countries for U.S. exporting firms was three in 2006. Moreover, for many top U.S. exporters, selling domestically produced goods to foreign consumers only counts as a limited part of their involvement in the world economy. For instance, it is common nowadays for a multinational corporation to own production facilities in several foreign countries.[1]

Motivated by the prevalence and increasing influence of multibusiness firms in the economy, this paper asks the following question: When local markets (defined by products, industries, geographic boundaries, demographics of targeted customers, etc.) feature uncertainty in their relative profitability, how should a multi-divisional organization optimally allocate decision-making authority to its managerial members? In particular, should decision rights be *centralized* to a headquarter manager who can coordinate the activities of different division contingent on the market prospects, or be *decentralized* to division managers who have advantages in collecting costly information about local market conditions?[2]

The uncertainty in relative market profitability is a highly relevant problem for many organizations. A broad set of economic and political conditions, which may be difficult to predict, can affect how rewarding it is to conduct activities in a product market or in a country. Thus, the value of success in a particular local market in terms of overall organizational performance is uncertain from the central management's perspective. For multinational firms, a major source of such uncertainty is the volatility of currency exchange rates.[3] With physically different products, uncertainty regarding the relative profitability or strategic importance of the markets may also be due to general shifts in consumer tastes or changes in market size.[4]

If the organizational activities in different local markets are unrelated, relative changes in market profitability may be of little relevance for central management. However, quite

[1]For an overview of the stylized facts about multinational firms documented in the international trade literature, see, e.g., the comprehensive survey by Antràs and Yeaple (2014).

[2]The design of multi-divisional organizations is an active area of research. See, e.g., Athey and Roberts (2001), Alonso, Dessein, and Matouschek (2008, 2015), Dessein, Garicano, and Gertner (2010), Friebel and Raith (2010), and Rantakari (2008, 2013). Section 2 reviews several prominent contributions in this growing literature. We refer interested readers to the excellent survey by Roberts and Saloner (2013).

[3]For example, even without considering the change in tariffs due to the U.S. - China trade war, the recent sharp depreciation of Chinese Yuan already implies a drop in the dollar value per sale that Ford can collect from its joint venture in China (Changan Ford).

[4]Making precise predictions about these factors can be challenging even for large firms. In 2013, soon after announcing his stepping down as the CEO of Microsoft, Steve Ballmer openly admitted that the company was too slow to recognize the importance of the smartphone market. He blamed this strategic failure on Microsoft's long-time focus on the business of its Windows operating system (see "Microsoft too slow on phones, admits boss Steve Ballmer", *BBC News*, 20 September 2013).

often an organization can benefit from synchronizing its activities across markets (e.g., because of economies of scale), though doing so may imply that these activities are less adapted to the local conditions of each market (e.g., product design is less fitted to the tastes of local consumers). When the prospect of each market cannot be perfectly forecasted, resolving the trade-off between *coordination* and *adaptation* is particularly challenging, because ex ante it is unclear whether the organization should adapt more to one market's local conditions and less to the other's. By centralizing the decision-making process, the headquarter manager can take into account the actual profitability conditions and make contingent decisions that are globally optimal for the organization. Yet, to the extent that local market information is privately acquired and observed by division managers, the flexibility granted by centralization also comes with a downside. That is, it may harm the incentives of the division managers by making them more skeptical about how their acquired information will be used.

The main insight of our paper is that whether the above cost of centralized decision-making may outweigh its benefit depends crucially on how important coordination is compared to adaptation, and in an unexpected way. In particular, due to a reinforcing interaction between the uncertainty in market profitability and the need for coordination, the optimality of a *decentralized* authority structure can be the result of a large coordination motive. In addition, as an important step toward establishing the optimality results, we show that if the information acquired by the division managers is verifiable, complete voluntary disclosure arises as a *unique* equilibrium outcome irrespective of the chosen authority structure.

We formalize our arguments by modeling an organization which needs to adapt and coordinate the strategic decisions of its two divisions. As in Alonso et al. (2008) and Rantakari (2008), a division's performance is determined by how close its action (e.g., the design of a product) is matched to an unobserved *local state*, and how well it is coordinated with the action of the other division. Specifically, any mismatch between division i's action and its local state or division j's action will result a quadratic loss in i's performance. Each division is run by an agent (e.g., a divisional manager, he) who can privately exert effort to acquire a signal about the local state, where more effort results in a better signal. The agents are led by a common and uninformed principal (e.g., a headquarter manager, she). While each agent cares only about the performance of his own division, possibly because of career concerns, the principal cares about the *overall* performance of the organization. The novel feature of our model is that the contribution of each division's success to the overall organizational performance need not be certain. This idea is formally captured by a pair of stochastic weights that the principal's payoff attaches to the divisions' performances. While these weights are observed by all players before the final actions are taken, they are unknown at the outset of the game and can be arbitrarily correlated. As mentioned, examples of such interrelated uncertainty include the exchange rate volatility incurred by multinational corporations, as well as the constantly changing prospects of different product markets. We refer to these stochastic weights as the *global states* of our model because, unlike the local states, they determine which actions are globally optimal for the organization rather than locally optimal for individual divisions.

We compare two widely-studied authority structures: centralization and decentralization. In both cases the agents first exert efforts to acquire information about the local states, and then they communicate their findings with the player(s) endowed with decision-making authority. Specifically, under centralization, the agents simultaneously report to the principal, who will subsequently dictate the actions of both divisions. Under decentralization, the agents can exchange messages with each other, after which they make independent decisions over the actions of their own divisions.[5] The communication between players is strategic and takes the form of verifiable disclosure, where an informed agent always has the option of certifying the outcome of his information experiment. This includes the persuasion game of Milgrom (1981) and Grossman (1981) and the evidence game of Dye (1985) as special cases. Previous works have argued that the quality of communication is important for explaining the relative performance of different organizational structures (e.g. Alonso et al., 2008; Aoki, 1986; Dessein and Santos, 2006; Rantakari, 2008). Our model predicts that if information is verifiable, the incentive constraints for communication are irrelevant in determining where the authority over decisions should be lodged in the organization. As we prove in Section 4, fully revealing communication arises as a unique equilibrium outcome regardless of which authority structure is chosen (Propositions 1 - 4). Thus, in equilibrium all the obtained information will be truthfully transmitted to the decision-making parties. The full-revelation result is not obvious, because it is known that costly information acquisition and/or uncertain information endowment can prevent complete voluntary disclosure (e.g., Shavell, 1994; Shin, 1994).[6]

While the resulting quality of communication does not differ between centralization and decentralization, the allocation of decision rights does have an impact on the agents' incentives for information gathering. In Section 5, we first establish a benchmark result (Theorem 1(i)) that if the local markets are always equally profitable, then, regardless of the importance of coordination, a centralized organization always outperforms its decentralized counterpart in motivating information acquisition. This optimality of centralized decision-making shows that the incentive view of delegation in Aghion and Tirole (1997) need not be valid in multi-agent settings with coordination motives. Intuitively, as the principal always deems the two divisions equally important, under centralization she acts as if she were a neutral party who aims to maximize the joint surplus of the agents. In contrast, the decentralized equilibrium outcome fails to achieve the same efficiency because of the conflicting interests between the agents. This

[5]The comparison between centralization and decentralization is only meaningful if contracts are incomplete as in Grossman and Hart (1986) and Hart and Moore (1990), because otherwise any decentralized allocation can be implemented centrally by a suitably designed mechanism. Thus, similar to Alonso et al. (2008) and Rantakari (2008), our analysis applies to situations where the organizational decisions of interests are sufficiently complex (e.g., product design), which renders ex ante contracting infeasible.

[6]Moreover, as shown in Section 6, if the space of the local states is unbounded and the agents are able to misrepresent their private information at a cost, then in addition to full revelation the equilibria under different authority structures will also feature language inflation (Kartik, 2009; Kartik, Ottaviani, and Squintani, 2007).

coordination failure lowers the marginal benefit of information and thus discourages information acquisition under decentralization. Hence, given the fully revealing communication equilibrium outcome, the principal is better off retaining the decision rights when doing so can motivate the agents to acquire more information.[7]

Since the local markets are symmetric ex ante, the equal-profitability condition in the benchmark result above is satisfied if and only if the global states are perfectly and positively correlated. This is a knife-edge case, and the picture changes as soon as we move away from it. A key finding of our paper, stated in Theorem 1(ii), is that provided there is *any* uncertainty in the relative profitability of the local markets, decentralization will outperform centralization in terms of information gathering if coordination is sufficiently important. To understand the result, note that under centralization the principal would prioritize the adaptation problem of the division that turns out to be more profitable. The information passed on by the less profitable division may thus receive little attention. Not surprisingly, since (i) the agents cannot perfectly forecast the relative profitability of the local markets ex ante and (ii) the loss from mis-adapting to one's local state is convex, the uncertainty in the ex post value of information tends to discourage information gathering. What is less obvious, perhaps, is the following reinforcing interaction between this negative effect and the need for coordination. As coordination becomes more important, knowledge about local market conditions plays less of a role in the principal's choice of actions. However, since the adaptation problem of the more profitable division is still relatively more important, the decrease of influence in decision-making is more substantial for the less profitable division. Hence, compared to the case where the agents are autonomous, a large coordination motive can be much more harmful for their information-gathering incentives when the principal is in charge.

Next, in Theorem 2, we show that if the distribution of the global states is sufficiently volatile, then the agents would also acquire more information under decentralization when coordination is of little importance relative to adaptation. Further, by fully characterizing the cases where the global states are binomially distributed, we demonstrate that with high volatility the comparative advantage of decentralization in motivating information acquisition may even hold regardless of the importance of coordination (Propositions 5 and 6). Thus, the more volatile the local markets, the larger the motivational benefits of decentralization.

Whether and when the above benefits of decentralization can outweigh the cost of losing control is not obvious, because it is exactly when the local markets are highly volatile that the principal would most appreciate the flexibility granted by centralization. We show in Theorems 3 and 4 that the answer depends largely on the convexity of the information cost . If the information cost is sufficiently convex, the additional gain in information quality from de-

[7]The optimality result of centralized decision-making echoes the recent experimental findings by Brandts and Cooper (2018). Assigning the subjects with different managerial roles and endowing them with *exogenous* information, their experimental design simulates how two parallel divisional decisions are made in an organization. They find that the subjects are surprisingly honest in communication and that the organizational performance is higher when decision rights are allocated to a central manager.

centralization is at most minor, so it would not be optimal for the principal to transfer the decision rights to the agents. However, if the cost of information is not too convex, the drop in information quality due to centralization is substantial enough to make decentralization optimal. Then, given that strong coordination motives widen the gap in information quality between centralization and decentralization, we arrive at the novel prediction that the importance of coordinating organizational activities can actually strengthen rather than weaken the optimality of decentralized decision-making.

Our results have direct implications on how formal authority over critical decisions should be allocated between organizational agents, which is a design architecture central to the stories of success (or failure) of many modern corporations.[8] More broadly, the results are also related to the core debate in economics on the role of (de)centralized systems in information aggregation and decision-making. Perhaps most famously and influentially, Friedrich von Hayek argued that the problem of rapid adaptation to local changes must be solved by "some form of decentralization" (Hayek, 1945, p. 524), since knowledge about local conditions is dispersed among individual agents rather than existing in concentrated form. However, if adaptation decisions are interdependent and their relative importance is uncertain, efficient use of information may also require some centralized coordination. We show that centralized decision-making need not suffer from the information asymmetry that Hayek criticized, and it is indeed more efficient in adapting to *existing* information. However, decentralization can still be optimal once the *endogeneity* of information is taken into account. This suggests that the fundamental advantage of decentralized systems is in information production.

Our results can shed light on some business cases of multi-divisional corporations. For example, it has been widely discussed that Japan's multinational mobile phone makers perform very poorly in the overseas markets.[9] In particular, many of them, such as NEC and Panasonic, were already struggling in the global competition in the 2000s, which was even before the smartphone era. Their unsatisfactory performance may be explained by the traditional centralized decision-making process of Japanese multinationals (Bartlett and Ghoshal, 2002; Bloom, Sadun, and Van Reenen, 2012) *and* the large economies of scale of the mobile phone industry. For instance, from their entry to China in 1995 until 2002, NEC and Panasonic "released only a few models that were adapted from their models in Japan" (Marukawa, 2009,

[8]See, for example, Freeland (2001) on the history of General Motors in the 1920's - 1960's. Narrative evidence supporting the importance of authority allocations for organizational performance includes the reform of decision-making structure that Louis Gerstner implemented soon after he became the CEO of IBM, which is considered to be a key factor that lead to the firm's success in the 1990's. See, for instance, Gerstner's memoir of his tenure in IBM (Gerstner, 2002) as well as the discussion by Malone (2004). Another case in point is the remarkable failure of the merger of Chrysler and Daimler in 1998. As convincingly argued by Garicano and Rayo (2016), a fatal problem of the merged company, DaimlerChrysler AG, was its poorly-designed allocation of authority. For more systematic empirical evidence, see, e.g., Aghion, Bloom, Lucking, Sadun, and Van Reenen (2017), and Thomas (2011).

[9]See, e.g., online articles "Why Japan's cellphones haven't gone global?", *New York Times*, 19 July 2009, "What happened to Japan's electronic giants?", *BBC News*, 2 April 2013, and "NEC and the sorrow of Japan", *Boy Genius Report*, 17 July 2013.

p. 428). This practice probably had a lot to do with the fact that many components of the Japanese phones were manufactured domestically rather than overseas in low-wage countries, so the cost-saving benefit from coordinating the choices of handset models across borders could be especially significant. But then, according to our theory it is not surprising that these phone makers were slow in learning some basic differences in the distribution structures and consumer tastes between Japan and China (see, e.g., Marukawa, 2009). Both NEC and Panasonic withdrew their mobile phone business from China around 2006, and from the entire global market around 2013.

Of course, centralization is not always inferior to decentralization. For example, for a company that sells both smartphones and personal computers (e.g., Apple and Microsoft), there may be very limited benefits from coordinating the design of these two products because they are supposed to serve different consumer needs. In that case, our theory suggests that it is unlikely that empowering the product managers would lead to much more informed decision-making. Thus, it is more important that the allocation of decision rights does not constrain the company's ability to react promptly to the changing prospects of different products (e.g., by allocating resources across divisions). This may help to understand why Steve Ballmer decided to massively reorganize Microsoft in 2013, moving the governance of the company closer to its very centralized peer Apple.[10]

2 Related Literature

The organizational problem of coordinated adaptation under dispersed information has a long intellectual history in organizational theory and economics (see, among many others, Barnard, 1938; Cyert and March, 1963; Simon, 1947; Williamson, 1975, 1996). Our paper belongs to a growing strand of this literature, which examines how an organization's decision-making structure can determine its ability to coordinate the activities of its sub-units while remaining responsive to changes in the local environments. Specifically, our model builds on the framework developed by Alonso et al. (2008) and Rantakari (2008), which are among the first papers to model strategic information transmission in the context of designing multi-divisional organizations. They focus on the case where information is "soft", meaning that communication between organizational members takes the form of cheap talk (Crawford and Sobel, 1982). One of their most insightful findings is that as the need for coordination increases, the communication of decision-relevant information under centralization (decentralization) becomes less (more) informative. This implies that the comparative advantage of an authority structure need not be monotone in the importance of coordination (Rantakari, 2008). In addition, if the interests of the local managers are sufficiently aligned, then the optimality of decentralized decision-making is not necessarily inconsistent with a large need for coordination (Alonso et al.,

[10]See "Microsoft overhauls, the Apple way", *New York Times*, 11 July 2013.

2008).[11] Our model departs from theirs mainly by (1) focusing on the case where information is "hard" (Grossman, 1981; Milgrom, 1981) or at least is costly to misrepresent (Kartik, 2009; Kartik et al., 2007), and (2) relaxing the (implicit) assumption that the local markets where the organization operates exhibit no uncertainty in their profitability conditions. More important than the modeling differences, we add to Alonso et al. (2008) and Rantakari (2008, 2013), and more generally to the literature of organizational design and coordinated adaptation, by showing that the importance of coordination can even make decentralized decision-making optimal. In particular, this result holds despite the fact that in our model the allocation of decision rights does not affect the informativeness of communication at all, and that the conflicts of interests between the local managers are maximal (as they only care about their own divisions).

Within the literature on organizational design and coordinated adaptation, our paper is further related to Dessein et al. (2010), Friebel and Raith (2010), and Alonso et al. (2015). In Dessein et al. (2010), the organization can better exploit the benefits of cost-saving standardization by integrating its manufacturing activities. Standardization, however, also comes with a loss in revenues because it impedes the organization's ability to tailor its marketing activities to local conditions. Dessein et al. (2010) find that a more decentralized authority structure can better incentivize the managerial members of the organization to exert division-specific effort, but it is still dominated by a more centralized one if the expected value of synergies (akin to the importance of coordination in our model) is sufficiently large. Thus, unlike in our paper, the advantage of decentralized decision-making in incentivizing effort provision is thwarted rather than strengthened by the importance of coordinating activities across organizational units.

In line with Alonso et al. (2008) and Rantakari (2008), both Friebel and Raith (2010) and Alonso et al. (2015) consider settings where the top management of the organization is constrained (and often also harmed) by its informational disadvantage compared to the division managers. In Friebel and Raith (2010), delegating resource-allocating rights to the division managers can be optimal since they control the information about the marginal return of their projects. But delegation can also be sub-optimal because sometimes it is more profitable to concentrate all resources on a single project. In Alonso et al. (2015), the headquarter may be better off by letting the division managers choose their production plans independently given that they know more about the demand conditions of each market, but the opposite may also occur since the costs of production are interdependent. Nevertheless, if the division managers were *non-strategic* in communication, then both the models of Friebel and Raith (2010) and Alonso et al. (2015) would conclude that it is always optimal to have the decisions centrally made. In contrast, in our model, even without the help of message-contingent transfers, the division managers are always incentivized to be truthful when communicating their private information to the decision-making parties.

[11]While both Alonso et al. (2008) and Rantakari (2008) assume that the private information of the managers is exogenous, their main results are subsequently shown to be robust to endogenous information acquisition (Rantakari, 2013). Their models have also been extended to more than two divisions (Yang and Zhang, 2017).

We also contribute to the literature on delegation as an instrument to motivate information acquisition. The seminal work of Aghion and Tirole (1997) argues that delegating formal authority over decisions to an agent can increase his initiative to acquire information, and for the principal of an organization this benefit may outweight the cost due to the "loss of control". However, recent contributions show that the incentive effect of delegation can be ambiguous if the communication between the principal and the agent is strategic.[12] For example, in Argenziano, Severinov, and Squintani (2016), the principal can benefit from retaining the decision-making authority while delegating the task of information acquisition to the agent. This is because the principal may either threaten the agent with a babbling off-path if information gathering is overt, or obstinately expect the information to be highly precise if it is acquired covertly. The finding that centralizing the authority to the principal can better motivate the agent to acquire information compared to delegation is shared by Che and Kartik (2009). A key driving force of their result is that the principal and the agent hold different priors about the state of nature ("opinions"), so under centralization whenever the latter fails to provide any evidence the former would make an adverse inference and take an unfavorable action.[13] All papers above focus on settings with a single agent, whereas ours feature multiple ones.[14] This modeling difference is not superfluous. As we show, the incentive effect of delegation (decentralization) crucially depends on the interaction between the need for coordinating the agents' actions and how their relative performance is valued by the principal.

3 The Model

An organization consists of two operating divisions, $i, j \in \{1, 2\}$, $i \neq j$. Division i's performance (e.g., profits/sales generated, number of patents obtained) is determined by its local conditions, described by $\theta_i \in \mathbb{R}$, and two actions $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$:

$$\pi_i(\mathbf{y}, \theta_i) = K - (y_i - \theta_i)^2 - \delta(y_1 - y_2)^2,$$

where $K > 0$ is some constant, and $\delta > 0$ measures the importance of coordinating actions within the organization. Each *local* state θ_i is independently and identically distributed ac-

[12]Abstracting from strategic communication, the incentive view of delegation is also discussed by Rantakari (2012). He shows that formal delgation is unlikely to be optimal when the quality of implementable projects is determined by both the principal's and the agent's effort choices. The reason is that an unconstrained agent would only be interested in improving the private return of his project. In contrast, under centralization, for his project to be implemented the agent would also need to make it sufficiently attractive to the principal.

[13]A similar persuasive motive of information acquisition under centralization is also present in Newman and Novoselov (2009). In their setting, the principal and the agent share a common prior about the state of nature, but they differ in the costs of committing different types of statistical errors.

[14]Kartik, Lee, and Suen (2017) show that if the principal cannot commit to decision rules ex ante, then having multiple agents compete with each other does not necessarily encourage information acquisition. In their setting, the efforts of the agents are (endogenously) strategic substitutes, whereas in ours, the equilibrium effort choices are strategically independent (see Propositions 2 and 4).

cording to a commonly known distribution Γ with support $\Theta \subseteq \mathbb{R}$. We normalize the mean of the distribution to zero ($\mathbb{E}[\theta_i] = 0$) and assume that it has a finite variance $\sigma_\theta^2 = \mathbb{E}[\theta_i^2] > 0$.

Each division i is run by an agent (e.g., a division manager, he), which we will refer to as agent i. Before any action is taken, each agent i can privately invest effort $e_i \in E = [0, 1]$ in acquiring information about the local state of his division. Specifically, by choosing an effort level $e_i \in E$, agent i incurs a cost of $c(e_i)$ and receives a perfectly revealing signal $s_i = \theta_i$ with probability e_i. With probability $1-e_i$, the agent receives a null signal $s_i = \emptyset$. Thus, the agent's effort enhances the probability that the true state will be revealed by the signal (Green and Stokey, 1981).[15] The realization of the signal is referred to as the agent's type. We assume a twice-differentiable, strictly increasing and strictly convex cost function $c : E \to \mathbb{R}_+$. Each agent cares about the performance of his own division. In particular, the ex post payoff of agent i is given by

$$u_i(\mathbf{y}, \theta_i, e_i) = q\pi_i(\mathbf{y}, \theta_i) - c(e_i),$$

where $q > 0$ captures the marginal benefit for the agent to increase his division's performance (e.g., price of sales, monetary bonus, promotion opportunities). For analytical convenience, we assume throughout the paper that the marginal cost of information is sufficiently small at $e = 0$ (e.g., $c'(0) < q\sigma_\theta^2/2$) and is sufficiently large at $e = 1$ (e.g., $c'(1) > q\sigma_\theta^2$) to ensure an interior solution $e_i \in (0, 1)$.

The agents are led by a common and uninformed principal (e.g., a headquarter manager, she), whose payoff depends on the performance of both divisions and a stochastic vector $\boldsymbol{\eta} = (\eta_1, \eta_2)$:

$$\pi_P(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\eta}) = \eta_1\pi_1(\mathbf{y}, \theta_1) + \eta_2\pi_2(\mathbf{y}, \theta_2). \tag{3.1}$$

Thus, η_i measures the marginal benefit for the principal from increasing division i's performance. As we have discussed in the introduction, there are many economic scenarios where the principal may care about the performance of different divisions to different extents (i.e., $\eta_1 \neq \eta_2$). In the context of multinational corporations, π_i can be profits measured in country i's currency, and η_i is the currency exchange rate between country i and the country where the headquarter is located. Another interpretation is that π_i is a measure (e.g., market share) which summarizes the firm's performance in market i *relative* to its competitors, while η_i reflects demand uncertainty such as changes in market size or in preference intensity ("fashion"). Alternatively, if π_i and q are the number and the price of sales in product market i, then we may have $\eta_i = q - \gamma_i$, where $\gamma_i \geq 0$ is the per unit cost for the headquarter to supply division i

[15] The assumption that an agent's can only acquire an "all-or-nothing" signal simplifies the analysis, but it is not crucial. Our main results can be extended to more general settings where the precision of the signal is increasing in the agent's effort, in the sense that the expectation of the conditional variance $\sigma^2_{\theta_i|s_i} = \mathbb{E}\left[(\theta_i - \mathbb{E}[\theta_i|s_i])^2\right]$ is decreasing in e_i.

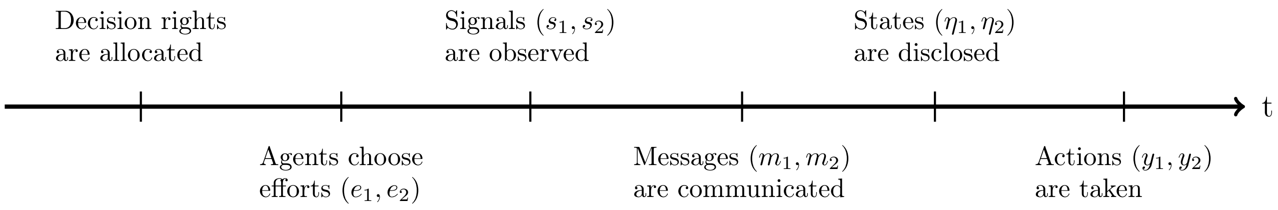


Figure 1: Timing of Events

with necessary resources. Finally, we may instead assume that $\eta_i = q + \gamma_i$, and then interpret γ_i as a parameter that captures (in reduced-form) the strategic importance of succeeding in product market i (e.g., gaining competitive advantage through brand-building or consumer habit-forming).[16]

We assume that the random variables η_1 and η_2 are drawn according to some symmetric and commonly known joint probability distribution $F(\eta_1, \eta_2)$ on the support $\left[\underline{\eta}, \bar{\eta}\right]^2$, where $\bar{\eta} > \underline{\eta} > 0$. The values of η_1 and η_2 are realized and publicly observed *after* the agents have acquired information about their local states (θ_1, θ_2) and *before* the decisions (y_1, y_2) are taken.[17] The uncertainty due to (η_1, η_2) is different from the uncertainty coming from the local states (θ_1, θ_2). First, unlike the local states, η_1 and η_2 are not required to be independently distributed, reflecting the observation that various economic environments are correlated, possibly in a rather complex way (e.g., when η_1 and η_2 are currency exchange rates). Second, from the principal's perspective, how the decision rules of different divisions should be optimally interlinked is determined by the *relative* value of η_1 and η_2. If, for example, $\eta_1 > \eta_2$, the principal would prefer agent 1 to adapt more aggressively towards his local state and agent 2 to focus more on coordination. In other words, η_1 and η_2 determine which actions are globally optimal for the organization rather than locally optimal for individual divisions. We will therefore refer to them as the *global* states of our game. All model parameters are common knowledge.

We complete the model description by specifying how exactly information is communicated and decisions are taken under centralization and decentralization, respectively. Under centralization, the principal takes the decisions (y_1, y_2) after communicating with both agents.[18] Under decentralization, each agent takes the decision of his own division after communicating with each other. Figure 1 summarizes the timing of events in our model.

Independent of the allocation of decision rights, we assume that in the communication stage the agents can credibly reveal their findings about the local states if they want to do so. In

[16]The specification of the utility function (3.1) is also open to the "behavioral" interpretation that the principal has some intrinsic biases and thus favors the two agents unequally. This interpretation relates our model to the growing literature on behaviorally biased managers/supervisors. See, e.g., Prendergast and Topel (1996), Giebe and Gürtler (2012), and Letina, Liu, and Netzer (2018).

[17]Information about (η_1, η_2) may also arrive exogenously before the agents have exerted efforts. Here, an implicit assumption is that such information can be summarized in the common prior F.

[18]As will become clear in Section 4.2, our main results hold regardless of whether the uncertainty of (η_1, η_2) resolves before or after the agents communicate with the principal under centralization.

particular, conditional on receiving a signal $s_i \in \mathcal{S} = \Theta \cup \{\emptyset\}$, agent i can send a message $m_i \in \mathcal{M}(s_i)$ to either agent j (under decentralization) or the principal (under centralization), where we denote $\mathcal{M} = \cup_{s_i \in \mathcal{S}} \mathcal{M}(s_i)$ and assume that the signal-dependent message spaces satisfy the condition below.[19]

(**A1**) $\emptyset \in \mathcal{M}(\emptyset)$, and $\forall s_i \neq \emptyset$, $\exists\, m^{s_i} \in \mathcal{M}(s_i) \setminus \cup_{s_i' \neq s_i} \mathcal{M}(s_i')$.

The essential requirement of assumption (A1) is that an *informed* agent can always certify his type (Seidmann and Winter, 1997). In particular, whenever the message m^{s_i} is communicated, the receiving party will know for sure that agent i has learned the value of the local state θ_i, which is equal to s_i. However, since the message $m = \emptyset$ is not necessarily only available to type $\emptyset$, (A1) allows the possibility that an agent may not be able to prove that he is uninformed. The assumption thus accomodates a large class of communication games. For example, it is satisfied by the evidence game introduced by Dye (1985), where $\mathcal{M}(s_i) = \{s_i, \emptyset\}\ \forall s_i \in \mathcal{S}$, i.e., the agents can always hide but cannot fake their findings about the local conditions. It is also satisfied by the persuasion game studied by Grossman (1981), Milgrom (1981), and Milgrom and Roberts (1986), where $\mathcal{M}(s_i) = \{S \subseteq \mathcal{S} : s_i \in S\}\ \forall s_i \in \mathcal{S}$, i.e., the agents cannot lie but they may send "vague" messages about their findings. Finally, while (A1) rules out pure cheap talk communication, it nevertheless permits the following game of cheap talk with certification: $\emptyset \in \mathcal{M}(\emptyset)$, and $\mathcal{M}(s_i) = \mathcal{S} \cup \{c^{s_i}\}$ if $s_i \neq \emptyset$, where $c^{s_i} \neq c^{s_i'}\ \forall s_i \neq s_i'$. The interpretation is that agent i can either send a non-verifiable message to claim that his type is $\tilde{s}_i \in \mathcal{S}$, or provide a certification to truthfully reveal the signal he has received. However, such a certification is not necessarily available when the agent has failed to obtain any informative signal.

We are interested in how the overall organizational performance is shaped by the interaction between authority allocation and the model's primitives, in particular δ and F (i.e, the coordination motive and the uncertainty/volatility of local market profitability). To answer this question, we first derive and analyze the respective perfect Bayesian equilibria (PBE; Fudenberg and Tirole, 1991, p. 333) of the games under centralization and decentralization (see Section 4). We show that under either of the two organizational structures, full revelation of agents' private signals can always be sustained as part of an equilibrium. Moreover, this is essentially the unique equilibrium outcome of the communication game. We then characterize (i) the agents' effort provision and (ii) the principal's expected payoff in the corresponding PBE, which are uniquely pinned down given the full-revelation communication, and use them to measure the performance of the organization. The main results on the optimal allocation of decision rights are presented in Section 5.

[19] Games with signal-independent message spaces ($\mathcal{M}_i(s_i) = \mathcal{S}\ \forall s_i \in \mathcal{S}$) and costly lying are studied in Section 6.

4 Equilibrium Analysis

4.1 Decentralized authority

We first analyze the game under decentralization. As mentioned, decentralization means that each agent has full control over the decision of his own division. Since the global states (η_1, η_2) only affect the principal's payoff, they are irrelevant for the agents' incentives under decentralization.

Formally, the strategy of each agent $i \in \{1, 2\}$ is a triple (e_i^d, m_i^d, y_i^d) where $e_i^d \in E$ is his effort to acquire decision-relevant information, m_i^d is a mapping that specifies for every given effort-signal pair (e_i, s_i) which message $m_i^d(e_i, s_i) \in \mathcal{M}(s_i)$ agent i will send to agent j, and y_i^d is a decision rule specifying the agent's action $y_i^d(e_i, s_i, m_i, m_j)$ conditional on the effort-signal pair (e_i, s_i) and the messages (m_i, m_j). In equilibrium, each agent i's choices of effort, messages and actions must be sequentially rational with respect to his beliefs (about θ_i, e_j and s_j), which are formed using Bayes' rule whenever applicable. In addition, since the message sets are signal-dependent, we further require that for every $m_j \in \mathcal{M}$ agent i's posterior belief about agent j's signal s_j, which we denote by $\mu_i^j(\cdot|m_j) \in \Delta(\mathcal{S})$, must be *consistent* (Milgrom and Roberts, 1986). Mathematically, this requires that $\mu_i^j(\mathcal{S}^{m_j}|m_j) = 1\ \forall m_j \in \mathcal{M}$, where $\mathcal{S}^{m_j} = \{s_j \in \mathcal{S} : m_j \in \mathcal{M}(s_j)\}$ is the set of signals which could possibly make the message m_j available to agent j.

Our first proposition shows that under decentralization there is essentially a unique equilibrium outcome of the communication stage: despite the conflicts of interests, both agents are incentivized to reveal all their private information.

Proposition 1. *Consider the decentralized authority structure.*

(i) Suppose that $\forall m \in \mathcal{M}$, $\mathcal{S}^m$ is closed. Then, there exists a fully revealing PBE in which $m_i^d(e_i, s_i) = m^{s_i}$ and $m_i^d(e_i, \emptyset) = \emptyset$, $\forall s_i \in \Theta$, $\forall e_i \in E, \forall i = 1, 2$.

(ii) If a communication strategy m_i^* is part of a PBE, then $\mu_j^i(\{s_i\}|m_i^*(e_i^*, s_i)) = 1$ for almost all $s_i \in \mathcal{S} \setminus \{0, \emptyset\}$ with respect to Γ.

The existence of a fully revealing equilibrium is not obvious. In particular, it is known that complete voluntary disclosure need not arise in equilibrium when information is costly to acquire (e.g., Shavell, 1994) and/or when the possibility that the sender is uninformed cannot be ruled out (e.g., Shin, 1994). To gain the intuition for Proposition 1(i), consider an agent i who observes $s_i > 0$ (and thus knows θ_i) and contemplates a deviation from the fully revealing strategy m_i^d. As we require in the proof, agent j always *assumes the worst* in the spirit of Milgrom and Roberts (1986): for every message $m_i \in \mathcal{M}$ observed, j would think that i's type is for sure $\underline{s}^{m_i} \in \arg\min_{s_i \in \mathcal{S}^{m_i}} \left|\mathbb{E}[\theta_i|s_i]\right|$, i.e., the one that minimizes the distance between j's posterior and prior expectations about θ_i among all types who have access to the message m_i. This implies that by deviating to any message $m_i \neq m^{s_i}$, i could only manipulate j to think that on average the local state θ_i is lower than s_i (i.e., $\mathbb{E}[\theta_i|m_i] \leq s_i$).

Now imagine, for the sake of the argument, that agent i knows that j has either received a signal $s_j = \emptyset$ or $s_j = \theta_j \leq \theta_i$. Given that the sequentially rational action for agent j is a weighted average of his posterior expectations of θ_j and y_i, the above manipulation is not profitable for agent i because it will mislead j to take an action even further away from what would have been ideal for i. In contrast, if agent j is known to have received a signal $s_j = \theta_j > \theta_i$, deceiving j to underestimate the value of θ_i could be tempting for agent i, since it may move j's action closer to i's local state θ_i than what j would have chosen otherwise. Of course, as the communication game is simultaneous, when deciding which message to send agent i does not know which of the above two cases j's signal falls into. However, since $\mathbb{E}[\theta_j] = 0$ and $\theta_i = s_i > 0$, agent i does know that either or both of the followings must hold: (i) $\Pr(\theta_j \leq \theta_i) \geq \Pr(\theta_j > \theta_i)$, i.e., a priori $\theta_j \leq \theta_i$ is a more likely scenario compared to $\theta_j > \theta_i$; (ii) $\left|\mathbb{E}[\theta_j|\theta_j \leq \theta_i]\right| \geq \mathbb{E}[\theta_j|\theta_j > \theta_i]$, i.e., the distribution Γ assigns a substantial weight to values that are far smaller than θ_i. Hence, on average the losses from mis-coordination and mis-adaptation are minimized when agent i reveals his type by sending the message m^{s_i} to j.[20]

The second part of Proposition 1 establishes that full revelation of private information is essentially the unique prediction of the communication game under decentralization. In any equilibrium, after the bilateral communication the agents can always be (almost) sure about each other's types, except possibly when the distribution Γ admits an atom at $\theta_i = 0$ and an agent may use the same message for types 0 and $\emptyset$. However, this exception is not payoff-relevant because knowing whether $s_i = 0$ or $s_i = \emptyset$ will not affect the subsequent decisions of the agents. If the distribution Γ is discrete, then the result can be proved by adapting the well-known unraveling argument (Grossman, 1981; Milgrom, 1981). More specifically, in our setting, if several types of agent i are using the same message $m \in \mathcal{M}$, then at least one of them, say s_i, would find that his finding is being understated ($|E[\theta_i|m]| < |E[\theta_i|m^{s_i}]| = s_i$). Thus, by deviating to the type-revealing message m^{s_i} agent i could convince j to take decisions that are more favorable to i in expectation. In the proof, we show how this intuitive argument can be generalized to arbitrary distributions, including the ones that are partly discrete and partly continuous.

Given that the private signals are truthfully revealed in equilibrium, the decision rules (y_1^d, y_2^d) are uniquely pinned down on the equilibrium path. Thus, when calculating the expected payoffs of the agents, the decision rules can be written as functions of the private signals $\mathbf{s} = (s_1, s_2)$ only. Taking these action functions and agent j's effort e_j as given, agent i then

[20] One may envision invoking the general results of Hagenbach, Koessler, and Perez-Richet (2014) to prove the existence of a fully revealing equilibrium in the current model. But their results are not directly applicable to our problem because they do not consider endogenous information acquisition.

solves the following maximization problem at the information acquisition stage:

$$\max_{e_i \in [0,1]} U_i^d(e_i, e_j) = \mathbb{E}_{\boldsymbol{\theta}} \left[\mathbb{E}_{\mathbf{s}} \left[u_i \left(y_1^d(\mathbf{s}), y_2^d(\mathbf{s}), \theta_i, e_i \right] \middle| e_i, e_j \right] \right]. \tag{4.1}$$

It turns out that (4.1) admits a unique solution $e_i^d \in (0,1)$, which is independent of the effort choice of agent j. Hence, the equilibrium outcome at the information acquisition stage is also unique under decentralization. The findings about the stages of decision-making and information acquisition are summarized in the next proposition.

Proposition 2. *In any fully revealing PBE under decentralization, the on-path equilibrium decisions are given by*

$$y_i^d(s_i, s_j) = \frac{1+\delta}{1+2\delta} \cdot \mathbb{E}[\theta_i | s_i] + \frac{\delta}{1+2\delta} \cdot \mathbb{E}[\theta_j | s_j], \ \forall i, j = 1, 2.$$

In addition, both agents exert the same effort

$$e_1^d = e_2^d = e^d \equiv (c')^{-1} \left(\left(1 - \frac{\delta^2 + \delta}{(1+2\delta)^2} \right) q \sigma_\theta^2 \right).$$

Thus, the equilibrium effort level e^d is increasing in q and σ_θ^2. Intuitively, this is because an increase in q or σ_θ^2 leads to a larger expected loss of being uninformed. Further, as formally shown in the Appendix, e^d is decreasing in δ. This is also intuitive: a higher need for coordination makes adaptation less important from the agents' perspective and thus decreases the value of information.

4.2 Centralized authority

In this section, we analyze the game under centralization. Recall that in this case the principal has full control over the decisions of both divisions. Thus, in contrast to the case of decentralization, when making their effort choices and communicating their signals, the agents take into account how the global states (η_1, η_2) may affect the principal's decisions.

Under centralization, each agent i's strategy is a pair (e_i^c, m_i^c), where $e_i^c \in E$ is his effort to acquire information about his local state θ_i and m_i^c is a mapping that specifies for every given effort-signal pair (e_i, s_i) which message $m_i^c(e_i, s_i)$ he reports to the principal. The principal's strategy is a pair of mappings (y_1^c, y_2^c), where $y_i^c(m_i, m_j, \eta_1, \eta_2)$ is the action that the principal takes for division i when receiving messages (m_i, m_j) from the agents and observing the global states (η_1, η_2). In equilibrium, each agent i chooses the effort level and signal-dependent messages that maximize his expected payoff, and the principal chooses actions that are sequentially rational with respect to his beliefs (about $\boldsymbol{\theta}$, $\boldsymbol{s}$ and $\boldsymbol{e}$), which are formed using Bayes' rule whenever applicable. Similar to the case of decentralization, we require that for every

$m_j \in \mathcal{M}$ and $j \in \{1,2\}$ the principal's posterior belief about agent j's type, which we denote by $\mu_p^j(\cdot|m_j) \in \Delta(\mathcal{S})$, must be consistent. That is, $\mu_p^j(\mathcal{S}^{m_j}|m_j) = 1 \ \forall m_j \in \mathcal{M}$.

The next result parallels Proposition 1 in the previous section. It shows that the principal need not be concerned about the agents strategically manipulating their reports under centralization, as they are incentivized to fully reveal their private information in equilibrium.

Proposition 3. *Consider the centralized authority structure.*

(i) Suppose that $\forall m \in \mathcal{M}$, $\mathcal{S}^m$ is closed. Then, there exists a fully revealing PBE in which $m_i^c(e_i, s_i) = m^{s_i}$ and $m_i^c(e_i, \emptyset) = \emptyset$, $\forall s_i \in \Theta$, $\forall e_i \in E, \forall i = 1, 2$.

(ii) If a communication strategy m_i^* is part of a PBE, then $\mu_p^i(\{s_i\}|m_i^*(e_i^*, s_i)) = 1$ for almost all $s_i \in \mathcal{S} \setminus \{0, \emptyset\}$ with respect to Γ.

The intuition of Proposition 3 is similar to that of Proposition 1. Together, our full-revelation results suggest that the allocation of decision rights does not affect the quality of communication in the organization. This finding can even be extended to settings where the message sets are type-independent. In Section 6 we show that fully revealing equilibria also arise in a communication game with $\mathcal{M}(s_i) = \mathcal{S} \ \forall s_i \in \mathcal{S}$ and costly exaggeration. Our results are in sharp contrast to Alonso et al. (2008) and Rantakari (2008), who show that if information about local states is dispersed and held by agents who communicate via cheap talk, the relative performance of different authority structures depends crucially on their endogenous quality of communication. Moreover, our results show that centralized decision-making does not suffer from the usual problem of information asymmetry, as the principal can elicit all the information from the agents even without the help of contingent transfers. Thus, different from related works such as Dessein (2002) and Deimen and Szalay (2018), strategic communication does not give rise to a trade-off between loss of control under delegation/decentralization and loss of information under centralization in our setting.

Given that the private signals are truthfully revealed in equilibrium, the principal's decision rules (y_1^c, y_2^c) are uniquely pinned down on the equilibrium path. Thus, when calculating the expected payoffs of the agents, the decision rules can be written as functions of the private signals $\mathbf{s} = (s_1, s_2)$ only. Especially, the action chosen by the principal for each division will be a weighted sum of the conditional expectations $\mathbb{E}[\theta_i|s_i]$ and $\mathbb{E}[\theta_j|s_j]$, while the weights will depend on the realization of the global states. Taking the principal's on-path decision rules and agent j's effort e_j as given, agent i then solves the following maximization problem at the information acquisition stage:

$$\max_{e_i \in [0,1]} U_i^c(e_i, e_j) = \mathbb{E}_{\boldsymbol{\theta}}\left[\mathbb{E}_{\mathbf{s}}\left[\mathbb{E}_{\boldsymbol{\eta}}\left[u_i(y_1^c(\mathbf{s}, \boldsymbol{\eta}), y_2^c(\mathbf{s}, \boldsymbol{\eta}), \theta_i, e_i)\right]|e_i, e_j\right]\right]. \tag{4.2}$$

Similar to the parallel problem under decentralization, (4.2) admits a unique solution $e_i^c \in (0,1)$, which is independent of e_j. Using the symmetry of the distribution of (η_1, η_2), we then obtain the following result:

Proposition 4. *In any fully revealing PBE under centralization, the on-path equilibrium decisions are given by*

$$y_i^c(s_i, s_j, \eta_i, \eta_j) = \frac{\frac{\eta_i}{\eta_i+\eta_j} \cdot \left(\frac{\eta_j}{\eta_i+\eta_j} + \delta\right) \mathbb{E}\left[\theta_i | s_i\right] + \frac{\delta \eta_j}{\eta_i+\eta_j} \mathbb{E}[\theta_j | s_j]}{\frac{\eta_i}{\eta_i+\eta_j} \cdot \frac{\eta_j}{\eta_i+\eta_j} + \delta}, \ \forall i, j = 1, 2.$$

In addition, both agents exert the same effort

$$e_1^c = e_2^c = e_F^c \equiv (c')^{-1} \left(\left(1 - \mathbb{E}_\lambda \left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2} \right] \right) q\sigma_\theta^2 \right),$$

where $\lambda \equiv \eta_1/(\eta_1+\eta_2)$.[21]

Similar to the case of decentralization, the equilibrium effort level e_F^c is unambiguously increasing in q and σ_θ^2. In addition, as we show in the Appendix, e_F^c is decreasing in δ. Intuitively, the value of information decreases in the importance of coordination because it makes adaptation less important from the perspective of all players. It is less clear, however, how the equilibrium effort level depends on the distribution of the global states (η_1, η_2). We investigate this question in the next section as we compare the effort provision under both organizational forms.

5 Comparing Organizational Structures

Having analyzed separately the fully revealing equilibria under centralization and decentralization, we now ask which allocation of decision rights is optimal for the organization. In our model, an immediate candidate for the criterion of optimality is the principal's expected payoff. Since communication is fully revealing and the principal directly controls the divisional decisions under centralization, a sufficient (necessary) condition for her to benefit more from a centralized (decentralized) authority structure is the extent of the agents' effort provision.[22] Hence, comparing agents' efforts under centralization and decentralization provides a useful

[21]Note that the distribution of λ can be derived from the joint distribution of (η_1, η_2):

$$\Pr(\lambda \le x) = \int_{[\underline{\eta}, \bar{\eta}]^2} \mathbb{1}_{\{\eta_1/(\eta_1+\eta_2) \le x\}} dF(\eta_1, \eta_2) \forall x \in \mathbb{R}.$$

[22]If the set of feasible effort choices is binary and is given by $E = \{0, \bar{e}\}$, where $\bar{e} \in (0,1]$, then $e^d > e_F^c$ is not only necessary but also sufficient for concluding that the principal is better off under decentralization.

stepping stone for answering the question of which allocation of decision rights is optimal for the principal. Moreover, the comparison of effort provision can be of interest per se, especially if one is concerned that our model may not capture *all* the benefits of learning for the organization. With these motivations in mind, in what follows we will start by analyzing the relative performance of the organization in terms of effort provision (Section 5.1). The analysis of the principal's expected payoff will then be presented in Section 5.2.

5.1 Effort provision

Propositions 2 and 4 directly imply that the equilibrium effort level is higher under decentralization than that under centralization ($e^d > e^c_F$) if and only if

$$D(\delta) = \frac{\delta^2 + \delta}{(1+2\delta)^2} < C_F(\delta) = \mathbb{E}\left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right], \tag{5.1}$$

where we recall that $\lambda = \eta_1/(\eta_1 + \eta_2)$, and we drop the subscript λ from the expectation operator for brevity.

To understand the above condition, note that when choosing his effort, an optimizing agent balances the marginal benefit and the marginal cost of information. Condition (5.1) is then equivalent to the statement that the marginal benefit of information is higher under decentralization than that under centralization. More specifically, if the agent had the right to choose actions for *both* divisions, he can always make sure that they are perfectly coordinated ($y_j = y_i$). In this case, the expected benefit of exerting an additional unit of effort will be $q\sigma_\theta^2$, which is exactly the payoff difference between taking an informed and ideal decision ($y_i = \theta_i$) and an uninformed one ($y_i = \mathbb{E}[\theta_i]$). When decision rights are decentralized, each agent only has control over his own division, and coordination is no longer perfect due to conflicting interests between the agents. Hence, from the agents' perspective, the value of information is impaired by the need of coordination, and so the marginal benefit of being better informed decreases to $\text{MB}^d = (1 - D(\delta))q\sigma_\theta^2$. Under centralization, the principal may better coordinate the actions of the divisions. However, because the *relative* profitability of the local markets (η_i/η_j) is uncertain, a forward-looking agent would be concerned that the principal may prioritize the adaptation problem of the other division and thus pay little attention to his acquired information. These effects jointly determine the marginal benefit of information under centralization, which is $\text{MB}^c = (1 - C_F(\delta))q\sigma_\theta^2$. It is then straightforward to check that $\text{MB}^d > \text{MB}^c$ if and only if (5.1) holds.

Exploiting the limiting properties of the functions $D(\delta)$ and $C_F(\delta)$ in (5.1), our first theorem below shows that, except for the knife-edge case where the global states are perfectly and positively correlated, a decentralized organization outperforms its centralized counterpart in terms of incentivizing effort provision (or information gathering) whenever coordination is sufficiently important.

Theorem 1. *Let $corr(\eta_1, \eta_2)$ be the correlation of the global states.*

(i) If $corr(\eta_1, \eta_2) = 1$, then $e^d < e_F^c \;\; \forall \delta > 0$.

(ii) If $corr(\eta_1, \eta_2) < 1$, then $\exists \bar{\delta} \in [0, +\infty)$, such that $e^d > e_F^c \; \forall \delta > \bar{\delta}$. In addition, the difference $e^d - e_F^c$ is increasing in $\delta \;\; \forall \delta > \bar{\delta}$.

To see the intuition, remember that under decentralization the agents are free to adjust their actions according to the acquired information. However, as in many settings with partial coordination motives and lack of commitment, the decentralized equilibrium outcome is not Pareto efficient for the agents. In contrast, whenever the global states are perfectly and positively correlated, under centralization the principal acts as a neutral party maximizing both agents' payoffs. In this case, a centralized authority structure effectively allows the agents to commit to the efficient action plans for *any* given information acquired. As a result, regardless of the importance of coordination, from an individual agent's perspective the marginal benefit of information is highest when decisions are centrally made by the principal.

However, the picture changes as we move away from the knife-edge case of a perfectly predictable profitability ratio $\eta_1/\eta_2 = 1$. Ex post, the performances of the two divisions may not be equally profitable/important for the principal, so she only aims to maximize a *weighted* sum of the agents' surplus. Thus, even though the principal values both divisions equally *on average* (i.e., $\mathbb{E}[\eta_1]/\mathbb{E}[\eta_2] = 1$), centralization is less valuable for the agents as a commitment device because of the uncertainty of the global states. Importantly, this negative uncertainty effect is further amplified by the need for coordination: as the latter increases, a biased principal gives substantially less consideration to her unfavored agent's report. This happens because she primarily wants the more profitable division to adapt more aggressively to its local state while minimizing the mis-coordination costs.[23] Eventually, the positive commitment effect of centralized decision-making is dominated by the negative effect due to the volatile profitability of the local markets, leading to an increasing gap in effort provision between decentralization and centralization.

Part (ii) of Theorem 1 establishes that a decentralized organization induces more efforts from the agents *if* the importance of coordination exceeds some cutoff value $\bar{\delta} \geq 0$. One may wonder whether the converse of this statement also holds, i.e., whether it is the case that a centralized organization is better in terms of effort provision whenever coordination is sufficiently *unimportant*. Note that this question is only meaningful if $\bar{\delta} \neq 0$. In the next subsection, we show that the lower bound $\bar{\delta} = 0$ can indeed be achieved by some distributions, implying that in those cases decentralization outperforms centralization in terms of effort provision

[23] To formalize this intuition, let $w = \frac{\eta_i}{\eta_i + \eta_i} \cdot \left(\frac{\eta_j}{\eta_i + \eta_j} + \delta \right) \Big/ \left(\frac{\eta_i}{\eta_i + \eta_j} \cdot \frac{\eta_j}{\eta_i + \eta_j} + \delta \right)$ be the strategic weight that the principal would assign to agent i's private information when making decision y_i under centralization (see Proposition 4). It can be shown that $\frac{\partial w}{\partial \eta_i} > 0$ and $\frac{\partial^2 w}{\partial \eta_i \partial \delta} > 0 \; \forall \delta > 0$ and $\forall \eta_1, \eta_2 \in \left[\underline{\eta}, \bar{\eta} \right]$, i.e., the principal's decision weights will respond more aggressively to the profitability conditions of the local markets as the need for coordination increases.

whenever coordination is of *any* importance. Nevertheless, as we will also show by example in the next subsection, when the cut-off is strictly positive it is not necessarily the case that $e^d < e^c_F \; \forall \delta \in (0, \bar{\delta})$. In particular, while for intermediate values of δ centralization may indeed outperform decentralization in terms of effort provision, it may fail to do so when the need for coordination is relatively small. The next result shows that this is likely to happen when the profitability conditions of the local markets are very volatile.

Theorem 2. *If* $\mathbb{E}\left[\frac{1}{\lambda^2}\right] > \mathbb{E}\left[\frac{2}{\lambda(1-\lambda)}\right] - 3$, *then* $\exists\, \underline{\delta} \in (0, +\infty]$, *such that* $e^d > e^c_F \; \forall \delta \in (0, \underline{\delta})$.

To understand Theorem 2, note that its condition is violated if $corr(\eta_1, \eta_2) = 1$, as this implies that $\Pr(\lambda = 0.5) = 1$. By continuity, it must also be violated if η_1 and η_2 are sufficiently positively correlated, meaning that the principal is unlikely to strongly bias her decisions in favor of the more profitable division ex post. As the profitability conditions become less and less positively correlated, the strategic weights η_1 and η_2 that the principal assigns to the two divisions are more likely to be extreme (i.e, λ is more likely to take values that are close to 0 and 1). This makes the condition of Theorem 2 more likely to be satisfied.[24] Thus, Theorem 2 captures the intuition that if the principal is likely to be highly biased ex post, then centralizing the decision rights can strongly discourage the agents from acquiring valuable information even when the motive of coordination is small. Therefore, the scope for decentralization to outperform centralization in effort provision is larger when the profitability conditions of the local markets are more volatile.

5.1.1 Binary distributions: characterizations

In this section, we use a class of binary distributions $\{F_\omega\}_{\omega \in [0,1)}$ to illustrate our main findings regarding the effect of decision right allocation on effort provision: for every $\omega \in [0, 1)$ the distribution F_ω is characterized by

$$\Pr(\eta_1 = 1 + \omega, \eta_2 = 1 - \omega) = \Pr(\eta_1 = 1 - \omega, \eta_2 = 1 + \omega) = \frac{1}{2}. \tag{5.2}$$

Thus, ω can be interpreted as a measure of both the volatility of the local markets' profitability conditions and the ex post bias of the principal: the larger ω, the more volatile are the local markets (since $\mathbb{E}[(\eta_i - \mathbb{E}[\eta_i])^2] = \omega^2$ and $Cov(\eta_1, \eta_2) = -\omega^2$) and the more biased is the principal ex post (as $|(\eta_i - \eta_j)/(\eta_i + \eta_j)| = \omega$).

For the above class of binary distributions, we fully characterize when a decentralized organization outperforms its centralized counterpart in providing incentives to the agents for exerting costly yet valuable effort. Fixing the volatility of the profitability conditions, or the

[24] While both $\mathbb{E}[\frac{1}{\lambda^2}]$ and $\mathbb{E}\left[\frac{2}{\lambda(1-\lambda)}\right]$ may increase if the distribution of λ puts more weight toward to endpoints of the interval $[0, 1]$, the first term increases much faster because of its quadratic form.

degree of the principal's ex post bias, the next result shows how this regime is shaped by the importance of promoting synergies in the organization.

Proposition 5. *Consider any binary distribution F_ω with $\omega \in [0, 1)$.*

(i) If $\omega \le \sqrt{2} - 1$, then $e^d > e^c_F$ if and only if $\delta \in (0, \max\{0, \underline{\delta}(\omega)\}) \cup (\bar{\delta}(\omega), +\infty)$, where

$$\underline{\delta}(\omega) \equiv -\frac{\omega^4 + 4\omega^2 - 1}{8\omega^2} - \frac{(1+\omega^2)\sqrt{\omega^4 - 6\omega^2 + 1}}{8\omega^2},$$

and

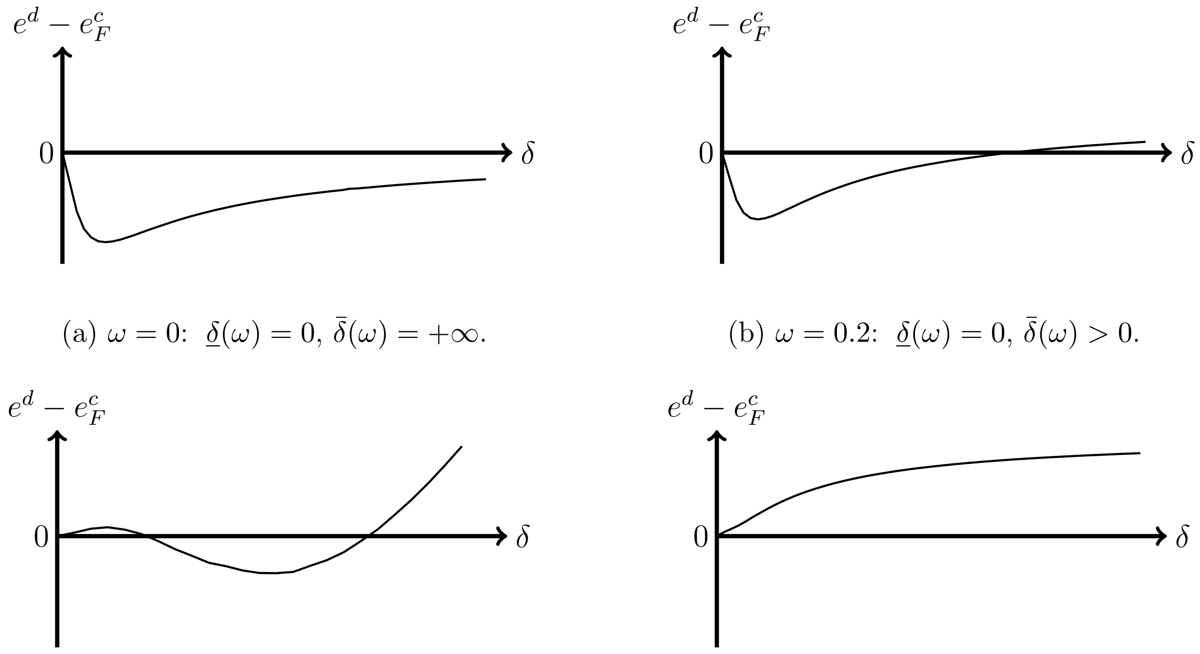
$$\bar{\delta}(\omega) \equiv -\frac{\omega^4 + 4\omega^2 - 1}{8\omega^2} + \frac{(1+\omega^2)\sqrt{\omega^4 - 6\omega^2 + 1}}{8\omega^2},$$

with $\underline{\delta}(\omega) = 0$ if and only if $\omega \le \sqrt{\frac{2\sqrt{3}}{3} - 1} \approx 0.393$, and $\lim_{\omega \to 0} \bar{\delta}(\omega) = +\infty$.

(ii) If $\omega > \sqrt{2} - 1$, then $e^d > e^c_F$ $\forall \delta > 0$.

Figure 2 provides an illustration of Proposition 5 as well as the key messages of Theorems 1 and 2. For the benchmark case of no uncertainty in the profitability conditions ($\omega = 0$), Figure 1(a) shows that the agents always work harder under centralization ($e^d - e^c_F < 0$). This illustrates Theorem 1(i) and the asymptotic result $\lim_{\omega \to 0} \bar{\delta}(\omega) = +\infty$ from Proposition 5(i). As we start introducing uncertainty to the profitability conditions (η_1, η_2), both Theorem 1(ii) and Proposition 5(i) suggest that a decentralized authority structure is superior in guaranteeing effort provision (and thus also in information production) whenever the need for coordination is sufficiently strong. Figure 2(b) demonstrates that a strong coordination motive is also necessary for the equilibrium effort level to be higher under decentralization if the uncertainty of profitability conditions is sufficiently small. If the degree of uncertainty takes an intermediate value, then additionally we have $e^d > e^c_F$ when coordination is sufficiently unimportant relative to adaptation ($\delta < \underline{\delta}(\omega)$). As depicted in Figure 2(c), in this case centralizing the decision rights improves the effort provision if and only if the need for coordination is also intermediate. This echoes the finding of Theorem 2.[25] Finally, when the degree of uncertainty becomes sufficiently large ($\omega > \sqrt{2} - 1$), the agents anticipate that the principal will be heavily biased when making decisions. This substantially impairs the marginal benefit of information under centralization from the agents' perspectives. Proposition 5(ii) and Figure 2(d) show that in such scenarios decentralization is optimal for guaranteeing effort provision regardless of the importance of coordination.

[25] With the binary distributions (5.2), it can be verified that the inequality condition in Theorem 2 is equivalent to $\omega > ((2\sqrt{3})/3 - 1)^{1/2}$, which is also necessary and sufficient for the cutoff $\underline{\delta}(\omega)$ in Proposition 5 to be strictly positive.



(a) $\omega = 0$: $\underline{\delta}(\omega) = 0$, $\bar{\delta}(\omega) = +\infty$.

(b) $\omega = 0.2$: $\underline{\delta}(\omega) = 0$, $\bar{\delta}(\omega) > 0$.

(c) $\omega = 0.41$: $\underline{\delta}(\omega) > 0$, $\bar{\delta}(\omega) > 0$.

(d) $\omega = 0.45$: $\underline{\delta}(\omega) = +\infty$, $\bar{\delta}(\omega) = 0$.

Figure 2: The dependence of effort difference on δ, with $c(e) = e^2$.

To further sharpen our understanding of the effort provision under both organizational forms we next fix the degree of the coordination requirement and ask, how does the effort provision depend on the volatility of the profitability conditions? As one may already expect, the equilibrium effort level is higher under decentralization if and only if the profitability conditions are sufficiently volatile (i.e., ω is sufficiently large). Perhaps less intuitively, the range of the volatility parameter ω for which decentralization provides more powerful incentives (i.e. the set $\{\omega \in (0,1) : e^d > e^c_F\}$) does not change monotonically with respect to the coordination motives. Starting with a situation where coordination is of little (large) importance, an increase in the need for coordination makes it more likely that the agents exert more effort under centralization (decentralization). These observations are summarized in the next proposition.

Proposition 6. *Given a binary distribution F_ω with $\omega \in (0,1)$ and $\delta > 0$, $e^d > e^c_F$ if and only if $\omega > \hat{\omega}(\delta)$, where*

$$\hat{\omega}(\delta) = \sqrt{\frac{(4\delta+2)\sqrt{4\delta^2+4\delta+3} - 2\delta - 3}{4\delta+3} - 2\delta}.$$

The cutoff $\hat{\omega}(\delta)$ is strictly increasing on $\left(0, \frac{\sqrt{2}-1}{2}\right)$, and it is strictly decreasing on $\left(\frac{\sqrt{2}-1}{2}, +\infty\right)$, with $\hat{\omega}\left(\frac{\sqrt{2}-1}{2}\right) = \sqrt{2}-1$ and $\lim_{\delta \to +\infty} \hat{\omega}(\delta) = 0$.

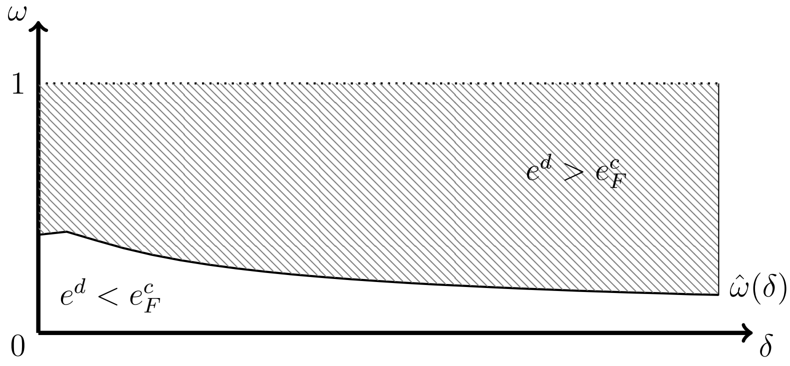


Figure 3: The cutoff $\hat{\omega}(\delta)$ and the regimes for $e^d > e^c_F$ and $e^d < e^F_c$.

The insight of Proposition 6 is further highlighted in Figure 3, where the hatched area indicates the regime of parameters for which the equilibrium effort level is higher under decentralization. Notably, this graphic representation does not require any specification of the effort cost function. This shows the generality and robustness of the qualitative results that we have obtained so far.

5.2 The principal's payoff

In this section, we turn to the question of when the principal can benefit from centralization (decentralization). The immediate implication of the full-revelation results (Propositions 1 and 3) is that centralization is optimal for the principal whenever it can better motivate the agents than decentralization ($e^d < e^c_F$), since this allows her to adjust the relevant organizational activities to better support the (ex post) more profitable division without sacrificing the (ex ante) informativeness of the decisions. As suggested by the characterization results Propositions 5 and 6, the principal is more likely to confront such a straightforward comparison between organizational forms when the need for coordination is small or intermediate and the local markets are not too volatile.

However, in the previous section we have also shown that the agents' incentives for information gathering are lower under centralization whenever the need for coordination is sufficiently large and/or the local markets are sufficiently volatile in their profitability conditions. If the disadvantage of centralization in motivating information gathering is substantial enough, having the flexibility to adapt decisions to the actual profitability conditions may not be so valuable for the principal after all.[26] The next result provides a sufficient condition under which the gap in effort provision between centralization and decentralization is large enough for the principal to prefer the latter. Specifically, we show that decentralization will outperform centralization in terms of the principal's expected payoff provided the effort cost function is not too convex

[26]This argument can be best understood by considering the extreme case where both agents exert very little effort under centralization: given the poor quality of information, the principal often have to take the uninformed decisions ($y_i = y_j = \mathbb{E}[\theta_i]$). Thus, the option of tailoring decisions to (η_1, η_2) is not quite useful.

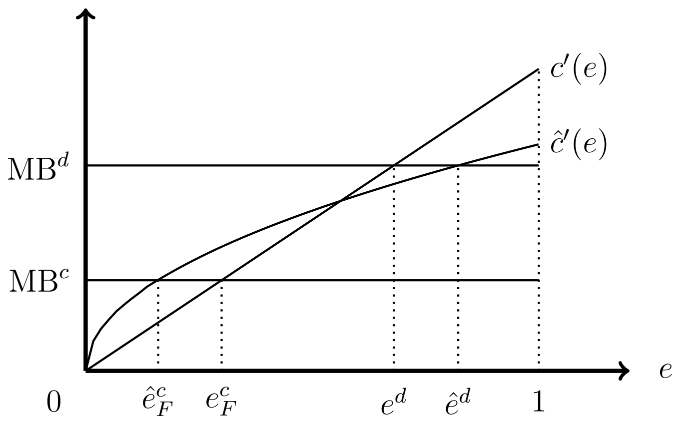


Figure 4: Differences in equilibrium effort levels with $c(e) = e^2$ and $\hat{c}(e) = e^{1.5}$.

and coordination is sufficiently important.

Theorem 3. *Suppose that* $corr(\eta_1, \eta_2) < 1$. *There exists* $\zeta > 0$, *such that if* $c''(e) \cdot e < \zeta$ $\forall e \in [0,1]$, *then we have* $\Pi_P^c < \Pi_P^d$ *for sufficiently large* δ.

To understand the above theorem, consider again an individual agent who is deciding on how much effort to invest in the task of information acquisition. As Theorem 1 shows, if the local markets exhibit *any* uncertainty in their relative profitability ($corr(\eta_1, \eta_2) < 1$) and coordination is sufficiently important, the marginal benefit of effort is higher when decision rights are allocated to the agents. The gap in the marginal benefits of effort between centralization and decentralization then translates into a gap in effort provision. Intuitively, this gap in effort provision will be larger if the derivative c' does not increase very fast, because the equilibrium effort levels are chosen to balance the corresponding marginal benefits and marginal costs. Figure 4 provides a graphical illustration of this intuition: Consider two cost functions $c(e) = e^2$ and $\hat{c}(e) = e^{1.5}$. The marginal cost is arguably increasing faster (on average) in the former case than in the latter (since $\mathbb{E}[c''(e)] > \mathbb{E}[\hat{c}''(e)]$). As we can see from the figure, for given marginal benefits of effort under centralization (MB^c) and decentralization (MB^d) with $\mathrm{MB}^d - \mathrm{MB}^c > 0$, the gap in effort provision is larger when the cost function is $\hat{c}$ than when it is c (i.e., $\hat{e}^d - \hat{e}_F^c > e^d - e_F^c$). In fact, in this case, the argument $\hat{e}^d - \hat{e}_F^c > e^d - e_F^c$ also follows from the observation that the marginal cost function $\hat{c}'(e)$ is a concave transformation of $c'(e)$. More generally, if the cost function takes the form $c(e) = ke^\alpha$, where $k > 0$ and $\alpha > 1$, then the "sufficiently small ζ" condition in Theorem 3 can be replaced by the requirement that the power parameter α is sufficiently close to one - in other words, the marginal cost function is sufficiently concave.[27] However, the proof of Theorem 3 shows that what is crucial is not the

[27]In fact, one can show that with the cost function $c(e) = ke^\alpha$, there exists a cutoff $\alpha^* > 1$, such that the conclusion of Theorem 3 holds if *and* only if $\alpha \leq \alpha^*$ (See Figure 5 for further illustration). More generally and similar to Theorem 3, if information cost is sufficiently convex, the motivational advantage of decentralization need not make it optimal for the principal even when coordination is extremely important.

concavity of the marginal cost function, but rather the bound of speed at which it grows.

We close this section with a result that parallels Theorem 2: if the profitability conditions of the local markets are sufficiently volatile and the cost function is not too convex, then, even when the need of coordination is relatively small, the resulting gap in effort provision can be substantial enough to make decentralization optimal for the principal.

Theorem 4. *Suppose that* $\mathbb{E}\left[\frac{1}{\lambda^2}\right] > \mathbb{E}\left[\frac{2}{\lambda(1-\lambda)}\right] - 3$*. For sufficiently small* $\delta > 0$*, there exists* $\zeta(\delta) > 0$*, such that if* $c''(e) \cdot e < \zeta(\delta)$ $\forall e \in [0, 1]$*, then* $\Pi_P^c < \Pi_P^d$*.*

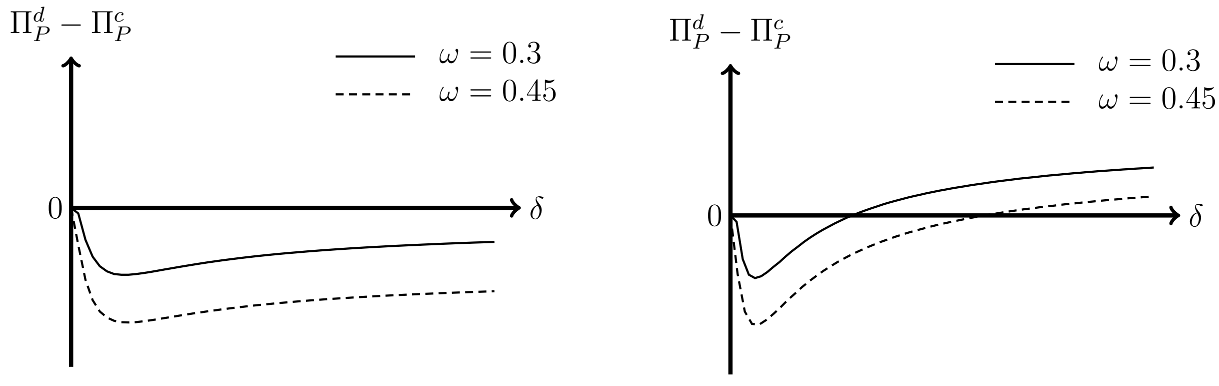
Unlike the uniform cutoff ζ in Theorem 3, the cutoff $\zeta(\delta)$ in Theorem 4 is δ-specific. From a technical point of view, this is because regardless of the distribution of the global states, the expected payoffs of the principal under both authority structures (i.e., Π_P^c and Π_P^d) converge to each other as δ goes to zero. A deeper insight we can gain from this exercise is that the optimal authority structure is more ambiguous when coordination is not so important *and* the local markets are highly volatile in their profitability conditions. In such cases, while decentralization can lead to more motivated agents (see Theorem 2 and Proposition 5), given the large uncertainty in relative market profitability the principal would also find the power of making contingent decisions especially valuable.

5.2.1 Binary distributions: examples

To sharpen our understanding on the role of the convexity of the cost function in determining the relative expected payoff of the principal under centralization and decentralization, we consider again the class of binary distributions $\{F_\omega\}_{\omega\in[0,1)}$ introduced in section 5.1.1. Note that assuming general cost functions makes it difficult to obtain characterization results that parallel Propositions 5 and 6. Thus, we look at particular cost functions and specify the degree of market volatility (ω) to illustrate how the principal's optimal organizational form depends on the coordination requirement.

Consider two cost functions, $c(e) = e^2$ and $c(e) = e^{1.25}$, and two situations of volatility, $\omega = 0.3$ and $\omega = 0.45$. The choices of ω are meant to be representative. According to Proposition 5, for $\omega = 0.3$ the agents exert higher effort under decentralization if and only if coordination is sufficiently important. In contrast, for $\omega = 0.45$ decentralization outperforms centralization for any degree of the coordination requirement.

In Figure 5(a), we let the cost function be $c(e) = e^2$. For both cases $\omega = 0.3$ and $\omega = 0.45$, the principal's expected payoff is always higher under centralization independent of the coordination parameter δ. Thus, with the quadratic cost function, the negative uncertainty effect of centralization on effort provision is not too severe a concern from the principal's perspective. Thus, centralization dominates decentralization by its advantage of allowing the principal to tailor the organizational activities to the actual profitability conditions across



(a) $c(e) = e^2$, $q\sigma_\theta^2 = 1.25$.

(b) $c(e) = e^{1.25}$, $q\sigma_\theta^2 = 1.25$.

Figure 5: The principal's payoff and cost functions with different degrees of convexity.

markets. Moreover, as Figure 5(a) shows, the value of such flexibility in decision-making is particularly high for the principal when the volatility of the profitability conditions is large. In the figure, the dashed curve lies strictly below the non-dashed one, meaning that regardless of the importance of coordination the payoff difference $\Pi_P^d - \Pi_d^c$ is more negative for $\omega = 0.45$ compared to $\omega = 0.3$.

The pattern that centralization is relatively more attractive to the principal when the volatility/bias measure ω is larger is shared by Figure 5(b), where we use a less convex cost function $c(e) = e^{1.25}$. However, unlike in the previous case, here the negative effect of centralization on effort provision is amplified *sufficiently* by the need for coordination. In both cases $\omega = 0.3$ and $\omega = 0.45$, as the coordination parameter δ increases, the difference in effort provision (and thus also in the quality of information) eventually becomes so large that the principal have to take the uninformed decisions much more often under centralization. Hence, confirming the finding of Theorem 3, when the effort cost is not too convex and the need for coordination is sufficiently large the principal is worse off by having the decisions centrally made.

6 Costly Exaggeration

So far, the communication stages under both centralization and decentralization have been modeled as a game with verifiable information: an agent can always send a certified message and reveal the finding of his information acquisition experiment to the receiving party. The crucial implication of this assumption is that in our model, the fundamental difference between centralization and decentralization is not the endogenous quality of communication - in both cases the messages communicated by agents will be truthful and fully informative - but rather the quality of information, which is endogenously determined by the effort of the agents.

The verifiability assumption is intended to capture situations where the decision-relevant information held by organizational agents is in the form of hard evidence, or at least it can be supported by objective measures. For instance, a division manager may conduct marketing research with statistical analysis to convey information about the consumer demand of his responsible market. However, there are certainly settings where one may view the assumption of perfect verifiability restrictive. For example, when a specialized manager provides marketing research showing that the consumer demand is high, others in the organization may not be able to tell for sure whether the conclusion is driven by a deliberate (and possibly biased) choice of statistical methods of analysis. If the manager wants to exaggerate the consumer demand by manipulating his data, the imperfect verifiability of information can be problematic because it seems conceivable that an exaggerated report may not be caught by his colleagues, especially when it is not too far away from the truth. In what follows, we will show that the insights from our full revelation results are robust provided that such exaggeration is not entirely costless.

Specifically, suppose that $\Theta = \mathbb{R}$, and when communicating (either with the principal or with each other) the agents are allowed to send any message $m_i \in \mathcal{M} = \mathbb{R} \cup \{\emptyset\}$, irrespective of the true findings of their experiments. However, given the true signal is $s_i \in \mathcal{S} = \mathbb{R} \cup \{\emptyset\}$, sending a message $m_i \in \mathcal{M} = \mathbb{R} \cup \{\emptyset\}$ will incur a non-negative cost $z(m_i, s_i)$ to agent i. This communication game converges to one with verifiable disclosure when the function z satisfies $z(m, s) = 0$ if $m \in \{s, \emptyset\}$, and $z(m, s) = +\infty$ otherwise. We now consider general cases which only require the following less restrictive assumption on the communication cost function z:

(**A2**) Function $z : \mathcal{M} \times \mathcal{S} \to \mathbb{R}_+ \cup \{+\infty\}$ satisfies

(i) $z(s, s) = 0 \quad \forall s \in \mathcal{S}$, and

(ii) $z(m, s) = \kappa(m - s)^2$ if either $m > s > 0$ or $m < s < 0$, where $\kappa > 0$.

In words, condition (i) states that telling the truth is always costless. However, as condition (ii) states, it is costly for the agents to exaggerate their findings. In particular, the further away an agent's message is from the truth, the more costly it is to send such a message. It may be natural to further assume that $z(m, s) = 0$ if $0 \leq m < s$, $s < m \leq 0$ or $m = \emptyset$ (i.e., understating or concealing one's finding is also costless), and $z(m, s) = +\infty$ if $m \cdot s < 0$ (i.e., lying about the *sign* of the local state is never feasible). One may also want to extend condition (ii) to the case where $s \in \{0, \emptyset\}$. None of these additional assumptions will be necessary for our analysis in this section.

Although assumption (A2) rules out pure cheap talk communication, it still provides arbitrarily rich possibilities to lie (under consideration of lying costs). The next proposition states that despite the non-verifiability of agents' private information, provided that (A2) is satisfied the endogenous quality of communication under both centralization and decentralization will be identical and maximal: just as in the main model, in either case a fully revealing equilibrium exists.

Proposition 7. *Suppose that the communication cost satisfies (A2).*

(i) Under decentralization, there exists a fully revealing PBE in which

$$\check{m}_i^d(e_i, s_i) = \begin{cases} t^d s_i & \textit{if } s_i \in \mathbb{R} \\ \emptyset & \textit{if } s_i = \emptyset \end{cases}, \ \forall e_i \in [0,1], \textit{ and } \forall i = 1, 2,$$

where $t^d = \frac{1}{2} + \sqrt{\frac{q\delta^2}{\kappa(1+2\delta)^2} + \frac{1}{4}}$. *In equilibrium, both agents exert the same effort*

$$\check{e}^d \equiv (c')^{-1}\left(\left(1 - \frac{\delta^2 + \delta}{(1+2\delta)^2}\right) q\sigma_\theta^2 - \kappa(t^d - 1)^2\sigma_\theta^2\right).$$

(ii) Under centralization, there exists a fully revealing PBE in which

$$\check{m}_i^c(e_i, s_i) = \begin{cases} t^c s_i & \textit{if } s_i \in \mathbb{R} \\ \emptyset & \textit{if} s_i = \emptyset \end{cases}, \ \forall e_i \in [0,1], \textit{ and } \forall i = 1, 2,$$

where $t^c = \frac{1}{2} + \sqrt{\mathbb{E}\left[\frac{q\lambda(1-\lambda)[2\delta^2+\delta(\lambda^2+(1-\lambda)^2)]}{2\kappa(\lambda(1-\lambda)+\delta)^2}\right] + \frac{1}{4}}$, *and* $\lambda = \eta_1/(\eta_1 + \eta_2)$. *In equilibrium, both agents exert the same effort*

$$\check{e}_F^c \equiv (c')^{-1}\left(\left(1 - \mathbb{E}_\lambda\left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2 - \kappa\left(t^c - 1\right)^2\sigma_\theta^2\right).$$

Proposition 7 shows that in equilibrium, the agents "lie" in such a way that their private signals can be perfectly inferred from the messages communicated. To relate it to our previous results in Section 4 (Propositions 1 - 4), note that as $\kappa \to +\infty$, both coefficients t^d and t^c converge to 1. In addition, the total expected communication costs $\kappa(t^d - 1)^2$ and $\kappa(t^c - 1)^2$ converge to zero. This implies that as exaggeration becomes infinitely costly, under both centralization and decentralization the agents simply disclose their acquired signals ($m_i^d(e_i, s_i), m_i^c(e_i, s_i) \to s_i \ \forall s_i \in \mathbb{R} \cup \{\emptyset\}$) and exert the same amount of effort as in the main model ($\check{e}^d \to e^d, \check{e}_F^c \to e_F^c$).

Perhaps a more interesting observation is that the equilibria under both centralization and decentralization feature language inflation ($t^c, t^d > 1$). This is reminiscent of the findings of Kartik et al. (2007) and Kartik (2009) on strategic communication with credulous receivers or with exogenous lying costs. Compared to the existing papers, a novelty of our language inflation results is that they hold in settings that feature either bilateral communication (in the case of decentralization) or competing senders with differentiated private information (in the case of centralization).[28]

7 Conclusion

When operating in multiple markets which exhibit uncertainty in their relative profitability, how should an organization optimally allocate decision-making authority to its managerial members? In this paper we addressed this question in a model where decision-relevant information is collected and transmitted by strategic and self-interested division managers, and the objective of the organization is to solve the problem of coordinated adaptation.

Our paper makes two main contributions. The first is that if information is verifiable or if lying is not costless, then the quality of communication is not affected by where the decision-making authority is lodged in the organization. Moreover, since the principal of the organization can elicit all private information from its local delegates, the fact that the principal is not well informed per se does not make centralized decision-making inferior. However, as a second contribution, we show that the quality of endogenously acquired information depends crucially on the allocation of decision rights. In particular, if the local markets exhibit any uncertainty in their relative profitability, a large coordination motive can strongly discourage information gathering under centralization, which in turn makes decentralized decision-making optimal. Yet it is also worth noting that when the need for coordination is small or intermediate, centralized decision-making is often optimal because it allows the organization to better cope with inter-market uncertainty, while not necessarily making the division managers less motivated. Overall, our results call for a more careful examination of the Delegation Principle, which is well-known in the management literature (see, e.g., Milgrom and Roberts, 1992) and emphasizes that "the power to make decisions should reside in the hands of those with relevant information" (Krishna and Morgan, 2008, p. 905).

Throughout the paper we have assumed that the division managers care only about their own performance. As noted by Athey and Roberts (2001) and Rantakari (2013), if performance measures are available, the organization designer may want to align the incentives of the managerial members by tying their compensation to each other's performance. A natural extension of our model is thus to allow for performance-based transfers. While in extreme cases an interdependent pay structure may discourage information acquisition (e.g., if agent i's reward is *primarily* determined by j's performance), an appropriate level of interdependence can lead to a more efficient use of information when decision rights are decentralized to the divisions. Under centralization, however, there is no room for such improvement given that a

[28]Emons and Fluet (2012) were the first to show that the feature of language information can also arise in a setting with multiple senders. However, the senders in their model (plaintiff and defendant) have perfectly correlated types (they share the private knowledge about the amount of damages). This is not the case in our model because the private types θ_i and θ_j are independently distributed.

central manager can elicit all information from the local ones for free. The implication of this preliminary analysis is that decentralized decision-making is even more likely to be optimal when performance-based transfers are available, echoing Milgrom and Roberts (1992)'s view that the alignment of incentives is complementary to the delegation of authority.

Finally, we suggest two venues for future research. First, given that the communication of decision-relevant information in organizations is often not entirely cheap talk (e.g., marketing reports must contain survey evidence or data analysis in order to be taken serious, lying to colleagues may result in retaliation or even being fired), it is worth reconsidering how essential the informational constraints are in various organizational design problems. A conjecture based on the analysis of our paper is that in settings with verifiable information, the incentive constraints for communication can be much less important than the physical or technological ones (Aoki, 1986; Dessein and Santos, 2006). Second, when the uncertainty in the profitability of different product markets is substantial, the principal of the organization may prefer a more moderate way to mitigate her commitment problem than unconditionally delegating the decisions to the division managers. It is an open question whether the principal can benefit from conditional delegation, e.g., committing to only execute her authority when it is reported that the local states take extreme values.

References

Aghion, P., N. Bloom, B. Lucking, R. Sadun, and J. Van Reenen (2017): "Turbulence, firm decentralization and growth in bad times," Working Paper 23354, National Bureau of Economic Research.

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

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.

Antràs, P. and S. R. Yeaple (2014): "Multinational firms and the structure of international trade," in *Handbook of International Economics*, Elsevier, vol. 4, 55–130.

Aoki, M. (1986): "Horizontal vs. vertical information structure of the firm," *American Economic Review*, 76, 971–983.

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

Athey, S. and J. Roberts (2001): "Organizational design: decision rights and incentive contracts," *American Economic Review*, 91, 200–205.

Barnard, C. I. (1938): *The Functions of the Executive*, Harvard University Press.

Bartlett, C. A. and S. Ghoshal (2002): *Managing Across Borders: The Transnational Solution*, Harvard Business Press.

Bloom, N., R. Sadun, and J. Van Reenen (2012): "The organization of firms across countries," *Quarterly Journal of Economics*, 127, 1663–1705.

Brandts, J. and D. J. Cooper (2018): "Truth be told: an experimental study of communication and centralization," Working Paper 1046, Barcelona GSE.

Che, Y.-K. and N. Kartik (2009): "Opinions as incentives," *Journal of Political Economy*, 117, 815–860.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50, 1431–1451.

Cyert, R. M. and J. G. March (1963): *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ.

Deimen, I. and D. Szalay (2018): "Delegated expertise, authority, and communication," Discussion Paper 12706, Centre for Economic Policy Research.

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

Dessein, W., L. Garicano, and R. Gertner (2010): "Organizing for synergies," *American Economic Journal: Microeconomics*, 2, 77–114.

Dessein, W. and T. Santos (2006): "Adaptive organizations," *Journal of Political Economy*, 114, 956–995.

Dye, R. A. (1985): "Disclosure of nonproprietary information," *Journal of Accounting Research*, 23, 123–145.

Emons, W. and C. Fluet (2012): "Adversarial versus inquisitorial testimony," Mimeo.

Freeland, R. F. (2001): *The Struggle for Control of the Modern Corporation: Organizational Change at General Motors, 1924-1970*, Cambridge University Press.

Friebel, G. and M. Raith (2010): "Resource allocation and organizational form," *American Economic Journal: Microeconomics*, 2, 1–33.

Fudenberg, D. and J. Tirole (1991): *Game Theory*, The MIT Press.

Garicano, L. and L. Rayo (2016): "Why organizations fail: models and cases," *Journal of Economic Literature*, 54, 137–92.

Gerstner, L. (2002): *Who Says Elephants Can't Dance? Inside IBM's Historic Turnaround*, New York: Harper Business.

Giebe, T. and O. Gürtler (2012): "Optimal contracts for lenient supervisors," *Journal of Economic Behavior and Organization*, 81, 403–420.

Green, J. R. and N. Stokey (1981): "The value of information in the delegation problem," Discussion Paper 776, Harvard Institute of Economic Research.

Grossman, S. J. (1981): "The informational role of warranties and private disclosure about product quality," *Journal of Law and Economics*, 24, 461–483.

Grossman, S. J. and O. D. Hart (1986): "The costs and benefits of ownership: a theory of vertical and lateral integration," *Journal of Political Economy*, 94, 691–719.

Hagenbach, J., F. Koessler, and E. Perez-Richet (2014): "Certifiable pre-play communication: full disclosure," *Econometrica*, 82, 1093–1131.

Hart, O. and J. Moore (1990): "Property rights and the nature of the firm," *Journal of Political Economy*, 98, 1119–1158.

Hayek, F. A. (1945): "The use of knowledge in society," *American Economic Review*, 35, 519–530.

Kartik, N. (2009): "Strategic communication with lying costs," *Review of Economic Studies*, 76, 1359–1395.

Kartik, N., F. X. Lee, and W. Suen (2017): "Investment in concealable information by biased experts," *RAND Journal of Economics*, 48, 24–43.

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, lies, and costly talk," *Journal of Economic Theory*, 134, 93–116.

Krishna, V. and J. Morgan (2008): "Contracting for information under imperfect commitment," *RAND Journal of Economics*, 39, 905–925.

Letina, I., S. Liu, and N. Netzer (2018): "Delegating performance evaluation," Mimeo.

Lincoln, W. F. and A. H. McCallum (2018): "The rise of exporting by US firms," *European Economic Review*, 102, 280–297.

Malone, T. (2004): "Making the decision to decentralize," *Harvard Business School–Working Knowledge for Business Leaders.*

Marukawa, T. (2009): "Why Japanese multinationals failed in the Chinese mobile phone market: a comparative study of new product development in Japan and China," *Asia Pacific Business Review*, 15, 411–431.

Milgrom, P. and J. Roberts (1986): "Relying on the information of interested parties," *RAND Journal of Economics*, 17, 18–32.

Milgrom, P. R. (1981): "Good news and bad news: representation theorems and applications," *The Bell Journal of Economics*, 12, 380–391.

Milgrom, P. R. and J. D. Roberts (1992): *Economics, Organization and Management*, Englewood Cliffs: Prentice-Hall.

Newman, D. P. and K. E. Novoselov (2009): "Delegation to encourage communication of problems," *Journal of Accounting Research*, 47, 911–942.

Prendergast, C. and R. H. Topel (1996): "Favoritism in organizations," *Journal of Political Economy*, 104, 958–978.

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

——— (2012): "Employee initiative and managerial control," *American Economic Journal: Microeconomics*, 4, 171–211.

——— (2013): "Organizational design and environmental volatility," *Journal of Law, Economics, and Organization*, 29, 569–607.

Roberts, J. and G. Saloner (2013): "Strategy and organization," in *Handbook of Organizational Economics*, ed. by R. Gibbons, Princeton, NJ: Princeton University Press, chap. 20, 799–852.

Seidmann, D. J. and E. Winter (1997): "Strategic information transmission with verifiable messages," *Econometrica*, 65, 163–169.

Shavell, S. (1994): "Acquisition and disclosure of information prior to sale," *RAND Journal of Economics*, 25, 20–36.

Shin, H. S. (1994): "News management and the value of firms," *RAND Journal of Economics*, 25, 58–71.

Simon, H. A. (1947): *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, New York: Macmillan.

Thomas, C. (2011): "Too many products: decentralized decision making in multinational firms," *American Economic Journal: Microeconomics*, 3, 280–306.

Williamson, O. E. (1975): *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press.

——— (1996): *The Mechanisms of Governance*, Oxford University Press.

Yang, H. and L. Zhang (2017): "Communication and the optimality of hierarchy in organizations," Mimeo.

A Appendix: Proofs

A.1 Proof of Proposition 1

Part (i) First, consider agent i's incentives in the decision-making stage. Taking (e_i, s_i, m_i, m_j) as given, in the decision-making stage agent i solves:

$$\max_{y_i \in \mathbb{R}} \; q\left(K - \mathbb{E}\left[(y_i - \theta_i)^2 \,|s_i\right] - \delta \mathbb{E}\left[(y_i - y_j^d(e_j, s_j, m_i, m_j))^2 | m_i, m_j\right]\right).$$

Sequential rationality then implies that agent i's should take the following action:

$$y_i = \frac{\mathbb{E}\left[\theta_i|s_i\right] + \delta \mathbb{E}\left[y_j^d(e_j, s_j, m_i, m_j)|m_i, m_j\right]}{1+\delta}.$$

Note that the best response of the agent does not depend on his sunk effort e_i. From now on, we drop (e_i, e_j) from the functions (y_i^d, y_j^d) as they play no role. Solving the best response functions through repeated substitution, we obtain the following decision rules which must be satisfied in any equilibrium:

$$y_i^d(s_i, m_i, m_j) = \frac{\mathbb{E}\left[\theta_i|s_i\right]}{1+\delta} + \frac{\delta^2 \mathbb{E}[\theta_i|m_i]}{(1+\delta)(1+2\delta)} + \frac{\delta \mathbb{E}[\theta_j|m_j]}{1+2\delta}, \; \forall i, j = 1, 2, i \neq j. \tag{A.1}$$

where the conditional expectations $\mathbb{E}[\theta_i|s_i]$ and $\mathbb{E}[\theta_j|m_j]$ ($\mathbb{E}[\theta_j|s_j]$ and $\mathbb{E}[\theta_i|m_i]$, resp.) are taken according to the agent i's (agent j's, resp.) posterior beliefs about the local states.

Now suppose that agent i anticipates that agent j will exert some arbitrary effort $e_j \in [0, 1]$, communicate his finding truthfully according to the strategy m_j^d specified in the proposition, and choose his action according to the mapping y_j^d specified in (A.1). Taking the the sequentially rational decision rule y_i^d as given, we consider agent i's incentive in the communication stage. Since by construction (m_i^d, m_j^d) are effort-independent, we drop the variables (e_i, e_j) from them. To ease notation, we also assume without loss of generality that $m^{s_i} = s_i \; \forall s_i \in \Theta$.

Let $s_i \in \mathcal{S}$ be the signal received by agent i. For any message $m_i \in \mathcal{M}(s_i)$, we have

$$\begin{aligned}
EL_a^d(s_i, m_i) &= \mathbb{E}_{s_j}\left[\mathbb{E}\left[\left(y_i^d(s_i, m_i, m_j^d(s_j)) - \theta_i\right)^2 \middle| s_i\right]\right] \\
&= \mathbb{E}_{s_j}\left[\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} + \frac{\delta^2\mathbb{E}[\theta_i|m_i]}{(1+\delta)(1+2\delta)} + \frac{\delta\mathbb{E}[\theta_j|m_j^d(s_j)]}{1+2\delta} - \theta_i\right)^2 \middle| s_i\right]\right] \\
&= \mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} + \frac{\delta^2\mathbb{E}[\theta_i|m_i]}{(1+\delta)(1+2\delta)} - \theta_i\right)^2 \middle| s_i\right] + \mathbb{E}_{s_j}\left[\left(\frac{\delta\mathbb{E}[\theta_j|s_j]}{1+2\delta}\right)^2\right],
\end{aligned} \tag{A.2}$$

where the last equality follows that $\mathbb{E}_{s_j}\left[\mathbb{E}[\theta_j|s_j]\right] = \mathbb{E}[\theta_j] = 0$.

Similarly, for the expected loss of mis-coordination resulted by any message m_i, we have

$$
\begin{aligned}
&EL_c^d(s_i, m_i) \\
=&\mathbb{E}_{s_j}\left[\mathbb{E}[(y_i^d(s_i, m_i, m_j^d(s_j)) - y_j^d(s_j, m_i, m_j^d(s_j)))^2|s_i]\right] \\
=&\mathbb{E}_{s_j}\left[\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} - \frac{\delta\mathbb{E}[\theta_i|m_i]}{(1+\delta)(1+2\delta)} - \frac{\mathbb{E}[\theta_j|s_j]}{1+\delta} + \frac{\delta\mathbb{E}[\theta_j|m_j^d(s_j)]}{(1+\delta)(1+2\delta)}\right)^2 \middle| s_j\right]\right] \\
=&\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} - \frac{\delta\mathbb{E}[\theta_i|m_i]}{(1+\delta)(1+2\delta)}\right)^2 \middle| s_i\right] + \mathbb{E}_{s_j}\left[\left(\frac{\mathbb{E}[\theta_j|s_j]}{1+\delta} - \frac{\delta\mathbb{E}[\theta_j|s_j]}{(1+\delta)(1+2\delta)}\right)^2\right] \\
=&\mathbb{E}\left[\left(\frac{(1+\delta)\mathbb{E}[\theta_i|s_i] + \delta(\mathbb{E}[\theta_i|s_i] - \mathbb{E}[\theta_i|m_i])}{(1+\delta)(1+2\delta)}\right)^2 \middle| s_i\right] + \mathbb{E}_{s_j}\left[\left(\frac{\mathbb{E}\left[\theta_j|s_j\right]}{1+2\delta}\right)^2\right]. \qquad \text{(A.3)}
\end{aligned}
$$

Since for every $(s_i, m_i) \in \mathcal{S} \times \mathcal{M}(s_i)$ the (interim) expected payoff of agent i is given by

$$\hat{\Pi}_i^d(s_i, m_i) = q\left(K - EL_a^d(s_i, m_i) - \delta EL_c^d(s_i, m_i)\right),$$

communicating according to m_i^d is incentive compatible for agent i *if* under some consistent posterior beliefs of agent j, we have

$$EL_a^d(s_i, s_i) \leq EL_a^d(s_i, m_i) \text{ and } EL_c^d(s_i, s_i) \leq EL_c^d(s_i, m_i), \ \forall m_i \in \mathcal{M}(s_i), \forall s_i \in \mathcal{S}. \qquad \text{(A.4)}$$

To construct the required posterior beliefs, for every $m_i \in \mathcal{M}$ we let agent j assign probability one to that agent i's type is $\underline{s}^{m_i} \in \arg\min_{s_i \in \mathcal{S}^{m_i}} \left|\mathbb{E}[\theta_i|s_i]\right|$, i.e., $\mu_j^i(\{\underline{s}^{m_i}|m_i\}) = 1$. If $\emptyset \in \mathcal{S}^{m_i}$, the existence of $\underline{s}^{m_i}$ is trivial. If $\emptyset \notin \mathcal{S}^{m_i}$, the existence of $\underline{s}^{m_i}$ is guaranteed by the assumption that $\mathcal{S}^{m_i}$ is closed. This is because $\min_{s_i \in \mathcal{S}^{m_i}} \left|\mathbb{E}[\theta_i|s_i]\right| = \min_{s_i \in \mathcal{S}^{m_i} \cap \left[-|s_i'|, |s_i'|\right]} \left|\mathbb{E}[\theta_i|s_i]\right|$, where s_i' is any element of $\mathcal{S}^{m_i}$, and the set $\mathcal{S}^{m_i} \cap [-|s_i'|, |s_i'|]$ is compact. In addition, by construction we have $\underline{s}^{s_i} = s_i \ \forall s_i \in \mathcal{S}$. Given the constructed beliefs, we have $\mathbb{E}[\theta_i|\emptyset] = \mathbb{E}[\theta_i|m_i] = 0$ $\forall m_i \in \mathcal{M}(\emptyset)$, $\mathbb{E}[\theta_i|s_i] = s_i \geq \mathbb{E}[\theta_i|m_i]$ $\forall s_i \geq 0$ and $m_i \in \mathcal{M}(s_i)$, and $\mathbb{E}[\theta_i|s_i] = s_i \leq \mathbb{E}[\theta_i|m_i]$ $\forall s_i < 0$ and $m_i \in \mathcal{M}(s_i)$. It is then straightforward to check that (A.4) is satisfied.

To complete the construction of a fully revealing PBE, we finally consider the information acquisition stage. Given the communication strategies (m_1^d, m_2^d), the decision rules (y_1^d, y_2^d), and any pair of efforts $(e_1, e_2) \in E^2$, agent i's expected payoff is

$$
\begin{aligned}
&U_i^d(e_i, e_j) \\
=&q\Bigg(K - (1-e_i)\left[(1-e_j) + e_j\left(\frac{\delta^2+\delta}{(1+2\delta)^2} + 1\right)\right]\sigma_\theta^2 \\
&- e_i\left[(1-e_j)\left(\frac{\delta^2+\delta}{(1+2\delta)^2}\right) + e_j\left(\frac{2\delta^2+2\delta}{(1+2\delta)^2}\right)\right]\sigma_\theta^2\Bigg) - c(e_i)
\end{aligned}
$$

$$
\begin{aligned}
&= q\left(K - (1-e_i)\left[1 + e_j\left(\frac{\delta^2+\delta}{(1+2\delta)^2}\right)\right]\sigma_\theta^2 - e_i(1+e_j)\left(\frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma_\theta^2\right) - c(e_i) \\
&= q\left(K - \left[1 + e_j\left(\frac{\delta^2+\delta}{(1+2\delta)^2}\right)\right]\sigma_\theta^2 + e_i\left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma_\theta^2\right) - c(e_i).
\end{aligned}
$$

Differentiating with respect to e_i, we obtain the following first-order condition:

$$
\frac{\partial U_i^d(e_i, e_j)}{\partial e_i} = \left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right) q\sigma_\theta^2 - c'(e_i) = 0. \tag{A.5}
$$

When the cost function c is strictly increasing, and satisfies $\lim_{e\to 0} c'(e) < \left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right) q\sigma_\theta^2 < c'(1)$, (A.5) will admit a unique interior solution $e_i^d \in (0,1)$, which is given by

$$
e_i^d = e^d \equiv (c')^{-1}\left(\left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right) q\sigma_\theta^2\right). \tag{A.6}
$$

In addition, when c is strictly convex, the function U_i^d will be strictly concave in e_i, and thus the solution $e_i = e^d$ is also the unique global maximizer of $U_i^d(e_i, e_j)$, $\forall e_j \in [0,1]$. We have assumed that the cost function c satisfies all these properties (see Section 3).

Similarly, choosing $e_j = e^d$ also maximizes the expected payoff of agent j independent of the effort choice of agent i. We can therefore conclude that, together with the "conservative" beliefs that we construct above for the agents, the symmetric strategy profile $((e^d, m_1^d, y_1^d), (e^d, m_2^d, y_2^d))$ constitutes a fully revealing PBE. ■

Part (ii) Let $((e_1^*, m_1^*, y_1^*), (e_2^*, m_2^*, y_2^*))$ be an equilibrium strategy profile under decentralization. Consider any $s_i \in \mathcal{S} \setminus \{0, \emptyset\}$. Repeating the calculations of (A.1), (A.2) and (A.3), it can be checked that agent i would strictly prefer the type-revealing message $m_i = m^{s_i}$ than the proposed equilibrium message $m_i^*(e_i^*, s_i)$ *if* both of the following two inequalities hold:

$$
\begin{aligned}
&\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} + \frac{\delta^2\mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)]}{(1+\delta)(1+2\delta)} - \theta_i\right)^2 \middle| s_i\right] \\
&> \mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} + \frac{\delta^2\mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)]}{(1+\delta)(1+2\delta)} - \theta_i\right)^2 \middle| s_i\right].
\end{aligned} \tag{A.7}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\left(\frac{(1+\delta)\mathbb{E}[\theta_i|s_i] + \delta(\mathbb{E}[\theta_i|s_i] - \mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)])}{(1+\delta)(1+2\delta)}\right)^2 \middle| s_i\right] \\
&> \mathbb{E}\left[\left(\frac{(1+\delta)\mathbb{E}[\theta_i|s_i] + \delta(\mathbb{E}[\theta_i|s_i] - \mathbb{E}[\theta_i|m^{s_i}])}{(1+\delta)(1+2\delta)}\right)^2 \middle| s_i\right].
\end{aligned} \tag{A.8}
$$

Note that for any $s_i \neq \emptyset$, (A.7) is further equivalent to

$$\left(\frac{\delta(1+\delta)s_i + \delta^2(s_i - \mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)])}{(1+\delta)(1+2\delta)}\right)^2 > \left(\frac{\delta(1+\delta)s_i + \delta^2(s_i - \mathbb{E}[\theta_i|m^{s_i}])}{(1+\delta)(1+2\delta)}\right)^2. \tag{A.9}$$

From (A.8) and (A.9), it is clear that if $s_i > 0$, then deviating to m^{s_i} is not profitable for agent i only if $s_i \leq \mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)]$. Similarly, if $s_i < 0$, then deviating to m^{s_i} is not profitable for agent i only if $s_i \geq \mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)]$. These arguments also imply that we must have $m_i^*(e_i^*, s_i) \neq m_i^*(e_i^*, s_i') \ \forall s_i, s_i' \in \mathcal{S} \setminus \{0, \emptyset\}$ such that $s_i \cdot s_i' < 0$.

Next, suppose, in contradiction to Proposition 1(ii), that there exist $i \in \{1, 2\}$ and a non-null subset $\hat{\mathcal{S}} \subseteq \mathcal{S} \setminus \{0, \emptyset\}$ with respect to Γ, such that $\mu_j^i(\{\hat{s}_i\}|m_i^*(e_i^*, \hat{s}_i)) < 1 \ \forall \hat{s}_i \in \hat{\mathcal{S}}$.[29] Since the beliefs must be consistent in equilibrium, we have $m_i^*(e_i^*, \hat{s}_i) \neq m^{\hat{s}_i} \ \forall \hat{s}_i \in \hat{\mathcal{S}}$. For every on-path equilibrium message $\hat{m}_i^*$ that is sent by some $\hat{s}_i \in \hat{\mathcal{S}}$, define $\hat{\mathcal{S}}(\hat{m}_i^*) = \{s_i \in \hat{\mathcal{S}} : m_i^*(e_i^*, s_i) = \hat{m}_i^*\}$. Let $\hat{\mathcal{M}}^*$ be the set of all such messages $\hat{m}_i^*$.

We claim that the set $\hat{\mathcal{S}}(\hat{m}_i^*)$ is null with respect to Γ for all $\hat{m}_i^* \in \hat{\mathcal{M}}^*$. This is because if $\hat{\mathcal{S}}(\hat{m}_i^*)$ is non-null with respect to Γ for some $\hat{m}_i^*$, the condition $\mu_j^i(\{\hat{s}_i\}|m_i^*(e_i^*, \hat{s}_i)) < 1 \, \forall \hat{s}_i \in \hat{\mathcal{S}}$ would imply that there exists $s_i \in \hat{\mathcal{S}}(\hat{m}_i^*)$ such that either $s_i > \max\{0, \mathbb{E}[\theta_i|\hat{m}_i^*]\}$ or $s_i < \min\{0, \mathbb{E}[\theta_i|\hat{m}_i^*]\}$ holds. This is not possible given our analysis of (A.8) and (A.9).

Since $\hat{\mathcal{S}}(\hat{m}_i^*)$ is null with respect to Γ for all $\hat{m}_i^* \in \hat{\mathcal{M}}^*$, Bayes' rule implies that for every $\hat{m}_i^* \in \hat{\mathcal{M}}^*$ there must exist an atom $s^{\hat{m}_i^*} \in \mathcal{S}$ in the distribution Γ, such that $m_i^*(e_i^*, s^{\hat{m}_i^*}) = \hat{m}_i^*$ and $\mu_j^i(\{s^{\hat{m}_i^*}\}|\hat{m}_i^*) = 1$. Note that by construction, each $\hat{m}_i^* \in \hat{\mathcal{M}}^*$ is associated with a different atom. However, since $\hat{\mathcal{S}} = \cup_{\hat{m}_i^*} \hat{\mathcal{S}}(\hat{m}_i^*)$ is non-null with respect to Γ, the set $\hat{\mathcal{M}}^*$ must be uncountable, and this would imply that the distribution Γ admits uncountably many atoms. We thus reach a contradiction. ■□

A.2 Proof of Proposition 2

Let $((e_1^*, m_1^*, y_1^*), (e_2^*, m_2^*, y_2^*))$ be a fully revealing equilibrium under decentralization. Given the full revelation, we have $\mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)] = \mathbb{E}[\theta_i|s_i] \ \forall s_i \in \mathcal{S}$ and $\forall i = 1, 2$. Then, (A.1) implies that the on-path equilibrium decision rules are uniquely pinned down by Bayes' rule and sequential rationality, and they are the exactly ones given by the proposition. As we have shown in Proposition 1(i), given the equilibrium decisions are taken according to $(y_1^d(\mathbf{s}), y_1^d(\mathbf{s}))$, e^d is the unique expected-payoff-maximizing effort level for both agents. Hence, we must have $e_i^* = e^d$ and $y_i^*(e_i^*, s_i, m_i^*(e_i^*, s_i), m_j^*(e_j^*, s_j)) = y_i^d(s_i, s_j)$, $\forall (s_i, s_j) \in \mathcal{S}^2$, $i = 1, 2$. □

[29] Formally, we say that a set $\hat{S} \subseteq \mathcal{S}$ is non-null with respect to Γ if $\int_{\mathcal{S}} \mathbb{1}_{\{s \in \hat{S}\}} d\Gamma > 0$, and it is null with respect to Γ if $\int_{\mathcal{S}} \mathbb{1}_{\{s \in \hat{S}\}} d\Gamma = 0$.

A.3 Proof of Proposition 3

Part (i) First, consider the principal's incentive in the decision-making stage. Taking (m_1, m_2) and (η_1, η_2) as given, in the decision-making stage the principal solves:

$$\max_{y_1, y_2 \in \mathbb{R}} (\eta_1 + \eta_2)\left(K - \delta(y_1 - y_2)^2\right) - \eta_1 \mathbb{E}\left[(y_1 - \theta_1)^2 | m_1\right] - \eta_2 \mathbb{E}\left[(y_2 - \theta_2)^2 | m_2\right].$$

The first-order conditions imply that at optimum the principal's actions (y_1, y_2) must solve the following system of equations:

$$\begin{aligned}
&-\delta(\eta_1 + \eta_2)(y_1 - y_2) - \eta_1 \left(y_1 - \mathbb{E}\left[\theta_1 | m_1\right]\right) = 0, \\
&-\delta(\eta_1 + \eta_2)(y_2 - y_1) - \eta_1 \left(y_2 - \mathbb{E}\left[\theta_2 | m_2\right]\right) = 0.
\end{aligned}$$

Solving the above equations, we obtain the following decision rules which must be satisfied in any equilibrium:

$$y_i^c(\boldsymbol{m}, \boldsymbol{\eta}) = \frac{\frac{\eta_i}{\eta_i + \eta_j} \cdot \left(\frac{\eta_j}{\eta_i + \eta_j} + \delta\right) \mathbb{E}\left[\theta_i | m_i\right] + \frac{\delta \eta_j}{\eta_i + \eta_j} \mathbb{E}[\theta_j | m_j]}{\frac{\eta_i}{\eta_i + \eta_j} \cdot \frac{\eta_j}{\eta_i + \eta_j} + \delta} \quad \forall i = 1, 2, \tag{A.10}$$

where the conditional expectations $\mathbb{E}[\theta_i | m_i]$ and $\mathbb{E}[\theta_j | m_j]$ are taken according to the principal posterior beliefs about the local states.

Next, we take the above decision rules (y_1^c, y_2^c) of the principal as given and consider the agents' incentives in the communication stage. We will only verify the equilibrium incentives of agent 1, as for agent 2 the problem is analogous. Suppose that agent 1 anticipates that agent 2 will exert some arbitrary effort $e_2 \in [0, 1]$ and communicate his finding truthfully according to the strategy m_2^c specified in the proposition. Since by construction (m_1^c, m_2^c) are effort-independent, we drop the variables (e_1, e_2) from them. To ease notation, we also assume without loss of generality that $m^s = s \ \forall s \in \Theta$.

Let $s_1 \in \mathcal{S}$ be the signal received by agent 1. Letting $\lambda = \eta_1/(\eta_1 + \eta_2)$, for every message $m_1 \in \mathcal{M}(s_1)$ and every $\boldsymbol{\eta} \in \left[\underline{\eta}, \bar{\eta}\right]^2$ we have

$$\begin{aligned}
&EL_a^c(s_1, m_1, \boldsymbol{\eta}) \\
=& \mathbb{E}_{s_2}\left[\mathbb{E}\left[\left(y_1^c(m_1, m_2^c(s_2), \boldsymbol{\eta}) - \theta_1\right)^2 \middle| s_1\right]\right] \\
=& \mathbb{E}_{s_2}\left[\mathbb{E}\left[\left(\frac{\lambda\left(1 - \lambda + \delta\right) \mathbb{E}\left[\theta_1 | m_1\right] + \delta(1 - \lambda)\mathbb{E}[\theta_2 | m_2^c(s_2)]}{\lambda(1 - \lambda) + \delta} - \theta_1\right)^2 \middle| s_1\right]\right] \\
=& \mathbb{E}\left[\left(\frac{(\lambda(1 - \lambda) + \lambda\delta)\,\mathbb{E}\left[\theta_1 | m_1\right]}{\lambda(1 - \lambda) + \delta} - \theta_1\right)^2 \middle| s_1\right] + \mathbb{E}_{s_2}\left[\left(\frac{\delta(1 - \lambda)\mathbb{E}[\theta_2 | s_2]}{\lambda(1 - \lambda) + \delta}\right)^2\right],
\end{aligned} \tag{A.11}$$

where the last equality follows that $\mathbb{E}_{s_2}\left[\mathbb{E}[\theta_2|s_2]\right] = \mathbb{E}[\theta_2] = 0$.

Similarly, for the expected loss of mis-coordination, we have for every $m_1 \in \mathcal{M}(s_1)$ and every $\boldsymbol{\eta} \in \left[\underline{\eta}, \bar{\eta}\right]^2$,

$$\begin{aligned} EL_c^c(s_1, m_1, \boldsymbol{\eta}) =& \mathbb{E}_{s_2}\left[\mathbb{E}[(y_1^c(m_1, m_2^c(s_2), \boldsymbol{\eta}) - y_2^c(m_1, m_2^c(s_2)))^2|s_1]\right] \\ =& \mathbb{E}_{s_2}\left[\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1] - \lambda(1-\lambda)\mathbb{E}[\theta_2|m_2^c(s_2)]}{\lambda(1-\lambda)+\delta}\right)^2\right] \\ =& \left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}\right)^2 + \mathbb{E}_{s_2}\left[\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_2|s_2]}{\lambda(1-\lambda)+\delta}\right)^2\right]. \end{aligned} \tag{A.12}$$

where the last equality follows that $\mathbb{E}_{s_2}\left[\mathbb{E}[\theta_2|s_2]\right] = \mathbb{E}[\theta_2] = 0$.

Since for every $(s_1, m_1) \in \mathcal{S} \times \mathcal{M}(s_1)$, the interim expected payoff of agent 1 is given by

$$\hat{\Pi}_i^c(s_1, m_1) = \mathbb{E}_{\boldsymbol{\eta}}\left[q\left(K - EL_a^c(s_1, m_1, \boldsymbol{\eta}) - \delta EL_c^c(s_1, m_1, \boldsymbol{\eta})\right)\right],$$

communicating according to m_1^c is incentive compatible for agent 1 *if* under some consistent posterior beliefs of the principal, we have

$$EL_a^c(s_1, s_1, \boldsymbol{\eta}) + \delta EL_c^c(s_1, s_1, \boldsymbol{\eta}) \leq EL_a^c(s_1, m_1, \boldsymbol{\eta}) + \delta EL_c^c(s_1, m_1, \boldsymbol{\eta}), \tag{A.13}$$

for all $s_1 \in \mathcal{S}$, $m_1 \in \mathcal{M}(s_1)$, $\boldsymbol{\eta} \in \left[\underline{\eta}, \bar{\eta}\right]^2$. We note that after some rearrangement, (A.13) is equivalent to

$$\begin{aligned} &\left(\frac{(\lambda(1-\lambda)+\lambda\delta)\mathbb{E}[\theta_1|\emptyset]}{\lambda(1-\lambda)+\delta}\right)^2 + \delta\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|\emptyset]}{\lambda(1-\lambda)+\delta}\right)^2 \\ \leq& \left(\frac{(\lambda(1-\lambda)+\lambda\delta)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}\right)^2 + \delta\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}\right)^2 \end{aligned} \tag{A.14}$$

if $s_1 = \emptyset$. If $s_1 \neq \emptyset$, then (A.13) is equivalent to

$$\begin{aligned} &\left(\frac{(\lambda(1-\lambda)+\lambda\delta)s_1}{\lambda(1-\lambda)+\delta} - s_1\right)^2 + \delta\left(\frac{\lambda(1-\lambda)s_1}{\lambda(1-\lambda)+\delta}\right)^2 \\ \leq& \left(\frac{(\lambda(1-\lambda)+\lambda\delta)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta} - s_1\right)^2 + \delta\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}\right)^2. \end{aligned} \tag{A.15}$$

Since $\mathbb{E}[\theta_1|\emptyset] = 0$, (A.14) always holds regardless of the principal's beliefs. To show (A.15), we construct the following consistent beliefs for the principal: for every $m_1 \in \mathcal{M}$ we let the principal assign probability one to that agent 1's type is $\underline{s}^{m_1} \in \arg\min_{s_1 \in \mathcal{S}^{m_1}} \left|\mathbb{E}[\theta_1|s_1]\right|$, i.e., $\mu_p^1(\{\underline{s}^{m_1}|m_1\}) = 1$. The existence of $\underline{s}^{m_1}$ is guaranteed by the assumption that $\mathcal{S}^{m_1}$ is closed. Also, by construction $\underline{s}^{s_1} = s_1\ \forall s_1 \in \mathcal{S}$. Given the constructed beliefs, we have

$\mathbb{E}[\theta_1|\emptyset] = \mathbb{E}[\theta_1|m_1] = 0 \; \forall m_1 \in \mathcal{M}(\emptyset)$, $\mathbb{E}[\theta_1|s_1] = s_i \geq \mathbb{E}[\theta_1|m_1] \; \forall s_1 \geq 0$ and $m_1 \in \mathcal{M}(s_1)$, and $\mathbb{E}[\theta_1|s_1] = s_1 \leq \mathbb{E}[\theta_1|m_1] \; \forall s_1 < 0$ and $m_1 \in \mathcal{M}(s_1)$.

Next, note that the RHS of (A.15) can be rewritten as the sum of the following two terms:

$$\frac{(\lambda(1-\lambda)+\lambda\delta)^2}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m_1]-s_1\right)^2 + \frac{(1-\lambda)^2\delta^2}{(\lambda(1-\lambda)+\delta)^2}s_1^2 \tag{A.16}$$

and

$$-\frac{2\delta(1-\lambda)(\lambda(1-\lambda)+\delta)}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m_1]-s_1\right)s_1 + \frac{\delta\lambda^2(1-\lambda)^2}{(\lambda(1-\lambda)+\delta)^2}\mathbb{E}[\theta_1|m_1]^2. \tag{A.17}$$

We claim that $\forall s_1 \in \Theta$, $\lambda \in [0,1]$ and $\delta > 0$, both of these two terms are minimized when $\mathbb{E}[\theta_1|m_1] = s_1$, which is in turn sufficient for (A.15) to hold for all $s_1 \neq \emptyset$. For the first term (A.16), this is straightforward. For the second term (A.17), we note that

$$\begin{aligned}
&-2\delta(1-\lambda)(\lambda(1-\lambda)+\delta)\mathbb{E}[\theta_1|m_1]s_1 + \delta\lambda^2(1-\lambda)^2\mathbb{E}[\theta_1|m_1]^2 \\
=&-\delta\lambda(1-\lambda)^2\left(2s_1-\lambda\mathbb{E}[\theta_1|m_1]\right)\mathbb{E}[\theta_1|m_1] - 2\delta^2(1-\lambda)\mathbb{E}[\theta_1|m_1]s_1,
\end{aligned}$$

and that the function $v(x) = -\delta\lambda(1-\lambda)^2 \cdot (2s_1 - \lambda x)x$ is decreasing in x when $x \leq \bar{x}$ and $s_1 \geq 0$, and it is increasing if $x \geq s_1$ and $s_1 < 0$. Hence, given the beliefs we construct for the principal, (A.17) is also minimized when $\mathbb{E}[\theta_1|m_1] = s_1$.

In sum, we have shown that the truthful-telling constraint (A.13) holds for every pair $\boldsymbol{\eta} \in \left[\underline{\eta}, \bar{\eta}\right]^2$. In other words, it is a best response for agent 1 to reveal his type even when the distribution of $\boldsymbol{\eta}$ is *deterministic*. Therefore, the same must also hold for arbitrary non-deterministic distribution of $\boldsymbol{\eta}$.

To complete the construction of a fully revealing PBE, we finally consider the information acquisition stage. Given the communication strategies (m_1^c, m_2^c), the decision rules (y_1^c, y_2^c), and any pair of efforts $(e_1, e_2) \in E^2$, agent 1's expected payoff is

$$\begin{aligned}
&U_1^c(e_1,e_2) \\
=&q\Bigg(K-(1-e_1)\left((1-e_2)+e_2\left(\mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]+1\right)\right)\sigma_\theta^2 \\
&-e_1\left((1-e_2)\mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]+e_2\mathbb{E}_\lambda\left[\frac{2(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma_\theta^2\Bigg)-c(e_1) \\
=&q\Bigg(K-\left(1+e_2\mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma_\theta^2 \\
&-e_1\left(1-\mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma_\theta^2\Bigg)-c(e_1).
\end{aligned}$$

Differentiating with respect to e_1, we obtain the following first-order condition:

$$\frac{\partial U_1^c(e_1,e_2)}{\partial e_1} = \left(1 - \mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2 - c'(e_1) = 0. \tag{A.18}$$

When the cost function c is strictly increasing and satisfies

$$\lim_{e\to 0} c'(e) < \left(1 - \mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2 < c'(1),$$

(A.18) will admit a unique interior solution $e_1^c \in (0,1)$, which is given by

$$e_1^c = (c')^{-1}\left(\left(1 - \mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2\right). \tag{A.19}$$

In addition, when c is strictly convex, the function U_1^c will be strictly concave in e_1, and thus the solution e_1^c is also the unique global maximizer of $U_1^c(e_1,e_2)$ $\forall e_2 \in [0,1]$. These properties of the cost function have all been assumed in Section 3.

By analogous arguments, one can show that choosing

$$e_2^c = (c')^{-1}\left(\left(1 - \mathbb{E}_\lambda\left[\frac{\lambda^2\left(\delta^2+\delta(1-\lambda)^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2\right)$$

will maximize the expected payoff of agent 2 regardless of the effort choice of agent 1. Further, since the distribution F is symmetric in its arguments (η_1,η_2), λ must be symmetrically distributed around 1/2. Exploiting this symmetry, we obtain

$$e_1^c = e_2^c = e_F^c \equiv (c')^{-1}\left(\left(1 - \mathbb{E}_\lambda\left[\frac{\delta^2(\lambda^2+(1-\lambda)^2)+2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2\right).$$

We can now conclude that, together with the degenerate posterior beliefs $\mu_p^i(e_i,m_i) = (e_F^c,m_i)$ (i.e., the principal assigns probability one to $e_i = e_F^c$ and $s_i = m_i$) $\forall i = 1,2$, the strategy profile $((e_F^c,m_1^c),(e_F^c,m_2^c),(y_1^c,y_2^c))$ constitutes a fully revealing PBE. ■

Part (ii) Let $((e_1^*,m_1^*,y_1^*),(e_2^*,m_2^*,y_2^*))$ be an equilibrium strategy profile under centralization. Without loss of generality, we focus on agent 1 and consider any $s_1 \in \mathcal{S}\setminus\{0,\emptyset\}$. Repeating the calculations of (A.10), (A.11) and (A.12), it can be checked that agent 1 would strictly prefer the type-revealing message $m_1 = m^{s_1}$ than the proposed equilibrium message $m_1^*(e_1^*,s_1)$ if *both* of the following two inequalities hold:

$$\begin{aligned}&\frac{(\lambda(1-\lambda)+\lambda\delta)^2}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]-s_1\right)^2 + \frac{(1-\lambda)^2\delta^2}{(\lambda(1-\lambda)+\delta)^2}s_1^2\\ &> \frac{(\lambda(1-\lambda)+\lambda\delta)^2}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m^{s_1}]-s_1\right)^2 + \frac{(1-\lambda)^2\delta^2}{(\lambda(1-\lambda)+\delta)^2}s_1^2,\end{aligned} \tag{A.20}$$

and

$$
\begin{aligned}
&-\frac{2\delta(1-\lambda)(\lambda(1-\lambda)+\delta)}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]-s_1\right)s_1+\frac{\delta\lambda^2(1-\lambda)^2}{(\lambda(1-\lambda)+\delta)^2}\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]^2 \\
&> -\frac{2\delta(1-\lambda)(\lambda(1-\lambda)+\delta)}{(\lambda(1-\lambda)+\delta)^2}\left(\mathbb{E}[\theta_1|m^{s_1}]-s_1\right)s_1+\frac{\delta\lambda^2(1-\lambda)^2}{(\lambda(1-\lambda)+\delta)^2}\mathbb{E}[\theta_1|m^{s_1}]^2. \qquad \text{(A.21)}
\end{aligned}
$$

Since $|\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]-s_1|\geq 0\ \forall s_1\in\mathcal{S}\setminus\{0,\emptyset\}$, (A.20) always holds. In addition, similar to what we have shown for (A.17), (A.21) will also hold if $\mathbb{E}[\theta_1|m^{s_1}]=s_1>\max\{\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)],0\}$ or $\mathbb{E}[\theta_1|m^{s_1}]=s_1<\min\{\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)],0\}$. Hence, for the proposed strategy profile to constitute an equilibrium, it is necessary that $\forall s_1\in\mathcal{S}\{0,\emptyset\}$, either $\mathbb{E}[\theta_1|m^{s_1}]=0<s_1\leq\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]$ or $\mathbb{E}[\theta_1|m^{s_1}]=0<s_1\leq\mathbb{E}[\theta_1|m_1^*(e_1^*,s_1)]$ must hold.

By replacing "the beliefs of agent j" (μ_j^i) with "the beliefs of the principal (μ_p^1)", the rest of the proof follows exactly the same steps as in the case of decentralization (see the proof of Proposition 1(ii)). ■□

A.4 Proof of Proposition 4

Analogous to the proof of Proposition 2. □

A.5 Comparative Statics of e^d and e_F^c

In this part of the Appendix, we will formally show that the equilibrium effort levels under decentralization and centralization (e^d and e_F^c) are both decreasing in δ. Let us define

$$D(\delta)\equiv\frac{\delta^2+\delta}{(1+2\delta)^2} \qquad \text{(A.22)}$$

and, for every $\lambda\in(0,1)$,

$$C(\delta,\lambda)\equiv\frac{\delta^2(\lambda^2+(1-\lambda)^2)+2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}. \qquad \text{(A.23)}$$

Differentiating with respect to δ, we have

$$D'(\delta)=\frac{(2\delta+1)(1+2\delta)-4(\delta^2+\delta)}{(1+2\delta)^3}=\frac{1}{(1+2\delta)^3}>0, \qquad \text{(A.24)}$$

and

$$
\begin{aligned}
&\frac{\partial C(\delta,\lambda)}{\partial\delta} \\
&=\frac{[2\delta(\lambda^2+(1-\lambda)^2)+2\lambda^2(1-\lambda)^2]\cdot(\lambda(1-\lambda)+\delta)-2[\delta^2(\lambda^2+(1-\lambda)^2)+2\delta\lambda^2(1-\lambda)^2]}{2(\lambda(1-\lambda)+\delta)^3}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\lambda^3(1-\lambda)^3 + \delta\lambda(1-\lambda)(\lambda^2 + (1-\lambda)^2 - \lambda(1-\lambda))}{(\lambda(1-\lambda)+\delta)^3} \\
&= \frac{\lambda^3(1-\lambda)^3 + \delta\lambda(1-\lambda)((2\lambda-1)^2 + \lambda(1-\lambda))}{(\lambda(1-\lambda)+\delta)^3} \\
&> 0. \qquad \text{(A.25)}
\end{aligned}
$$

Thus, both functions $D(\delta)$ and $C(\delta, \lambda)$ are increasing in δ, for all $\lambda \in (0,1)$. This further implies that both e^d and e^c_F are decreasing in δ, because

$$
e^d = (c')^{-1}\left((1 - D(\delta))\, q\sigma_\theta^2\right), \quad e^c_F = (c')^{-1}\left((1 - \mathbb{E}_\lambda\left[C(\delta,\lambda)\right])\, q\sigma_\theta^2\right), \qquad \text{(A.26)}
$$

and the cost function c is strictly increasing and convex.

A.6 Proof of Theorem 1

First, suppose that $corr(\eta_1, \eta_2) = 1$. Since the distribution F is symmetric in η_1 and η_2, for the global states to be perfectly and positively correlated, we must have $\Pr(\eta_1 = \eta_2) = 1$, and thus $\Pr\left(\lambda = \frac{1}{2}\right) = 1$, where $\lambda = \eta_1/(\eta_1 + \eta_2)$. In this case, the RHS of condition (5.1) becomes

$$
C_F(\delta) = \frac{\delta^2\left(\frac{1}{4} + \frac{1}{4}\right) + 2\delta \cdot \frac{1}{4} \cdot \frac{1}{4}}{2\left(\frac{1}{4} + \delta\right)^2} = \frac{4\delta^2 + \delta}{(1+4\delta)^2} = \frac{\delta}{1+4\delta}.
$$

$\forall \delta > 0$, we have

$$
\begin{aligned}
\frac{\delta^2 + \delta}{(1+2\delta)^2} > \frac{\delta}{1+4\delta} \quad &\iff \quad \frac{(1+\delta)(1+4\delta)}{(1+2\delta)^2} > 1 \\
&\iff \quad \frac{1 + 5\delta + 4\delta^2}{1 + 4\delta + 4\delta^2} > 1,
\end{aligned}
$$

which always holds. Therefore, when $corr(\eta_1, \eta_2) = 1$, we have $D(\delta) > C_F(\delta)\ \forall \delta > 0$, i.e., condition (5.1) is always violated. From the arguments in the main text, this immediately implies that $e^d < e^c_F\ \forall \delta > 0$.

Next, consider the case $corr(\eta_1, \eta_2) < 1$. Taking the limit of both sides of (5.1) with respect to δ, we obtain

$$
\lim_{\delta \to +\infty} D(\delta) = \lim_{\delta \to +\infty} \frac{1 + \frac{1}{\delta}}{\left(\frac{1}{\delta} + 2\right)^2} = \frac{1}{4},
$$

and

$$\lim_{\delta\to+\infty} C_F(\delta) = \lim_{\delta\to+\infty} \mathbb{E}\left[\frac{(\lambda^2+(1-\lambda)^2)+\frac{2\lambda^2(1-\lambda)^2}{\delta}}{2\left(\frac{\lambda(1-\lambda)}{\delta}+1\right)^2}\right] = \mathbb{E}\left[\frac{\lambda^2+(1-\lambda)^2}{2}\right] = \mathbb{E}\left[\lambda^2\right],$$

where the last equality follows that the distribution of λ must be symmetric around 1/2. Since $corr(\eta_1,\eta_2) < 1$, the distribution of λ cannot be degenerate, and thus by Jensen's inequality we further have

$$\lim_{\delta\to+\infty} C_F(\delta) > (\mathbb{E}[\lambda])^2 = \frac{1}{4} = \lim_{\delta\to+\infty} D(\delta).$$

Therefore, by continuity there must exist $\bar{\delta}_1 < +\infty$, such that $D(\delta) < C_F(\delta)\ \forall \delta > \bar{\delta}_1$. Since $e^d > e_F^C \iff D(\delta) < C_F(\delta)$, it immediately follows that $e^d > e_F^c\ \forall \delta > \bar{\delta}_1$.

To show that the effort difference $e^d - e_F^c$ is increasing in δ for sufficiently large δ, note that

$$\frac{\partial(e^d - e_F^c)}{\partial\delta} = \frac{-D'(\delta)q\sigma_\theta^2}{c''\left((c')^{-1}\left((1-D(\delta))\,q\sigma_\theta^2\right)\right)} - \frac{-C_F'(\delta)q\sigma_\theta^2}{c''\left((c')^{-1}\left((1-C_F(\delta))\,q\sigma_\theta^2\right)\right)}.$$

Since the cost function c is strictly convex, and both $D'(\delta)$ and $C_F'(\delta)$ are strictly positive (see Section A.5), the above partial derivative is strictly positive if and only if

$$\frac{C_F'(\delta)}{D'(\delta)} > \frac{c''\left((c')^{-1}\left((1-C_F(\delta))\,q\sigma_\theta^2\right)\right)}{c''\left((c')^{-1}\left((1-D(\delta))\,q\sigma_\theta^2\right)\right)}. \tag{A.27}$$

For the RHS of (A.27), we have

$$\lim_{\delta\to+\infty} \frac{c''\left((c')^{-1}\left((1-C_F(\delta))\,q\sigma_\theta^2\right)\right)}{c''\left((c')^{-1}\left((1-D(\delta))\,q\sigma_\theta^2\right)\right)} = \frac{c''\left((c')^{-1}\left((1-\mathbb{E}[\lambda^2])\,q\sigma_\theta^2\right)\right)}{c''\left((c')^{-1}\left(\frac{3}{4}\cdot q\sigma_\theta^2\right)\right)} < +\infty.$$

Using the calculation results from Section A.5 (see (A.24) and (A.25)), we also have

$$\begin{aligned}
\lim_{\delta\to+\infty} \frac{C_F'(\delta)}{D'(\delta)} &= \lim_{\delta\to+\infty} \mathbb{E}\left[\frac{\lambda(1-\lambda)(1+2\delta)^3}{(\lambda(1-\lambda)+\delta)^3}\cdot\left(\lambda^2(1-\lambda)^2+\delta((2\lambda-1)^2+\lambda(1-\lambda))\right)\right] \\
&= \lim_{\delta\to+\infty} \mathbb{E}\left[\frac{\lambda(1-\lambda)(\frac{1}{\delta}+2)^3}{\left(\frac{\lambda(1-\lambda)}{\delta}+1\right)^3}\cdot\left(\lambda^2(1-\lambda)^2+\delta((2\lambda-1)^2+\lambda(1-\lambda))\right)\right] \\
&= \lim_{\delta\to+\infty} \mathbb{E}\left[8\lambda(1-\lambda)\cdot\left(\lambda^2(1-\lambda)^2+\delta((2\lambda-1)^2+\lambda(1-\lambda))\right)\right] \\
&= \mathbb{E}\left[8\lambda^3(1-\lambda)^3\right] + \mathbb{E}\left[\lambda(1-\lambda)(2\lambda-1)^2+\lambda^2(1-\lambda)^2\right]\cdot\lim_{\delta\to+\infty}\delta \\
&= +\infty.
\end{aligned}$$

Therefore, by continuity, there must exist $\bar{\delta}_2 < +\infty$, such that (A.27) holds for all $\delta > \bar{\delta}_2$. Equivalently, the effort difference $e^d - e^c_F$ must be increasing in δ for all $\delta > \bar{\delta}_2$.

Finally, we complete the proof of the theorem by letting $\bar{\delta} \equiv \max\{\bar{\delta}_1, \bar{\delta}_2\}$. $\square$

A.7 Proof of Theorem 2

To simplify the algebra, let us define

$$\alpha \equiv \frac{\eta_1\eta_2}{(\eta_1+\eta_2)^2} = \lambda(1-\lambda),\ \beta \equiv \frac{\eta_1^2+\eta_2^2}{(\eta_1+\eta_2)^2} = \lambda^2 + (1-\lambda)^2 \tag{A.28}$$

and

$$\Delta_F(\delta) \equiv C_F(\delta) - D(\delta) = \mathbb{E}\left[\frac{\delta^2\beta + 2\delta\alpha^2}{2(\alpha+\delta)^2}\right] - \frac{\delta^2+\delta}{(1+2\delta)^2}.$$

From (A.24) and (A.25), we have

$$\Delta'_F(\delta) = \mathbb{E}\left[\frac{\alpha^3 + \delta\alpha(\beta-\alpha)}{(\alpha+\delta)^3}\right] - \frac{1}{(1+2\delta)^3}.$$

Further, the second derivative of $\Delta_F(\delta)$ is given by

$$\Delta''_F(\delta) = \mathbb{E}\left[\frac{\alpha^2\beta - 4\alpha^3 - 2\alpha(\beta-\alpha)\delta}{(\alpha+\delta)^4}\right] + \frac{6}{(1+2\delta)^4}.$$

Therefore,

$$\Delta_F(0) = 0,\ \Delta'_F(0) = \mathbb{E}\left[\frac{\alpha^3}{\alpha^3}\right] - 1 = 0,$$

and

$$\begin{aligned}\Delta''_F(0) &= \mathbb{E}\left[\frac{\alpha^2\beta - 4\alpha^3}{\alpha^4}\right] + 6\\ &= \mathbb{E}\left[\frac{1}{\lambda^2} + \frac{1}{(1-\lambda)^2} - \frac{4}{\lambda(1-\lambda)}\right] + 6\\ &= \mathbb{E}\left[\left(\frac{1}{\lambda} - \frac{1}{1-\lambda}\right)^2 - \frac{2}{\lambda(1-\lambda)}\right] + 6\\ &= \mathbb{E}\left[\frac{2}{\lambda^2} - \frac{4}{\lambda(1-\lambda)}\right] + 6,\end{aligned}$$

where the last equality follows that λ is symmetrically distributed around $1/2$. Note that

$$\Delta''_F(0) > 0 \iff \mathbb{E}\left[\frac{1}{\lambda^2}\right] > \mathbb{E}\left[\frac{2}{\lambda(1-\lambda)}\right] - 3.$$

Hence, if the condition of Theorem 2 is satisfied, then $\Delta_F''(0) > 0$. Since $\Delta_F'(0) = 0$, by continuity, there must exists $\tilde{\delta} > 0$ such that $\Delta_F'(\delta) > 0$ for all $\delta \in \left(0, \tilde{\delta}\right)$. Since $\Delta_F(0) = 0$, and Δ_F is strictly increasing on $\left(0, \tilde{\delta}\right)$, then again by continuity there must exist $\underline{\delta} \in (0, +\infty]$, such that $\Delta_F(\delta) > 0$ for all $\delta \in (0, \underline{\delta})$. This immediately implies that $e^d > e_F^c \ \forall \delta \in (0, \underline{\delta})$. $\square$

A.8 Proof of Proposition 5

By part (i) of Theorem 1, we know that if $\omega = 0$, then $e^d < e_F^c \ \forall \delta > 0$. In this case, we let $\underline{\delta}(\omega) = 0$ and $\bar{\delta}(\omega) = +\infty$.

Now consider the functions $D(\delta)$ and $C_F(\delta)$ as defined in (5.1). For every binary distribution F_ω with $\omega \in [0, 1)$, define

$$\Delta(\delta, \omega) \equiv C_{F_\omega}(\delta) - D(\delta) = \frac{\delta^2 \cdot \frac{1+\omega^2}{2} + 2\delta \cdot \left(\frac{1-\omega^2}{4}\right)^2}{2\left(\frac{1-\omega^2}{4} + \delta\right)^2} - \frac{\delta^2 + \delta}{(1 + 2\delta)^2}. \tag{A.29}$$

Since $\Delta(0, \omega) = 0$ for all $\omega \in [0, 1)$, the equation $\Delta(\delta, \omega) = 0$ always has a root $\delta = 0$. To ease the exposition of the algebra, we again use the variables defined in (A.28), which are now given by $\alpha = (1 - \omega^2)/4$ and $\beta = (1 + \omega^2)/2$. Provided that $\delta > 0$, we have for all $\omega \in (0, 1)$,

$$\begin{aligned}
& \Delta(\delta, \omega) = 0 \\
\iff & \frac{(\beta\delta + 2\alpha^2)(1 + 2\delta)^2 - 2(\delta + 1)(\alpha + \delta)^2}{2(\alpha + \delta)^2(1 + 2\delta)^2} = 0 \\
\iff & (\beta\delta + 2\alpha^2)(4\delta^2 + 4\delta + 1) - (2\delta + 2)(\alpha^2 + \delta^2 + 2\alpha\delta) = 0 \\
\iff & (4\beta - 2)\delta^3 + (8\alpha^2 - 4\alpha + 4\beta - 2)\delta^2 + (6\alpha^2 - 4\alpha + \beta)\delta = 0 \\
\iff & (4\beta - 2)\delta^2 + (8\alpha^2 - 4\alpha + 4\beta - 2)\delta + (6\alpha^2 - 4\alpha + \beta) = 0 \\
\iff & \left(\delta + \frac{4\alpha^2 - 2\alpha + 2\beta - 1}{4\beta - 2}\right)^2 = \frac{(4\alpha^2 - 2\alpha + 2\beta - 1)^2 - (6\alpha^2 - 4\alpha + \beta)(4\beta - 2)}{(4\beta - 2)^2} \\
\iff & \left(\delta + \frac{4\alpha^2 - 2\alpha + 2\beta - 1}{4\beta - 2}\right)^2 = \frac{(1 - 2\alpha)^2(4\alpha^2 - 2\beta + 1)}{(4\beta - 2)^2},
\end{aligned} \tag{A.30}$$

where the fifth equivalence follows that $4\beta - 2 = 2 + 2\omega^2 - 2 = 2\omega^2 > 0$. In addition, we can verify that the RHS of (A.30) is strictly negative if $\omega > \sqrt{2} - 1$. This is because $(1 - 2\alpha)^2 = (1 + \omega^2)^2/4 > 0$, and

$$4\alpha^2 - 2\beta + 1 = \frac{(1 - \omega^2)^2}{4} - \omega^2 = \left(\frac{1 - \omega^2}{2} + \omega\right)\left(\frac{1 - \omega^2}{2} - \omega\right),$$

which, given that $\omega \in (0, 1)$, will be positive if and only if $1 - \omega^2 - 2\omega \geq 0$, or, equivalently, $\omega \leq \sqrt{2} - 1$. Hence, if $\omega > \sqrt{2} - 1$, the equation $\Delta(\delta, \omega) = 0$ does not have any non-zero root on $[0,+\infty)$, and Theorem 1 implies that we must have $\Delta(\delta,\omega)>0$ for all $\delta>0$. Since $e^d > e_F^c \iff \Delta(\delta,\omega)>0$, part (ii) of the proposition immediately follows.

Next, suppose that $\omega\in(0,\sqrt{2}-1]$. In this case, the equation $\Delta(\delta,\omega)=0$ admits the following two non-zero roots

$$\begin{aligned}
\underline{\delta}(\omega) &= -\frac{4\alpha^2-2\alpha+2\beta-1}{4\beta-2}-\frac{(1-2\alpha)\sqrt{4\alpha^2-2\beta+1}}{4\beta-2}, \text{ and}\\
\bar{\delta}(\omega) &= -\frac{4\alpha^2-2\alpha+2\beta-1}{4\beta-2}+\frac{(1-2\alpha)\sqrt{4\alpha^2-2\beta+1}}{4\beta-2}.
\end{aligned}$$

In addition, we note that

$$4\alpha^2-2\alpha+2\beta-1=\frac{(1-\omega^2)^2}{4}-\frac{1-\omega^2}{2}+1+\omega^2-1=\frac{\omega^4+4\omega^2-1}{4},$$

which is clearly increasing in ω, and it is approximately equal to -0.07 when $\omega=\sqrt{2}-1$. Thus, the term $-(4\alpha^2-2\alpha+2\beta-1)/(4\beta-2)$ must be strictly positive for all $\omega\in(0,\sqrt{2}-1]$. This implies that if $\omega=\sqrt{2}-1$, the equation $\Delta(\delta,\omega)=0$ will actually admit two identical and strictly positive roots, i.e., $\bar{\delta}(\omega)=\underline{\delta}(\omega)>0$. By continuity, we must have $\Delta(\delta,\sqrt{2}-1)>0$ (and thus $e^d>e_F^c$) for all $\delta\in(0,\underline{\delta}(\omega))\cup(\bar{\delta}(\omega),+\infty)$.

If $\omega<\sqrt{2}-1$, from the above analysis we know that $\bar{\delta}(\omega)>\max\{\underline{\delta}(\omega),0\}$. Thus, by continuity and part (ii) of Theorem 1, it follows that $\Delta(\delta,\omega)>0$ for all $\delta>\bar{\delta}(\omega)$. In addition, since $\lim_{\omega\to0}4\beta-2=\lim_{\omega\to0}\omega^2=0$, it is straightforward to verify that $\lim_{\omega\to0}\bar{\delta}(\omega)=+\infty$. As for the interval $[\max\{0,\underline{\delta}(\omega)\},\bar{\delta}(\omega)]$, because we have $\bar{\delta}(\omega)>\underline{\delta}(\omega)$, $4\alpha^2-2\beta+1>0$, and $\Delta(\delta,\omega)<0$ if

$$\left(\delta-\bar{\delta}(\omega)+\frac{(1-2\alpha)\sqrt{4\alpha^2-2\beta+1}}{4\beta-2}\right)^2<\frac{(2\alpha-1)^2(4\alpha^2-2\beta+1)}{(4\beta-2)^2},$$

it is necessarily the case that $\Delta(\delta,\omega)<0$ for $\delta=\bar{\delta}(\omega)-\varepsilon>0$, where $\varepsilon>0$ is sufficiently small. Hence, we must have $\Delta(\delta,\omega)\le0$ for all $\delta\in[\max\{0,\underline{\delta}(\omega)\},\bar{\delta}(\omega)]$.

It remains to show that $\Delta(\delta,\omega)>0$ for all $\delta\in(0,\max\{0,\underline{\delta}(\omega)\})$. We note that

$$\begin{aligned}
\underline{\delta}(\omega)\le0 &\iff -\frac{4\alpha^2-2\alpha+2\beta-1}{4\beta-2}\le\frac{(1-2\alpha)\sqrt{4\alpha^2-2\beta+1}}{4\beta-2}\\
&\iff (4\alpha^2-2\alpha+2\beta-1)^2\le(1-2\alpha)^2(4\alpha^2-2\beta+1)\\
&\iff 6\alpha^2-4\alpha+\beta\le0\\
&\iff \frac{3(1-\omega^2)^2}{8}-\frac{1-3\omega^2}{2}\le0,\\
&\iff (1-\omega^2)^2-4(1-\omega^2)+4-\frac{4}{3}\frac{4}{3}\le0
\end{aligned}$$

$$
\begin{aligned}
&\iff (1-\omega^2-2)^2 - \frac{4}{3} \le 0 \\
&\iff (1+\omega^2)^2 - \frac{4}{3} \le 0,
\end{aligned}
$$

where the second equivalence holds because, as we have shown above, $4\alpha^2 - 2\alpha + 2\beta - 1 < 0$ and $4\beta - 2 > 0$ for all $\omega \in \big(0, \sqrt{2}-1\big)$. Clearly, the equation $(1+\omega^2)^2 - 4/3 = 0$ has a unique real root on $(0,1)$, which is given by

$$
\hat{\omega} = \sqrt{\frac{2\sqrt{3}}{3} - 1} \approx 0.393.
$$

It is also straightforward to check that $\underline{\delta}(\omega) < 0$ if $\omega < \hat{\omega}$, and $\underline{\delta}(\omega) > 0$ if $\omega \in (\hat{\omega}, \sqrt{2}-1)$. Thus, the claim that $\Delta(\delta,\omega) > 0$ for all $\delta \in (0, \max\{0, \underline{\delta}(\omega)\})$ holds trivially if $\omega \in (0, \hat{\omega})$. As for the case $\omega \in (\hat{\omega}, \sqrt{2-1})$, note that the inequality $\Delta(\delta,\omega) > 0$ can be rewritten as

$$
\left(\delta - \underline{\delta}(\omega) - \frac{(1-2\alpha)\sqrt{4\alpha^2 - 2\beta + 1}}{4\beta - 2}\right)^2 > \frac{(2\alpha-1)^2(4\alpha^2 - 2\beta + 1)}{(4\beta-2)^2}.
$$

But then, given that we have shown $1 - 2\alpha > 0$ and $4\alpha^2 - 2\beta + 1 > 0$, it immediately follows that we also have $\Delta(\delta,\omega) > 0$ for all $\delta \in (0, \max\{\underline{\delta}(\omega), 0\})$ in this case. $\square$

A.9 Proof of Proposition 6

Consider the function $\Delta(\delta,\omega)$ defined in (A.29). For every $\delta > 0$, we have

$$
\begin{aligned}
&\frac{\partial \Delta(\delta,\omega)}{\partial \omega} \\
&= \frac{[8\delta^2\omega + \delta(4\omega^3 - 4\omega)](1-\omega^2+4\delta) - [4\delta^2(1+\omega^2) + 2\delta(\omega^4 - 2\omega^2 + 1)](-4\omega)}{(1-\omega^2+4\delta)^3} \\
&\ge \frac{\delta(4\omega^3 - 4\omega)(1-\omega^2) + 8\delta\omega(\omega^4 - 2\omega^2 + 1)}{(1-\omega^2+4\delta)^3} + \frac{4\delta^2(4\omega^3 - 4\omega) + 16\delta^2\omega(1+\omega^2)}{(1-\omega^2+4\delta)^3} \\
&= \frac{4\delta\omega(1-\omega^2)^2}{(1-\omega^2+4\delta)^3} + \frac{32\delta^2\omega^3}{(1-\omega^2+4\delta)^3},
\end{aligned}
$$

which is strictly positive for all $\omega \in (0,1)$. Thus, $\Delta(\delta,\omega)$ is strictly increasing in ω. Since

$$
\Delta(\delta, 1) = \frac{1}{2} - \frac{\delta^2 + \delta}{(1+2\delta)^2} > 0 > \frac{4\delta^2 + \delta}{(1+4\delta)^2} - \frac{\delta^2 + \delta}{(1+2\delta)^2} = \Delta(\delta, 0),
$$

there must exist a unique cutoff $\hat{\omega}(\delta) \in (0,1)$, such that $\Delta(\delta,\omega) > 0$ if and only if $\omega > \hat{\omega}(\delta)$.

To obtain the exact analytic form of the cutoff, we expand the equation $\Delta(\delta, \omega) = 0$:

$$\begin{aligned}
\Delta(\delta, \omega) = 0 &\Longleftrightarrow 2\omega^2\delta^2 + \frac{\omega^4 + 4\omega^2 - 1}{2} \cdot \delta + \frac{3\omega^4 + 6\omega^2 - 1}{8} = 0 \\
&\Longleftrightarrow (4\delta + 3)\omega^4 + (16\delta^2 + 16\delta + 6)\omega^2 - 4\delta - 1 = 0 \\
&\Longleftrightarrow \omega^4 + \frac{16\delta^2 + 16\delta + 6}{4\delta + 3}\omega^2 - \frac{4\delta + 1}{4\delta + 3} = 0 \\
&\Longleftrightarrow \left(\omega^2 + 2\delta + \frac{2\delta + 3}{4\delta + 3}\right)^2 = \frac{4\delta + 1}{4\delta + 3} + \left(2\delta + \frac{2\delta + 3}{4\delta + 3}\right)^2 \\
&\Longleftrightarrow \left(\omega^2 + 2\delta + \frac{2\delta + 3}{4\delta + 3}\right)^2 = \frac{64\delta^4 + 128\delta^3 + 128\delta^2 + 64\delta + 12}{(4\delta + 3)^2} \\
&\Longleftrightarrow \left(\omega^2 + 2\delta + \frac{2\delta + 3}{4\delta + 3}\right)^2 = \frac{(4\delta + 2)^2(4\delta^2 + 4\delta + 3)}{(4\delta + 3)^2}. \qquad \text{(A.31)}
\end{aligned}$$

For every $\delta > 0$, equation (A.31) has a unit root on $[0, 1]$, which is given by

$$\hat{\omega}(\delta) = \sqrt{\frac{(4\delta + 2)\sqrt{4\delta^2 + 4\delta + 3} - 2\delta - 3}{4\delta + 3} - 2\delta}.$$

To prove the remaining claims of the theorem, let us denote

$$\begin{aligned}
Z(\delta) &= \frac{(4\delta + 2)\sqrt{4\delta^2 + 4\delta + 3} - 2\delta - 3}{4\delta + 3} - 2\delta \\
&= \left(1 - \frac{1}{4\delta + 3}\right)\sqrt{4\delta^2 + 4\delta + 3} - \frac{2\delta + 3}{4\delta + 3} - 2\delta,
\end{aligned}$$

and thus $\hat{\omega}(\delta) = \sqrt{Z(\delta)}$. Differentiating with respect to δ, we obtain

$$Z'(\delta) = \frac{4\sqrt{4\delta^2 + 4\delta + 3}}{(4\delta + 3)^2} + \frac{(4\delta + 2)^2}{4\delta + 3} \cdot \left(4\delta^2 + 4\delta + 3\right)^{-\frac{1}{2}} + \frac{6}{(4\delta + 3)^2} - 2,$$

It is easy to verify that $Z'(0) > 0$. In addition, using Mathematica, one can also check that there is a unique solution to $Z'(\delta) = 0$ on $(0, +\infty)$, which is $\delta^* = \frac{\sqrt{2}}{2} - \frac{1}{2}$, and it satisfies $\hat{\omega}(\delta^*) = \sqrt{2} - 1$. Hence, by continuity, we must have $Z'(\delta) > 0$ $\forall \delta \in (0, \delta^*)$, and $Z'(\delta) > 0$ $\forall \delta \in (\delta^*, +\infty)$. This further implies that the cutoff $\hat{\omega}(\delta)$ must be strictly increasing on $(0, \delta^*)$, and strictly decreasing on $(\delta^*, +\infty)$. It is also straightforward to verify that $\lim_{\delta \to +\infty} Z(\delta) = 0$, and thus $\lim_{\delta \to +\infty} \hat{\omega}(\delta) = 0$. $\square$

A.10 Proof of Theorem 3

Using Propositions 1 and 2, we can compute the expected performance of each division $i \in \{1, 2\}$ in the fully revealing equilibrium under decentralization, which is given by

$$\Pi_i^d(e^d, y_i^d, y_j^d) = K - \sigma_\theta^2 + e^d \left(1 - \frac{2\delta^2 + 2\delta}{(1+2\delta)^2}\right)\sigma_\theta^2.$$

Exploiting that F is symmetric in $\boldsymbol{\eta}$ and the decision rules $\mathbf{y}^d = (y_1^d, y_2^d)$ are independent of $\boldsymbol{\eta}$, we then obtain the expected payoff of the principal under decentralization:

$$\begin{aligned}
\Pi_P^d &= \mathbb{E}\left[\eta_1 \Pi_1^d(e^d, \mathbf{y}^d) + \eta_2 \Pi_2^d(e^d, \mathbf{y}^d)\right] \\
&= 2\mathbb{E}[\eta_i]\Pi_i^d(e^d, \mathbf{y}^d) \\
&= 2\mu \left(K - \sigma_\theta^2 + e^d \left(1 - \frac{2\delta^2 + 2\delta}{(1+2\delta)^2}\right)\sigma_\theta^2\right),
\end{aligned}$$

where $\mu \equiv \mathbb{E}[\eta_i] > 0$, $\forall i = 1, 2$.

We next derive the equilibrium payoff of the principal under centralization, which we will denote as Π_P^c. Under decentralization, each agent invests $e_i = e_F^c$ in acquiring information, and the decision rules are $\mathbf{y}^c = (y_1^c, y_2^c)$ as described in Proposition 4. Hence, for each agent i and a given pair of global states (η_1, η_2), the expected performance of the two divisions are

$$\begin{aligned}
\Pi_1^c(e_F^c, y^c, \boldsymbol{\eta}) &= K - \sigma_\theta^2 + e_F^c \left(1 - \frac{2\delta^2 + 2\delta\lambda^2}{\left(\lambda + \frac{\delta}{1-\lambda}\right)^2}\right)\sigma_\theta^2, \\
\Pi_2^c(e_F^c, y^c, \boldsymbol{\eta}) &= K - \sigma_\theta^2 + e_F^c \left(1 - \frac{2\delta^2 + 2\delta(1-\lambda)^2}{\left(1 - \lambda + \frac{\delta}{\lambda}\right)^2}\right)\sigma_\theta^2,
\end{aligned}$$

where we recall that $\lambda = \eta_1/(\eta_1 + \eta_2)$. Exploiting the symmetry of F, we have

$$\begin{aligned}
\Pi_P^c &= \mathbb{E}\left[\eta_1 \Pi_1^c(e_F^c, y^c, \boldsymbol{\eta}) + \eta_2 \Pi_2^c(e^c, y^c, \boldsymbol{\eta})\right] \\
&= 2\mu \left[K - \sigma_\theta^2 + e_F^c \left(1 - \frac{1}{\mu}\left(\mathbb{E}\left[\eta_1 \cdot \frac{\delta^2 + \delta\lambda^2}{\left(\lambda + \frac{\delta}{1-\lambda}\right)^2}\right] + \mathbb{E}\left[\eta_2 \cdot \frac{\delta^2 + \delta(1-\lambda)^2}{\left(1-\lambda+\frac{\delta}{\lambda}\right)^2}\right]\right)\right)\sigma_\theta^2\right] \\
&= 2\mu \left[K - \sigma_\theta^2 + e_F^c \left(1 - \frac{1}{\mu}\mathbb{E}\left[\frac{\delta^2 \cdot \frac{\eta_1\eta_2}{\eta_1+\eta_2} + \delta \cdot \frac{\eta_1^2\eta_2^2}{(\eta_1+\eta_2)^3}}{\left(\frac{\eta_1\eta_2}{(\eta_1+\eta_2)^2} + \delta\right)^2}\right]\right)\sigma_\theta^2\right].
\end{aligned}$$

Therefore, $\Pi_P^d > \Pi_P^c$ if and only if the following inequality holds:

$$\frac{e^d}{e_F^c} > R_F(\delta) \equiv \left(1 - \frac{1}{\mu}\mathbb{E}\left[\frac{\delta^2 \cdot \frac{\eta_1\eta_2}{\eta_1+\eta_2} + \delta \cdot \frac{\eta_1^2\eta_2^2}{(\eta_1+\eta_2)^3}}{\left(\frac{\eta_1\eta_2}{(\eta_1+\eta_2)^2} + \delta\right)^2}\right]\right) \Bigg/ \left(1 - \frac{2\delta^2 + 2\delta}{(1+2\delta)^2}\right). \tag{A.32}$$

Note that

$$\lim_{\delta\to+\infty} R_F(\delta) = 2 - \frac{2}{\mu}\mathbb{E}\left[\frac{\eta_1\eta_2}{\eta_1+\eta_2}\right] < 2,$$

and

$$\lim_{\delta\to+\infty} \frac{e^d}{e_F^c} = \lim_{\delta\to+\infty} \frac{(c')^{-1}\left((1-D(\delta))\,q\sigma_\theta^2\right)}{(c')^{-1}\left((1-C_F(\delta))q\sigma_\theta^2\right)} = \frac{(c')^{-1}\left(0.75q\sigma_\theta^2\right)}{(c')^{-1}\left(\left(1-\mathbb{E}\left[\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^2\right]\right)q\sigma_\theta^2\right)} > 1,$$

where $D(\delta)$ and $C_F(\delta)$ are defined in (5.1), and the last inequality follows Theorem 1. Thus, the value of $\lim_{\delta\to} e^d/e_F^c$ is strictly large than 1, and it is increasing as the term $\mathbb{E}\left[(\eta_1/(\eta_1+\eta_2))^2\right]$ increases. Let $\varepsilon \equiv \left(\mathbb{E}\left[(\eta_1/(\eta_1+\eta_2))^2\right] - 0.25\right) q\sigma_\theta^2$, which is strictly positive given the assumption $corr(\eta_1,\eta_2) < 1$. Using Taylor's theorem, we obtain

$$(c')^{-1}\left(\left(1-\mathbb{E}\left[\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^2\right]\right)q\sigma_\theta^2\right) = (c')^{-1}\left(0.75q\sigma_\theta^2\right) - \frac{\varepsilon}{c''((c')^{-1}(0.75q\sigma_\theta^2))} + o\left(\varepsilon^2\right).$$

Since $c''(e)\cdot e < \zeta\ \forall e\in[0,1]$, we further have

$$(c')^{-1}\left(\left(1-\mathbb{E}\left[\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^2\right]\right)q\sigma_\theta^2\right) \leq \frac{\zeta-\varepsilon}{c''((c')^{-1}(0.75q\sigma_\theta^2))} + o\left(\varepsilon^2\right). \tag{A.33}$$

When ε is sufficiently small, the value of the higher order terms in $o(\varepsilon^2)$ can be neglected. Thus, if ζ is sufficiently close to (but still larger than) ε, the LHS of (A.33) becomes arbitrarily close to zero (but it is still strictly positive). Hence, if the bound ζ is sufficiently close to $\varepsilon > 0$, then for the case with small enough ε (i.e., $\varepsilon \leq \bar{\varepsilon}$, where $\bar{\varepsilon}$ is some strictly positive cutoff) we must have $\lim_{\delta\to+\infty} e^d/e_F^c > \lim_{\delta\to+\infty} R_F(\delta)$.[30] But then, because the value of $\lim_{\delta\to+\infty} e^d/e_F^c$ is strictly increasing in $\mathbb{E}[(\eta_1/(\eta_1+\eta_2))^2]$, and thus also in ε, with the same bound ζ we will also have $\lim_{\delta\to+\infty} e^d/e_F^c > \lim_{\delta\to+\infty} R_F(\delta)$ for all $\varepsilon \geq \bar{\varepsilon}$. By continuity, it follows that $e^d/e_F^c > R_F(\delta)$ for sufficiently large δ. We can conclude that if ζ is sufficiently small, then there must exist $\bar{\delta} > 0$, such that $\Pi_P^d > \Pi_P^c$ if $\delta > \bar{\delta}$. □

[30] The assumption that "the marginal cost is sufficiently small at $e=0$ ... so that the agents will endogenously choose to be partially informed ($e_i \in (0,1)$ in equilibrium", which is stated in Section 3, implicitly restricts that ζ cannot be too small. Otherwise, we may have the RHS of (A.33) being negative, which implies that in the limit the agents would actually choose to be not informed at all under centralization ($\lim_{\delta\to+\infty} e_F^c = 0$). In this case, if we still have $\lim_{\delta\to+\infty} e^d = (c')^{-1}(0.75q\sigma_\theta^2) > 0$, i.e., the agents would still want to exert some effort under decentralization, then obviously the statement $\lim_{\delta\to+\infty} e^d/e_F^c > \lim_{\delta\to+\infty} R_F(\delta)$ will also hold.

A.11 Proof of Theorem 4

In the proof of Theorem 2, it is shown that if the condition $\mathbb{E}\left[\frac{1}{\lambda^2}\right] > \mathbb{E}\left[\frac{2}{\lambda(1-\lambda)}\right] - 3$ is satisfied, then there exists $\underline{\delta} > 0$, such that $C_F(\delta) > D(\delta)\ \forall\delta \in (0, \underline{\delta})$. Using arguments that are analogous to those in the proof of Theorem 3, we can further show that, for every of such δ there must exist a cutoff $\zeta(\delta) > 0$, such that if $c''(e) \cdot e < \zeta(\delta)\ \forall e \in [0,1]$, then we will also have

$$\frac{e^d}{e^c_F} = \frac{(c')^{-1}\left(\left(1 - D(\delta)\right) q\sigma_\theta^2\right)}{(c')^{-1}\left(\left(1 - C_F(\delta)\right) q\sigma_\theta^2\right)} > R_F(\delta),$$

which, according to (A.32), is both necessary and sufficient for $\Pi^d_P > \Pi^c_P$. $\square$

A.12 Proof of Proposition 7

Decentralization Let us first consider part (i) of the proposition. We argue that the proposed equilibrium effort profile $(e_1^*, e_2^*) = (\check{e}^d, \check{e}^d)$, the fully revealing communication strategies $(\check{m}_1^d, \check{m}_2^d)$, and the decision rules $(\check{y}_1^d, \check{y}_2^d)$ described below constitute a PBE, where the beliefs of the players will be fully pinned down Bayes' rule. Specifically, the decision rules are similar to the ones in Proposition 1, and they are given by

$$\check{y}_i^d\left(e_i, s_i, m_i, m_j\right) = \frac{\mathbb{E}[\theta_i|s_i]}{1+\delta} + \frac{\delta^2\mathbb{E}[\theta_i|m_i]}{(1+2\delta)(1+\delta)} + \frac{\delta\mathbb{E}[\theta_j|m_j]}{1+2\delta},$$

$\forall e_i \in E, \forall s_i, m_i, m_j \in \mathbb{R}\cup\{\emptyset\}$, and $\forall i, j = 1, 2$, where for each $i \in \{1, 2\}$ $\mathbb{E}[\theta_i|m_i]$ is the posterior expectation of the local state θ_i conditional on $s_i = m_i/t^d$ if $m_i \neq \emptyset$, and it is conditional on $s_i = \emptyset$ otherwise. To prove Proposition 7(i), we will suppose that agent j is playing according to $(\check{e}^d, \check{m}_j^d, \check{y}_j^d)$, and then show that it is a best response for agent i to also adopt the proposed strategy.

Taking the first stage effort e_i, the signal s_i received, and the message m_i sent as given, we can solve agent i's utility-maximizing problem at the decision-making stage as in the proof of Proposition 1, and obtain the decision rule $\check{y}_i^d$ as a solution. Thus, given agent j's strategy and the corresponding beliefs, the decision rule $\check{y}_i^d$ is sequentially rational for agent i.

Next, we take the decision rule $\check{y}_i^d$, effort e_i and signal s_i as given, and consider agent i's strategic incentives when communicating his private information with agent j. We start by showing that when agent i receives a *non-null* signal $s_i \in \mathbb{R}$, he will prefer to send message $m_i^d(e_i, s_i) = t^d s_i$ than any other message $m_i \in \mathbb{R}$. Note that since agent j will follow the proposed fully revealing communicating strategy, agent i can always infer the realization of agent j's signal s_j (which is equal to m_j/t^d if $m_j \neq \emptyset$, and it is equal to $\emptyset$ otherwise). Thus for every message $m_i \in \mathbb{R}$ sent by agent i (which may or may not equal to $\check{m}_i^d(e_i, s_i)$), sequential rationality implies that the agents will choose the following actions:

$$y_i = \frac{\theta_i}{1+\delta} + \frac{\delta^2(m_i/t^d)}{(1+2\delta)(1+\delta)} + \frac{\delta\mathbb{E}[\theta_j|s_j]}{1+2\delta},\ y_j = \frac{(1+\delta)\mathbb{E}[\theta_j|s_j]+\delta(m_i/t^d)}{1+2\delta},$$

where $\mathbb{E}[\theta_j|s_j] = 0$ if $s_j = \emptyset$, and $\theta_j = s_j$ otherwise. Thus, given any effort $e_j \in E$ chosen by agent j, conditional on sending the message $m_i \in \mathbb{R}$, the expected performance of division i is

$$\tilde{\Pi}_i^d(m_i, \theta_i, e_j) = e_j \mathbb{E}_{\theta_j}[\tilde{\pi}_1^d(m_i, \theta_i, \theta_j)] + (1-e_j)\tilde{\pi}_i^d(m_i, \theta_i, \emptyset),$$

where

$$\begin{aligned}\tilde{\pi}_i^d(m_i, \theta_i, \theta_j) = K &- \left(\frac{(\delta+2\delta^2)\theta_i - (m_i/t^d)\delta^2}{(1+2\delta)(1+\delta)} - \frac{\delta}{1+2\delta}\cdot\theta_j\right)^2 \\ &- \delta\left(\frac{(1+2\delta)\theta_i - (m_i/t^d)\delta}{(1+2\delta)(1+\delta)} - \frac{1}{1+2\delta}\cdot\theta_j\right)^2\end{aligned}$$

and $\tilde{\pi}_i^d(m_i, \theta_i, \emptyset) = \tilde{\pi}_i^d(m_i, \theta_i, \theta_j)|_{\theta_j=0}$. Differentiating with respect to m_i, we have, $\forall \theta_j \in \mathbb{R}$,

$$\begin{aligned}&\frac{\partial \mathbb{E}_{\theta_j}[\tilde{\pi}_i^d(m_i, \theta_i, \theta_j)]}{\partial m_i} \\ =& -\mathbb{E}_{\theta_j}\left[2\left(\frac{(\delta+2\delta^2)\theta_i - (m_i/t^d)\delta^2}{(1+2\delta)(1+\delta)} - \frac{\delta}{1+2\delta}\cdot\theta_j\right)\left(\frac{-\delta^2/t^d}{(1+2\delta)(1+\delta)}\right)\right] \\ &-\delta\mathbb{E}_{\theta_j}\left[2\left(\frac{(1+2\delta)\theta_i - (m_i/t^d)\delta}{(1+2\delta)(1+\delta)} - \frac{1}{1+2\delta}\cdot\theta_j\right)\left(\frac{-\delta/t^d}{(1+2\delta)(1+\delta)}\right)\right] \\ =& \frac{2\delta^2}{(1+2\delta)(1+\delta)}\cdot\left(\frac{(\delta+2\delta^2)\theta_i - (m_i/t)\delta^2 + (1+2\delta)\theta_i - (m_i/t)}{(1+2\delta)(1+\delta)}\right)\cdot\frac{1}{t^d},\end{aligned}$$

which is also equal to $\partial\tilde{\pi}_i^d(m_i, \theta_i, \emptyset)/\partial m_i$. Hence, we further have

$$\begin{aligned}&\frac{\partial\tilde{\Pi}_i^d(m_i, \theta_i, e_j)}{\partial m_i} \\ =& \frac{2\delta^2}{(1+2\delta)(1+\delta)}\cdot\left(\frac{(\delta+2\delta^2)\theta_i - (m_i/t^d)\delta^2 + (1+2\delta)\theta_i - (m_i/t^d)\delta}{(1+2\delta)(1+\delta)}\right)\cdot\frac{1}{t^d},\end{aligned}$$

which is independent of e_j. This implies that the strategic communication incentives of agent i is independent of his belief about the effort exerted by agent j. Thus, from now on we drop the variable e_j from the function $\tilde{\Pi}_i^d$. We distinguish the following two cases:

Case 1: $s_i = \theta_i = 0$. It is straightforward to verify that

$$\left.\frac{\partial[q\tilde{\Pi}_i^d(m_i, \theta_i)]}{\partial m_i}\right|_{m_i=\theta_i=0} = 0.$$

Since $z(\hat{m}_i, 0) \geq z(0,0) = 0 \; \forall \hat{m}_i \in \mathbb{R}$, it immediately follows that

$$q\tilde{\Pi}_i^d(0,0) - z(0,0) \geq q \cdot \tilde{\Pi}_i^d(\hat{m}_i, 0) - z(\hat{m}_i, 0) \, \forall \hat{m}_i \in \mathbb{R}.$$

Hence, when agent i learns that $\theta_i = 0$, he always prefer to send $m_i = 0$ than any other $\hat{m}_i \in \mathbb{R}$.

Case 2: $s_i = \theta_i \neq 0$. Suppose first that $\theta_i > 0$. Consider any message that $m_i \geq \theta_i$. By (A2), sending this message will incur a cost $z(m_i, \theta_i) = \kappa(m_i - \theta_i)^2$. Note that by construction, we have $t^d \equiv \frac{1}{2} + \sqrt{\frac{q\delta^2}{\kappa(1+2\delta)^2} + \frac{1}{4}}$, which is larger than 1 for all $\delta \geq 0$, and

$$\left[\frac{\partial[q\tilde{\Pi}_i^d(m_i, \theta_i)]}{\partial m_i} - \frac{\partial z(m_i, \theta_i)}{\partial m_i} \right] \Bigg|_{m_i = t^d\theta_i} = 0,$$

i.e., the first-order condition is satisfied exactly at $m_i = t^d\theta_i$. This implies that if agent i learns that the true state is $\theta_i > 0$, he will prefer to send the message $m_i = t^d\theta_i$ than any other messages $\hat{m}_i \in [\theta_i, +\infty)$.

It remains to show that agent i will also prefer $m_i = t^d\theta_i$ than any message $\hat{m}_i \in [0, \theta_i)$. This is indeed the case, because $\frac{\partial\tilde{\Pi}_i^d(m_i,\theta_i)}{\partial m_i} > 0 \, \forall m_i \in [0, \theta_i)$, and, thus, $\forall \theta_i > 0$ and $\hat{m}_i < \theta_i$,

$$\begin{aligned} q\tilde{\Pi}_i^d\left(t^d\theta_i, \theta_i\right) - z(t^d\theta_i, \theta_i) &\geq q\tilde{\Pi}_i^d\left(\theta_i, \theta_i\right) - z(\theta_i, \theta_i) \\ &= q\tilde{\Pi}_i^d\left(\theta_i, \theta_i\right) \\ &> q\tilde{\Pi}_i^d\left(\hat{m}_i, \theta_i\right) \\ &\geq q\tilde{\Pi}_i^d\left(\hat{m}_i, \theta_i\right) - z(\hat{m}_i, \theta_i). \end{aligned}$$

In sum, we have shown that for all $s_i = \theta_i > 0$, agent i would prefer sending $m_i = t^d\theta_i$ than any other message $\hat{m}_i \in \mathbb{R}$. By symmetry, the same conclusion also holds for all $s_i = \theta_i < 0$.

Since the effect on the actions chosen by the agents are the same for $\hat{m}_i = 0$ and $\hat{m}_i = \emptyset$, and we allow the cost $z(m, 0)$ to be fully general, it also follows that agent i will not find it profitable to send $\hat{m}_i = \emptyset$ when $s_i \neq \emptyset$. As for the remaining scenario $s_i = \emptyset$, it is straightforward to verify that the expected performance of division i conditional on agent i sending any message $m_i \in \mathbb{R} \cup \emptyset$ is the same as when he actually receives a non-null signal that $s_i = \theta_i = 0$. Hence, this expected performance is maximized when $m_1 = \emptyset$. Trivially, the communication cost is also minimized at $m_1 = \emptyset$. Therefore, it must be optimal for agent i to send $m_i = \emptyset$ whenever $s_i = \emptyset$ is observed.

Finally, we take both the decision rules $(\check{y}_1^d, \check{y}_2^d)$ and communication strategies $(\check{m}_1^d, \check{m}_2^d)$ as given and consider the information acquisition problem for agent i. Given an arbitrary effort profile (e_1, e_2), the expected payoff of agent i is now given by

$$U_i^d(e_i, e_j) - e_i \mathbb{E}\left[\kappa(t^d\theta_i - \theta_i)^2\right],$$

where $U_i^d(e_i, e_j)$ was defined in the proof of Proposition 1. Thus, with costly exaggeration, the first-order condition at the information acquisition stage is

$$\left(1 - \frac{\delta^2 + \delta}{(1+2\delta)^2}\right) q\sigma_\theta^2 - c'(e_i) - \kappa(t^d - 1)^2 \sigma_\theta^2 = 0.$$

Solving the above equation, we obtain the proposed equilibrium effort level $\check{e}^d$ as a unique solution. Therefore, given the above-mentioned decision rules and communication strategies, choosing $e_i = \check{e}^d$ is indeed optimal for agent i. ■

Centralization For part (ii) of the proposition, we consider first the principal's incentive at the decision-making stage. Given the communication strategy profile $(\check{m}_1^c, \check{m}_2^c)$, the relevant information held by both agents will be perfectly revealed to the principal. In particular, whenever the principal observes that $m_i \neq \emptyset$, she can infer that agent i has learned about his local state, which is given by $\theta_i = m_i/t^c$. Thus, sequential rationality implies that the decision rules of the principal should be similar to the ones in the proof of Proposition 3:

$$\check{y}_i^c(m_i, m_j, \eta_i, \eta_j) = \frac{\frac{\eta_i}{\eta_i+\eta_j} \cdot \left(\frac{\eta_j}{\eta_i+\eta_j} + \delta\right) \mathbb{E}\left[\theta_i | m_i\right] + \frac{\delta\eta_j}{\eta_i+\eta_j} \mathbb{E}[\theta_j | m_j]}{\frac{\eta_i}{\eta_i+\eta_j} \cdot \frac{\eta_j}{\eta_i+\eta_j} + \delta},$$

$\forall m_i, m_j \in \Theta \cup \{\emptyset\}$, $\forall \eta_i, \eta_j \in [\underline{\eta}, \bar{\eta}]$, and $\forall i, j = 1, 2$, where for each $i \in \{1, 2\}$ $\mathbb{E}[\theta_i | m_i]$ is the posterior expectations of the local state θ_i conditional on $s_i = m_i/t^c$ if $m_i \neq \emptyset$, and it is conditional on $s_i = \emptyset$ otherwise.

Next, we take the above decision rules of the principal as given, and consider the strategic incentives of the agents at the communication stage. Suppose that agent 2 plays the fully revealing communication strategy m_2^c. We start by showing that when receiving a non-null signal $s_1 \in \mathbb{R}$, agent 1 will prefer to send $m_1 = t^c s_1$ than any other message $\hat{m}_1 \in \mathbb{R}$. In particular, suppose that agent 1 learns that his local state is $\theta_1 \in \mathbb{R}$, then by sending an arbitrary message $m_1 \in \mathbb{R}$ he will induce the following contingent actions of the principal:

$$y_1^c = \frac{(\lambda(1-\lambda) + \lambda\delta) \cdot (m_1/t^c) + (1-\lambda)\delta \mathbb{E}[\theta_2 | s_2]}{\lambda(1-\lambda) + \delta}$$

and

$$y_2^c = \frac{(\lambda(1-\lambda) + (1-\lambda)\delta) \mathbb{E}[\theta_2 | s_2] + \lambda\delta \cdot (m_1/t^c)}{\lambda(1-\lambda) + \delta},$$

where $\lambda = \eta_1/(\eta_1+\eta_2)$, $\mathbb{E}[\theta_2|s_2] = 0$ if $s_2 = \emptyset$, and $\theta_2 = s_2$ (and thus $\mathbb{E}[\theta_2|s_2] = s_2$) otherwise. Note that we can write the action of the principal as a function of agent 2's private signal because agent 2 s communication strategy is fully revealing.

Given any effort $e_2 \in E$ chosen by agent 2, and any realization of the profitability conditions $\eta_1, \eta_2 \in [\underline{\eta}, \bar{\eta}]$, conditional on sending a message $m_i \in \mathbb{R}$ the expected performance of agent 1 is

$$\tilde{\Pi}_1^c(m_1, \theta_1, e_2, \boldsymbol{\eta}) = e_2 \mathbb{E}_{\theta_2}\left[\tilde{\pi}_1^c(m_1, \theta_2, \boldsymbol{\eta})\right] + (1-e_2)\tilde{\pi}_1^c(m_1, \emptyset, \boldsymbol{\eta}),$$

where

$$\begin{aligned}\tilde{\pi}_1^c(m_1, \theta_2, \boldsymbol{\eta}) =& K - \left(\frac{(\lambda(1-\lambda)+\delta)\theta_1 - (\lambda(1-\lambda)+\lambda\delta)\cdot(m_1/t^c)}{\lambda(1-\lambda)+\delta} - \frac{(1-\lambda)\delta}{\lambda(1-\lambda)+\delta}\cdot\theta_2\right)^2 \\ & - \delta\left(\frac{\lambda(1-\lambda)\cdot(m_1/t^c - \theta_2)}{\lambda(1-\lambda)+\delta}\right)^2\end{aligned}$$

and $\tilde{\pi}_1^c(m_1, \emptyset, \boldsymbol{\eta}) = \tilde{\pi}_1^c(m_1, \theta_2, \boldsymbol{\eta})|_{\theta_2=0}$. Differentiating with respect to m_1, we have, $\forall \theta_2 \in \mathbb{R}$,

$$\begin{aligned}&\frac{\partial \mathbb{E}_{\theta_2}\left[\tilde{\pi}_1^c(m_1, \theta_2, \boldsymbol{\eta})\right]}{\partial m_1} \\ =& \mathbb{E}_{\theta_2}\left[2\left(\frac{(\lambda(1-\lambda)+\delta)\theta_1 - (\lambda(1-\lambda)+\lambda\delta)\cdot(m_1/t^c) - (1-\lambda)\delta\theta_2}{\lambda(1-\lambda)+\delta}\right)\left(\frac{(\lambda(1-\lambda)+\lambda\delta)/t^c}{\lambda(1-\lambda)+\delta}\right)\right] \\ & - \mathbb{E}_{\theta_2}\left[2\delta\left(\frac{\lambda(1-\lambda)\cdot(m_1/t^c - \theta_2)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{\lambda(1-\lambda)/t^c}{\lambda(1-\lambda)+\delta}\right)\right] \\ =& 2\left(\frac{(\lambda(1-\lambda)+\delta)\theta_1 - (\lambda(1-\lambda)+\lambda\delta)\cdot(m_1/t^c)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{(\lambda(1-\lambda)+\lambda\delta)/t^c}{\lambda(1-\lambda)+\delta}\right) \\ & - 2\delta\left(\frac{\lambda(1-\lambda)\cdot(m_1/t^c)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{\lambda(1-\lambda)/t^c}{\lambda(1-\lambda)+\delta}\right),\end{aligned}$$

which is also equal to $\partial\tilde{\pi}_1^c(m_1, \emptyset, \boldsymbol{\eta})/\partial m_1$. Hence, we further have

$$\begin{aligned}&\frac{\partial \tilde{\Pi}_1^c(m_1, \theta_1, e_2, \boldsymbol{\eta})}{\partial m_1} \\ =& 2\left(\frac{(\lambda(1-\lambda)+\delta)\theta_1 - (\lambda(1-\lambda)+\lambda\delta)\cdot(\hat{m}_1/t^c)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{(\lambda(1-\lambda)+\lambda\delta)/t^c}{\lambda(1-\lambda)+\delta}\right) \\ & - 2\delta\left(\frac{\lambda(1-\lambda)\cdot(\hat{m}_1/t^c)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{\lambda(1-\lambda)/t^c}{\lambda(1-\lambda)+\delta}\right),\end{aligned}$$

which is independent of e_2. This implies that the strategic communication incentives of agent 1 is independent of his belief about the effort exerted by agent 2. Thus, from now on we drop the variable e_2 from the function $\tilde{\Pi}_1^c$. We distinguish the following two cases:

Case 1: $s_i = \theta_i = 0$. It is straightforward to verify that

$$\left.\frac{\partial[q\tilde{\Pi}_1^c(m_1,\theta_1,\boldsymbol{\eta})]}{\partial m_1}\right|_{m_1=\theta_1=0} = 0 \quad \forall \eta_1,\eta_2 \in [\,\underline{\eta},\bar{\eta}\,].$$

Since $z(\hat{m}_1, 0) \geq z(0,0) = 0 \; \forall \hat{m}_1 \in \mathbb{R}$, it immediately follows that

$$q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(0,0,\boldsymbol{\eta})\right] - z(0,0) \geq \mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(\hat{m}_1,0,\boldsymbol{\eta})\right] - z(\hat{m}_1,0) \, \forall \hat{m}_1 \in \mathbb{R}.$$

Hence, when agent 1 learns that $\theta_i = 0$, he always prefer to send $m_1 = t^c\theta_1 = 0$ than any other message $\hat{m}_1 \in \mathbb{R}$.

Case 2: $s_i = \theta_i \neq 0$. Suppose first that $\theta_i > 0$. Consider any message that $m_1 \geq \theta_1$. By (A2), sending this message will incur a cost $z(m_1,\theta_1) = \kappa(m_1 - \theta_1)^2$. Note that by construction, we have $t^c = \frac{1}{2} + \sqrt{\mathbb{E}\left[\frac{q\lambda(1-\lambda)[2\delta^2+\delta(\lambda^2+(1-\lambda)^2)]}{\kappa(\lambda(1-\lambda)+\delta)^2}\right] + \frac{1}{4}}$, which is larger than 1 for all $\delta \geq 0$, and

$$\left.\left[\frac{\partial \mathbb{E}_{\boldsymbol{\eta}}\left[q\tilde{\Pi}_1^c(m_1,\theta_1,\boldsymbol{\eta})\right]}{\partial m_1} - \frac{\partial z(m_1,\theta_1)}{\partial m_1}\right]\right|_{m_1=t^c\theta_1} = 0,$$

i.e., the first-order condition is satisfied exactly at $m_i = t^c\theta_i$.[31] This implies that if agent 1 learns that the true state is $\theta_1 > 0$ and the principal expects him to send messages according to the fully revealing communication rule $\tilde{m}_1^c$, then agent 1 will indeed prefer to send the message $m_1 = t^c\theta_1$ than any other messages $\hat{m}_1 \in [\theta_1, +\infty)$.

It remains to show that agent 1 will also prefer $m_1 = t^c\theta_1$ than any message $\hat{m}_1 \in [0,\theta_1)$. This is indeed the case, because $\forall m_1 \in [0,\theta_1)$ and $\forall \eta_1,\eta_2 \in [\,\underline{\eta},\bar{\eta}\,]$,

$$\frac{\partial \tilde{\Pi}_1^c(m_1,\theta_1,\boldsymbol{\eta})}{\partial m_1} \geq 2\left(\frac{(\lambda(1-\lambda)+\delta)(\theta_1 - m_1/t^c)}{\lambda(1-\lambda)+\delta}\right)\left(\frac{(\lambda(1-\lambda)+\lambda\delta)/t^c}{\lambda(1-\lambda)+\delta}\right) > 0,$$

and, thus, $\forall \theta_1 > 0$ and $\hat{m}_1 < \theta_1$,

$$\begin{aligned}
q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(t^c\theta_1,\theta_1,\boldsymbol{\eta})\right] - z(t^c\theta_1,\theta_1) &\geq q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(\theta_1,\theta_1,\boldsymbol{\eta})\right] - z(\theta_1,\theta_1) \\
&= q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(\theta_1,\theta_1,\boldsymbol{\eta})\right] \\
&> q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(\hat{m}_1,\theta_1,\boldsymbol{\eta})\right] \\
&\geq q\mathbb{E}_{\boldsymbol{\eta}}\left[\tilde{\Pi}_1^c(\hat{m}_1,\theta_1,\boldsymbol{\eta})\right] - z(\hat{m}_1,\theta_1).
\end{aligned}$$

In sum, we have shown that for all $s_1 = \theta_i > 0$, agent 1 would prefer sending $m_1 = t^c s_1$ than any other message $\hat{m}_1 \in \mathbb{R}$. By symmetry, the same conclusion also holds for all $s_1 < 0$.

[31] To arrive at the expression of t^c, we exploit that the symmetry of the distribution of λ and observe that $\mathbb{E}_\lambda\left[\frac{\lambda(1-\lambda)\delta^2+\lambda(1-\lambda)^3\delta}{(\lambda(1-\lambda)+\delta)^2}\right] = \mathbb{E}\left[\frac{\lambda(1-\lambda)[2\delta^2+\delta(\lambda^2+(1-\lambda)^2)]}{2(\lambda(1-\lambda)+\delta)^2}\right]$.

Since the effect on the actions chosen by the principal are the same for $\hat{m}_1 = 0$ and $\hat{m}_1 = \emptyset$, and we allow the cost $z(m, 0)$ to be fully general, it also follows from the argument in Case 1 that agent 1 will not find it profitable to send $\hat{m}_1 = \emptyset$ when $s_1 \neq \emptyset$. As for the remaining scenario $s_1 = \emptyset$, it is straightforward to verify that the expected performance of division 1 conditional on agent 1 sending any message $\hat{m}_1$ is the same as when he actually receives a non-null signal of $s_1 = \theta_1 = 0$. Hence, this expected performance is maximized when $\hat{m}_1 = \emptyset$. Trivially, the communication cost is also minimized when $\hat{m}_1 = \emptyset$. Therefore, it is optimal for agent 1 to report $m_1 = \emptyset$ to the principal whenever $s_1 = \emptyset$ is observed.

By the symmetry of distribution F, the incentive problem of agent 2 is analogous. Thus, given that agent 1 will be fully revealing his private information to the principal, it is also a best response for agent 2 to follow the communication strategy $\check{m}_2^c$.

Finally, we take both the decision rules $(\check{y}_1^c, \check{y}_2^c)$ and the communication strategies $(\check{m}_1^c, \check{m}_2^c)$ as given and consider the information acquisition problem for agent i. Given an arbitrary effort profile (e_1, e_2), the expected payoff of agent i is now given by

$$U_i^c(e_i, e_j) - e_i \mathbb{E}\left[\kappa(t^c\theta_i - \theta_i)^2\right],$$

where $U_i^c(e_i, e_j)$ was defined in the proof of Proposition 3. Thus, with costly exaggeration, the first-order condition at the information acquisition stage is

$$\left(1 - \mathbb{E}\left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]\right) q\sigma_\theta^2 - c'(e_i) - \kappa(t^c - 1)^2\sigma_\theta^2 = 0.$$

Solving the above equation, we obtain the proposed equilibrium effort level $\check{e}_F^c$ as a unique solution. Therefore, given the above-mentioned decision rules and communication strategies, choosing $e_i = \check{e}_F^c$ is indeed optimal for agent i. ■□