When does centralization undermine adaptation?

Citation for published version:
Liu, S & Migrow, D 2022, 'When does centralization undermine adaptation?', *Journal of Economic Theory*, vol. 205, 105533, pp. 1-35. https://doi.org/10.1016/j.jet.2022.105533

Digital Object Identifier (DOI):
10.1016/j.jet.2022.105533

Published In:
Journal of Economic Theory

When Does Centralization Undermine Adaptation?

Shuo Liu and Dimitri Migrow*

August 19, 2022

Abstract

We revisit the classic problem of optimally allocating decision rights in a multi-divisional organization. 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 division managers are certain about how the headquarter (HQ) weights each division's performance, centralization may always dominate decentralization in generating information, and therefore even lead to more adaptative decisions. However, with uncertainty in HQ's decision criterion, centralization can perform poorly in motivating information acquisition, and particularly so when it is highly important to coordinate the activities of different divisions. As a result, decentralization can be optimal even with an arbitrarily strong coordination motive.

Keywords: centralization, decentralization, coordinated adaptation, information acquisition, verifiable disclosure

JEL Classification: D82, M52

---

*Liu: Guanghua School of Management, Peking University, China. Migrow: School of Economics, University of Edinburgh, United Kingdom. Emails: shuo.liu@gsm.pku.edu.cn and dimitri.migrow@ed.ac.uk. This paper was previously circulated under the title "Desigining Organizations in Volatile Markets." We thank Dan Bernhardt, Julia Grünseis, Yingni Guo, Alexander Jakobsen, Navin Kartik, Alexey Kushnir, Igor Letina, Longjian Li, Nick Netzer, Christian Oertel, Rob Oxoby, Harry Pei, Michael Powell, Marek Pycia, Heikki Rantakari, Armin Schmutzler, Alastair Smith, Jakub Steiner, Kremena Valkanova, Alex Whalley, Mengxi Zhang, the editor Alessandro Pavan, three anonymous referees, and seminar participants at Beijing, Bern, Bonn, Calgary, Shanghai, Vienna, Warwick and Zurich for many useful comments and suggestions. Shuo Liu would like to acknowledge the financial support from the National Natural Science Foundation of China (Grant No. 72103006) and the Forschungskredit of the University of Zurich (grant no. FK-17-018). Dimitri Migrow would like to acknowledge the financial support from the Social Sciences and Humanities Research Council, Canada (SSHRC 2021 Insight Development Grant competition). Declarations of interest: none.

1 Introduction

Uncertainty in managerial decisions, as recognized by the management literature, is endemic to many organizations (e.g. Jackall, 1988; Mintzberg, 1983; Pfeffer, 1981; Stevenson, Cruikshank, and Moldoveanu, 1998). In particular, Stevenson et al. (1998) emphasize that uncertainty in the objectives of top executives can have severe ramifications for the incentives of their subordinates. Likewise, Mintzberg notes that personal concentration of authority can make an organization susceptible to the "whim of the individual manager" (Mintzberg (1983), p. 223), threatening its overall performance.[1]

William Durant's second presidency at General Motors (GM) provides a vivid case in point. Amid signs of an economic downturn, Durant's GM aggressively expanded its production and stocked up on inventory in the first months of 1920, aggravating enormously its loss from the subsequent deflationary crisis that led to the collapse of the automobile market. A major culprit for GM's disastrous performance, according to many business historians, was the chaotic decision-making structure under Durant's regime. As Rae (1958) describes, Durant "provided only spasmodic and haphazard coordination among the component parts of General Motors, and he subjected his subordinates to capricious and unpredictable interference."[2] Moreover, Durant frustrated even the ablest of his associates by downplaying their thoughtful advice.[3] Thus, it was probably unsurprising that when the 1920 crisis began the company did not have the necessary information to put a stop to expansion and further loan acquisitions, but instead pursued inefficient policies such as a price freeze (Chandler (1962), p. 127 -129).

A possible remedy could therefore be decentralization. Indeed, extending decision-making power to mid-level managers (while enhancing cross-divisional communication) was a key part of Alfred Sloan's celebrated reform of GM following Durant's departure from the company.[4] However, to the extent that divisional managers may be more interested in using received information to maximize their own performances rather than the performance of the *entire* firm, the success of Sloan's reform was not obvious ex ante.

[1]Indeed, in a *Harvard Business Review* article Stevenson and Moldoveanu warn that "[division managers] will go mad if punishment and reward are doled out randomly and if they cannot know in advance whether a given outcome will be a win or a loss." (Stevenson and Moldoveanu, 1995).

[2]Chandler (1962) documents how major corporate decisions, such as plant expansion, capital investment, and pricing, were frequently made ad hoc by Durant either in sporadic conferences or personal conversations. Rae (1958) provides a detailed story that Durant repeatedly intervened Walter Chrysler's management of Buick (e.g., moving people in and out of it, making new factory-building plan), an operating division responsible for half of GM's revenue at that time, without bothering to consult the latter man at all.

[3]For instance, in 1919, Durant rejected Chrysler's blunt warning of entering the farm machinery business, which was proved right at a cost to GM of $30 million and put the last straw on Chrysler's resignation.

[4]Both Chandler (1962) and Freeland (1996) emphasize the importance of Sloan's decision to uphold various committees that brought together division managers, where they were encouraged to communicate information and coordinate accordingly with each other. Although GM's owners vetoed Sloan's proposal for granting formal authority to these committees, evidence shows that they were deeply involved in shaping fundamental policies and the strategic planning of the company (see Freeland (1996), p.494 - 497).

Motivated by the above discussion in the management literature and the case of GM, this paper studies a novel trade-off in authority allocation. On the one hand, the information received under decentralization may be used suboptimally from the perspective of the center. On the other hand, centralization leads to a motivational cost arising from the uncertainty surrounding the principal's use of the information provided by divisional agents. Such uncertainty is critical in our theory: without this feature, centralization would always dominate decentralization in incentivizing information acquisition. In that case, centralization may even achieve *both* better coordination (e.g., promoting synergies across divisions) *and* adaptation (e.g., tailoring each division's product to the tastes of local consumers) than decentralization.

Our key result is that under uncertainty in the principal's decision criterion, the resolution of the above trade-off depends primarily on how much the organization values coordination relative to adaptation, and in an unexpected way. The optimality of decentralization can be the result of a large coordination motive. The driving force for this result is that although a large coordination motive can decrease the inefficiency as to how information acquired by autonomous agents is translated into divisional decisions, it cannot eliminate the uncertainty related to the principal's decision-making under centralization. As coordination becomes sufficiently important, the advantage of decentralization in motivating information acquisition eventually dominates.

Our paper contributes to a better understanding of how an organization's performance is affected by its authority structure. This is a central question in organizational economics as suggested by both the abundant research on related topics (Roberts and Gibbons, 2013) and the stories of many modern corporations like GM (Garicano and Rayo, 2016). Our paper also makes a normative contribution by shedding light on how authority over critical decisions should be allocated within a multi-divisional organization, for instance a multiproduct or multinational firm. Such firms often experience uncertainty in the *relative* return of conducting activities in different markets due to, for example, unforeseen changes in consumer tastes or market size or the volatility of currency exchange rate. In volatile environments such as these, the principal may want to centralize the decision-making process so that she can take into account the actual profitability conditions of each local market and make contingent decisions that are globally optimal for the firm.[5] Our findings, however, suggest that this may lead to insufficient incentives for the firm's local delegates to acquire relevant information, because they are uncertain about the decision weights that the principal would attach to that information.

We formalize our arguments by modeling an organization that needs to adapt and coordinate the strategic decisions of its two divisions. As in Alonso, Dessein, and Matouschek (2008) and Rantakari (2008), a division's performance is determined by how close its action (e.g., the

[5]For instance, Cheung and Sengupta (2013), Héricourt and Nedoncelle (2018), and Héricourt and Poncet (2015) provide evidence that the trade flows of multinational firms are substantially affected by exchange rate volatility (e.g., reallocating trade flows towards destinations favorably affected by exchange rate shocks).

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 in 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 higher effort leads to a more informative 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 due to 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 performance to the success of the entire organization (or to the principal's satisfaction) 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. A key variable in our setting is the ratio of the above weights, which determines the *relative importance* of each division in the eyes of the principal.

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 subsequently chooses 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.[6]

In our setting the players exchange private but verifiable ("hard") information (e.g. Dye, 1985; Milgrom and Roberts, 1986; Tirole, 1986). Typical examples of hard information in organizations include revenue data, cash flow and capital expenditures (Bertomeu and Marinovic, 2015; Udell, 2009). They are relevant in our setting if a division manager wants, for instance, to convince the headquarter of the good financial standing of his own division.

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. This is in sharp contrast to previous studies (e.g. Alonso et al., 2008; Aoki, 1986; Dessein and Santos, 2006; Rantakari, 2008) that have argued that the quality of communication is important for determining the relative performance of different organizational structures. As we show in Section 3, fully revealing communication arises as a unique equilibrium outcome

[6]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 interest are sufficiently complex (e.g., product design), which renders ex ante contracting infeasible.

regardless of which authority structure is chosen.

While the quality of communication does not differ between centralization and decentralization, the allocation of decision rights does have an impact on the agents' information-gathering incentives. The pattern of the impact crucially depends on the volatility of the decision weights in the principal's payoff function (which is the source of uncertainty in the principal's decision-making process under centralization). More specifically, in Section 4, we first establish a benchmark result (Theorem 1(i)): If both operating divisions are always equally important to the organization, then centralization always induces higher efforts towards information acquisition, independent of the strength of coordination motive.

The main driving force of the above sharp result is the comparative disadvantage of decentralization in internalizing the *externality* of agents' actions. Due to the coordination requirement, an agent's adaptation behavior (i.e., being responsive to his local information) reduces the profit of the other division, which further decreases the expected benefit that the other agent can gain from acquiring better information. While such externality of action choices is present under both decentralization and centralization, the principal has the ability to internalize it in the latter case as she controls *both* decisions. In particular, if the principal is balanced, she would choose the actions that maximize the joint profits of the agents, which will lead to a higher marginal value of information than the decentralized outcome. Hence, centralization is optimal for motivating information acquisition if both divisions will be equally important to the principal for sure. In that scenario, it is conceivable that centralization may actually improve adaptation, which is in contrast to the received wisdom.

However, the situation changes once we introduce uncertainty about the relative importance of the divisions. In Theorem 1 (ii), we show that if the two (ex ante identical) divisions can be unequal in their importance to the principal ex post, then decentralization outperforms centralization in terms of information gathering whenever the coordination motive is sufficiently strong. Intuitively, the decision weights that the principal assigns to different divisions have no effect under decentralization as the agents are autonomous. However, these weights are highly relevant under centralization, because they vary how the principal wants to coordinate agents' actions based on received information. Since the losses from mis-adaptation are convex, variability in the principal's use of information is painful for the agents, and therefore discourages information acquisition. Importantly, this motivational damage due to the principal's uncertain decision criterion is present under centralization at any level of the coordination motive, while the externality problem under decentralization – that the agents are *too responsive* to their local information – becomes progressively less relevant as coordination becomes progressively more important. Consequently, decentralization induces higher efforts than centralization whenever coordination is sufficiently important (in which case centralization necessarily undermines adaptation).

Next, our Theorem 2 shows that even if the coordination motive is weak, decentralization can result in higher efforts. For this to be true, the decision weights of the principal have to be sufficiently volatile. The reason is that, although the motivational damages due to the externality effect of decentralization and to the uncertainty effect of centralization both vanish as the need for coordination diminishes (because divisions' performances are independent in the limit), the *speed* at which the latter effect vanishes is slower when the uncertainty in the relative importance of the divisions is larger. Furthermore, by fully characterizing the cases where the principal's decision weights are binomially distributed, we demonstrate that with substantially high volatility the motivational advantage of decentralization may hold regardless of the importance of coordination.

Whether and under which conditions the motivational advantage of decentralization outweighs the cost of losing control are more subtle because it is exactly when the managerial priorities are highly uncertain that the principal values most the control granted by centralization. In Theorems 3 and 4, we identify some intuitive properties of the information technology that ensures that the gap in information quality between decentralization and centralization is large enough, so that the principal's expected payoff is also higher under decentralization. Specifically, we note that if the marginal cost of effort increases very fast relative to the associated change in the precision of the signal, the additional gain in information quality due to decentralization will be relatively minor. In this case it would not be optimal for the principal to extend authority to the agents. In contrast, if the marginal cost of information does not increase very fast (for example, if the effort cost function is not too convex when it is of the power form), the drop in information quality due to centralization can be substantial enough to make decentralization optimal for the principal. Finally, we use the case with binomial distribution to illustrate that the scope for centralization to be optimal is the largest when *both* coordination motive *and* the uncertainty in principal's decision weights are small to intermediate.

The paper proceeds as follows. Section 2 introduces the model. Section 3 characterizes the equilibria under different organizational structures and Section 4 studies the optimal organizational structure. Section 5 contains three extensions that illustrate the robustness of our main results. Section 6 discusses the related literature, and Section 7 concludes. All proofs are relegated to Appendices A and A.6.

2 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 according 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 $e \in E = [0, 1]$ to acquire an informative signal s_i^e about the local state of his division. We use $\mathcal{S}^e$ and $G^e(\cdot|\theta)$ to denote the support and the conditional distribution of the signal, respectively. Importantly, we assume that a higher effort always leads to a more precise signal, in the sense of being associated with a higher variance of conditional expectations:

$$\forall e, e' \in E,\ e < e' \Longrightarrow \sigma^2_{E[\theta|s]}(e) := \int_{\theta\in\Theta} \int_{s\in S^e} \left(\mathbb{E}[\theta|s]\right)^2 dG^e(s|\theta) d\Gamma(\theta) \leq \sigma^2_{E[\theta|s]}(e'). \quad (2.1)$$

Note that an equivalent condition of (2.1) is that the *mean squared error* for the Bayesian estimator of the local state, $\int_{\theta\in\Theta} \int_{s\in S^e} \left(\mathbb{E}[\theta|s] - \theta\right)^2 dG^e(s|\theta) d\Gamma(\theta)$, is decreasing in effort.[7] For simplicity, we also assume that $\sigma^2_{E[\theta|s]}(\cdot)$ is twice-differentiable and concave in effort, the effort cost function $c(\cdot)$ is twice-differentiable, strictly increasing and strictly convex, and $\lim_{e\to 0}((\sigma^2_{E[\theta|s]})'(e)/c'(e))$ is sufficiently large and $\lim_{e\to 1}((\sigma^2_{E[\theta|s]})'(e)/c'(e))$ is sufficiently small. These regularity conditions guarantee that an interior effort level will be chosen in equilibrium.

Each agent cares about the performance of his own division. In particular, the ex post payoff of agent i is given by

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

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 $\boldsymbol{\eta} =$

[7]Given the quadratic payoff structure, it is clear that condition (2.1) will hold if the signals can be ranked according to Blackwell's informativeness measure (Blackwell, 1953). More generally, it will also be sufficient if the set of available signals satisfies two common assumptions in the literature of information acquisition (see, e.g., Bergemann and Välimäki, 2002; Dewatripont, Jewitt, and Tirole, 1999; Persico, 2000): First, for each $e \in E$, the distribution of the signal satisfies the *monotone likelihood ratio property* (Milgrom, 1981). Second, the *Lehmann-precision* (Lehmann, 1988) of the acquired signal is increasing in effort. Examples of information acquisition technology satisfying both assumptions include Gaussian learning ($s_i^e = \theta_i + \varepsilon^e$, where ε^e is independently and normally distributed and has a variance decreasing in e), and the "truth-or-noise" signal structure (with probability increasing in effort, the signal fully reveals the state, while with the remaining probability it is drawn from some distribution Γ' on Θ independently from the true state).

$(\eta_1, \eta_2) \in \Psi \subseteq \mathbb{R}^2_+$:

$$\pi_P(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\eta}) = \eta_1 \pi_1(\mathbf{y}, \theta_1) + \eta_2 \pi_2(\mathbf{y}, \theta_2). \tag{2.2}$$

Thus, η_i measures the marginal benefit for the principal from increasing division i's performance. We assume that the random variables η_1 and η_2 are drawn according to some *symmetric* and commonly known joint probability distribution $F_{\boldsymbol{\eta}}$. Our formulation permits the possibility that ex post the principal may assign different weights to the two divisions (i.e., $\eta_1 \neq \eta_2$), which captures the uncertainty in managerial priority that we discussed in the introduction.[8]

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.[9] 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 principal's objectives with respect to the organizational divisions can be interdependent in a rather complex way. 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 model. Note that the global states affect the principal's incentive only through their relative weight $\lambda \equiv \eta_1/(\eta_1 + \eta_2) \in [0, 1]$ (we adopt the convention $0/0 = 0$), which can be interpreted as the *relative importance* of division 1 to the organization. We use F_λ to denote the distribution function of λ, which is induced by $F_{\boldsymbol{\eta}}$ and thus is symmetric around $1/2$.

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.[10] 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.

[8] Our model is agnostic about the precise source of such uncertainty in managerial priority. It may reflect the volatility of the relevant economic variables, in which case the interest of the principal would indeed be aligned with the organization's. For instance, in the context of multinational corporations, π_i can represent 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. Alternatively, the uncertainty could also stem from some intrinsic biases or "unpredictable characteristics" of the principal (see Durant's story in the introduction). This interpretation relates our model to the growing literature on behaviorally biased supervisors (e.g. Giebe and Gürtler, 2012; Letina, Liu, and Netzer, 2020; Prendergast and Topel, 1996), and it also provides a rationale for why the *absolute* importance of a division's performance may be different for the corresponding agent and the principal.

[9] Agastya, Bag, and Chakraborty (2014) study similar interim uncertainty in the degree of conflict of interests between a sender and a receiver, and show how that affects the informativeness of communication.

[10] As will become clear in Section 3.1, 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.

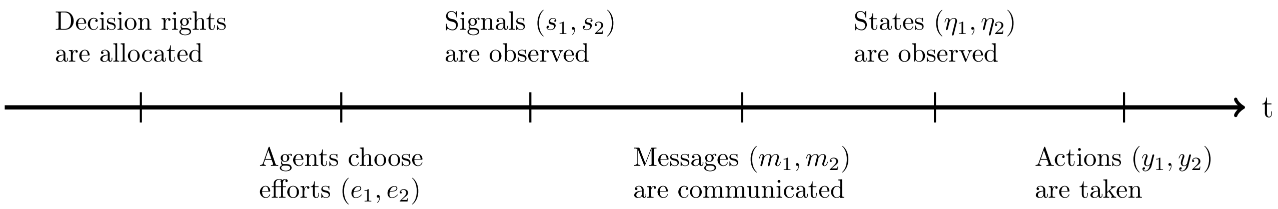


Figure 1: Timing of Events

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 particular, we assume that conditional on receiving a signal s_i, 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 $\mathcal{S} = \cup_{e \in E} S^e$ and $\mathcal{M} = \cup_{s_i \in \mathcal{S}} \mathcal{M}(s_i)$ and assume that the signal-dependent message spaces satisfy the conditions below.[11]

Assumption 1. *Let $\Gamma(\cdot|s_i)$ be the conditional distribution function of the local state given signal s_i. The messages spaces satisfy (i) $\forall t$, $\exists\, m^t$ such that $m^t \in \mathcal{M}(s_i)$ if and only if $\Gamma(\cdot|s_i) \neq \Gamma(\cdot)$ and $\mathbb{E}[\theta_i|s_i] = t$, and (ii) $\exists\, m_i = \emptyset$ such that $\emptyset \in \mathcal{M}(s_i)$ if $\Gamma(\cdot|s_i) = \Gamma(\cdot)$.*

Assumption 2. *$\forall m \in \mathcal{M}$, there exists $\underline{s}^m \in \mathcal{S}^m$ such that $\mathbb{E}[\theta_i|\underline{s}^m] = \min_{s_i \in \mathcal{S}^m} \big|\mathbb{E}[\theta_i|s_i]\big|$, where $\mathcal{S}^m = \{s_i \in \mathcal{S} : m \in \mathcal{M}(s_i)\}$ is the set of signals which could possibly make the message m available to the agent.*

The essential requirement of Assumption 1 is that an *informed* agent can always certify his *updated* estimation of the local state, which will be shown to be a sufficient statics for the players' strategic interaction in the rest of the game and hence can be labelled as the agent's *type* (Seidmann and Winter, 1997). In particular, whenever the message m^t is communicated, the receiving party will know for sure that agent i's posterior expectation about θ_i equals t, obviating the need to know what information structure (or effort level) was actually chosen by the agent. This property greatly simplifies our equilibrium analysis. Meanwhile, as in most setups of certifiable information, Assumption 1 allows for the possibility that an agent who has received an uninformative signal may not be able to prove that he is uninformed (note that the message $\emptyset$ may be available to both informed and uninformed type). Assumption 2 is mainly technical. We will be explicit about its role in the analysis later (see discussion following Proposition 1).[12]

[11] In an earlier working paper version (Liu and Migrow, 2019), we study an extension with signal-independent message spaces and costly lying (Emons and Fluet, 2020; Kartik, 2009; Kartik, Ottaviani, and Squintani, 2007).

[12] Our assumptions accommodate a large class of communication games. For example, suppose that information acquisition takes the "all-or-nothing" form, where $\mathcal{S}^e = \Theta \cup \{\emptyset\}\ \forall e$ and Θ is compact. The assumptions about the message spaces are then satisfied by the evidence game introduced by Dye (1985), where $\mathcal{M}(s_i) = \{s_i, \emptyset\}$

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 $F_{\boldsymbol{\eta}}$ (or $F_{\boldsymbol{\lambda}}$), i.e, the coordination motive and the uncertainty of the stochastic states. To answer this question, we first derive and analyze the respective pure strategy *perfect Bayesian equilibria* (PBE; Fudenberg and Tirole, 1991, p. 333) of the games under centralization and decentralization (see Section 3). 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 fully revealing 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 4.

3 Equilibrium Analysis

We begin our analysis with the formalization of the players' strategies and the belief-updating under different organizational structures. In a decentralized organization, each agent has full control over the decision of his own division and his incentives will be independent of the global states (which only affect the principal's payoff). Hence, the strategy of each agent $i \in \{1, 2\}$ is a triple $(e_i^d, m_i^d(\cdot), y_i^d(\cdot))$ where $e_i^d \in E$ is his effort to acquire decision-relevant information, $m_i^d(\cdot)$ 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 $y_i^d(\cdot)$ 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 $\mu_i^j(\mathcal{S}^{m_j}|m_j) = 1 \ \forall m_j \in \mathcal{M}$.

Under centralization, the principal has full control over the decisions of both divisions. Thus, in contrast to decentralization, when making their effort choices and communicating their signals, the agents take into account how the global states may affect the principal's

$\forall s_i \in \mathcal{S}$, 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 Assumption 1 rules out pure cheap talk communication, it nevertheless permits the following game of cheap talk with certification: $\emptyset \in \mathcal{M}(\emptyset)$, and $\mathcal{M}(s_i) = \Theta \cup \{\emptyset, c^{s_i}\}$ 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 (if it is informative).

decisions. As a result, each agent i's strategy is a pair $(e_i^c, m_i^c(\cdot))$, where $e_i^c \in E$ is his effort to acquire information about his local state θ_i and $m_i^c(\cdot)$ 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 $(y_1^c(\cdot), y_2^c(\cdot))$, where $y_i^c(m_i, m_j, \lambda)$ is the action that the principal takes for division i when receiving messages (m_i, m_j) from the agents and observing the relative importance of the divisions λ. 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 private signal, which we denote by $\mu_p^j(\cdot|m_j) \in \Delta(\mathcal{S})$, must be consistent. That is, $\mu_p^j(\mathcal{S}^{m_j}|m_j) = 1\ \forall m_j \in \mathcal{M}$.

3.1 Information transmission and decision-making

Our first proposition shows that despite the conflicts of interests, all relevant information can be incentive-compatibly transmitted to the decision-making parties in equilibrium irrespective of the choice of organizational structure.

Proposition 1. *Truth-telling can be achieved under both centralization and decentralization:* $\forall z \in \{c, d\}$, *there exists a PBE in which both agents* $i = 1, 2$ *adopt the communication strategy* $m_i^z(e_i, s_i) = m^{\mathbb{E}[\theta_i|s_i]}$ *if* $\Gamma(\cdot|s_i) \neq \Gamma(\cdot)$, *and* $m_i^z(e_i, s_i) = \emptyset$ *otherwise.*

Note that given the quadratic payoff structure, the posterior estimation $\mathbb{E}[\theta_i|s_i]$ is the only part of agents' private information that is relevant for the subsequent action choices. In Appendix B.1, we further show that full disclosure is essentially the unique equilibrium outcome at the communication stage (see Proposition B.1). Taken together, our full-revelation results suggest that with verifiable information, the allocation of decision rights does not affect the quality of communication in the organization.

Intuitively, Proposition 1 follows from the incentives to misrepresent the state: each agent i would prefer other decision makers (agent j or the principal) to overestimate the *absolute value* of θ_i, so that *on average* they would more aggressively coordinate the other action toward θ_i. More specifically, consider the case of decentralization and, without loss of generality, an agent i who observes a signal s_i, estimates $\mathbb{E}[\theta_i|s_i] \geq 0$, and considers a deviation from the fully revealing strategy m_i^d. As common in disclosure games, 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 $\underline{s}^{m_i}$, 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 (which is well-defined given Assumption 2). This implies that by deviating to any message $m_i \neq m^{\mathbb{E}[\theta_i|s_i]}$, i could only make j believe that i's own estimation about θ_i is lower than $\mathbb{E}[\theta_i|s_i]$.

Now imagine, for the sake of the argument, that agent i knows that j has received a signal s_j with $\mathbb{E}[\theta_j|s_j] \leq \mathbb{E}[\theta_i|s_i]$. 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 can only 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 with $\mathbb{E}[\theta_j|s_j] > \mathbb{E}[\theta_i|s_i]$, deceiving j to underestimate the value of θ_i could be beneficial for agent i, since it might 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. Nevertheless, since $\mathbb{E}[\theta_j] = 0 \leq \mathbb{E}[\theta_i|s_i]$, agent i does know that either or both of the followings must hold: (i) a priori $\mathbb{E}[\theta_j|s_j] \leq \mathbb{E}[\theta_i|s_i]$ is a more likely scenario compared to $\mathbb{E}[\theta_j|s_j] > \mathbb{E}[\theta_i|s_i]$; (ii) the ex ante distribution of agent j's signal assigns a substantial weight to the events in which $\mathbb{E}[\theta_j|s_j]$ is far smaller than $\mathbb{E}[\theta_i|s_i]$. Thus, truthful disclosure minimizes the average distance between agent j's action and θ_i, so is optimal for agent i. The intuition for the case of centralization is similar.

Given that private signals are truthfully revealed in equilibrium, the decision rules (of the agents under decentralization, and of the principal under centralization) are uniquely pinned down on the equilibrium path. In particular, it is straightforward to show that the equilibrium actions of the agents under decentralization can be written as functions of the private signals $\mathbf{s} = (s_1, s_2)$ only:

$$y_i^d(s_1, s_2) = \frac{\tau^d}{1+\tau^d} \cdot \frac{\mathbb{E}[\theta_1|s_1] + \mathbb{E}[\theta_2|s_2]}{2} + \frac{1}{1+\tau^d} \cdot \mathbb{E}[\theta_i|s_i], \ \forall i = 1, 2, \tag{3.1}$$

where $\tau^d = 2\delta$. Analogously, the principal's equilibrium action choices under centralization are characterized by the following mappings:

$$y_i^c(s_1, s_2, \lambda) = \frac{\tau^c}{1+\tau^c} \cdot (\lambda\mathbb{E}[\theta_1|s_1] + (1-\lambda)\mathbb{E}[\theta_2|s_2]) + \frac{1}{1+\tau^c} \cdot \mathbb{E}[\theta_i|s_i], \ \forall i = 1, 2, \tag{3.2}$$

where $\tau^c = \delta/(\lambda(1-\lambda))$. In both cases, the equilibrium actions are weighted sums of (i) a "common action" that would have been chosen if coordination is all important, and (ii) a deviation from that common action. The decision weights are chosen to minimize the decision-makers' expected losses from mis-coordination and mis-adaptation, given *all* acquired information. Since $\tau^d < \tau^c \ \forall \delta > 0$, it is clear that autonomous agents underestimate the need for coordination from the principal's perspective. Nevertheless, as $\delta \to +\infty$, both organizational structures converge to perfect coordination.

Remark. The truthful disclosure result (Proposition 1) continues to hold when the local states do not have the same variance. In Appendix Section B.2, we show that the result can also be extended to settings where the local states are correlated. What is really critical for our sharp result is the assumption that the local states are equal in expectation ($\mathbb{E}[\theta_i] = \mathbb{E}[\theta_j]$), meaning that the interests of the players are ex ante aligned. If this assumption does not hold, then for some signal realizations the agents may prefer not to fully reveal their types, and the size of the pooling regime will depend on the degree of asymmetry. We chose to focus on the case with symmetric means because it allows us to clearly identify the impact of organizational structure on information acquisition, which we view as a more substantial contribution of our paper to the literature.

3.2 Information acquisition

Taking the equilibrium decision rules (3.1) (which are independent of effort choices) and agent j's effort e_j as given, agent i then solves the following optimization problem at the information acquisition stage under decentralization:

$$\max_{e_i \in [0,1]} U_i^d(e_i, e_j) = \mathbb{E}_{\boldsymbol{\theta}} \left[\mathbb{E}_{\mathbf{s}} \left[u_i \left(y_1^d(\mathbf{s}), y_2^d(\mathbf{s}), \theta_i, e_i \right] |e_i, e_j \right] \right]. \tag{3.3}$$

Similarly, by taking the principal's decision rules (3.2) and agent j's effort choice e_j as given, agent i's would solve the following optimization problem at the information acquisition stage under centralization:

$$\max_{e_i \in [0,1]} U_i^c(e_i, e_j) = \mathbb{E}_{\boldsymbol{\theta}} \left[\mathbb{E}_{\mathbf{s}} \left[\mathbb{E}_\lambda \left[u_i(y_1^c(\mathbf{s}, \lambda), y_2^c(\mathbf{s}, \lambda), \theta_i, e_i) \right] |e_i, e_j \right] \right]. \tag{3.4}$$

It turns out that both (3.3) and (3.4) admit a unique solution, which is independent of the effort choice of agent j. This leads to the following characterization of the equilibrium outcome at the information acquisition stage.

Proposition 2. *The equilibrium outcome at the information acquisition stage is unique under both organizational structures. Specifically, under decentralization, both agents exert the effort*

$$e^d \equiv \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'} \right)^{-1} \left(1 - \frac{\delta^2 + \delta}{(1 + 2\delta)^2} \right).$$

Under centralization, both agents exert the effort

$$e_F^c \equiv \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'} \right)^{-1} \left(1 - \mathbb{E}_\lambda \left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda) + \delta)^2} \right] \right).$$

The equilibrium efforts are pinned down by equating the marginal benefit and the marginal cost of information under different organizational structures. As we formally show in Appendix B.3, both e^d and e^c_F are decreasing in δ. This is intuitive: as the relative importance of coordination increases, tailoring actions to local states becomes less important from the perspectives of all players. Hence, the value of information also decreases. It is less clear, however, how the effort level e^c_F under centralization depends on the distribution of λ. We investigate this question in the next section as we compare the effort provision under both organizational forms.

4 Comparing Organizational Structures

Following the analysis of the communication outcomes, we now study 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. Hence, comparing agents' efforts under centralization and decentralization provides a useful 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, we will start by analyzing the relative performance of the organization in terms of effort provision in Section 4.1. The analysis of the principal's expected payoff will be deferred to Section 4.2.

4.1 Effort provision

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

$$D(\delta) := \frac{\delta^2 + \delta}{(1+2\delta)^2} < C_F(\delta) := \mathbb{E}\left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]. \tag{4.1}$$

Exploiting the limiting properties of the functions $D(\delta)$ and $C_F(\delta)$ in (4.1), our first theorem below shows that a decentralized organization outperforms its centralized counterpart in terms of incentivizing effort provision (or information gathering) whenever there is some uncertainty in the relative weight λ and coordination is sufficiently important.

Theorem 1. *Consider the fully revealing equilibria under centralization and decentralization.*

(i) If the operating divisions are always equally important to the principal (i.e., $\lambda = 0.5$ for sure), then $e^d < e^c_F$ $\forall \delta > 0$.

(ii) If the operating divisions are not always equally important to the principal (i.e., the distribution of λ is symmetric but not deterministic), then $\exists\,\bar{\delta} \in [0, +\infty)$, such that $e^d > e^c_F\ \forall \delta > \bar{\delta}$. In addition, the difference $e^d - e^c_F$ is increasing in $\delta\ \forall \delta > \bar{\delta}$.

To understand the intuition, notice that because the principal's payoff is a weighted sum of the divisions' performances, an agent's preferences will be more aligned with the principal's than the other agent's. This implies that allocating the decision-making power to the principal can mitigate the following externality problem: Whenever an agent's action is adjusted to better match the local state, it reduces the payoff of the other agent as the coordination requirement becomes more stringent. Hence, centralization can benefit the agents by restricting their deviations from a common action, which in turns has a positive impact on the marginal value of information. In particular, if the principal is balanced, she would always target the common action that maximizes the joint profits of both divisions (i.e., the average local state), making centralization superior in motivating information acquisition. As a result of the increase in information quality, centralization may even improve adaptation on average.

However, in instances in which the principal could treat the two divisions unequally, centralization also has a negative effect on the agents' incentives: As the relative importance of the operating divisions varies, the common action in the principal's decision rule becomes uncertain. Given that the loss functions of the agents are convex, such uncertainty in principal's decision-making process creates an asymmetry that is painful for the agents: a shift going in the direction against an agent's ideal action hurts him more than the extra gain from the same shift towards his ideal action. Thus, the value of information is impaired even though the principal's decisions remain balanced on average.

The above analysis implies that determining which organizational structure can better incentivize effort provision amounts to comparing the motivational damages due to the *externality effect* under decentralization and to the *uncertainty effect* under centralization. Crucially, the former effect vanishes as the need for coordination is sufficiently large, because in that case the agents themselves would already make sure that their decisions are always well-coordinated. In contrast, the latter effect under centralization is present even when coordination is all important: As the principal still needs to choose which action to coordinate on, the uncertainty in how much weight she would assign to each agent remain relevant for the value of information. Thus, by continuity, a positive gap in the effort provision between decentralization and centralization necessarily emerges whenever coordination is sufficiently important.

One may wonder whether the converse of Theorem 1(ii) 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 the cutoff value $\bar{\delta}$ in Theorem 1(ii) is strictly positive. 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 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 (i.e., the regime where the agents' interests are most conflicted) 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 which division's performance is more important to the principal is highly uncertain ex ante.

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

Note that the condition in Theorem 2 is likely to be satisfied if the decision weights that the principal assigns to the two divisions tend to be extreme (i.e, λ is likely to take values that are close to 0 and 1). Intuitively, *both* the externality effect associated with decentralization *and* the uncertainty effect associated with centralization vanish when the need for coordination becomes sufficiently small, because the decisions of divisions are no longer interdependent in the limit. Nevertheless, the *speed* at which the uncertainty effect vanishes is slower when the relative importance of the divisions is more volatile. Consequently, when the coordination motive is weak, whether decentralization can better motivate information gathering primarily depends on the degree of uncertainty in the principal's decision-making process under centralization.

4.1.1 Effort provision with binary distributions

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

$$\Pr(\eta_1 = 1 + \omega, \eta_2 = 1 - \omega) = \Pr(\eta_1 = 1 - \omega, \eta_2 = 1 + \omega) = \frac{1}{2}. \tag{4.2}$$

This uniquely maps into a binary distribution of the relative weight λ, where λ takes the values of $(1+\omega)/2$ and $(1-\omega)/2$ with equal probabilities, which is all that matters for the players' incentives in our model. Thus, the severity of the shocks ω can be interpreted as a measure of both the volatility of the global states and how unequally the principal would treat the two divisions ex post. In particular, the larger ω, the more volatile are the global states (since $\mathbb{E}[(\eta_i - \mathbb{E}[\eta_i])^2] = \omega^2$ and $Cov(\eta_1, \eta_2) = -\omega^2$) and the more important is one division relative to the other from the principal's perspective (as $|(\lambda - (1-\lambda)| = \omega$).

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

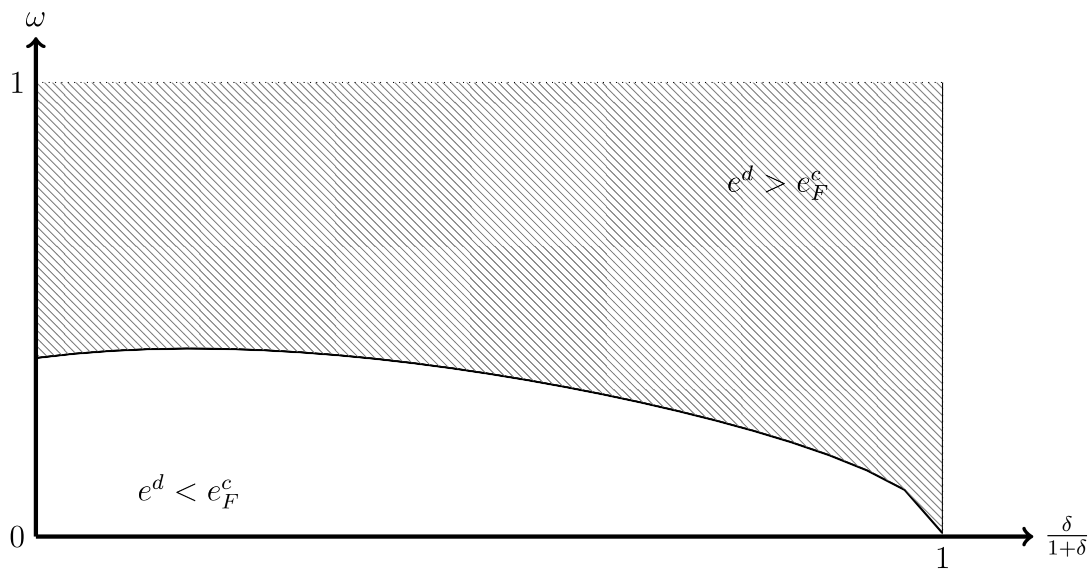


Figure 2: The parametric regimes for $e^d > e_F^c$ and $e^d < e_F^c$.

exerting costly yet valuable effort. Fixing the volatility of the global states, or the relative importance of the two operating divisions ex post, the next result shows how this regime is shaped by the importance of promoting synergies in the organization.

Proposition 3. *For the class of binary distributions* (4.2)*, there exist* $\underline{\omega}, \bar{\omega} \in (0, 1)$ *such that:*

(i) If $\omega = 0$*, then* $e^d < e_F^c$ $\forall \delta > 0$.

(ii) If $0 < \omega \leq \underline{\omega}$*, then there exists* $\bar{\delta} > 0$ *such that* $e^d > e_F^c$ *if and only if* $\delta > \bar{\delta}$.

(iii) If $\underline{\omega} < \omega \leq \bar{\omega}$*, then there exist* $\underline{\delta}, \bar{\delta} > 0$ *such that* $e^d > e_F^c$ *if and only if* $\delta \in (0, \underline{\delta}) \cup (\bar{\delta}, +\infty)$.

(iv) If $\omega > \bar{\omega}$*, then* $e^d > e_F^c$ $\forall \delta > 0$.

Figure 2 provides a graphical illustration of Proposition 3 as well as the key messages of Theorems 1 and 2, which does not require any specification of the information acquisition technology and the effort cost function. Most notably, the figure shows that as δ increases, the cutoff of ω that determines the effort-maximizing organizational structure first increases and then decreases, and it converges to zero as $\delta \to +\infty$. Hence, except for the benchmark case $\omega = 0$, we necessarily enter the parametric regime with $e^d > e_F^c$ when δ is sufficiently large, confirming the main prediction of Theorem 1. Moreover, as the degree of uncertainty ω surpasses a cutoff $\underline{\omega}$, we will also find ourselves in the regime with $e^d > e_F^c$ whenever δ is sufficiently small, which is consistent with the prediction of Theorem 2. Finally, we can see from the figure that all pairs (ω, δ) with ω larger than some cutoff $\bar{\omega}$ are contained in the dashed area, meaning that a sufficiently large degree of uncertainty can make decentralization optimal for incentivizing effort provision regardless of the importance of coordination.

4.2 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 result (Proposition 1) 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) favored division without sacrificing the (ex ante) informativeness of the decisions. As suggested by the previous analysis, 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 relative importance of the operating divisions does not vary a lot.

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 decision weights that the principal assigns to different divisions are highly uncertain. If the disadvantage of centralization in motivating information gathering is substantial enough, having the flexibility to adapt decisions to the realized objective may not be so valuable for the principal after all. 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.

Theorem 3. *Suppose that the operating divisions are not always equally important to the principal (i.e., the distribution of λ is symmetric but not deterministic). There exists $\zeta > 0$, such that if*

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}(x)\right)}{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}(x')\right)} > \left(\frac{x}{x'}\right)^{\zeta} \quad \forall x, x' \in (0,1) \text{ with } x > x', \text{ where } g(\cdot) = \frac{c'(\cdot)}{{\sigma^2_{\mathbb{E}[\theta|s]}}'(\cdot)}, \tag{4.3}$$

then we have $\Pi^c_P < \Pi^d_P$ for sufficiently large δ.

In essence, the sufficient condition (4.3) requires that a small increase in the marginal benefit of information can result in a large increase in the quality of signal acquired in equilibrium. To understand the economic intuition, note that by increasing his effort marginally, an agent would incur a direct cost $c'(e)$ and increases the informativeness of his signal by ${\sigma^2_{\mathbb{E}[\theta|s]}}'(e)$. As Theorem 1 shows, if the divisions exhibit any uncertainty in their relative values to the principal and coordination is sufficiently important, the marginal benefit of a further increase in the informativeness of the signal is higher when decision rights are allocated to the agents. The gap in the marginal benefits of information between centralization and decentralization then translates into a gap in effort provision. Therefore, from the principal's perspective, the increase in signal quality due to decentralization will be substantial enough to outweigh the benefit of centralized decision-making if (i) the ultimate gap in effort provision between the two organizational structures is large, which will be the case when the marginal cost $c'(\cdot)$ does not increase very fast relative to the marginal quality change of the signal ${\sigma^2_{\mathbb{E}[\theta|s]}}'(\cdot)$, as the equilibrium effort levels are chosen to balance these two marginal effects; and (ii) the informativeness of signal $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ is very sensitive to effort.

To make the idea more concrete, one may further consider a parametric example with $c(e) = \kappa e^\alpha$ and $\sigma^2_{\mathbb{E}[\theta|s]}(e) = e\sigma^2_\theta$, where $\kappa > 0$ and $\alpha > 1$.[13] With some algebra, we can show that condition (4.3) is equivalent to $1/(\alpha - 1) > \zeta$, which, for any given ζ, is satisfied if α is sufficiently close to one. In other words, provided that the effort cost is not too convex, Theorem 3 holds and decentralization is optimal for the principal when coordination motive is sufficiently large. However, the proof of Theorem 3 shows that what is crucial is not the convexity of the cost function, but rather the bound on the speed at which the marginal cost grows relative to the change of signal quality with respect to agent's effort.

We close this section with a result that parallels Theorem 2, which provides sufficient conditions under which decentralization is optimal for the principal when the need of coordination is relatively small.

Theorem 4. *Suppose that* $\mathbb{E}\left[\left(\frac{\lambda}{1-\lambda} - 1\right)^2\right] > 1$. *For sufficiently small* $\delta > 0$, *there exists* $\zeta(\delta) > 0$, *such that if*

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}(x)\right)}{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}(x')\right)} > \left(\frac{x}{x'}\right)^{\zeta(\delta)} \quad \forall x, x' \in (0,1) \text{ with } x > x', \text{ where } g(\cdot) = \frac{c'(\cdot)}{{\sigma^2_{\mathbb{E}[\theta|s]}}'(\cdot)}, \tag{4.4}$$

then $\Pi^c_P < \Pi^d_P$.

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 the expected payoffs of the principal under both authority structures (i.e., Π^c_P and Π^d_P) always converge to each other as δ goes to zero. A deeper insight we can gain from Theorem 2 and its proof is that the optimal authority structure is more ambiguous when coordination is not so important *and* the principal is highly uncertain about her ex post realized objective. Intuitively, if the principal's objective is sufficiently volatile, the power of making contingent decisions is especially valuable to her.[14]

[13]This functional form of $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ can be micro-founded by assuming that information is acquired via the "success-enhancing effort" technology introduced by Green and Stokey (1981): Given the effort level $e_i \in [0,1]$, agent i receives a perfectly revealing signal $s_i = \theta_i$ with probability e_i. With probability $1 - e_i$, the agent receives a null signal $s_i = \emptyset$. In this case, the set of available information structures is Blackwell-ordered.

[14]It would be interesting to study to what extent Theorems 3 and 4 are robust if each agent intrinsically cares about his *actual* contribution to the principal's payoff, i.e., $\eta_i \pi_i(\mathbf{y}, \theta_i)$. In this case, the analysis will be more involved because the absolute levels of (η_1, η_2) will also matter and they can be arbitrarily correlated. However, we note that if each agent i only observes η_i but not η_j, the uncertainty effect under centralization should still be present as he cannot be sure about the principal's decision criterion at the interim stage.

4.2.1 Optimal organizational structure with binary distributions

To sharpen our understanding on the roles of the volatility in the principal's objective and information costs in determining the relative expected payoff of the principal under centralization and decentralization, we consider again the class of binary distributions $\{F(\cdot|\omega)\}_{\omega\in[0,1]}$ introduced in Section 4.1.1. In addition, we assume, for analytical tractability, that the functions of information cost and quality are $c(e) = \kappa e^{\alpha}$ and $\sigma^2_{\mathbb{E}[\theta|s]}(e) = e\sigma^2_{\theta}$, respectively, where $\alpha > 1$ and $\kappa > 0$ are chosen in a way that the equilibrium effort levels are ensured to be interior. Given the specification, we obtain the following result: provided that the effort cost function is not too convex, decentralization is optimal for the principal if either (i) the need for synchronizing the activities across divisions is sufficiently large, or (ii) the relative importance of the divisions to the principal is sufficiently uncertain.

Proposition 4. *Consider the class of binary distributions $\{F(\cdot|\omega)\}_{\omega\in[0,1]}$ and assume that $\sigma^2_{\mathbb{E}[\theta|s]}(e) = e\sigma^2_{\theta}$ and $c(e) = \kappa e^{\alpha}$, where $\kappa > 0$ and $\alpha > 1$.*

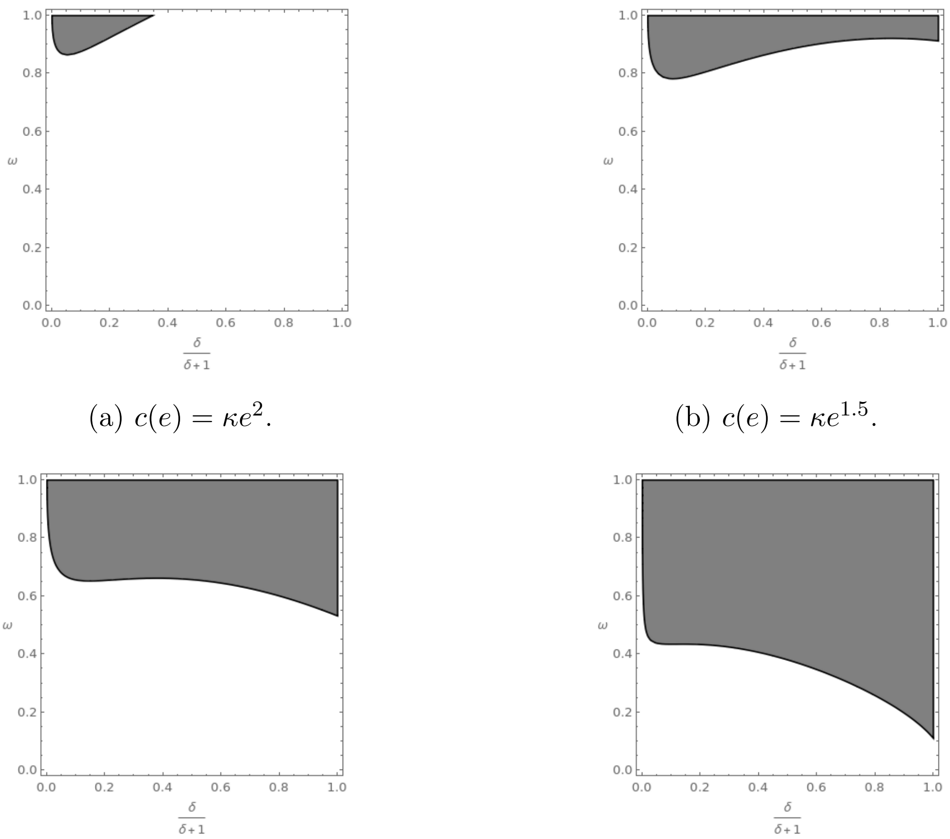
(i) For every $\delta > 0$, there exists $\check{\alpha}(\delta) > 1$, such that if $\alpha < \check{\alpha}(\delta)$, then $\Pi^d_P > \Pi^c_P$ whenever ω is sufficiently large.

(ii) For every $\omega > 0$, there exists $\hat{\alpha}(\omega) > 1$, such that if $\alpha < \hat{\alpha}(\omega)$, then $\Pi^d_P > \Pi^c_P$ whenever δ is sufficiently large.

To better illustrate our findings, we further consider four specific cost functions with different degrees of convexity ($\alpha = 2, 1.5, 1.2$ and 1.01), and then use numerical method to fully characterize the optimal organizational structure. For each of these four cases, Figures 3 (a), (b), (c) and (d) depict the corresponding parametric regimes of (δ, ω) where decentralization is optimal for the principal (the gray area). Compared to Figure 2, Figure 3 reveals that the evaluation of organizational performance is more subtle when we consider the overall payoff of the principal, rather than just agents' effort provision. In particular, with arbitrary information cost, neither a large coordination motive nor a large uncertainty in the principal's decision rule can guarantee that decentralization is optimal for the principal. This reflects the idea that under centralization, the principal has the ability to respond to the fluctuation of her decision rule, even if information can be better in the absence of that responsiveness. In fact, both the (ex ante) demotivating effect and the (ex post) value of flexibility of the responsiveness under centralization can be increasing in the degree of global volatility.[15]

Nevertheless, as a general implication from Theorems 3 and 4 and Proposition 4, the scope of decentralization being optimal should be larger for less convex cost functions. Indeed, the

[15]Given the assumptions of Proposition 4, it is easy to show that (i) the difference of equilibrium effort levels $e^d - e^c_F$ is increasing in ω; and (ii) with the same effort level, the principal's payoff difference between centralization and decentralization $\Pi^c_P - \Pi^d_P$ is also increasing in ω.



(a) $c(e) = \kappa e^2$.

(b) $c(e) = \kappa e^{1.5}$.

(c) $c(e) = \kappa e^{1.2}$.

(d) $c(e) = \kappa e^{1.01}$.

Figure 3: Optimal organizational structure with binomially distributed λ, $\sigma^2_{\mathbb{E}[\theta|s]}(e) = e\sigma^2_\theta$, and different cost functions $c(\cdot)$. The gray area indicates the parametric regime of (δ, ω) where decentralization is optimal for the principal (i.e., $\Pi^d_P - \Pi^c_P > 0$).

gray area in Figure 3 expands as the cost function becomes less and less convex (i.e., from the quadratic one in case (a) to the almost linear one in case (d)). Moreover, in all cases except for (a) in the figure, decentralization is optimal for the principal whenever she is sufficiently uncertain about the relative importance of the divisions. This validates the prediction of Proposition 4(i). In fact, the figure shows that when degree of uncertainty is sufficiently large, we may have decentralization as the optimal organizational structure regardless of the importance of coordination. Finally, the irregular shape of the gray area in the figure clearly shows that the relative performance of the two organizational structures is not necessarily monotone in the coordination requirement. Nevertheless, consistent with the insight formalized by Proposition 4(ii), a large coordination motive generally nudges the principal to have the decisions independently made by the agents when information cost is not too convex (e.g., cases (b) - (d) in the figure).

5 Extensions

5.1 Alternative payoff specification

In Appendix Section B.6, we study an alternative specification of agents' performances:

$$\tilde{\pi}_i(\mathbf{y}, \theta_i) = K - (1-q)(y_i - \theta_i)^2 - q(y_i - y_j)^2, \tag{5.1}$$

where $q := \delta/(1+\delta) \in (0,1)$ measures the the relative importance of adaptation and coordination. Compared to our main specification, a key difference here is that adaptation becomes completely irrelevant as $\delta \to +\infty$. We show that all the main findings of Theorems 1 - 4 continue to hold in this setup. First, when there is uncertainty in the relative importance of the operating divisions, the effort level is higher under decentralization if δ is sufficiently large (i.e., $q \to 1$). Second, if the uncertainty is substantial, decentralization also induces higher effort when δ is sufficiently small (i.e., $q \to 0$). In both of these two cases, the principal's expected payoff will also be higher under decentralization if the marginal cost of effort does not increase very fast relative to the associated change in signal precision.[16]

5.2 Asymmetric organizations

In the baseline model, we assumed that the global states are symmetrically distributed, meaning that the organizational management is ex ante balanced about the outcomes of different divisions. Naturally, this simplifying assumption does not always hold in reality as the principal can be systematically biased towards one of the divisions (e.g., due to a biased personal preference or asymmetric distributions of the variables that are relevant to the managerial priority). In our setting, such ex ante asymmetry can be captured by an asymmetric distribution of the relative weight λ. To illustrate the trade-offs, and the implications for information production in such asymmetric organizations, we consider a binary distribution of λ that takes values in the support $\{\frac{1+\omega+\upsilon}{2}, \frac{1-\omega+\upsilon}{2}\}$ with equal probabilities, where $\omega \in [0,1]$ and $\upsilon \in [0, 1-\omega]$ are known parameters. This distribution reduces to the symmetric one that we considered in Section 4.1.1 when $\upsilon = 0$. More generally, the degree of the organizational asymmetry is captured by the parameter υ: as its value increases, the weight that the principal would assign to the performance of division 1 becomes higher on average.

We first note that the full disclosure result continues to hold with arbitrary asymmetry in the distribution of the relative weight λ. The reason is that, as we showed in the proof of Proposition

[16] However, note that with the payoff specification (5.1), the gap in effort provision between decentralization and centralization necessarily converges to zero as $\delta \to +\infty$, because the value of information relative to its costs decreases to zero under both organizational structures. For the same reason, the ζ-cutoff for the relative speed of change between $c'(\cdot)$ and $\sigma^{2}_{\mathbb{E}[\theta|s]}{}'(\cdot)$ will be contingent on δ even when δ is sufficiently large, which is slightly different from our main setup.

1, truth-telling is incentive compatible for the agents even conditional on the *realization* of λ. Thus, truth-telling can be achieved under both centralization and decentralization with arbitrary distribution of λ, irrespective of whether it is symmetric or not.

As for the results on how information production is affected by the allocation of decision rights (in particular, Theorems 1 and 2), simple continuity arguments imply that they are robust whenever the degree of asymmetry v is sufficiently small. Below, we further show that the *aggregate* equilibrium effort under centralization can actually be uniformly decreasing in v. This is a strong robustness result, as it implies that whenever decentralization can induce more effort with a symmetric distribution of the global states, introducing asymmetry will only strengthen the motivational advantage of decentralization.

Proposition 5. *Suppose that the ratio of marginal cost and marginal quality change of signal, $g(\cdot) = c'(\cdot)/{\sigma^2_{\mathbb{E}[\theta|s]}}'(\cdot)$, is concave in effort. As the degree of asymmetry v increases, the sum of the agents' efforts under centralization decreases, and therefore the motivational advantage of decentralization increases.*

As we show in the proof, a higher degree of asymmetry increases (decreases) the value of information for agent 1 (agent 2), because his acquired signal will have more (less) influence on the principal's decision. As a result, agent 1's effort increases and agent 2's effort decreases in equilibrium. However, due to the convexity of the loss functions, the increase in the value of information for agent 1 will not be large enough to compensate the decrease in the value of information for agent 2. Under the additional assumption about the marginal cost/precision ratio (which is satisfied, for example, by the cost function $c(e) = e^{\alpha}$, where $1 < \alpha \leq 2$, and the information technology $\sigma^2_{\mathbb{E}[\theta|s]} = e\sigma^2_{\theta}$), the unbalanced changes in the value of information will further imply a decrease in the aggregate effort under centralization. We therefore conclude that the main insights from the Theorems 1 and 2 continue to hold in this setting. In particular, decentralization is likely to result in higher (aggregate) effort when either the coordination motive is strong, or the relative importance of divisions for the principal is highly uncertain.

We now briefly discuss the principal's expected payoffs under the two organizational forms. Naturally, the principal prefers to centralize if the aggregate effort under centralization is higher than under decentralization, and prefers to decentralize if the effort provision under decentralization is sufficiently larger relative to centralization. Proposition 5 implies that in a highly asymmetric organization, the effort provision by the (ex ante) less favored division will be very low, and it cannot be compensated through the effort provision by the other division. At the same time, an increase in organizational asymmetry also means that the principal increasingly disregards the outcome in the less favored division. These countervailing forces make the analysis of optimal organizational structure more nuanced. As an illustration,

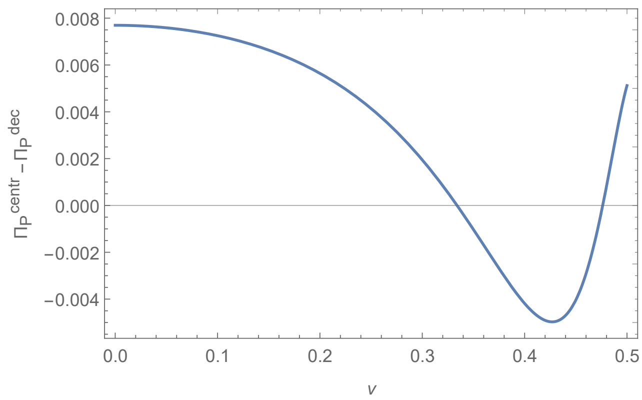


Figure 4: The principal's payoff difference between centralization and decentralization as a function of the degree of asymmetry v, with $c(e) = e^{1.1}$, $\sigma_\theta^2(e) = e$, $\delta = 0.05$ and $\omega = 0.5$.

Figure 4 provides an example where the principal's preference over organizational forms is non-monotonic in the degree of asymmetry. It shows, first, that with an intermediate level of global volatility and a low coordination motive, centralization is preferred for low levels of asymmetry. For intermediate levels of asymmetry, the principal switches to prefer decentralization (due to the substantial decrease in the aggregate effort under centralization). However, as the degree of asymmetry increases further, the principal eventually finds it optimal to centralize.[17]

5.3 Introducing transfers

In this section, we discuss some general implications of introducing transfers for the optimal choice of organizational structure. We focus on two prominent types of conditional transfers: pay for performance and pay for information.

So far, we have assumed that the agents care only about their own performance. This can be interpreted as that an agent only gets paid based on the performance of his own division. However, as Athey and Roberts (2001) and Rantakari (2013) point out, due to informational externalities the principal may want to align the incentives of the managerial members by tying their compensation to each other's performance. Indeed, 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 (from the principal's perspective) when decision rights are decentralized to the divisions. Under centralization, however, there is no room for such an improvement given that a central manager can elicit all information from the local managers for free. The implication of this analysis is that decentralized decision-making is even more likely

[17]It is conceivable that when the degree of asymmetry is substantial, the principal may find an asymmetric allocation of decision rights superior. While the formal analysis goes beyond the scope of this paper, we note here that if the principal wants to allocate all decision rights to a single agent, she would clearly choose the agent that she values ex ante more than the other.

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.

Since the central trade-off of our model comes from strategic information acquisition rather than strategic communication, one may also envision improving the organization's performance by directly rewarding information collection. As a simple illustration, consider the "all-or-nothing" signal structure and success-enhancing information acquisition technology with $\sigma^2_{\mathbb{E}[\theta|s]}(e) = e$ (see footnote 13). Suppose that the principal can commit to pay a fixed bonus $b \geq 0$ to an agent provided that he credibly discloses that his information experiment is successful ($s_i \neq \emptyset$). In general, allowing for such information-based transfers will make centralization more likely to be optimal. This is because, other things equal, additional effort will be more valuable for the principal when she can decide how to use the resulting information. Thus, in contrast to performance-based transfers, information-based transfers are complimentary to centralization. However, it is worth noting that ex ante it may be optimal for the principal *not* to provide any direct rewards for information collection (i.e., $b^* = 0$). For instance, under centralization, implementing an effort level $\tilde{e} > e^c_F$ would require the principal to choose the bonus $b(\tilde{e}) = c'(\tilde{e}) - (1 - C_F(\delta)) > 0$, where $C_F(\delta)$ is defined as in (4.1). Repeating the algebra used in the proof of Theorem 3, we show that the principal can strictly benefit from paying the information-based bonus if and only if

$$\Xi_F(\delta) \cdot (\tilde{e} - e^c_F) - 2\left(c'(\tilde{e}) - (1 - C_F(\delta))\right)\tilde{e} > 0, \tag{5.2}$$

where $\Xi_F(\delta) := 1 - \mathbb{E}\left[\frac{\lambda(1-\lambda)\delta}{\lambda(1-\lambda)+\delta}\right]$. Therefore, other things equal, (5.2) is more likely to be violated if the term Ξ_F is small or if $c'(\tilde{e})$ is large. In particular, if $\Xi_F(\delta)$ is sufficiently small, then (5.2) will not hold for any $\tilde{e} > e^c_F$.[18] In those cases, it would be optimal for the principal to choose $b = 0$ under centralization.

6 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 upon the framework developed by Alonso et al. (2008) and Rantakari (2008), which are among the first

[18]For example, this is the case when the cost function is $c(e) = \kappa e^\alpha$ with $\alpha > 1.5$, and the distribution of λ is sufficiently concentrated around its mean 0.5 (so that $\Xi_F(\delta) < 2c''(\tilde{e})\tilde{e} = 2(\alpha - 1)c'(\tilde{e})\ \forall \tilde{e} > e^c_F$).

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).[19] 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 centralization need not be optimal even when coordination is extremely important (Alonso et al., 2008).[20] Our model departs from theirs mainly by (1) focusing on the case where information is "hard" (Grossman, 1981; Milgrom, 1981), and (2) relaxing the (implicit) assumption that the decision rule exhibits no uncertainty in the ex post treatment of local markets. More important than the modeling differences, we add to Alonso et al. (2008) and Rantakari (2008, 2013) by showing that it can be that decentralized decision-making is optimal precisely because the coordination motive is sufficiently strong. 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 each local manager cares only about the performance of his own division.

Within the literature on organizational design and coordinated adaptation, our paper is further related to Dessein, Garicano, and Gertner (2010), Friebel and Raith (2010), and Alonso, Dessein, and Matouschek (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

[19]Other theoretical works on cheap-talk communication in organizations include Deimen and Szalay (2019); Dessein (2002). For experimental studies, see Brandts and Cooper (2018); Evdokimov and Garfagnini (2019).

[20]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, 2019).

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. Nonetheless, 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.

We also contribute to the literature on delegation as an instrument to motivate information acquisition. The seminal work of Aghion and Tirole (1997) introduces an important trade-off between employee initiative and the loss of control. In their framework, an agent has to acquire decision-relevant information and has better incentives to do so when being able to formally control the decision. With multiple agents and partial coordination motives, the ability of an agent to influence the decisions is restricted by the optimal behavior of the other agents. In fact, in our multi-agent setting, absent the uncertainty in the principal's (interim) decision rule, the agents' incentives for information acquisition are always weaker under delegation. Nevertheless, we show that this pessimistic view of delegation need not hold once some uncertainty over the principal's decision rule is introduced. Thus, the driving force of the motivational advantage of delegation in our model is different from that in Aghion and Tirole (1997).

More recent contributions show that the incentive effect of delegation can be ambiguous if the communication between the principal and the agent is strategic.[21] 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 that under centralization whenever the latter fails to provide any evidence the former would make an adverse inference and take an unfavorable action.[22]

[21]Abstracting from strategic communication, the incentive view of delegation is also discussed by Rantakari (2012). He shows that formal delegation 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.

[22]A similar persuasive motive of information acquisition under centralization is also present in Newman and

The above papers analyze single-agent settings while we model multiple agents.[23] As an implication, we show that 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.

7 Conclusion

How should an organization optimally allocate decision-making authority to its managerial members, if there is uncertainty in the relative importance of the performances of its operating divisions? 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.

We have shown that if information is verifiable, the quality of communication may not be affected by where the decision-making authority is lodged in the organization. Further, since the principal of the organization can elicit all private information from its local managers, the fact that the principal is not well informed per se does not make centralized decision-making inferior. However, as our main contribution, we have also shown that the quality of endogenously acquired information crucially depends on the allocation of decision rights. In particular, if the relative value of each operating division's performance to the organization is uncertain, a large coordination motive can strongly discourage information gathering under centralization, which can make decentralization the optimal organizational form. Yet it is also worth noting that when the need for coordination is small or intermediate, centralized decision-making is often optimal for the principal because it grants flexibility in decision-making once the principal's objective is realized, 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).

We envisage 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

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.

[23] 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 Proposition 2).

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 relative importance of different divisions 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

Agastya, M., P. K. Bag, and I. Chakraborty (2014): "Communication and authority with a partially informed expert," *The RAND Journal of Economics*, 45, 176–197.

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.

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.

Bergemann, D. and J. Välimäki (2002): "Information acquisition and efficient mechanism design," *Econometrica*, 70, 1007–1033.

Bertomeu, J. and I. Marinovic (2015): "A theory of hard and soft information," *The Accounting Review*, 91, 1–20.

Blackwell, D. (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, 24, 265–272.

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

Chandler, A. D. (1962): *Strategy and Structure: Chapters in the History of the American Enterprise*, MIT Press.

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

Cheung, Y.-W. and R. Sengupta (2013): "Impact of exchange rate movements on exports: an analysis of Indian non-financial sector firms," *Journal of International Money and Finance*, 39, 231–245.

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

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.

Dewatripont, M., I. Jewitt, and J. Tirole (1999): "The economics of career concerns, part I: Comparing information structures," *The Review of Economic Studies*, 66, 183–198.

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

Emons, W. and C. Fluet (2020): "Adversarial versus inquisitorial testimony," *Revue Economique*, 71, 429–457.

Evdokimov, P. and U. Garfagnini (2019): "Communication and behavior in organizations: An experiment," *Quantitative Economics*, 10, 775–801.

Freeland, R. F. (1996): "The myth of the M-form? Governance, consent, and organizational change," *American Journal of Sociology*, 102, 483–526.

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.

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.

Héricourt, J. and C. Nedoncelle (2018): "Multi-destination firms and the impact of exchange-rate risk on trade," *Journal of Comparative Economics*, 46, 1178–1193.

Héricourt, J. and S. Poncet (2015): "Exchange rate volatility, financial constraints, and trade: Empirical evidence from Chinese firms," *The World Bank Economic Review*, 29, 550–578.

Jackall, R. (1988): "Moral mazes: The world of corporate managers," *International Journal of Politics, Culture, and Society*, 1, 598–614.

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.

Lehmann, E. L. (1988): "Comparing location experiments," *Annals of Statistics*, 16, 521–533.

Letina, I., S. Liu, and N. Netzer (2020): "Delegating performance evaluation," *Theoretical Economics*, 15, 477–509.

Liu, S. and D. Migrow (2019): "Designing organizations in volatile markets," University of Zurich, Department of Economics, Working Paper No. 319.

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.

Mintzberg, H. (1983): *Power In and Around Organisations*, NJ: Englewood Cliffs.

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

Persico, N. (2000): "Information acquisition in auctions," *Econometrica*, 68, 135–148.

Pfeffer, J. (1981): *Power in Organizations*, Massachusetts: Pitman.

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

Rae, J. B. (1958): "The fabulous Billy Durant," *Business History Review*, 32, 255–271.

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 R. Gibbons (2013): *The Handbook of Organizational Economics*, Princeton University Press.

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

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

Stevenson, H. H., J. L. Cruikshank, and M. C. Moldoveanu (1998): *Do Lunch or Be Lunch: The Power of Predictability in Creating Your Future*, Harvard Business Press.

Stevenson, H. H. and M. C. Moldoveanu (1995): "The power of predictability." *Harvard Business Review*, 73, 140–143.

Tirole, J. (1986): "Hierarchies and bureaucracies: On the role of collusion in organizations," *Journal of Law, Economics, and Organization*, 2, 181.

Udell, G. F. (2009): "Financial innovation, organizations, and small business lending," in *The Changing Geography of Banking and Finance*, ed. by P. Alessandrini, M. Fratianni, and A. Zazzaro, Springer, 15–26.

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 (2019): "Communication and the Optimality of Hierarchy in Organizations," *Journal of Law, Economics, and Organization*, 35, 154–191.

A Proofs of Main Results

A.1 Proof of Proposition 1

Decentralization. First, consider agent i's incentive at the decision-making stage under decentralization. Taking (e_i, s_i, m_i, m_j) as given, agent i solves:

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

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

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

Solving the best response functions through repeated substitution, we obtain the following decision rules which must be satisfied in any equilibrium:

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

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. Since efforts (e_i, e_j) only affect the agents' decisions though the distributions of signals, we omit them from the expressions whenever it does not create confusion.

Now suppose that agent i anticipates that agent j will exert some arbitrary effort $e_j \in [0, 1]$, communicate his finding 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.

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

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

where the last equality follows that s_i and s_j are independent, $\mathbb{E}[\theta_j|m_j^d(s_j)] = \mathbb{E}[\theta_j|s_j]$ and $\mathbb{E}\left[\mathbb{E}[\theta_j|s_j]\right] = \mathbb{E}[\theta_j] = 0$.

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

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

In sum, for every $(s_i, m_i) \in \mathcal{S} \times \mathcal{M}(s_i)$ the (interim) expected payoff of agent i (subtracting the sunk information cost) is given by

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

It is straightforward to check that the above interim expected payoff has increasing differences in $(\mathbb{E}[\theta_i|s_i], \mathbb{E}[\theta_i|m_i])$, so the *asyclic masquerade property* in Hagenbach, Koessler, and Perez-Richet (2014) is satisfied (see Theorem 2, p. 1107). In addition, Assumption 1 implies that our communication game admits an *evidence base* as defined in their paper. Hence, we can invoke the general result of Hagenbach et al. (2014) (Theorem 1, p. 1103) and conclude that there exists a fully revealing equilibrium. In equilibrium, the agents' posterior beliefs are extremal: for every $m_i \in \mathcal{M}$, agent j assign probability one to that agent i's type is $\underline{s}^{m_i} \in \arg\min_{s_i \in \mathcal{S}^{m_i}} \left|\mathbb{E}[\theta_i|s_i]\right|$ (which exists by assumption), i.e., $\mu_j^i(\{\underline{s}^{m_i}|m_i\}) = 1$.

To complete the construction of the 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

$$
\begin{aligned}
U_i^d(e_i, e_j) =& K - \mathbb{E}\left[\left(\frac{(1+\delta)\mathbb{E}[\theta_i|s_i]}{1+2\delta} - \theta_i\right)^2 \middle| e_i\right] - \mathbb{E}\left[\left(\frac{\delta\mathbb{E}[\theta_j|s_j]}{1+2\delta}\right)^2 \middle| e_j\right] \\
& - \delta\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i]}{1+2\delta}\right)^2 \middle| e_i\right] - \delta\mathbb{E}\left[\left(\frac{\mathbb{E}\left[\theta_j|s_j\right]}{1+2\delta}\right)^2 \middle| e_j\right] - c(e_i) \\
=& K + \left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma^2_{\mathbb{E}[\theta|s]}(e_i) - \sigma_\theta^2 - \frac{\delta^2+\delta}{(1+2\delta)^2}\sigma^2_{\mathbb{E}[\theta|s]}(e_j) - c(e_i),
\end{aligned}
$$

where the second equality follows that

$$\mathbb{E}\left[\theta_i\mathbb{E}[\theta_i|s_i]|e_i\right] = \mathbb{E}\left[\mathbb{E}\left[\theta_i\mathbb{E}[\theta_i|s_i]|s_i\right]|e_i\right] = \mathbb{E}\left[\mathbb{E}\left[\theta_i|s_i\right] \cdot \mathbb{E}[\theta_i|s_i]|e_i\right] = \mathbb{E}\left[\left(\mathbb{E}[\theta_i|s_i]\right)^2|e_i\right].$$

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

$$\frac{\partial U_i^d(e_i, e_j)}{\partial e_i} = \left(1 - \frac{\delta^2 + \delta}{(1+2\delta)^2}\right) {\sigma^2_{\mathbb{E}[\theta|s]}}'(e_i) - c'(e_i) = 0. \tag{A.4}$$

Given our assumptions on the functions $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ and $c(\cdot)$ (see Section 2), (A.4) admits a unique interior solution $e_i^d \in (0,1)$, which is given by

$$e_i^d = e^d := \left(\frac{c'}{{\sigma^2_{\mathbb{E}[\theta|s]}}'}\right)^{-1} \left(1 - \frac{\delta^2 + \delta}{(1+2\delta)^2}\right), \tag{A.5}$$

where the inverse of the function

$$\left(\frac{c'}{{\sigma^2_{\mathbb{E}[\theta|s]}}'}\right)(e) = \frac{c'(e)}{{\sigma^2_{\mathbb{E}[\theta|s]}}'(e)}$$

is well-defined for all $e \in (0,1)$ because $c(\cdot)$ is strictly convex and $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ is weakly concave. The convexity conditions also imply that the function U_i^d will be strictly concave in e_i, and thus the solution $e_i = e^d$ is indeed the unique global maximizer of $U_i^d(e_i, e_j)$, $\forall e_j \in [0,1]$.

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. ■

Centralization. Consider the principal's incentive at the decision-making stage under centralization. Taking (m_1, m_2) and the realization relative weight λ as given, in the decision-making stage the principal solves:

$$\max_{y_1, y_2 \in \mathbb{R}} K - \delta(y_1 - y_2)^2 - \lambda \mathbb{E}\left[(y_1 - \theta_1)^2 | m_1\right] - (1-\lambda)\mathbb{E}\left[(y_2 - \theta_2)^2 | m_2\right].$$

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

$$-\delta(y_1 - y_2) - \lambda\left(y_1 - \mathbb{E}\left[\theta_1 | m_1\right]\right) = 0, \quad -\delta(y_2 - y_1) - (1-\lambda)\left(y_2 - \mathbb{E}\left[\theta_2 | m_2\right]\right) = 0.$$

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

$$y_1^c(\boldsymbol{m}, \lambda) = \frac{\lambda(1 - \lambda + \delta)\,\mathbb{E}\left[\theta_1 | m_1\right] + (1-\lambda)\delta\mathbb{E}[\theta_2 | m_2]}{\lambda(1-\lambda) + \delta}, \text{ and}$$

$$y_2^c(\boldsymbol{m}, \lambda) = \frac{(1-\lambda)(\lambda+\delta)\,\mathbb{E}[\theta_2|m_2] + \lambda\delta\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}, \tag{A.6}$$

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. Suppose that agent 1 anticipates that agent 2 will exert some arbitrary effort $e_2 \in [0,1]$ and communicate his finding 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.

Let $s_1 \in \mathcal{S}$ be the signal received by agent 1. For every message $m_1 \in \mathcal{M}(s_1)$ and every realized weight $\lambda \in [0,1]$ we have

$$\begin{aligned}
&EL_a^c(s_1, m_1, \lambda)\\
=&\mathbb{E}\left[\left(y_1^c(m_1, m_2^c(s_2), \lambda) - \theta_1\right)^2 \middle| s_1\right]\\
=&\mathbb{E}\left[\left(\frac{\lambda(1-\lambda+\delta)\,\mathbb{E}[\theta_1|m_1] + \delta(1-\lambda)\mathbb{E}[\theta_2|m_2^c(s_2)]}{\lambda(1-\lambda)+\delta} - \theta_1\right)^2 \middle| s_1\right]\\
=&\mathbb{E}\left[\left(\frac{(\lambda(1-\lambda)+\lambda\delta)\,\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta} - \theta_1\right)^2 \middle| s_1\right] + \mathbb{E}\left[\left(\frac{\delta(1-\lambda)\mathbb{E}[\theta_2|s_2]}{\lambda(1-\lambda)+\delta}\right)^2\right],
\end{aligned} \tag{A.7}$$

where the last equality follows that $\mathbb{E}\left[\mathbb{E}[\theta_2|m_2^c(s_2)]\right] = \mathbb{E}\left[\mathbb{E}[\theta_2|s_2]\right] = \mathbb{E}[\theta_2] = 0$. Similarly, for the expected loss of mis-coordination, we have for every $m_1 \in \mathcal{M}(s_1)$ and every $\lambda \in [0,1]$,

$$\begin{aligned}
EL_c^c(s_1, m_1, \lambda) =& \mathbb{E}[(y_1^c(m_1, m_2^c(s_2), \lambda) - y_2^c(m_1, m_2^c(s_2), \lambda))^2 | s_1]\\
=&\mathbb{E}\left[\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1] - \lambda(1-\lambda)\mathbb{E}[\theta_2|m_2^c(s_2)]}{\lambda(1-\lambda)+\delta}\right)^2\right]\\
=&\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_1|m_1]}{\lambda(1-\lambda)+\delta}\right)^2 + \mathbb{E}\left[\left(\frac{\lambda(1-\lambda)\mathbb{E}[\theta_2|s_2]}{\lambda(1-\lambda)+\delta}\right)^2\right].
\end{aligned} \tag{A.8}$$

where the last equality also follows that $\mathbb{E}\left[\mathbb{E}[\theta_2|m_2^c(s_2)]\right] = \mathbb{E}\left[\mathbb{E}[\theta_2|s_2]\right] = \mathbb{E}[\theta_2] = 0$.

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

$$\hat{\Pi}_1^c(s_1, m_1) = \mathbb{E}_\lambda\left[K - EL_a^c(s_1, m_1, \lambda) - \delta EL_c^c(s_1, m_1, \lambda)\right].$$

It is straightforward to check that the above interim expected payoff of agent 1 has increasing differences in $(\mathbb{E}[\theta_1|s_1], \mathbb{E}[\theta_1|m_1])$. One can show that the interim expected payoff of agent 2 takes a similar form, and it has increasing differences in $(\mathbb{E}[\theta_2|s_2], \mathbb{E}[\theta_2|m_2])$. Together with Assumption 1, this allows us to again invoke the result by Hagenbach et al. (2014), and conclude that for arbitrary *deterministic value of* λ, a fully revealing equilibrium exists under centralization. Therefore, a fully revealing equilibrium must also exist when agents only know that λ follows some arbitrary and possibly non-deterministic distribution (which is the case according to the timeline of our game). Analogous to the agents' beliefs under decentralization, the principal's equilibrium beliefs under centralization are extremal: for every $m_i \in \mathcal{M}$ she assigns probability one to that agent i's type is $\underline{s}^{m_i} \in \arg\min_{s_i \in \mathcal{S}^{m_i}} \left|\mathbb{E}[\theta_i|s_i]\right|$ (which exists by assumption), i.e., $\mu_p^i(\{\underline{s}^{m_i}|m_i\}) = 1$.

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

$$\begin{aligned}
&U_1^c(e_1, e_2) \\
&= K - \mathbb{E}\left[\left(\frac{\lambda\,(1-\lambda+\delta)\,\mathbb{E}\,[\theta_1|s_1] + (1-\lambda)\delta\mathbb{E}[\theta_2|s_2]}{\lambda(1-\lambda)+\delta} - \theta_1\right)^2 \middle| e_1, e_2\right] \\
&\quad - \delta\mathbb{E}\left[\left(\frac{\lambda(1-\lambda)\,(\mathbb{E}[\theta_1|s_1] - \mathbb{E}[\theta_1|s_2])}{\lambda(1-\lambda)+\delta}\right)^2 \middle| e_1, e_2\right] \\
&= K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\frac{(1-\lambda)^2\,(\delta^2 + \delta\lambda^2)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma^2_{\mathbb{E}[\theta|s]}(e_1) \\
&\quad - \left(\mathbb{E}\left[\frac{(1-\lambda)^2\,(\delta^2 + \delta\lambda^2)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma^2_{\mathbb{E}[\theta|s]}(e_2) - c(e_1).
\end{aligned}$$

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

$$\frac{\partial U_1^c(e_1, e_2)}{\partial e_1} = \left(1 - \mathbb{E}_\lambda\left[\frac{(1-\lambda)^2\,(\delta^2 + \delta\lambda^2)}{(\lambda(1-\lambda)+\delta)^2}\right]\right)\sigma^2_{\mathbb{E}[\theta|s]}{}'(e_1) - c'(e_1) = 0. \tag{A.9}$$

Given our assumptions on the functions $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ and $c(\cdot)$ (see Section 2), (A.9) admits a unique interior solution $e_i^c \in (0,1)$, which is given by

$$\begin{aligned}
e_1^c = e_F^c &:= \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}\left(1 - \mathbb{E}\left[\frac{(1-\lambda)^2\,(\delta^2 + \delta\lambda^2)}{(\lambda(1-\lambda)+\delta)^2}\right]\right) \\
&= \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}\left(1 - \mathbb{E}\left[\frac{\delta^2(\lambda^2 + (1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]\right),
\end{aligned} \tag{A.10}$$

where the last equality follows that λ is symmetrically distributed around 1/2. The convexity conditions of functions $c(\cdot)$ and $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ imply that the function U_1^c is strictly concave in e_1, and thus the solution $e_1 = e_F^c$ is indeed the unique global maximizer of $U_1^c(e_1, e_2)$, $\forall e_2 \in [0,1]$. By analogous arguments, one can show that choosing $e_2^c = e_F^c$ will maximize the expected payoff of agent 2 regardless of the effort choice of agent 1.

To conclude, together with the above constructed posterior beliefs $(\mu_p^1(\cdot), \mu_p^2(\cdot))$ of the principal, 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. ■□

A.2 Proof of Proposition 2

Combining Propositions 1 and B.1, we learn that full disclosure of private information is essentially the unique equilibrium prediction for the communication game under both organizational structures. Let $((e_1^*, m_1^*, y_1^*), (e_2^*, m_2^*, y_2^*))$ be an equilibrium under decentralization. Given the full disclosure result, we have $\mathbb{E}[\theta_i | m_i^*(e_i^*, s_i)] = \mathbb{E}[\theta_i | s_i, e_i^*]$ $\forall s_i \in \mathcal{S}$ 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 (3.1). As we have shown in Proposition 1, given the equilibrium decisions are taken according to $(y_1^d(e_1, \mathbf{s}), y_2^d(e_2, \mathbf{s}))$, e^d is the unique expected-payoff-maximizing effort level for both agents. Hence, we must have $e_i^* = e^d$ and $y_i^*(e_i^*, s_i, m_i^*(e_i^*, s_i), m_j^*(e_j^*, s_j)) = y_i^d(e_i^*, s_i, s_j)$, $\forall (s_i, s_j) \in \mathcal{S} \times \mathcal{S}$ and $\forall i = 1, 2$.

The proof for the case of centralization is analogous. □

A.3 Proof of Theorem 1

First, suppose that the distribution of λ is deterministic, so by symmetry, $\Pr\left(\lambda = \frac{1}{2}\right) = 1$. In this case, the RHS of condition (4.1) 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

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

which always holds. Therefore, when there is no uncertainty in λ, we have $D(\delta) > C_F(\delta)$ $\forall \delta > 0$, i.e., condition (4.1) is always violated. From the arguments in the main text, this immediately implies that $e^d < e_F^c$ $\forall \delta > 0$.

Next, consider the case where the distribution of λ is not deterministic. Taking the limit of both sides of (4.1) 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. 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 $e^d > e^c_F \iff D(\delta) < C_F(\delta)$, it immediately follows that $e^d > e^c_F\ \forall \delta > \bar{\delta}_1$.

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

$$\frac{\partial(e^d - e^c_F)}{\partial\delta} = \frac{-D'(\delta)}{g'\left(g^{-1}\left(1-D(\delta)\right)\right)} - \frac{-C'_F(\delta)}{g'\left(g^{-1}\left(1-C_F(\delta)\right)\right)},$$

where the function $g(\cdot)$ is defined by $g(e) = c'(e)/((\sigma^2_{\mathbb{E}[\theta|s]})'(e))\ \forall e \in [0,1]$.

Since function $c(\cdot)$ is strictly convex and function $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ is concave, 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

$$\frac{C'_F(\delta)}{D'(\delta)} > \frac{g'\left((g)^{-1}\left(1-C_F(\delta)\right)\right)}{g'\left((g)^{-1}\left(1-D(\delta)\right)\right)}. \tag{A.11}$$

For the RHS of (A.11), we have

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

Using the calculation results from Section A.5 (see (B.11) and (B.12)), 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.11) holds for all $\delta > \bar{\delta}_2$. Equivalently, the effort difference $e^d - e_F^c$ 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.4 Proof of Theorem 2

To simplify the algebra, let us define

$$\alpha \equiv \lambda(1-\lambda), \;\; \beta \equiv = \lambda^2 + (1-\lambda)^2 \tag{A.12}$$

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 (B.11) and (B.12), 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

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

where the third and the fifth equalities follow that λ is symmetrically distributed around 1/2. 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 $e^d > e_F^c$ $\forall \delta \in (0, \underline{\delta})$. □

A.5 Proof of Theorem 3

Using Proposition 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

$$\Pi_i^d(e^d, y_i^d, y_j^d) = K - \sigma_\theta^2 + \left(1 - \frac{2\delta^2 + 2\delta}{(1 + 2\delta)^2}\right) \sigma^2_{\mathbb{E}[\theta|s]}(e^d).$$

Exploiting that $\mathbb{E}[\lambda] = 0.5$ and the decision rules $\mathbf{y}^d = (y_1^d, y_2^d)$ are independent of λ, we then obtain the expected payoff of the principal under decentralization:

$$\Pi_P^d = \mathbb{E}\left[\lambda \Pi_1^d(e^d, \mathbf{y}^d) + (1 - \lambda)\Pi_2^d(e^d, \mathbf{y}^d)\right] = K - \sigma_\theta^2 + \left(1 - \frac{2\delta^2 + 2\delta}{(1 + 2\delta)^2}\right) \sigma^2_{\mathbb{E}[\theta|s]}(e^d).$$

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 (3.2). Hence, for each agent i and a given decision weight λ, the expected performance of the two divisions are

$$\Pi_1^c(e_F^c, y^c, \lambda) = K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\frac{(1 - \lambda)^2 \left(\delta^2 + \delta\lambda^2\right)}{(\lambda(1 - \lambda) + \delta)^2}\right]\right) \sigma^2_{\mathbb{E}[\theta|s]}(e_F^c),$$
$$\Pi_2^c(e_F^c, y^c, \lambda) = K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\frac{\lambda^2 \left(\delta^2 + \delta(1 - \lambda)^2\right)}{(\lambda(1 - \lambda) + \delta)^2}\right]\right) \sigma^2_{\mathbb{E}[\theta|s]}(e_F^c).$$

Exploiting the symmetry of the distribution of λ, we have

$$\begin{aligned}
\Pi_P^c =& \mathbb{E}\left[\lambda \Pi_1^c(e_F^c, y^c, \lambda) + (1 - \lambda)\Pi_2^c(e_F^c, y^c, \lambda)\right] \\
=& K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\lambda \cdot \frac{(1 - \lambda)^2 \left(\delta^2 + \delta\lambda^2\right)}{(\lambda(1 - \lambda) + \delta)^2}\right]\right. \\
&\left. - \mathbb{E}\left[(1 - \lambda) \cdot \frac{\lambda^2 \left(\delta^2 + \delta(1 - \lambda)^2\right)}{(\lambda(1 - \lambda) + \delta)^2}\right]\right) \sigma^2_{\mathbb{E}[\theta|s]}(e_F^c) \\
=& K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\frac{\lambda(1 - \lambda)\delta^2 + \lambda^2(1 - \lambda)^2\delta}{(\lambda(1 - \lambda) + \delta)^2}\right]\right) \sigma^2_{\mathbb{E}[\theta|s]}(e_F^c) \\
=& K - \sigma_\theta^2 + \left(1 - \mathbb{E}\left[\frac{\lambda(1 - \lambda)\delta}{\lambda(1 - \lambda) + \delta}\right]\right) \sigma^2_{\mathbb{E}[\theta|s]}(e_F^c).
\end{aligned}$$

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

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}(e^d)}{\sigma^2_{\mathbb{E}[\theta|s]}(e^c_F)} > R_F(\delta) \equiv \frac{1 - \mathbb{E}\left[\frac{\lambda(1-\lambda)\delta}{\lambda(1-\lambda)+\delta}\right]}{1 - \frac{2\delta^2+2\delta}{(1+2\delta)^2}}. \tag{A.13}$$

Note that

$$\lim_{\delta\to+\infty} R_F(\delta) = 2 - 2\mathbb{E}\left[\lambda(1-\lambda)\right].$$

Also, we have

$$\lim_{\delta\to+\infty} \frac{\sigma^2_{\mathbb{E}[\theta|s]}(e^d)}{\sigma^2_{\mathbb{E}[\theta|s]}(e^c_F)} = \lim_{\delta\to+\infty} \frac{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(1-D(\delta)))}{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(1-C_F(\delta)))} \geq \lim_{\delta\to+\infty} \left(\frac{1-D(\delta)}{1-C_F(\delta)}\right)^{\zeta}.$$

Since $\lim_{\delta\to+\infty} D(\delta) < \lim_{\delta\to+\infty} C_F(\delta)$, for sufficiently large $\zeta > 0$, we must have

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

implying that (A.13) will hold in the limit. By continuity, it follows that (A.13) holds for sufficiently large δ. We can conclude that if ζ is sufficiently large, then there must exist $\bar{\delta} > 0$, such that $\Pi_P^d > \Pi_P^c$ if $\delta > \bar{\delta}$. $\square$

A.6 Proof of Theorem 4

In the proof of Theorem 2, it is shown that if the condition $\mathbb{E}\left[\left(\frac{\lambda}{1-\lambda} - 1\right)^2\right] > 1$ 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 for all $x, x' \in (0,1)$ with $x > x'$,

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(x))}{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(x'))} \geq \left(\frac{x}{x'}\right)^{\zeta}$$

for some $\zeta \geq \zeta(\delta)$, then the following will necessarily hold:

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(1-D(\delta)))}{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(1-C_F(\delta)))} \geq \left(\frac{1-D(\delta)}{1-C_F(\delta)}\right)^{\zeta(\delta)} > R_F(\delta).$$

According to (A.13), we equivalently have $\Pi_P^d > \Pi_P^c$. $\square$

Online Appendix: Additional Results and Proofs

B.1 Equilibrium uniqueness

In this part of the appendix, we establish that full revelation of private estimation is essentially the unique prediction of the communication game under both organizational structures. In particular, we will show that in any PBE under decentralization, after the bilateral communication the agents can always be (almost) sure about each other's types. Similarly, in any PBE under centralization, the principal can always be (almost) sure about the agents' types after receiving their reports.

Proposition B.1. *If a communication strategy $m_i^*(\cdot)$ of agent $i \in \{1, 2\}$ is part of a PBE under either decentralization or centralization, then $\mathbb{E}[\theta_i|m_i^*(e_i^*, s_i)] = \mathbb{E}[\theta_i|s_i, e_i^*]$ for almost all $s_i \in \mathcal{S} \setminus \{s : \mathbb{E}[\theta_i|s, e_i^*] = 0\}$ with respect to the ex ante distribution of the signal in equilibrium.*

Proof. If the distributions Γ and $G^e(\cdot|\theta)$ are discrete, then the result can be proved by adapting the well-known unraveling argument (Grossman, 1981; Milgrom, 1981). More specifically, in our setting, discreteness of type space would imply that 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 ($|E[\theta_i|m]| < |E[\theta_i|s_i]|$). Thus, by deviating to the type-revealing message $m^{E[\theta_i|s_i]}$ agent i could convince j to take decisions that are more favorable to i in expectation. In what follows, we show how this intuitive argument can be generalized to arbitrary distributions, including the ones that are partly discrete and partly continuous.

Decentralization. Let $((e_1^*, m_1^*(\cdot), y_1^*(\cdot)), (e_2^*, m_2^*(\cdot), y_2^*(\cdot)))$ 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 $m_i = m^{\mathbb{E}[\theta_i|s_i]}$ than the proposed equilibrium message $m_i^*(e_i^*, s_i)$ *if* both of the following two inequalities hold:

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

and

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

Note that (B.1) is further equivalent to

$$\left(\mathbb{E}\left[\theta_i|m_i^*(e_i^*,s_i)\right]-\mathbb{E}[\theta_i|s_i]\right)\left((2+4\delta)\mathbb{E}[\theta_i|s_i]+\delta^2\left(\mathbb{E}\left[\theta_i|m_i^*(e_i^*,s_i)\right]+\mathbb{E}[\theta_i|s_i]\right)\right)>0. \tag{B.3}$$

From (B.2) and (B.3), it is clear that if $\mathbb{E}[\theta_i|s_i]>0$, then deviating to $m^{\mathbb{E}[\theta_i|s_i]}$ is not profitable for agent i only if $\mathbb{E}[\theta_i|s_i]\leq\mathbb{E}[\theta_i|m_i^*(e_i^*,s_i)]$. Similarly, if $\mathbb{E}[\theta_i|s_i]<0$, then deviating to $m^{\mathbb{E}[\theta_i|s_i]}$ is not profitable for agent i only if $\mathbb{E}[\theta_i|s_i]\geq\mathbb{E}[\theta_i|m_i^*(e_i^*,s_i)]$. These arguments also imply that we must have $m_i^*(e_i^*,s_i)\neq m_i^*(e_i^*,s_i')$ $\forall s_i,s_i'\in\mathcal{S}$ such that $\mathbb{E}[\theta_i|s_i]\cdot\mathbb{E}[\theta_i|s_i']<0$.

Next, suppose, in contradiction to the current proposition, that there exist $i\in\{1,2\}$ and a non-null subset $\hat{\mathcal{S}}\subseteq\mathcal{S}\setminus\{s:\mathbb{E}[\theta_i|s,e_i^*]=0\}$ with respect to the distribution H^{e^*}, such that $\mu_j^i(\{\hat{s}_i\}|m_i^*(e_i^*,\hat{s}_i))<1$ $\forall\hat{s}_i\in\hat{\mathcal{S}}$.[24] Since the beliefs must be consistent in equilibrium, we have $m_i^*(e_i^*,\hat{s}_i)\neq m^{\mathbb{E}[\theta_i|\hat{s}_i]}$ $\forall\hat{s}_i\in\hat{\mathcal{S}}$. 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 H^{e^*} 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 $\mathbb{E}\left[\theta_i|s_i\right]>\max\{0,\mathbb{E}[\theta_i|\hat{m}_i^*]\}$ or $\mathbb{E}\left[\theta_i|s_i\right]<\min\{0,\mathbb{E}[\theta_i|\hat{m}_i^*]\}$ holds. This is not possible in equilibrium given our analysis of (B.2) and (B.3).

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 H^{e^*}, 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. ■

Centralization. Let $((e_1^*,m_1^*(\cdot),y_1^*(\cdot)),(e_2^*,m_2^*(\cdot),y_2^*(\cdot)))$ be an equilibrium strategy profile under centralization. Without loss of generality, we focus on agent 1 and consider any signal $s_1\in\mathcal{S}$ such that $\mathbb{E}[\theta_1|s,e_1^*]\neq0$. Repeating the calculations of (A.6), (A.7) and (A.8), it can be checked that agent 1 would strictly prefer the type-revealing message $m_1=m^{\mathbb{E}[\theta_1|s_1]}$ than

[24]Formally, we define H^{e^*} by letting $H^{e^*}(s)=\int_\Theta G^{e^*}(s|\theta)d\Gamma$ $\forall s\in\mathcal{S}$. We say that a set $\hat{S}\subseteq\mathcal{S}$ is non-null with respect to H^{e^*} if $\int_{\mathcal{S}}\mathbb{1}_{\{s\in\hat{S}\}}dH^{e^*}>0$, and it is null with respect to H^{e^*} if $\int_{\mathcal{S}}\mathbb{1}_{\{s\in\hat{S}\}}dH^{e^*}=0$.

the proposed equilibrium message $m_1^*(e_1^*, s_1)$ if *both* of the following two inequalities hold:

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{\lambda\,(1-\lambda+\delta)\,\mathbb{E}\left[\theta_1|m^{\mathbb{E}[\theta_1|s_1]}\right]}{\lambda(1-\lambda)+\delta}-\theta_1\right)^2\middle|s_1\right]\\
>&\,\mathbb{E}\left[\left(\frac{\lambda\,(1-\lambda+\delta)\,\mathbb{E}\left[\theta_1|m_1^*(e_1^*,s_1)\right]}{\lambda(1-\lambda)+\delta}-\theta_1\right)^2\middle|s_1\right],
\end{aligned} \tag{B.4}$$

and

$$\mathbb{E}\left[\left(\frac{\lambda\,(1-\lambda)\,\mathbb{E}\left[\theta_1|m^{\mathbb{E}[\theta_1|s_1]}\right]}{\lambda(1-\lambda)+\delta}\right)^2\middle|s_1\right]>\mathbb{E}\left[\left(\frac{\lambda\,(1-\lambda)\,\mathbb{E}\left[\theta_1|m_1^*(e_1^*,s_1)\right]}{\lambda(1-\lambda)+\delta}\right)^2\middle|s_1\right]. \tag{B.5}$$

A sufficient condition for (B.4) and (B.5) to hold simultaneously is that we always have either $\mathbb{E}[\theta_1|s_1] > \max\{\mathbb{E}[\theta_1|m_1^*(e_1^*, s_1)], 0\}$ or $\mathbb{E}[\theta_1|s_1] < \min\{\mathbb{E}[\theta_1|m_1^*(e_1^*, s_1)], 0\}$. Hence, for the proposed strategy profile to constitute an equilibrium, it is necessary that $\forall s_1 \in \mathcal{S} \setminus \{s : \mathbb{E}[\theta_1|s, e_1^*] = 0\}$, either $0 < \mathbb{E}[\theta_1|s_1] \leq \mathbb{E}[\theta_1|m_1^*(e_1^*, s_1)]$ or $\mathbb{E}[\theta_1|m_1^*(e_1^*, s_1)] \leq \mathbb{E}[\theta_1|s_1] < 0$ 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. ■□

B.2 Correlated local states

In this section, we extend the truthful disclosure result to settings where the local states are correlated. We consider a Gaussian environment where the local states are given by

$$\theta_1 = z + \varepsilon_1,\ \ \theta_2 = z + \varepsilon_2.$$

Here, $z \sim \mathcal{N}(0, \sigma_z^2)$ and $\varepsilon_1, \varepsilon_2 \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ are normally and independently distributed variables with finite variances. Thus, the local states θ_1 and θ_2 are positively correlated, with a correlation coefficient $\rho = \sigma_z^2/(\sigma_z^2 + \sigma_\varepsilon^2)$. Note that joint normality and the Bayes' rule imply that

$$\mathbb{E}[\theta_j|\theta_i] = \rho\theta_i, \quad \forall i, j = 1, 2. \tag{B.6}$$

For simplicity, we assume that (i) information acquisition takes the "all-or-nothing" form (i.e., by exerting effort $e_i \in [0, 1]$, agent i receives a signal $s_i = \theta_i$ with probability e_i, and he receives a null signal with the remaining probability $(1 - e_i)$, and (ii) the message spaces are given by $\mathcal{M}(s_i) = \{\emptyset\} \cup \{S \subset \mathbb{R} : s_i \in S\}$.[25]

Now consider the case of decentralization (the analysis under centralization is analogous),

[25] Alternatively, at the cost of more restricted message spaces, we can also work with the Gaussian learning technology, where each private signal is drawn from a normal distribution with effort-dependent variance.

and the communication strategy profile $m_i^d(e_i, s_i) = s_i$, $\forall e_i \in [0,1]$, $s_i \in \mathbb{R} \cup \{\emptyset\}$ and $i = 1, 2$. We also specify that, upon observing a message from agent i, agent j would update his belief to be such that $\mu_j^i(\{\underline{s}^{m_i}\}) = 1$ if $m_i \neq \emptyset$, where $\underline{s}^{m_i}$ is as defined in the main text, and $\mu_j^i(\{\emptyset\}) = 1$ if $m_i = \emptyset$. Note that this belief satisfies both the Bayes' rule and the consistency requirement, and we have $\underline{s}^{m_i} = s_i$ when $m_i = s_i \neq \emptyset$. We will show that, regardless of the effort choices, no agent will have the incentive to unilaterally deviate from the proposed strategy profile.

The disclosure incentive for an uninformed agent is trivial, so we suppose that agent i has drawn an informative signal $s_i = \theta_i$ and thus learned that his local state is θ_i. If agent i chooses to send a message m_i to agent j, sequential rationality and the beliefs specified above will imply the following action choices:

$$\begin{aligned} y_i(m_i, \theta_i, s_j) &= \frac{1}{1+\delta} \cdot \theta_i + \frac{\delta}{1+\delta} \cdot y_j(m_i, s_j), \text{ and} \\ y_j(m_i, s_j) &= \frac{1+\delta}{1+2\delta} \cdot \mathbb{E}[\theta_j | \underline{s}^{m_i}, s_j] + \frac{\delta}{1+2\delta} \cdot \mathbb{E}[\theta_i | \underline{s}^{m_i}, s_j]. \end{aligned}$$

Hence, if agent i sends a message $m_i \in \mathcal{M}(\theta_i)$, he will incur the following expected losses from mis-adaptation and mis-coordination:

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

and

$$EL_c^d(\theta_i, m_i) = \frac{\delta}{(1+\delta)^2} \cdot \mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta} \cdot \mathbb{E}[\theta_j | \underline{s}^{m_i}, s_j] - \frac{\delta}{1+2\delta} \cdot \mathbb{E}[\theta_i | \underline{s}^{m_i}, s_j] \right)^2 \middle| s_i = \theta_i \right].$$

We argue that both $EL_a^d(\theta_i, m_i)$ and $EL_c^d(\theta_i, m_i)$ are minimized when $m_i = \theta_i$. For this purpose, it suffices to show that both

$$\begin{aligned} &\mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta} \cdot \mathbb{E}[\theta_j | \underline{s}^{m_i}, s_j] - \frac{\delta}{1+2\delta} \cdot \mathbb{E}[\theta_i | \underline{s}^{m_i}, s_j] \right)^2 \middle| s_i = \theta_i, s_j = \emptyset \right] \\ =&\mathbb{E}\left[\left(\theta_i - \left(\frac{1+\delta}{1+2\delta} \cdot \rho + \frac{\delta}{1+2\delta} \right) \cdot \mathbb{E}[\theta_i | \underline{s}^{m_i}] \right)^2 \middle| s_i = \theta_i \right] \end{aligned} \tag{B.7}$$

and

$$\mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta} \cdot \mathbb{E}[\theta_j | \underline{s}^{m_i}, s_j] - \frac{\delta}{1+2\delta} \cdot \mathbb{E}[\theta_i | \underline{s}^{m_i}, s_j] \right)^2 \middle| s_i = \theta_i, s_j \neq \emptyset \right]$$

$$
= \mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta}\cdot\theta_j - \frac{\delta}{1+2\delta}\cdot\mathbb{E}[\theta_i|\underline{s}^{m_i},\theta_j]\right)^2 \middle| s_i=\theta_i\right] \tag{B.8}
$$

are both minimized when $m_i = \theta_i$. This clearly holds for (B.7), because $|\mathbb{E}[\theta_i|\underline{s}^{m_i}]| \leq |\theta_i|$ $\forall\theta_i \in \Theta_i$ and $m_i \in \mathcal{M}(\theta_i)$.

As for (B.8), for all $m_i \neq \emptyset$, we have

$$
\mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta}\cdot\theta_j - \frac{\delta}{1+2\delta}\cdot\underline{s}^{m_i}\right)^2 \middle| s_i=\theta_i\right],
$$

which, given that $\max\{|\mathbb{E}[\theta_j|\theta_i|, |\underline{s}^{m_i}|\} \leq |\theta_i|$ $\forall m_i \in \mathcal{M}(\theta_i)$ and $\forall\theta_i \in \Theta$, will attain a lower value at $m_i = \theta_i$ than any other $m_i \neq \emptyset$. Finally, we compare the cases $m_i = \theta_i$ and $m_i = \emptyset$:

$$
\begin{aligned}
&\mathbb{E}\left[\left(\theta_i - \frac{1+\delta}{1+2\delta}\cdot\theta_j - \frac{\delta}{1+2\delta}\cdot\theta_i\right)^2 - \left(\theta_i - \frac{1+\delta}{1+2\delta}\cdot\theta_j - \frac{\delta}{1+2\delta}\cdot\rho\theta_j\right)^2 \middle| s_i=\theta_i\right] \\
=&\frac{\delta}{1+2\delta}\cdot\mathbb{E}\left[(\rho\theta_j-\theta_i)\left(\frac{2+3\delta}{1+2\delta}\cdot\theta_i - \frac{2+2\delta+\rho\delta}{1+2\delta}\cdot\theta_j\right)\middle| s_i=\theta_i\right] \\
=&\frac{\delta}{(1+2\delta)^2}\cdot\mathbb{E}\left[-\delta\left(\rho\theta_j-\theta_i\right)^2 + (2+2\delta)(\rho\theta_j-\theta_i)(\theta_i-\theta_j)|s_i=\theta_i\right] \\
\leq&\frac{2\delta+2\delta^2}{(1+2\delta)^2}\cdot\mathbb{E}\left[(\rho\theta_j-\theta_i)(\theta_i-\theta_j)|s_i=\theta_i\right] \\
=&\frac{2\delta+2\delta^2}{(1+2\delta)^2}\cdot\mathbb{E}\left[-\rho(\theta_i-\theta_j)^2 + (\rho\theta_i-\theta_i)(\theta_i-\theta_j)|s_i=\theta_i\right] \\
\leq&\frac{2\delta+2\delta^2}{(1+2\delta)^2}\cdot\mathbb{E}\left[(\rho-1)\theta_i(\theta_i-\theta_j)|s_i=\theta_i\right] \\
=&-\frac{2\delta+2\delta^2}{(1+2\delta)^2}\cdot(1-\rho)^2\theta_i^2 \\
\leq&\, 0,
\end{aligned}
$$

where the second inequality follows that $\rho \geq 0$. In sum, we have shown that both $EL_a^d(\theta_i, m_i)$ and $EL_c^d(\theta_i, m_i)$ are minimized when agent i reports truthfully. In particular, this is the case regardless of the agents' choice of efforts.

When agents fully reveal their private signals, the on-path equilibrium decision rules are then given by

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

We therefore obtain that, for every pair of effort levels, the following expected payoff agent i

under decentralization:

$$\begin{aligned}
U_i^d(e_i,e_j) =& K - \mathbb{E}\left[\left(\frac{(1+\delta)\mathbb{E}[\theta_i|s_i,s_j]}{1+2\delta} - \theta_i\right)^2 \middle| e_i,e_j\right] - \mathbb{E}\left[\left(\frac{\delta\mathbb{E}[\theta_j|s_j,s_i]}{1+2\delta}\right)^2 \middle| e_i,e_j\right] \\
&- \delta\mathbb{E}\left[\left(\frac{\mathbb{E}[\theta_i|s_i,s_j]}{1+2\delta}\right)^2 \middle| e_i,e_j\right] - \delta\mathbb{E}\left[\left(\frac{\mathbb{E}\left[\theta_j|s_j,s_i\right]}{1+2\delta}\right)^2 \middle| e_i,e_j\right] - c(e_i) \\
=& K + \left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma^2_{\mathbb{E}[\theta|s]}(e_i,e_j) - \sigma^2_\theta - \frac{\delta^2+\delta}{(1+2\delta)^2}\sigma^2_{\mathbb{E}[\theta|s]}(e_j,e_i) - c(e_i) \\
=& K + (1 - D(\delta))\,\sigma^2_{\mathbb{E}[\theta|s]}(e_i,e_j) - \sigma^2_\theta - D(\delta)\sigma^2_{\mathbb{E}[\theta|s]}(e_j,e_i) - c(e_i),
\end{aligned}$$

where the residual variance of the local states are defined by

$$\sigma^2_{E[\theta|s]}(e_i,e_j) := \int_{\theta_i,\theta_j\in\Theta}\int_{s_i\in S^{e_i},s_j\in\in S^{e_j}} \left(\mathbb{E}[\theta_i|s_i,s_j]\right)^2 dG^{e_i,e_j}(s_i,s_j|\theta_i,\theta_j)d\Gamma(\theta_i,\theta_j),$$

and $G^{e_i,e_j}(\cdot,\cdot|\theta_i,\theta_j)$ is the conditional joint distribution of the private signals. Similarly, one can show that, for every pair of effort levels, the expected payoff of agent i under centralization is given by

$$U_i^c(e_i,e_j) = K + (1 - C_F(\delta))\,\sigma^2_{\mathbb{E}[\theta|s]}(e_i,e_j) - \sigma^2_\theta - C_F(\delta)\sigma^2_{\mathbb{E}[\theta|s]}(e_j,e_i) - c(e_i).$$

Clearly, unlike in the setting where the local states are independent, the effort choices of the agents under both organizational structures will now be interdependent. Nevertheless, if we restrict attention to symmetric equilibria (and assume that the first-order conditions are sufficient), it will be again straightforward to determine the sign of the effort difference between the two organizational structures by comparing the marginal benefits of acquiring information (which amounts to check whether we have $D(\delta) < C_F(\delta)$ or the opposite). In particular, similar to Theorems 1 and 2, when there is uncertainty in the relative importance λ, decentralization always leads to a higher effort level if either (i) coordination is sufficiently important, or (ii) coordination is relative unimportant but the degree of global uncertainty is sufficiently large.

B.3 Comparative statics of e^d and e^c_F

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

$$D(\delta) := \frac{\delta^2+\delta}{(1+2\delta)^2} \tag{B.9}$$

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

$$C(\delta,\lambda) := \frac{\delta^2(\lambda^2+(1-\lambda)^2)+2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}. \tag{B.10}$$

Differentiating with respect to δ, we have

$$D'(\delta) = \frac{(2\delta+1)(1+2\delta)-4(\delta^2+\delta)}{(1+2\delta)^3} = \frac{1}{(1+2\delta)^3} > 0, \tag{B.11}$$

and

$$\begin{aligned}
&\frac{\partial C(\delta,\lambda)}{\partial \delta} \\
&= \frac{[2\delta(\lambda^2+(1-\lambda)^2)+2\lambda^2(1-\lambda)^2]\cdot(\lambda(1-\lambda)+\delta)-2\delta[\delta(\lambda^2+(1-\lambda)^2)+2\lambda^2(1-\lambda)^2]}{2(\lambda(1-\lambda)+\delta)^3} \\
&= \frac{\lambda^3(1-\lambda)^3+\delta\lambda(1-\lambda)(\lambda^2+(1-\lambda)^2-\lambda(1-\lambda))}{(\lambda(1-\lambda)+\delta)^3} \\
&= \frac{\lambda^3(1-\lambda)^3+\delta\lambda(1-\lambda)((2\lambda-1)^2+\lambda(1-\lambda))}{(\lambda(1-\lambda)+\delta)^3} \\
&> 0.
\end{aligned} \tag{B.12}$$

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

$$e^d = \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}(1-D(\delta)), \quad e_F^c = \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}(1-\mathbb{E}\left[C(\delta,\lambda)\right]), \tag{B.13}$$

and $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ is concave and $c(\cdot)$ is strictly convex.

B.4 Proof of Proposition 3

The case with $\omega = 0$ immediately follows from Theorem 1(i), so we focus on the cases with $\omega \in (0,1]$. For every $\omega \in (0,1]$, define

$$\Delta(\delta,\omega) \equiv C_{F(\cdot|\omega)}(\delta) - D(\delta) = \frac{\delta^2\cdot\frac{1+\omega^2}{2}+2\delta\cdot\left(\frac{1-\omega^2}{4}\right)^2}{2\left(\frac{1-\omega^2}{4}+\delta\right)^2} - \frac{\delta^2+\delta}{(1+2\delta)^2}. \tag{B.14}$$

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.12), which are now given by $\alpha = (1-\omega^2)/4$ and $\beta = (1+\omega^2)/2$. Provided that $\delta > 0$, we have

$$\Delta(\delta,\omega)=0$$

$$
\begin{aligned}
&\iff \frac{(\beta\delta+2\alpha^2)(1+2\delta)^2-2(\delta+1)(\alpha+\delta)^2}{2(\alpha+\delta)^2(1+2\delta)^2}=0\\
&\iff (\beta\delta+2\alpha^2)(4\delta^2+4\delta+1)-(2\delta+2)(\alpha^2+\delta^2+2\alpha\delta)=0\\
&\iff (4\beta-2)\delta^3+(8\alpha^2-4\alpha+4\beta-2)\delta^2+(6\alpha^2-4\alpha+\beta)\delta=0\\
&\iff (4\beta-2)\delta^2+(8\alpha^2-4\alpha+4\beta-2)\delta+(6\alpha^2-4\alpha+\beta)=0\\
&\iff \left(\delta+\frac{4\alpha^2-2\alpha+2\beta-1}{4\beta-2}\right)^2=\frac{(4\alpha^2-2\alpha+2\beta-1)^2-(6\alpha^2-4\alpha+\beta)(4\beta-2)}{(4\beta-2)^2}\\
&\iff \left(\delta+\frac{4\alpha^2-2\alpha+2\beta-1}{4\beta-2}\right)^2=\frac{(1-2\alpha)^2(4\alpha^2-2\beta+1)}{(4\beta-2)^2}, \qquad \text{(B.15)}
\end{aligned}
$$

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 (B.15) 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(ii) 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 (iv) 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 and Theorem 1(ii), 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 Theorem 1(ii), it follows that $\Delta(\delta,\omega)>0$ for all $\delta>\bar{\delta}(\omega)$. 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 $\Delta(\delta, \omega) \leq 0$ for all $\delta \in (\max\{0, \underline{\delta}(\omega)\}, \bar{\delta}(\omega)]$. Note that

$$\begin{aligned}
\underline{\delta}(\omega) \leq 0 &\iff -\frac{4\alpha^2 - 2\alpha + 2\beta - 1}{4\beta - 2} \leq \frac{(1-2\alpha)\sqrt{4\alpha^2 - 2\beta + 1}}{4\beta - 2} \\
&\iff (4\alpha^2 - 2\alpha + 2\beta - 1)^2 \leq (1-2\alpha)^2(4\alpha^2 - 2\beta + 1) \\
&\iff 6\alpha^2 - 4\alpha + \beta \leq 0 \\
&\iff \frac{3(1-\omega^2)^2}{8} - \frac{1 - 3\omega^2}{2} \leq 0, \\
&\iff (1-\omega^2)^2 - 4(1-\omega^2) + 4 - \frac{4}{3} \leq 0 \\
&\iff (1 - \omega^2 - 2)^2 - \frac{4}{3} \leq 0 \\
&\iff (1 + \omega^2)^2 - \frac{4}{3} \leq 0,
\end{aligned}$$

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 $\omega \in \left(0, \sqrt{2} - 1\right)$. Clearly, the equation $(1 + \omega^2)^2 - 4/3 = 0$ has a unique real root on $(0, 1)$, which is given by

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

Therefore, when $\omega \leq \hat{\omega}$, we have $\underline{\delta}(\omega) \leq 0$, and consequently $e_d > e_F^c$ if and only if $\delta > \bar{\delta}(\omega)$. This proves part (ii) of the proposition.

Finally, suppose that $\omega \in (\hat{\omega}, \sqrt{2} - 1)$. In this case, it is clear from the above analysis that we have $\underline{\delta}(\omega) > 0$, and 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 $\delta \in (0, \underline{\delta}(\omega))$. In addition, Theorem 1(ii) and continuity again imply that $\Delta(\delta, \omega) > 0$ $\forall \delta > \bar{\delta}(\omega)$. We can therefore conclude that part (iii) of the proposition also holds. □

B.5 Proof of Proposition 4

Given the specification of $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ and $c(\cdot)$ and that λ is binomially distributed, the optimality condition for decentralization (A.13) becomes equivalent to the following:

$$\left(\frac{1-\frac{x}{(1+x)^2}}{1-\frac{(3+6\omega^2-\omega^4)x^2+(1-\omega^2)^2x}{((3+\omega^2)x+1-\omega^2)^2}}\right)^{\frac{1}{\alpha-1}} > \frac{1-\frac{(1-\omega^2)x}{(3+\omega^2)x+1-\omega^2}}{1-\frac{2x}{(1+x)^2}}, \tag{B.16}$$

where $x = \delta/(1+\delta)$.

For any given $\delta > 0$, by letting $\omega \to 1$ (B.16) will become equivalent to

$$\left(\frac{1-\frac{x}{(1+x)^2}}{0.5}\right)^{\frac{1}{\alpha-1}} > \frac{1}{1-\frac{2x}{(1+x)^2}}. \tag{B.17}$$

For all $x \in (0,1)$, the LHS of (B.17) is strictly larger than one and converges to $+\infty$ as $\alpha \to 1$. Hence, for any given $x \in (0,1)$ (or, equivalently, any given $\delta > 0$), (B.17) holds whenever α is sufficiently close to one. Then, by continuity, (B.16) must also hold when ω is sufficiently close to one. This proves part (i) of the proposition.

Similarly, for any given $\omega \in (0,1]$, by taking $\delta \to +\infty$ (B.16) will become equivalent to

$$\left(\frac{0.75}{0.75-0.25\omega^2}\right)^{\frac{1}{\alpha-1}} > \frac{0.75+0.25\omega^2}{0.5}. \tag{B.18}$$

For every $\omega \in (0,1]$, the LHS of (B.18) is strictly larger than one and converges to $+\infty$ as $\alpha \to 1$. Hence, the inequality (B.18) holds whenever α is sufficiently close to one. Then, by continuity, (B.16) must also hold when δ is sufficiently large. This proves part (ii) of the proposition. $\square$

B.6 Alternative specification of agents' performances

Suppose that each division i's performance is now given by

$$\tilde{\pi}_i(\mathbf{y}, \theta_i) = K - (1-q)(y_i - \theta_i)^2 - q(y_i - y_j)^2, \tag{B.19}$$

where $q := \delta/(1+\delta) \in (0,1)$. Thus, agent i's payoff becomes $\tilde{u}_i(\mathbf{y}, \theta_i, e_i) = \tilde{\pi}_i(\mathbf{y}, \theta_i) - c(e_i)$. In what follows, we will first show that with this alternative specification of divisional performance, the agents would exert higher effort under decentralization if and only if they would want to do so before, i.e., when his division's performance was given by $\pi_i(\cdot)$ as in the main text. Then, we provide sufficient conditions (similar to those in Theorems 3 and 4) under which decentralization will lead to a higher expected payoff for the principal than centralization.

Effort provision. Consider the case of decentralization. Because

$$\tilde{u}_i(\cdot) + c(\cdot) - K = \left(u_i(\cdot) + c(\cdot) - K\right)/(1+\delta),$$

the incentive problems facing each agent i in the communication and the decision-making stages are qualitatively the same as before. Hence, truthful disclosure can again be achieved, and the on-path equilibrium decision rules are given by (3.1). Then, for a given pair of efforts (e_1, e_2), agent 1's expected payoff is

$$\tilde{U}_i^d(e_i, e_j) + c(e_i) - K = \frac{1}{1+\delta} \cdot \left(U_i^d(e_i, e_j) + c(e_i) - K\right),$$

where the expression of $U_i^d(e_i, e_j)$ can be found in the proof of Proposition 1. Hence, we have

$$\tilde{U}_i^d(e_i, e_j) = K + \frac{1}{1+\delta}\left[\left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma^2_{\mathbb{E}[\theta|s]}(e_i) - \sigma^2_\theta - \frac{\delta^2+\delta}{(1+2\delta)^2}\sigma^2_{\mathbb{E}[\theta|s]}(e_j)\right] - c(e_i).$$

Differentiating $\tilde{U}_i^d(e_i, e_j)$ with respect to e_i, we obtain the first-order condition

$$\frac{1}{1+\delta}\left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\sigma^{2\,\prime}_{\mathbb{E}[\theta|s]}(e_i) - c'(e_i) = 0.$$

Therefore, the equilibrium effort level under decentralization is given by

$$\tilde{e}^d := \left(\frac{c'}{\sigma^{2\,\prime}_{\mathbb{E}[\theta|s]}}\right)^{-1}\left(\frac{1}{1+\delta}\left(1 - \frac{\delta^2+\delta}{(1+2\delta)^2}\right)\right). \tag{B.20}$$

For the case of centralization, the analysis is analogous. In particular, the corresponding equilibrium effort level under centralization is given by

$$\tilde{e}^c_F := \left(\frac{c'}{\sigma^{2\,\prime}_{\mathbb{E}[\theta|s]}}\right)^{-1}\left(\frac{1}{1+\delta}\left(1 - \mathbb{E}\left[\frac{\delta^2(\lambda^2+(1-\lambda)^2) + 2\delta\lambda^2(1-\lambda)^2}{2(\lambda(1-\lambda)+\delta)^2}\right]\right)\right). \tag{B.21}$$

Clearly, for all $\delta > 0$, we have

$$\tilde{e}^d > \tilde{e}^c_F \iff \frac{1-D(\delta)}{1+\delta} > \frac{1-C_F(\delta)}{1+\delta} \iff D(\delta) < C_F(\delta) \iff e^d > e^c_F,$$

where $D(\delta)$ and $C_F(\delta)$ are as defined in (4.1), and e^d and e^c_F are the equilibrium effort levels with the previous payoff function $u_i(\cdot)$. Therefore, although both $\tilde{e}^d$ and $\tilde{e}^c_F$ converge to zero as $\delta \to +\infty$, we must have $\tilde{e}^d - \tilde{e}^c_F > 0$ for sufficiently large δ if the condition of Theorem 1 is satisfied. Similarly, if the condition of Theorem 2 is satisfied, then $\tilde{e}^d - \tilde{e}^c_F > 0$ also holds when δ is sufficiently small.

The principal's payoff. With payoff specification (B.19), the principal's ex post payoff is equal to $\lambda\tilde{\pi}_1(\mathbf{y},\theta_1)+(1-\lambda)\tilde{\pi}_2(\mathbf{y},\theta_2)$. Thus, given the equilibrium effort level $\tilde{e}^d$, the principal's expected payoff under decentralization is

$$\tilde{\Pi}_P^d = K - \frac{\sigma_\theta^2}{1+\delta} + \frac{1}{1+\delta}\left(1 - \frac{2\delta^2+2\delta}{(1+2\delta)^2}\right)\sigma^2_{\mathbb{E}[\theta|s]}\left(\tilde{e}^d\right).$$

Similarly, given the equilibrium effort level $\tilde{e}_F^c$, the principal's expected payoff under centralization is

$$\tilde{\Pi}_P^c = K - \frac{\sigma_\theta^2}{1+\delta} + \frac{1}{1+\delta}\left(1 - \mathbb{E}\left[\frac{\lambda(1-\lambda)\delta}{\lambda(1-\lambda)+\delta}\right]\right)\sigma^2_{\mathbb{E}[\theta|s]}\left(\tilde{e}_F^c\right).$$

Therefore, we have $\tilde{\Pi}_P^d > \tilde{\Pi}_P^c$ if and only if the follow inequality holds:

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}\left(\tilde{e}^d\right)}{\sigma^2_{\mathbb{E}[\theta|s]}\left(\tilde{e}_F^c\right)} > R_F(\delta), \tag{B.22}$$

where $R_F(\delta)$ is the same as defined in (A.13).

Now, similar to Theorems 3 and 4, suppose that either of the following conditions holds: (i) the two operating divisions are not always equally important to the principal and δ is sufficiently large, or (ii) $\mathbb{E}\left[(\lambda/(1-\lambda)-1)^2\right] > 1$ and δ is sufficiently small. As argued before, in either case we will have $D(\delta) < C_F(\delta)$, which implies that

$$\left(\frac{1-D(\delta)}{1-C_F(\delta)}\right)^{\zeta} > R_F(\delta)$$

necessarily holds whenever ζ is larger than some cutoff $\zeta(\delta) > 0$. Hence, if the following condition (which also appears in Theorem 4) is further satisfied:

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(x))}{\sigma^2_{\mathbb{E}[\theta|s]}(g^{-1}(x'))} \geq \left(\frac{x}{x'}\right)^{\zeta(\delta)} \quad \forall x, x' \in (0,1) \text{ with } x > x', \text{ where } g(\cdot) = \frac{c'(\cdot)}{\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)},$$

then we immediately obtain

$$\frac{\sigma^2_{\mathbb{E}[\theta|s]}(\tilde{e}^d)}{\sigma^2_{\mathbb{E}[\theta|s]}(\tilde{e}_F^c)} = \frac{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}\left(\frac{1-D(\delta)}{1+\delta}\right)\right)}{\sigma^2_{\mathbb{E}[\theta|s]}\left(g^{-1}\left(\frac{1-C_F(\delta)}{1+\delta}\right)\right)} \geq \left(\frac{1-D(\delta)}{1-C_F(\delta)}\right)^{\zeta(\delta)} > R_F(\delta).$$

According to (B.22), this allows us to conclude that the principal's expected payoff is strictly higher under decentralization.

B.7 Proof of Proposition 5

From the proof of Proposition 1, we know that the agents' equilibrium efforts under centralization are given by

$$
\begin{aligned}
e_1^c &= \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}\left(1-\mathbb{E}\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right), \text{ and} \\
e_2^c &= \left(\frac{c'}{\sigma^2_{\mathbb{E}[\theta|s]}{}'}\right)^{-1}\left(1-\mathbb{E}\left[\frac{\lambda^2\left(\delta^2+\delta(1-\lambda)^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right]\right).
\end{aligned}
$$

Given the binary distribution of λ, we define

$$
\begin{aligned}
C_1(v) &:= \mathbb{E}\left[\frac{(1-\lambda)^2\left(\delta^2+\delta\lambda^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right] \\
&= \frac{1}{2}\delta\left(\frac{(1-v+\omega)^2\left(4\delta+(v-\omega+1)^2\right)}{\left((v-\omega)^2-4\delta-1\right)^2}+\frac{(v+\omega-1)^2\left(4\delta+(v+\omega+1)^2\right)}{((v+\omega-1)(v+\omega+1)-4\delta)^2}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
C_2(v) &:= \mathbb{E}\left[\frac{\lambda^2\left(\delta^2+\delta(1-\lambda)^2\right)}{(\lambda(1-\lambda)+\delta)^2}\right] \\
&= \frac{1}{2}\delta\left(\frac{(v-\omega+1)^2\left(4\delta+(-v+\omega+1)^2\right)}{\left((v-\omega)^2-4\delta-1\right)^2}+\frac{(v+\omega+1)^2\left(4\delta+(v+\omega-1)^2\right)}{((v+\omega-1)(v+\omega+1)-4\delta)^2}\right).
\end{aligned}
$$

We first show that $C_1'(v)<0$. After some rearrangement, we obtain

$$
\begin{aligned}
C_1'(v) = -4\delta^2\Bigg(&\frac{3(v-\omega-1)}{\left(4\delta+1-(v-\omega)^2\right)^2}-\frac{4(4\delta+1)(v-\omega-1)}{\left(4\delta+1-(v-\omega)^2\right)^3} \\
&+\frac{3(v+\omega-1)}{(4\delta+1-(v+\omega)^2)^2}-\frac{4(4\delta+1)(v+\omega-1)}{(4\delta+1-(v+\omega)^2))^3}\Bigg).
\end{aligned}
$$

To show that $C_1'(v)>0$, it suffices to prove that

$$
\frac{3(v-\omega-1)}{\left(4\delta+1-(v-\omega)^2\right)^2}>\frac{4(4\delta+1)(v-\omega-1)}{\left(4\delta+1-(v-\omega)^2\right)^3} \tag{B.23}
$$

and

$$
\frac{3(v+\omega-1)}{(4\delta+1-(v+\omega)^2)^2}>\frac{4(4\delta+1)(v+\omega-1)}{(4\delta+1-(v+\omega)^2))^3}. \tag{B.24}
$$

Since $\omega\in[0,1]$ and $v\in[0,1-\omega]$, it is straightforward to show that (B.23) is equivalent to

$$
-4\delta-1-3(v-\omega)^2<0, \tag{B.25}
$$

while (B.23) is equivalent to

$$-4\delta - 1 - 3(v+\omega)^2 < 0. \tag{B.26}$$

Clearly, both (B.25) and (B.26) always hold. This proves that $C_1'(v) < 0$.

Next, we argue that $C_2'(v) > 0$. Consider the derivative

$$\begin{aligned}
C_2'(v) = &- 4\delta^2 \Bigg(\frac{3(v-\omega+1)}{(4\delta+1-(v-\omega)^2)^2} - \frac{4(4\delta+1)(v-\omega+1)}{(4\delta+1-(v-\omega)^2)^3} \\
&+ \frac{3(v+\omega+1)}{(4\delta+1-(v+\omega)^2)^2} - \frac{4(4\delta+1)(v+\omega+1)}{(4\delta+1-(v+\omega)^2)^3} \Bigg).
\end{aligned}$$

To show that $C_2'(v) > 0$ it suffices to show that

$$\frac{3(v-\omega+1)}{(4\delta+1-(v-\omega)^2)^2} < \frac{4(4\delta+1)(v-\omega+1)}{(4\delta+1-(v-\omega)^2)^3} \tag{B.27}$$

and

$$\frac{3(v+\omega+1)}{(4\delta+1-(v+\omega)^2)^2} < \frac{4(4\delta+1)(v+\omega+1)}{(4\delta+1-(v+\omega)^2)^3}. \tag{B.28}$$

Since $\omega \in [0,1]$ and $v \in [0, 1-\omega]$, it is straightforward to show that (B.27) is equivalent to (B.25), while (B.28) is equivalent to (B.26). Therefore, similar to the case of proving $C_1'(v) < 0$, we also have $C_2'(v) > 0$.

We now proceed to prove that $C_1'(v) + C_2'(v) > 0$. The sum of the corresponding first derivatives is:

$$C_1'(v) + C_2'(v) = -\frac{16\delta^2 v}{(\omega^4 - 2\omega^2(4\delta+v^2+1) + (v^2-4\delta-1)^2)^3} \cdot Z(\delta,\omega,v),$$

where

$$\begin{aligned}
Z(\delta,\omega,v) \equiv &- 9\omega^8 + 4\omega^6\left(4\delta+6v^2+1\right) - 12(4\delta+1)\omega^2\left(4\delta-v^2+1\right)\left(4\delta+3v^2+1\right) \\
&- \left(4\delta-v^2+1\right)^3\left(4\delta+3v^2+1\right) + 2\omega^4\left(-16(4\delta+1)v^2+9(4\delta+1)^2-9v^4\right).
\end{aligned}$$

To determine the sign of $C_1'(v) + C_2'(v)$, we first note that

$$\omega^4 - 2\omega^2\left(4\delta+v^2+1\right) + \left(v^2-4\delta-1\right)^2 > 0. \tag{B.29}$$

To see why, consider the derivative

$$\frac{\partial(\omega^4 - 2\omega^2\left(4\delta + v^2 + 1\right) + \left(v^2 - 4\delta - 1\right)^2)}{\partial v} = -4v(1 + 4\delta - v^2 + \omega^2) < 0$$

and so, letting $v = 1 - \omega$, we obtain the minimum of the LHS of (B.29):

$$-2\omega^2\left(4\delta + v^2 + 1\right) + \left(-4\delta + v^2 - 1\right)^2 + \omega^4\Big|_{v=1-\omega} = 16\delta(\delta + \omega(1-\omega)) > 0.$$

Next, consider the derivative

$$\begin{aligned}\frac{\partial Z(\delta,\omega,v)}{\partial \delta} =& 16\Big(w^4\left(36\delta - 8v^2 + 9\right) - \left(v^2 - 4\delta - 1\right)^2\left(4\delta + 2v^2 + 1\right) \\ &- 3\omega^2\left(4(4\delta+1)v^2 + 3(4\delta+1)^2 - 3v^4\right) + \omega^6\Big) \\ <& \, 0.\end{aligned}$$

Thus, regardless of the values of ω and v, the minimum of the function Z is achieved at $\delta = 0$:

$$\begin{aligned}Z(0,\omega,v) =& -9\omega^8 + 4\left(6v^2+1\right)\omega^6 + 12\left(v^2-1\right)\left(3v^2+1\right)\omega^2 \\ &+ \left(v^2-1\right)^3\left(3v^2+1\right) - 2\left(9v^4 + 16v^2 - 9\right)\omega^4 \\ <& \, 0,\end{aligned}$$

which then implies that $Z(\delta,\omega,v)$ is always strictly negative. Together with (B.29), we can now conclude that $C_1'(v) + C_2'(v) > 0$.

Finally, we show that when the condition of the proposition is satisfied, we further have $\partial(e_1^c + e_2^c)/\partial v < 0$. To show this, we consider the derivative

$$\frac{\partial(e_1^c + e_2^c)}{\partial v} = -\left(g^{-1}\right)'(1 - C_1(v)) \cdot C_1'(v) - \left(g^{-1}\right)'(1 - C_2(v)) \cdot C_2'(v),$$

Since $C_1'(v) < 0$, $C_2'(v) > 0$ and $g(\cdot)$ is strictly increasing (recall that $c(\cdot)$ is strictly convex and $\sigma^2_{\mathbb{E}[\theta|s]}(\cdot)$ is concave), the above derivative is strictly negative if and only if

$$-\frac{C_1'(v)}{C_2'(v)} < \frac{(g^{-1})'(1 - C_2(v))}{(g^{-1})'(1 - C_1(v))}. \tag{B.30}$$

Note that the LHS of (B.30) is strictly less than one. When g is concave, its inverse is also concave, which implies that the function $(g^{-1})'$ is decreasing. Then, given that $C_2(v) > C_1(v)$, the RHS of (B.30) must be larger than one. Hence, if g is concave, it always holds that the aggregate effort $e_1^c + e_2^c$ is strictly decreasing in v. □