Authority in a theory of the firm

Inga Deimen
University of Arizona
and CEPR

and

Dezső Szalay*
University of Bonn
and CEPR

July 6, 2020

Abstract

We study a simple model of the firm comprised of a production unit, a sales unit, and an owner with interests in both units. The owner has the right to adapt the production quantity to changes in demand and costs. Whether the owner effectively assumes this right or delegates decision-making depends on the relative uncertainty about demand and costs, on the division of surplus in the firm, and on the riskiness of the environment the firm faces. We characterize conditions that make acquiring ownership rights feasible and efficient. The same conditions determine the boundaries of the firm in our model.

JEL: D83, D82

*Deimen: University of Arizona and CEPR, Eller College of Management, University of Arizona, 1130 E. Helen St, Tucson, AZ 85721, ideimen@arizona.edu. Szalay: University of Bonn and CEPR, Institute for Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany, szalay@uni-bonn.de.
We have benefitted enormously from discussions with numerous colleagues and audiences. We would like to thank in particular Charles Angelucci, Wouter Dessein, Hülya Eraslan, Amanda Friedenberg, Bob Gibbons, Renato Gomes, Yingni Guo, Matt Jackson, Navin Kartik, Niko Matouschek, Stephen Morris, Joel Sobel, and Konrad Stahl. Financial support through the Hausdorff Center of Mathematics, by the DFG in form of a SFB-TR15 grant and the BGSE is gratefully acknowledged. All remaining errors are our own.

Keywords: ownership, authority, strategic communication, delegation, organizations, noisy information, common and private values, risk

"Divide et impera"

1 Introduction

Understanding the boundaries of the firm and what happens inside these boundaries is one of the most important questions in economics.[1] Theories of incomplete contracting offer important guidance towards understanding why and when firms integrate, and how they operate after integration. Grossman and Hart (1986) and Hart and Moore (1990) explain the optimal allocation of ownership through *residual* rights of control.[2] We propose a model of the firm built around *actual* rights of control as in Coase (1937) and Simon (1951), who think of the firm as a power relationship (Coase) and the employment contract as an authority relationship (Simon). In our theory, ownership of a firm provides the owner with *formal* authority – the right to make decisions on behalf of the entire firm. As emphasized by Aghion and Tirole (1997), formal authority need not confer *real* authority – the effective control over decisions. Indeed, once in place, the owner may prefer to put someone else in charge. We build a formal model to disentangle the resulting trade-offs. We show that acquiring formal authority through ownership rights in the firm is only optimal for a potential owner if the owner expects to assume real authority this way. In other words, formal and real authority are reunited in the equilibrium of our model.

Envision the firm as a supply chain comprised of a producer, P, and a seller, S. An owner, M, holds shares in P and S. If perfect information were available from the start or a comprehensive mechanism were in place to make efficient use of information later on, then neither the notion of a firm nor the notion of authority would be meaningful. Market based transactions between independent units would

[1]Lafontaine and Slade (2007) report that the volume of transactions inside firms in the US is roughly equal to the volume of transactions over markets.

[2]The owner of an asset can use the asset in any way if the relationship falls apart; these residual rights of control shape incentives to invest and, thereby, determine efficient ownership structures.

achieve efficiency; likewise, with perfectly aligned objectives of S, P, and M, it would not matter who makes decisions. This is no longer true when information is privately observed and no sophisticated monetary mechanism is in place to deal with it. This is our starting point. In particular, we consider a situation in which S and P privately observe information about their own costs, and in addition, S – who is closer to customers – gets private information about demand. The new information is relevant to S, P and M to different extents: Since S, P and M would all want to respond to information about demand in the same way, demand corresponds to a *common value* component. In contrast, the individual costs correspond to *private value* components. S and P want to adapt to changes in their own cost but not to changes in the other one's costs. M, who has an interest in the payoffs of S and P, wants to consider changes in all costs but to a smaller degree than S and P do. In this situation, everybody favors different choices and, hence, authority matters.

As the owner of the firm, M has the formal authority to decide or to delegate the decision. Whoever makes the decision can take advice from the others. Because of differences in preferences, information transmission is strategic in our problem, as in Crawford and Sobel (1982). When M makes the decision, M communicates with the units beforehand. We term this *consultative decision-making*. When M delegates decision-making to P, then P can take advice from S. We term this *delegation to P*. When M delegates decision-making to S (*delegation to S*), then there will be no meaningful communication. This is the essence of private value information. Common value information, by contrast, can be shared.

Before we study ownership, we analyze optimal decision-making for a given ownership structure. We determine the optimal mode of decision-making – the allocation of real authority – as a function of three factors: the information about common and private values, the distribution of surplus in the organization, and the riskiness of the environment.

To understand the role of these factors step-by-step, we first look at the pure private value case where demand is commonly known. For a large set of payoff division rules around equal sharing, M ends up making decisions at the optimum; delegating to S or P is optimal only for relatively unequal surplus division rules. Adding an element of common values makes it relatively more attractive to delegate to S and relatively less attractive to delegate to P. The reason is a pecking order in the use of common value information. S, who has direct access to common value information, makes efficient use of that piece of information. M is next in line since M has a comparative advantage relative to P at extracting common value information from S, because M's preferences are better aligned with those of S than P's. By implication, real authority is redistributed from P to M and from M to S. When the common value component gets very large – in an *almost common value* environment – this redistribution is pushed to its limit. By monotonicity, any point of indifference between those arising in the extreme cases of pure private value and almost common value information can arise in some informational environment.

With the allocation of real authority pinned down, we can take one step back and discuss the incentives to acquire ownership – formal authority – in our model. We envision that M makes a take it or leave it offer to P and S, who both need to accept to transfer ownership rights to M. This pins down the equilibrium payoff sharing rule. Two questions arise. When is *some* form of formal authority better than a purely market based transaction? When does an integrated structure with M on top arise? We find that in a pure private value environment, S and P would never integrate on their own, but they may agree to integrate with M on top. In particular, if S and P face the same amount of uncertainty with respect to their costs, then M can acquire the right to make decisions on behalf of all three of them in exchange for a fraction of the payoffs of S and P. The resulting payoff division rule effectively makes M indeed acquire real authority through this arrangement. We also show a partial converse to this result: If S and P benefit to sufficiently different extents from adapting to changes in costs, then there does not exist any arrangement that puts M in charge and is individually rational for both of them. The model can also rationalize which integrated structure is most efficient, if several ones are feasible. In particular, in an environment with quite a lot of common value uncertainty, some form of integration will always occur, either with S or with M on top. The integrated structure with M on top is efficient if and only if it is socially efficient to have M assume real authority.

In sum, our theory provides a connection from real authority to formal authority. In particular, acquiring formal authority is optimal only if formal authority confers real authority to the owner. Hence, we obtain a theory of ownership based on actual rights of control – as opposed to residual rights of control as in Grossman and Hart (1986) and Hart and Moore (1990).

We build on the notion of real authority first analyzed in Aghion and Tirole (1997). The main focus of that work is that delegating formal authority may benefit a principal by providing better incentives for information acquisition for a subordinate.[3] We focus on aspects that are not present in Aghion and Tirole (1997). In particular, we explain the allocation of real authority in the presence of common and private value information in an environment in which some – strategic – communication of that information is possible and explore the consequences for the acquisition of formal authority – i.e., ownership rights. Dessein (2002) has first studied authority in the strategic communication model of Crawford and Sobel (1982) with one sender and one receiver. He shows that whenever meaningful communication is possible at all, the receiver prefers to delegate decision-making to the sender. Even though control is lost, the informational loss through strategic communication is more severe. Our main point of departure from that literature is our take on the conflicts between the participants. Developing our model from scratch in Section 2, we show that in the context of coordinating arrangements in a supply chain, the systematic conflicts assumed in Crawford and Sobel (1982) do not arise. Instead, conflicts based on ex ante known information are eliminated; ex post, conflicts may still arise depending on the realizations of shocks.

The papers that are most closely related to the present one are Alonso et al. (2008) and Rantakari (2008). This work studies the allocation of decision-authority in a multi-divisional organization in which each division needs to take an action – thus, there is a desire for coordinating actions.[4] Such coordination motives arise, e.g., in a *horizontally* integrated firm that is organized in multiple divisions. Each division is represented by a division-manager, so there are no incentive problems within the divisions. In contrast, we analyze how management wants to organize decision-

[3]See Hidir and Migrow (2019) for a more recent contribution showing that a principal may benefit from delegating to a relatively less able agent.

[4]See also Dessein (2013), which views organizational architectures, as we do, as incomplete contracts. Blume et al. (2020) study the interplay of incomplete contracts and cheap talk. See Gibbons et al. (2012), for an overview of approaches to choosing organizational architectures.

making in a single *vertical* chain. In terms of a broader organizational picture, this complements their approach by looking at the incentive problems within a single division. Clearly, the reasons we find for or against keeping decision-authority in the hands of management are unrelated to coordination motives. An interesting next step is to integrate both approaches to study organizations that have both breadth and depth.[5]

We start from the smallest organizational form that may be called a firm and arguably the simplest decision-problem it might face: how much to produce. The problem mandates several ingredients, some – but not all combined – have been studied in the literature, some are entirely new. There are three pieces of information, one common value and two private value components;[6] S has multi-dimensional information: S observes some information about the common value component and some information about the own private value component;[7] and information is noisy and dispersed in the organization and communicated strategically.[8] Moreover, information is collected from many senders with uncorrelated pieces of information.[9]

Technically, our own prior work (Deimen and Szalay (2019)) is most closely connected. In that paper, we look at incentives for information acquisition and communication in a sender-receiver game. Conflicts arise and get eliminated depending on the type of information that the sender acquires. Since communication under con-

[5]See also Liu and Migrow (2019) for an analysis of uncertainty about relative division profits in the framework of Alonso et al. (2008) and Rantakari (2008). Unlike we do, they focus on verifiable information.

[6]This relates to Li and Madarász (2008) and Li (2010), where the sender's bias is unknown.

[7]This contrasts with Levy and Razin (2007) and Chakraborty and Harbaugh (2007) which study one-sender models with high-dimensional actions and information.

[8]Blume et al. (2007) look at the effect of exogenous noise in communication.

[9]Wolinsky (2002) looks at aggregating independent, partially verifiable signals in a setting with and without commitment. Battaglini (2002), Ambrus and Takahashi (2008), and Meyer et al. (forthcoming) study problems with multidimensional states and actions where the senders have the same information. In McGee and Yang (2013), the senders have partial and non-overlapping private information, and their information transmissions exhibit strategic complementarity. Hagenbach and Koessler (2010) study a pure common value problem with known sender-biases in the context of networks with a coordination motive. In contrast, we look at the mixed private and common value case in a three-player game, in which the private value components impact the biases.

flicts works very badly in the fat-tailed environments considered there, the sender prefers to acquire information that aligns incentives completely. To deal with the updating of noisy, multidimensional information in a concise way, we assume an elliptically contoured joint distribution of the state. In the current paper, we focus on thin-tailed environments in the same class of distributions. These are much more conducive to mechanisms involving communication: in fat-tailed environments, M would have less authority. Importantly, breaking up information into common and private value components is new here.

Elliptical distributions are well known in statistics and have been found useful in finance theory, roughly because they are accessible by similar methods as the Normal distribution but are more flexible (see Fang et al. (1990) for an in-depth analysis). There is a huge advantage over the Normal for the analysis of strategic communication, due to the partial pooling such communication induces. For a class of distributions, expected payoffs arising from such partial pooling can be computed in closed form. The most prominent members of this class are the uniform and the Laplace distribution (see, e.g., Kotz et al. (2001)); a single parameter, capturing the mass in the tails, spans all the distributions between them.[10] We have characterized this class in Deimen and Szalay (2019). Since there – for fat-tailed environments – biased communication occurs only off equilibrium path, no proof of existence of equilibria is necessary. We provide this proof here, for the entire class of elliptical distributions with logconcave marginal densities.

The remainder of the paper is organized as follows. In Section 2, we first motivate our setup formally and then present the reduced form model in Section 3. We analyze strategic communication in Section 4. In Section 5, we derive the optimal mode of decision-making and discuss different private and common value constellations. Section 6 discusses ownership. Finally, Section 7 extends to endogenous information and the role of risk. We show that if M can choose what S and P get to know, then real authority is redistributed systematically towards M. In particular, M always decides if surplus is divided equally within the firm. Moreover, we provide comparative statics

[10]While the use of tail risk in economic theory is not yet wide spread, actuarial scientists have long been using the tail conditional expectation function of a distribution – the expected value conditional on truncations to the tail – as a consistent measure of risk (Artzner et al. (1999)).

of the set of payoff division rules for which M has real authority, as a function of the riskiness of the environment. In riskier environments, communication is less effective and hence M has less authority because consultative decision-making relies heavily on communication. Section 8 concludes. Proofs of theorems are in the Appendix, technical proofs of lemmas and propositions are in the Online Appendix.

2 In a nutshell

In this section, we justify the reduced form of our model from scratch that we analyze from Section 3 onwards. The algebraic details matter only for the analysis of optimal ownership in Section 6.

Consider a vertically integrated firm: a parent company, M, owns shares in two subsidiaries, S and P, where S is the sales unit and P is the production unit. Let k_P and k_S denote the marginal costs of units P and S, respectively. The firm faces a linear market demand function, $P(y) = A - y$, where y denotes the quantity that the firm brings to the market. Contracts are in place to coordinate the units as follows. The units P and S split the revenue with shares ω_P and ω_S, $\omega_P + \omega_S = 1$. A transfer price governs trade between the units. As has been shown by Cachon and Lariviere (2005), if the transfer price is set appropriately, then this arrangement is efficient.[11] In particular, it is easy to verify that a transfer price of $\omega_S k_P - \omega_P k_S$ completely aligns the incentives of S and P. The arrangement *coordinates* the supply chain: it does not matter whether S or P decides how much to supply, the quantity that maximizes aggregate profits maximizes also the individual payoffs. Thus, the allocation of decision-rights is irrelevant given an efficient and perfectly coordinated arrangement.

After contracts are fixed, however, demand or cost conditions may change. Let $a = A - k_P - k_S$ denote everything that is commonly known ex ante. Let demand and cost parameters change to $a + \Delta_A$, $k_P - \Delta k_P$, and $k_S - \Delta k_S$, respectively. Suppose that M receives shares of $1 - \delta_P$ and $1 - \delta_S$ of the units' profits. Then, P's and S's objectives are $\delta_P \left(\omega_P y \left(a + \Delta_A - y\right) + \Delta k_P y\right)$ and $\delta_S \left(\omega_S y \left(a + \Delta_A - y\right) + \Delta k_S y\right)$, respectively;

[11] See also the seminal work by Holmström and Tirole (1991).

M receives the residual, $(1-\delta_S)(\omega_S y(a+\Delta_A - y)+\Delta k_S y)+(1-\delta_P)(\omega_P y(a+\Delta_A - y)+\Delta k_P y)$.

The allocation of decision rights now matters. With complete information about demand and cost shocks, S, P, and M would all favor different quantity choices. More precisely, they would all respond to the demand shock in the same way but would all respond to cost shocks differently – demand corresponds to a *common value* component, costs correspond to *private value* components. Authority matters in a second way if the realizations of the shocks are observed by different parties: As is natural, let S observe the realizations of the demand shock and the own cost shock; and let P observe the realization of the own cost shock. The allocation of authority determines what the one who makes the decision knows to begin with and how well this decision-maker can communicate with others.

We turn to the formal analysis, starting from the assumption that a structure with M on top with a given payoff sharing rule is in place. We justify these assumptions in Section 6.

3 Model

There are three players, M, P, and S. A decision to change production by $\Delta y \in \mathbb{R}$ has to be made. The state of the world (x_C, x_P, x_S) is the realization of a random variable (X_C, X_P, X_S) that decomposes into a common value component X_C and two private value components X_P and X_S of P and S, respectively. In the linear demand and cost environment, $x_C = \frac{\Delta_A}{2}$, $x_P = \frac{\Delta k_P}{2\omega_P}$, and $x_S = \frac{\Delta k_S}{2\omega_S}$, and payoffs are constant fractions of

$$\pi_S(\Delta y, x_C, x_S) = \pi_S^* - (\Delta y - (x_C + x_S))^2$$

and

$$\pi_P(\Delta y, x_C, x_S) = \pi_P^* - (\Delta y - (x_C + x_P))^2,$$

where π_S^* and π_P^* correspond to the maximal unit profits that arise if the individually ideal adaptation decisions to change production by $\Delta y_S^* = x_C + x_S$ or $\Delta y_P^* = x_C + x_P$ are taken.

M obtains a constant fraction of a weighted average of the units' payoffs

$$\pi_M(\Delta y, x_C, x_P, x_S) = \lambda \pi_S(\Delta y, x_C, x_S) + (1-\lambda)\pi_P(\Delta y, x_C, x_P),$$

where $\lambda \in [0,1]$ is determined by the payoff sharing rule; it measures the relative importance of S's payoff for M. For the special case in which $\lambda = 1$ ($\lambda = 0$), the interests of M and S (P) coincide. We skip constant factors of proportionality for parsimony, since they do not affect preferences over decisions. We reintroduce these factors below when we discuss ownership in Section 6.[12]

M has the right to choose Δy. M's first-best ideal choice is the weighted sum of S's and P's ideal choices and given by $\Delta y^{fb} = x_C + \lambda x_S + (1-\lambda)\, x_P$. If M knew the state, then M would choose the first-best action. However, M does not know the state. Instead, P and S privately receive some information about the state. In particular, P privately observes a noisy signal $s_P = x_P + \varepsilon_P$ about P's private value component, and S privately observes noisy signals $(s_C, s_S) = (x_C + \varepsilon_C, x_S + \varepsilon_S)$ about the common value component and S's private value component. The noise terms $(\varepsilon_C, \varepsilon_P, \varepsilon_S)$ are the realizations of the random variables (E_C, E_P, E_S).

We assume that the random vector $\boldsymbol{Z} = (X_C, X_P, X_S, E_C, E_P, E_S)$ follows a joint elliptically contoured distribution with finite first and second moments, and with a logconcave marginal density f on appropriate interval supports $\mathcal{S}_i \subseteq \mathbb{R}$, $i \in \{X_C, X_P, X_S, E_C, E_P, E_S\}$. Elliptical distributions owe their name to the fact that the level curves of their densities are elliptical; they have convenient symmetry and linearity properties (that we summarize in Lemma A.1 in the Appendix).[13] Prominent members of the class of elliptical distributions include the Normal distribution, the Laplace distribution, the Uniform distribution, and many more. As we demonstrate below, there is a huge benefit to working with this larger set rather than the Normal only: for other members in this class, the trade-offs in strategic information

[12]We attribute the exact mapping from the linear demand and cost environment to the quadratic loss functions to Alonso and Matouschek (2008), where we have first seen it in a regulation context. To the best of our knowledge, the application to coordinating arrangements in a supply chain is new to the literature.

For completeness, the factors of proportionality are $\delta_S \omega_S$, $\delta_P \omega_P$ and $((1-\delta_S)\,\omega_S + (1-\delta_P)\,\omega_P)$. A natural case is $\delta_S = \delta_P$ and $\omega_S = \omega_P$ in which case $\lambda = \frac{1}{2}$.

[13]The joint distribution is defined by the characteristic function, so that all marginal distributions have the same characteristic (generator) function. The characteristic function is a quadratic form, $\psi(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{\mu})\phi(\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t})$, with covariance matrix $\boldsymbol{\Sigma}$ and some scalar function ϕ. For more details on elliptical distributions see, e.g., Fang et al. (1990).

transmission environments can be analyzed with closed form expressions. Due to the coordinating arrangement, the first moment is zero, $\mathbb{E}[\boldsymbol{Z}] = \mathbf{0}$. The second moments are given by the covariance matrix of $\boldsymbol{Z}$, $\boldsymbol{\Sigma} = diag\left(\sigma_C^2, \sigma_P^2, \sigma_S^2, \sigma_{\varepsilon_C}^2, \sigma_{\varepsilon_P}^2, \sigma_{\varepsilon_S}^2\right) \in \mathbb{R}_+^6$; thus, the state components and the noise terms are all uncorrelated.[14] We assume that the joint distribution is commonly known. Only the signal realizations are private information.

The game unfolds as follows. At the outset, M chooses between three modes of decision-making: consultative, delegation to P, or delegation to S (explained below). Then, the state is realized, and the signals are observed. The remainder of the game depends on the chosen mode of decision-making.

Consultative decision-making. P privately observes the signal realization $s_P = x_P + \varepsilon_P \in \mathcal{S}_{S_P}$, and S privately observes $(s_C, s_S) = (x_C + \varepsilon_C, x_S + \varepsilon_S) \in \mathcal{S}_{S_C} \times \mathcal{S}_{S_S}$. P and S choose what messages $m_P, m_S \in \mathbb{M}$ to send to M. We do not impose restrictions on the message space $\mathbb{M}$. Formally, P's message strategy is a function $M_P : \mathcal{S}_{S_P} \to \mathbb{M}$ and S's message strategy is a function $M_S : \mathcal{S}_{S_C} \times \mathcal{S}_{S_S} \to \mathbb{M}$. After observing the messages, M takes an action. Thus M's action strategy is a function $Y_M : \mathbb{M}^2 \to \mathbb{R}$. We assume that there is no cost of sending messages and M is unable to commit to the strategy Y_M as a function of the information received – i.e., communication is modeled as cheap talk in the sense of Crawford and Sobel (1982).

Delegation to P. Given the privately observed signal realizations $(s_C, s_S) \in \mathcal{S}_{S_C} \times \mathcal{S}_{S_S}$, S chooses what message $m \in \mathbb{M}$ to send to P. Formally, S's strategy is a function $M_s : \mathcal{S}_{S_C} \times \mathcal{S}_{S_S} \to \mathbb{M}$. P's strategy is to choose an action as a function the own signal realization and S's message, $Y_P : \mathcal{S}_{S_P} \times \mathbb{M} \to \mathbb{R}$. As under consultative decision-making, we assume cheap talk communication.

Delegation to S. Given the privately observed signal realizations $(s_C, s_S) \in \mathcal{S}_{S_C} \times \mathcal{S}_{S_S}$, S chooses what action to take, $Y_S : \mathcal{S}_{S_C} \times \mathcal{S}_{S_S} \to \mathbb{R}$. There is no communication, because P has no relevant information from S's perspective.

We solve for Bayesian equilibria of the game. For each message, the receiving party forms a belief over the types who might have sent the message. The belief

[14] We define the generalized inverse of a degenerate covariance matrix as inverse of the full rank matrix while keeping the zero entries zero.

is derived from the prior and the sending parties' strategies. The receiving party's strategy maximizes the payoff given the belief and the sending parties' strategy. Likewise, the sending parties' message strategies maximize their payoffs given the receiving party's strategy.

4 Strategic communication

Under consultative decision-making, P and S communicate with M; under delegated decision-making to P, S communicates with P. Naturally, pure common value information could be shared easily, but the private value components give rise to conflicts. As long as interests are only partially aligned, we expect that the information transmitted is garbled. We have to deal with two types of garbling. First, the action is one-dimensional but S obtains a two-dimensional signal; S optimally aggregates the signals to send a one-dimensional message. Second, as is standard in strategic communication, information is partially pooled into an interval partition of the state space and recommendations are coarse: instead of revealing the precise realizations of the signals, P and S only communicate the respective partition intervals.

4.1 Merging signals into recommendations

It is intuitive that P and S only reveal information that matters for their optimal actions. Since we assume quadratic loss functions, the optimal actions are the posterior means. Moreover, in our statistical environment, the posterior means are linear functions of the signals.[15] We define S's and P's posterior means as

$$\theta_S := \mathbb{E}\left[X_C + X_S \middle| (S_C, S_S) = (s_C, s_S)\right] = \frac{\sigma_C^2}{\sigma_C^2 + \sigma_{\varepsilon_C}^2} s_C + \frac{\sigma_S^2}{\sigma_S^2 + \sigma_{\varepsilon_S}^2} s_S,$$
$$\theta_P := \mathbb{E}\left[X_P | S_P = s_P\right] = \frac{\sigma_P^2}{\sigma_P^2 + \sigma_{\varepsilon_P}^2} s_P.$$

S's (P's) interim expected utility satisfies the single-crossing condition in Δy and

[15] Lemma A.1 in the Appendix gathers the properties of elliptical distributions that we refer to in this section.

θ_S (θ_P).[16] Moreover, the level of the posterior mean $\theta = \theta_P, \theta_S$ is the only statistic of the posterior distribution that interacts with the action Δy. Hence, it is natural – and proven formally in Deimen and Szalay (2019) – that without loss of generality we can describe all equilibria of the communication games in which S communicates, in terms of communication about θ_S only; naturally, P communicates about θ_P only. The following observation is useful to understand the informational content of communication.

Observation 1 *θ_S and θ_P follow an elliptical distribution with the same characteristic generator with mean vector zero as $\boldsymbol{Z}$.*

Variances and covariances play an important role in our analysis, as they measure informational contents. For readability, we introduce the following notation:

$$c := \frac{\sigma_C^2}{\sigma_C^2 + \sigma_{\varepsilon_C}^2}\sigma_C^2, \quad s := \frac{\sigma_S^2}{\sigma_S^2 + \sigma_{\varepsilon_S}^2}\sigma_S^2, \quad p := \frac{\sigma_P^2}{\sigma_P^2 + \sigma_{\varepsilon_P}^2}\sigma_P^2,$$

such that c represents the amount of common value information and s and p the private value information of S and P, respectively. The variances of Θ_S and Θ_P are given by $var(\Theta_S) = c + s$ and $var(\Theta_P) = p$; the covariances can be calculated from $cov\,(X_i, \Theta_i) = var\,(\Theta_i)$ for $i = S, P$ and by recalling non-correlation. For example, the covariance of M's optimal action and S's posterior Θ_S is given by $cov\,(X_C + \lambda X_S + (1-\lambda)X_P, \Theta_S) = c + \lambda s$, it measures the informational content of Θ_S for M.

4.2 Conflicting interests

Consider now the optimal actions from the receiver of some communication. Let $\theta = \theta_S, \theta_P$ be the realization of the state $\Theta = \Theta_S, \Theta_P$ that is communicated. To fix ideas, consider truthful non-strategic communication of conditional means. Under

[16] S's interim expected utility takes the following form $\mathbb{E}\left[-\left(\Delta y - (X_C + X_S)\right)^2 \mid (S_C, S_S) = (s_C, s_S)\right] = -(\Delta y)^2 + 2\Delta y \theta_S - \mathbb{E}\left[X_C^2 + X_S^2 \mid (S_C, S_S) = (s_C, s_S)\right]$. Similarly, for P, $\mathbb{E}\left[-\left(\Delta y - (X_C + X_P)\right)^2 \mid S_P = s_P\right] = -(\Delta y)^2 + 2\Delta y \theta_P - \mathbb{E}\left[X_C^2 + X_P^2 \mid S_P = s_P\right]$.

consultative decision-making, the recommendations θ_P, θ_S would induce an ideal action of M of

$$\mathbb{E}\left[X_C + \lambda X_S + (1-\lambda)X_P\right| (\Theta_S, \Theta_P) = (\theta_S, \theta_P)] = \beta_{MS} \cdot \theta_S + \beta_{MP} \cdot \theta_P,$$

where

$$\beta_{MS} := \frac{c + \lambda s}{c+s} \text{ and } \beta_{MP} := (1-\lambda). \tag{1}$$

We term the regression coefficients $\beta_{MS} = \frac{cov(X_C+\lambda X_S, \Theta_S)}{var(\Theta_S)}$ and $\beta_{MP} = \frac{cov(X_C+(1-\lambda)X_P, \Theta_P)}{var(\Theta_P)}$ the *sensitivities* of M relative to S and P, respectively. The sensitivities reflect how much a receiver is inclined to follow a sender's advice.

Likewise, under delegated decision-making to P, P's optimal decision function based on S's recommendation θ_S and the own information s_P is

$$\mathbb{E}\left[X_C + X_P | \Theta_S = \theta_S, S_P = s_P\right] = \beta_{PS} \cdot \theta_S + \theta_P,$$

where

$$\beta_{PS} := \frac{c}{c+s}. \tag{2}$$

Here, P's sensitivity with respect to S's recommendation is the regression coefficient $\beta_{PS} = \frac{cov(X_C+X_P, \Theta_S)}{var(\Theta_S)}$.[17] Note that the players agree in expectation on the optimal choice only for $\beta = 1$, where $\beta \in \{\beta_{MS}, \beta_{MP}, \beta_{PS}\}$. Thus, the sensitivities measure the conflicts of interest and how much information can be communicated when communication is strategic. In the terminology of the literature, we have a state-dependent bias $(1-\beta) \cdot \theta$.

Due to single crossing, all strategic communication equilibria either take the form of interval partitions of the Θ-space or are fully revealing, mixtures of these two cannot occur. Conveniently, updating conditional on a partition interval remains a linear function in our model. For example,

$$\mathbb{E}\left[X_C \middle| \Theta_S \in \left[\underline{\theta}, \overline{\theta}\right]\right] = \beta_{PS} \cdot \mathbb{E}\left[\Theta_S \middle| \Theta_S \in \left[\underline{\theta}, \overline{\theta}\right]\right].$$

Since the states are uncorrelated, S's signals are uninformative about x_P and P's signal is uninformative about x_S and x_C.

[17] Under delegated decision-making to S, no meaningful communication from P to S is possible, since $\beta_{SP} = 0$.

4.3 Communication equilibria

As is standard in cheap talk, partitional equilibria are characterized by a sequence of indifference types, $\{a_i^n\}_i$, with $a_{i-1}^n < a_i^n$ and n relating to the number of induced actions. Types strictly within an interval, $\left(a_{i-1}^n, a_i^n\right)$, induce the same expected action; types on the boundaries are indifferent between inducing the action in the interval below or the action in the interval above. For any finite number of induced actions, equilibria are symmetric in our model.[18] For notational simplicity, we, therefore take $a_i^n \geq 0$ and denote the indifference types below zero by $-a_i^n$ for all i and n. The description of communication equilibria (Proposition 1) does not depend on who exactly communicates. We, therefore, skip the indices for P and S whenever possible, in this section. We define the conditional expectations for a given interval $\left[a_{i-1}^n, a_i^n\right)$ by

$$\mu_i^n := \mathbb{E}\left[\Theta \middle| \, \Theta \in \left[a_{i-1}^n, a_i^n\right)\right] \quad \text{for } i = 1, \ldots, n \text{ and } \quad \mu_{n+1}^n := \mathbb{E}\left[\Theta \middle| \, \Theta \geq a_n^n\right].$$

Thus, the expected action in the $i - th$ interval above zero is $\beta \cdot \mu_i^n$, with $\beta \in \{\beta_{MS}, \beta_{MP}, \beta_{PS}\}$ and $\mu_i^n \in \left\{\mu_{\theta_S,i}^n, \mu_{\theta_P,i}^n\right\}$. We define the random variables μ_S and μ_P of truncated expectations that have supports $\left\{\mu_{\theta_S,i}^n\right\}_i$ and $\left\{\mu_{\theta_P,i}^n\right\}_i$. These discrete random variables are important for the calculation of the value of communication: to compute expected payoffs, we need to determine the moments $var\left(\mu_S\right)$ and $var\left(\mu_P\right)$ from the marginal distributions of Θ_S and Θ_P and the equilibrium characterization.

The indifference conditions of marginal types that determine partitional equilibria are

$$a_i^n - \beta \cdot \mu_i^n = \beta \cdot \mu_{i+1}^n - a_i^n, \quad \text{for } i = 1, \ldots, n. \tag{3}$$

Symmetric equilibria come in two classes, depending on whether the total number of induced actions is even or odd. In an equilibrium with an even number of actions, type $\theta = 0$ must be a threshold type. We call this type of equilibrium a *Class I* equilibrium, and the characterization uses $a_0^n = 0$. If the total number of induced

[18] By symmetry of distributions and loss functions, symmetric equilibria exist. Moreover, logconcavity implies that the equilibrium partition of S types is unique (see Szalay (2012)). Hence, only symmetric equilibria exist. For a related argument – that logconcavity implies uniqueness – in a dynamic context, see Meyer-ter Vehn et al. (2018).

actions is odd, then a symmetric interval around zero is part of the equilibrium. We call this a *Class II* equilibrium. In this case, we omit a_0^n from the construction.

Proposition 1 *For all elliptical distributions with a logconcave marginal density the following holds.*
i) For all n, there exist an essentially unique equilibrium, which is symmetric and induces $2(n+1)$ actions (Class I), and an essentially unique equilibrium, which is symmetric and induces $2n+1$ actions (Class II).
ii) For $n \to \infty$, the limits of the finite Class I and Class II equilibria exist and correspond to equilibria inducing infinitely many actions (limit equilibria).
iii) Within any of the two classes of equilibria, the sequence of first thresholds above zero $\{a_1^n\}_n$ satisfies $\lim_{n\to\infty} a_1^n = 0$.

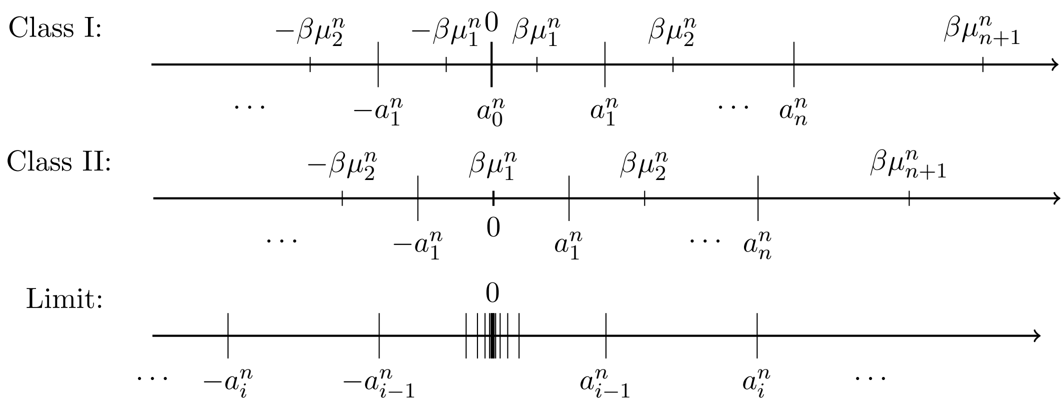

Figure 1: Partitional equilibria. Intervals around the prior mean $\mathbb{E}[\Theta] = 0$ get arbitrarily small as $n \to \infty$.

Proposition 1 proves the existence of partitional equilibria for arbitrary n.[19] Moreover, it proves that the limit as $n \to \infty$ also is an equilibrium. While the partitional

[19]For the proof, we take equilibria as a combination of a "forward solution" and a "closure condition". A forward solution that starts at a_0, takes the length of the first interval, say t, as given, and computes the "next" threshold, $a_2(t)$, as a function of the preceding two, t and a_0. Likewise, all following thresholds are constructed using their two predecessors. The closure condition for an equilibrium with n positive thresholds requires that t is such that type $a_n^n(t)$ satisfies the indifference

form of equilibria is known from the literature, (e.g., Crawford and Sobel (1982), Gordon (2010)), it is typically assumed that the state space is a compact interval. In contrast, we allow for the case of an unbounded state space. For the Laplace distribution, a characterization of partitional equilibria is shown in Deimen and Szalay (2019). Proposition 1 generalizes the characterization to all elliptical distributions with a logconcave marginal density. For an illustration, see Figure 1.

The take-away for the analysis that follows is that limit equilibria always exist. Moreover, it is standard in the literature to focus on the equilibrium with the highest number of partition elements, because all players unanimously prefer this equilibrium over any other equilibrium from an ex ante perspective. We can, therefore, meaningfully compare different forms of decision-making, since we can stick to the same type of equilibrium in the communication subgames.

We now turn to the moments $var\left(\mu_S\right)$ and $var\left(\mu_P\right)$, which are needed for expected equilibrium payoffs. To determine these values in closed form, we make the following assumption.

Assumption 1 *The marginal density f has linear tail conditional expectations: let $\alpha \in \left[\frac{1}{2}, 1\right]$, for any $\overline{\theta} \in \left[0, \overline{\mathcal{S}}_\Theta\right]$, we have*

$$\mathbb{E}\left[\Theta|\Theta \geq \overline{\theta}\right] = \mathbb{E}\left[\Theta|\Theta \geq 0\right] + \alpha \cdot \overline{\theta}. \tag{4}$$

The class of elliptical distributions with linear tail conditional expectations is well-defined for our purposes for $\alpha \in (0, 2)$.[20] For $\alpha \in \left[\frac{1}{2}, 1\right]$, the distribution has a logconcave density. Note that the uniform distribution features $\alpha = \frac{1}{2}$, at one extreme, and the Laplace distribution $\alpha = 1$, at the other extreme.

condition. We prove existence of an equilibrium for any n, and logconcavity of the density ensures uniqueness.

[20]The distribution per se is well-defined more generally, the second moment is finite only for $\alpha < 2$. The covariance matrix is finite if and only if the variances of the marginals are finite. The joint distribution is defined using the characteristic function of the marginal distribution, extended to the multivariate case. Deimen and Szalay (2019) introduce this class of distributions. Importantly, in that paper, the focus is on distributions with logconvex half-support distributions, $\alpha \geq 1$.

Proposition 2 *Under Assumption 1, the variance of* μ_θ *in a limit equilibrium is given by*

$$var(\mu_\theta) = \frac{2-\alpha}{2-\beta\cdot\alpha} var(\Theta). \tag{5}$$

The proposition follows from the conjunction of two results: first, the existence of equilibria for logconcave distributions is given in Proposition 1; second, Deimen and Szalay (2019) show that expression (5) provides an upper bound for the variance of μ_θ. The exact value is attained in the most informative equilibrium and this exists if the distribution has a logconcave density.

Naturally, $var(\mu_\theta) \leq var(\Theta)$; the fraction $\frac{2-\alpha}{2-\beta\cdot\alpha}$ – that we call *effectiveness of biased communication* – reaches unity exactly if $\beta = 1$, that is, if interests are perfectly aligned. For a given $\beta < 1$, the effectiveness is decreasing in α. The parameter α is a measure for the mass in the tails of the distribution. Hence, the variance is smaller and communication is less effective if the distribution has more mass in the tails – i.e., has more tail-risk.[21] The intuition for this argument is the following: more mass in the tail of the distribution means that extreme realizations of the state are more likely. The bias is increasing in the state and, thus, very large at extreme realizations of the state. Hence, under a distribution with high tail risk, large disagreement is more likely and communication is less effective. For an illustration, see Figure 2.

We maintain Assumption 1 for the remainder of the paper.

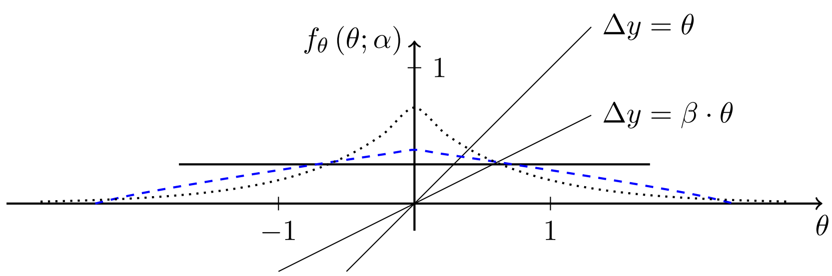

Figure 2: The one-dimensional density depicted for $\alpha = 0.5$ uniform (solid), $\alpha = 0.65$ (dashed), and $\alpha = 1$ Laplace (dotted), all for a variance of one.

[21] The tail conditional expectation is a well known risk measure (see Artzner et al. (1999)).

5 The optimal mode of decision-making

M's payoff, π_M, can be written as an *expected payoff gain*, $\Delta\pi_M$, net of prior uncertainty, $\sigma_C^2 + \lambda\sigma_S^2 + (1-\lambda)\,\sigma_P^2$. To compare modes of decision-making, we focus on the expected payoff gains.

While unrealistic as an institution, the first-best scenario is a useful benchmark. M would like to observe the signal realization (s_C, s_P, s_S) directly and, then, take an action that maximizes the payoff, $\Delta y^{fb} = \frac{\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2} s_C + \lambda \frac{\sigma_S^2}{\sigma_S^2+\sigma_{\varepsilon_S}^2} s_S + (1-\lambda)\,\frac{\sigma_P^2}{\sigma_P^2+\sigma_{\varepsilon_P}^2} s_P$. It is straightforward to show that M's expected payoff gain in this first-best scenario is

$$\Delta\pi_M^{fb} = c + \lambda^2 s + (1-\lambda)^2 p.$$

Under delegation to S, S's optimal adaptation Δy^{delS}, determined by the own signals s_C, s_S, is simply S's posterior mean. Under delegation to P, P's optimal adaptation Δy^{delP} is derived conditional on the own signal s_P and on a message send by S that reveals that $\theta_S \in [a_{S,i-1}, a_{S,i}]$. Under consultative decision-making, M makes the adaptation Δy^{con} based on recommendations from P and S: P reveals that Θ_P has realized in some interval, $\theta_P \in [a_{P,j-1}, a_{P,j}]$, and S reveals that Θ_S has realized in some interval, $\theta_S \in [a_{S,i-1}, a_{S,i}]$. Formally, the respective optimal adaptations are

$$\begin{aligned}
\Delta y^{delS} &= \theta_S, \\
\Delta y^{delP} &= \beta_{PS} \cdot (\mu_S)_i + \theta_P, \\
\Delta y^{con} &= \beta_{MS} \cdot (\mu_S)_i + \beta_{MP} \cdot (\mu_P)_j\,,
\end{aligned}$$

where the sensitivities β_{MS} and β_{MP}, and β_{PS} are defined in equations (1) and (2), and the realizations of the truncated expectations are given by $(\mu_S)_i$ and $(\mu_P)_j$.

These optimal adaptation decisions result in the following expected payoff gains for M.

Lemma 1 *M's expected payoff gains under delegation to S, consultative decision-*

making, and delegation to P are

$$\begin{aligned}
\Delta\pi_M^{delS} =& c + (2\lambda - 1)s, \\
\Delta\pi_M^{con} =& \beta_{MS}^2 \frac{2-\alpha}{2-\alpha\beta_{MS}}(c+s) + (1-\lambda)^2 \frac{2-\alpha}{2-\alpha\beta_{MP}} p, \\
\Delta\pi_M^{delP} =& (2\lambda + (1-2\lambda)\beta_{PS}) \frac{2-\alpha}{2-\alpha\beta_{PS}} c + (1-2\lambda)p.
\end{aligned}$$

To get an intuition for these expressions, consider, first, the payoff gain from delegation to S. Since everyone benefits from common value information and S has direct access to this information, $\Delta\pi_M^{delS}$ increases one for one with c. In contrast, the amount of private value information s is a gain to S which counts with weight λ, but a loss from P's perspective which counts with weight $1-\lambda$, adding up to the factor $2\lambda - 1$. Consider, next, the gain from consultative decision-making. The total amount of information held by S is captured by the variance of Θ_S; M would like to follow that information with sensitivity β_{MS}, thus generating a maximal value of $\beta_{MS}^2(c+s)$. However, due to strategic information transmission, only the fraction $\frac{(2-\alpha)}{2-\alpha\beta_{MS}}$ of this information gets through to M (see equation (5)). Finally, the value of delegating to P can be decomposed by the same logic. This is a bit more intricate due to the fact that S benefits to a higher degree than P, if P is convinced to base the decision on information provided by S.

Suppose that we only have a common value but no private values, $s = p = 0$. In this case, we can directly see from Lemma 1 that the payoffs from all three modes of decision-making coincide with the first-best payoff. For the remainder of the paper, we assume

Assumption 2 *There is some private value element, $s > 0$ or $p > 0$.*

The comparison of the three modes of decision-making yields the following characterization.

Proposition 3 *There exist $\lambda_P(c, s, p, \alpha), \lambda_S(c, s, p, \alpha)$ with $0 \leq \lambda_P \leq \lambda_S \leq 1$ such that*

i) Delegation to P is optimal if and only if we have $\lambda \in [0, \lambda_P]$ and $\lambda_S > 0$.

ii) Consultative decision-making is optimal if and only if we have $\lambda \in [\lambda_P, \lambda_S] \cup \{1\}$ *or* $\lambda = 0$ *and* $\lambda_S > 0$.

iii) Delegation to S is optimal if and only if $\lambda \in [\lambda_S, 1]$.

We first compare the modes of decision-making pairwise; then, we determine the overall winner as a function of λ. Key properties that we use are that $\Delta\pi_M^{delS}$ is linearly increasing, $\Delta\pi_M^{delP}$ is linearly decreasing, and $\Delta\pi_M^{con}$ is convex in λ. Moreover, consultative decision-making coincides with delegation to P in $\lambda = 0$ and with delegation to S in $\lambda = 1$.

The points of indifference of M between delegation to P (S) and consultative decision-making at λ_P (λ_S) are the objects of our interest, in what follows. It is important to note that the regions described in Proposition 3 can be empty. If, for example, $\lambda_S = 0$, then delegation to S is the only mode that is chosen by M.

We now turn to the comparative statics of the indifference points with respect to the amount of common value information. We trace out the consequences of the informational environment for the allocation of real authority – effective control over decisions – in our model. In turn, this provides the key building block for the optimal acquisition of formal authority – ownership rights.

5.1 Private values

In the pure private value environment, with $c = 0$, the payoff gains under delegation to S and P simplify to $(2\lambda - 1)s$ and $(1 - 2\lambda)\,p$, respectively. Information used by S constitutes pure noise from the perspective of P and vice versa. In the absence of a common value component, β_{MS} reduces to λ, and under consultative decision-making, the payoff gain from communicating with S simplifies to $\lambda^2 \frac{2-\alpha}{2-\alpha\lambda} s$. This is isomorphic to the payoff gain from communicating with P, $(1-\lambda)^2 \frac{2-\alpha}{2-\alpha(1-\lambda)} p$. Clearly, there is no fundamental difference between S and P, and we expect the comparison between delegating to any one of them and consultative decision-making to be symmetric. Formally, the choice between the modes of decision-making depends on

$$\left(\Delta\pi_M^{con} - \Delta\pi_M^{delS}\right)\Big|_{c=0} = (1-\lambda)\left(\frac{2\,(1-\lambda) - \alpha\lambda}{2-\alpha\lambda} s + \frac{2\,(1-\lambda) - \alpha\,(1-\lambda)}{2-\alpha\,(1-\lambda)} p\right)$$

and

$$\left(\Delta\pi_M^{con} - \Delta\pi_M^{delP}\right)\big|_{c=0} = \lambda\left(\frac{2\lambda - \alpha\lambda}{2 - \alpha\lambda}s + \frac{2\lambda - \alpha\,(1-\lambda)}{2 - \alpha\,(1-\lambda)}p\right).$$

The comparison yields the following result.

Theorem 1 *Consider the pure private-value environment with $c = 0$. If $p, s > 0$, then $0 < \lambda_P < \lambda_S < 1$. Moreover, λ_P and λ_S are increasing in p and decreasing in s. If $p = 0$, then $\lambda_P = 0$ and $\lambda_S = \frac{2}{2+\alpha}$. If $s = 0$, then $\lambda_S = 1$ and $\lambda_P = \frac{\alpha}{2+\alpha}$.*

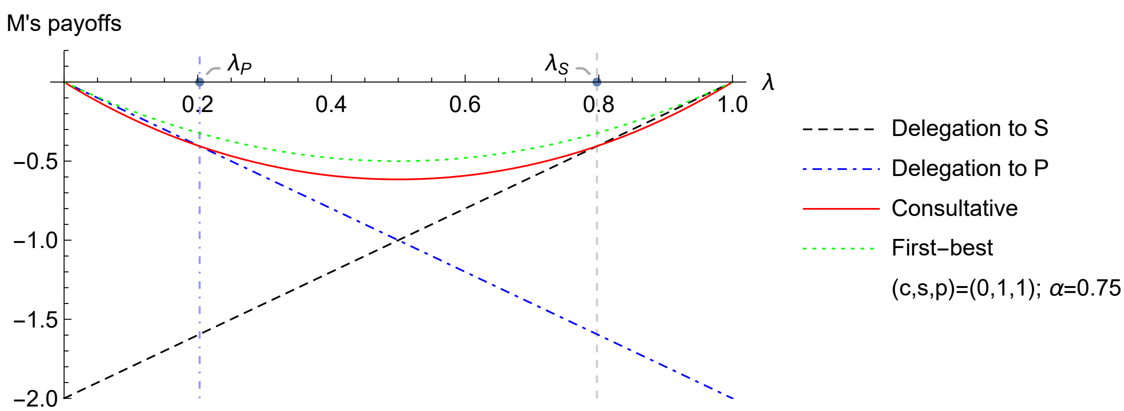

Figure 3: Payoffs for pure private values with $c = 0, s = p = 1$, $\alpha = 0.75$.

We illustrate M's payoffs as functions of the payoff weight λ, in Figure 3. The vertical lines in the figure depict the intersections $\lambda_P < \lambda_S$ of the payoffs under consultative decision-making and delegation to P and S, respectively – i.e., the switching points of the optimal modes of decision-making.

The allocation of real authority is jointly determined by the division of surplus within the organization, λ, and the information held by P and S. At the optimum, P makes the decision if λ is close to zero, S makes the decision if λ is close to one, and M decides if the surplus division rule is relatively balanced. The precise arrangement depends on the relative amount of information held by P and S, and on the riskiness of the environment, α. Naturally, the set of payoff weights for which P (S) is in charge is larger – P (S) has more real authority – if P (S) provides relatively more information. The pure private value setting is quite conducive to real authority on the part of M. M always has real authority in a substantial set of environments, around a relatively equal division of surplus within the organization: regardless of the information held by the units P and S, M optimally decides in the range $\lambda \in \left(\frac{\alpha}{2+\alpha}, \frac{2}{2+\alpha}\right)$. This range depends on the effectiveness of communication that is determined by the risk parameter α. In the uniform environment ($\alpha = \frac{1}{2}$), in which communication works very well, M has real authority in a range $\left(\frac{1}{5}, \frac{4}{5}\right)$. In the Laplace environment ($\alpha = 1$), in which communication works very badly, this range shrinks to $\left(\frac{1}{3}, \frac{2}{3}\right)$.[22]

5.2 Mostly private values and some common value

Suppose now that there is in addition uncertainty about the common value component. Moreover, suppose for now that the common value uncertainty is relatively low. In this environment, each party retains some real authority. The optimal switching points λ_P, λ_S depend monotonically on the amount of common value information:

Theorem 2 *Suppose that* $c > 0$ *and that* $0 < \lambda_P < \lambda_S < 1$. *Then,* λ_P *and* λ_S *are decreasing in c: common values shift authority from P to M and from M to S.*

Intuitively, more common value information crowds out mechanisms that have a comparative disadvantage at extracting this type of information. S is the one who directly observes the common value signal. Therefore, S has a comparative advantage relative to M and, as a result, delegation to S dominates consultative decision-making more often. Similarly, the conflict between M and S is smaller than the conflict between P and S so that consultative decision-making works relatively better than delegation to P. As a result of these forces, M gains more authority relative to P but loses ground relative to S, who observes the common value component directly.

[22]The private value case is focal in the existent literature. The trade-offs we just described are those identified by Dessein (2002). For λ close to one, the interests of S and M, while for λ close to zero, the interests of P and M get closely aligned. For a small bias, delegating real authority to the informed party is preferred to communicating with the informed party: the loss of control – associated with giving away decision rights – is smaller than the loss of information – associated with communicating strategically with the informed party.

5.3 Almost common values

If we increase the common value component c sufficiently, this becomes the most important piece of information. As a consequence, the organizational performance predominantly depends on how well common value information impacts decision-making. We find that

$$\lim_{c\to\infty}\left(\Delta\pi_M^{con}-\Delta\pi_M^{delS}\right)=(1-\lambda)\left(-\frac{\alpha}{2-\alpha}s+\frac{2(1-\lambda)-\alpha(1-\lambda)}{2-\alpha(1-\lambda)}p\right) \tag{6}$$

and

$$\lim_{c\to\infty}\left(\Delta\pi_M^{con}-\Delta\pi_M^{delP}\right)=\lambda\left(\frac{\alpha}{2-\alpha}s+\frac{2\lambda-\alpha(1-\lambda)}{2-\alpha(1-\lambda)}p\right). \tag{7}$$

Any institution makes efficient use of common value information in the limit, since in the limit as c gets large, the conflicts in any communication vanish. Since, at the same time, the value of using this information grows large, by l'Hôpital's rule, the relative performance of the institutions is driven by the difference in speed at which the institutions approach efficient use of the information c. Delegating decision-making to S evidently makes the best use of common-value information and beats consultative decision-making in speed by $\frac{\alpha}{2-\alpha}(1-\lambda)s$, in the limit. Likewise, consultative decision-making makes a better use of common-value information than delegating to P and beats the latter in speed by $\frac{\alpha}{2-\alpha}\lambda s$, in the limit. We note that the differences disappear when the objective of M coincides with the objectives of S or P. Moreover, the differences are more pronounced in environments with less effective communication – i.e., a high value of α amplifies the differences in performance. The comparative advantage of delegating to S or P relative to consultative decision-making is determined by the terms in brackets – a weighted difference of private information of S and P. Delegating to S (P) becomes relatively more (less) attractive compared to consultative decision-making if λ is higher. It is even possible that institutions get completely crowded out.

Theorem 3 *Consider the almost common value environment in which $p, s > 0$ and $c \to \infty$.*

i) For $p \leq \frac{\alpha}{2-\alpha}s$, we have $\lambda_S = 0$.

ii) For $\frac{\alpha}{2-\alpha}s < p \leq s$*, we have* $\lambda_P = 0$ *and* $\lambda_S = \frac{1}{\frac{s}{p}\frac{\alpha^2}{(2-\alpha)^2}+1}\left(1 - \frac{s}{p}\frac{\alpha}{2-\alpha}\right) > 0.$

iii) For $p > s$*, we have* λ_S *as in (ii) and* $\lambda_P = \frac{1}{\frac{s}{p}\frac{\alpha}{2-\alpha}+\frac{\alpha+2}{\alpha}}\left(1 - \frac{s}{p}\right) > 0.$

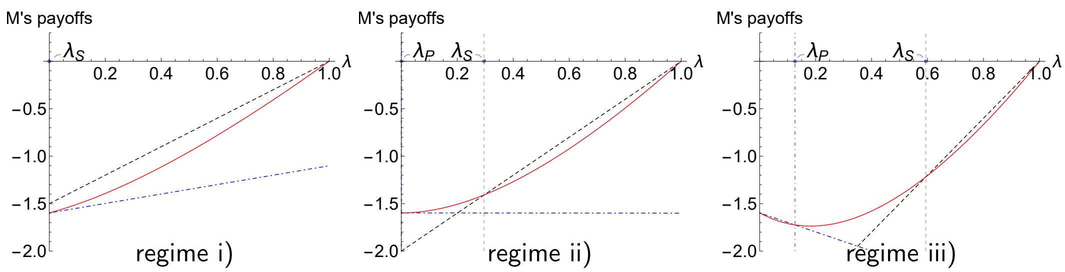

Figure 4: M's payoffs under delegation to P (black, dashed), delegation to S (blue, dash-dotted), and consultative decision-making (red, solid) for almost common values with $c = 1000$, $s = 1$, and $p = 0.5$ (left panel), $p = 1$ (central panel), and $p = 2$ (right panel); $\alpha = 0.75$.

We depict the three regimes in Figure 4. Again, the relative endowment of S and P with private value information determines the optimal allocation of real authority. There are two effects. First, the relative endowment determines parameter bounds beyond which some mechanisms get crowded out completely (extensive margins). Second, the points of indifference between any two mechanisms are affected by the relative endowment with information (intensive margins). Consider first the extensive margins. As p increases relative to s, the optimum changes from regime i) – only delegation to S – to regime ii) – including consultative decision-making – to regime iii) – also including delegation to P. The extensive margins depend on $\frac{\alpha}{2-\alpha}$ which in turn is determined by the tail-risk in the distribution. In regime i), S is essentially an informational monopolist and, therefore, S optimally has real authority for all λ. Any attempt to temper with S's authority results in a less efficient use of common value information, and the resulting loss cannot be offset by gains from using P's information in decision-making. In regime ii), P has more substantial information to contribute. Since the conflicts between S and P are too pronounced relative to the informational gain, the only one who can step in to reach a compromise is M. Finally, if P really has substantial information to contribute, in regime iii), it becomes optimal to delegate to P for λ low enough.

Consider now the intensive margins. In regimes ii) and iii), there is an interior point of indifference between delegation to S and consultative decision-making. Delegation to P is part of the picture only in regime iii). Gradually increasing p redistributes real authority from S to M and from M to P. The first effect seems more relevant: P's authority never extends beyond the upper bound obtained in the pure private value case, $\frac{\alpha}{2+\alpha}$, which is strictly below one half. Hence, P is never in charge for balanced surplus division rules.[23]

In sum, whether M finds it optimal to make decisions or to delegate to someone else depends crucially on the informational environment and the payoff sharing rule. Ultimately, the payoff sharing rule itself depends on the informational environment. Therefore, we next address whether and in which environments M wants to acquire ownership rights.

6 Endogenous ownership

When is it optimal for M to acquire ownership rights? Our analysis reveals two key economic roles that ownership by M can play. First, by creating an integrated structure, M can make adaptation possible where it would otherwise not occur. Second, M can improve efficiency by creating an integrated structure, where integration would be possible already without M. Finally, our analysis also reveals when integration is not feasible.

We impose two assumptions for the analysis, in this section. First, we assume that P and S have the same bargaining power, which implies that their initial revenue sharing arrangement is fifty-fifty. Any alternative assumption would imply that we explain the power institution "ownership" by another notion of power. We believe that this would merely shift the locus of ignorance but would not really explain more. Second, for simplicity, we assume that the signals are perfect.

[23]Note that our findings are different from Aghion and Tirole (1997), where more information increases the real authority of a party. Here, more information in the hands of P redistributes real authority primarily from S to M.

6.1 M makes integration possible

For $\sigma_C^2 < \sigma_S^2$, S and P cannot get to terms without the involvment of M. For convenience, we illustrate this insight in a pure private values environment, $\sigma_C^2 = 0$. Here, the status quo option for S and P is non-integration. The reason is easy to see. The initial agreement is optimal, based on ex ante known information. The information that each of them receives later on concerns the own costs but – by the nature of private value information – not the costs of the other one. Imagine that S (P) is allowed to change the amount of production later on. From the perspective of P (S), this makes the actual production choice more variable. Due to the concave loss functions, this amounts to increased risk with no benefit to it, because the change in production is unrelated to variation in P's (S's) costs.

Suppose that M makes a take-it-or-leave-it offer to S and P individually: M acquires the right to make decisions and obtains profit shares of S and P. Suppose, moreover, that S and P each have a veto right – they both have to accept for the deal to go along.

Theorem 4 *Consider a pure private-value environment $\sigma_C^2 = 0$, fully revealing signals $\sigma_{\varepsilon_S}^2 = \sigma_{\varepsilon_P}^2 = 0$, equal uncertainty $\sigma_S^2 = \sigma_P^2 = \sigma^2$, and a fifty-fifty revenue sharing arrangement between S and P, $\omega_S = \omega_P$. M optimally acquires the right to make decisions with shares $\delta_S = \delta_P = \delta^*$, where δ^* satisfies*

$$\frac{\delta^*}{1-\delta^*} = \frac{\Pi}{\frac{1}{2}\frac{2-\alpha}{2-\alpha\frac{1}{2}}\sigma^2},$$

where Π *is firm profit with ex ante information.*
The optimal surplus division rule, δ^*, implies that $\lambda^* = \frac{1}{2}$, so that real authority is transferred to M.

In the pure private value environment with symmetric uncertainty, the optimal organizational structure is an integrated structure with M on top. Moreover, M's objective is endogenously balanced, $\lambda^* = \frac{1}{2}$. M plays a vital role: M is the reason why S and P can integrate. Bringing in M, serves as a commitment device for S and P to adapt to changes in their costs to some extent. Crucially, the commitment is credible, because the equilibrium sharing arrangement makes M want to keep real authority once endowed with formal authority.

Conceptually, it is easy to generalize this insight to environments with $0 < \sigma_C^2 < \sigma_S^2$. The technical difficulty is that in such environments, one needs to verify whether there exist acceptable deals for S and P and whether such deals indeed put M in charge. This is a relatively complex task. Since the economic role of M is the same as in the pure private value case – making integration possible – we abstain from a detailed discussion.

6.2 Efficient Integration: with or without M

Suppose that $\sigma_C^2 > \sigma_S^2$. In this case, S and M compete for the right to make decisions. Both of them can offer a better deal than the status quo – no integration – to P. With S in charge, P benefits from adapting to the common value shock, but is harmed by adapting to S's costs; since the common value component dominates, the positive effect outweighs the negative effect. With M in charge, P gets a more favorable deal, because production is adapted also to P's cost shock. Without specifying a precise game that determines the division of surplus, it is in general not clear what the outcome is, because S clearly prefers to be in charge himself over having M in charge. For the class of procedures that result in an efficient outcome, we have the following result.

Theorem 5 *Consider a mixed common- and private-value environment with fully revealing signals, $\sigma_{\varepsilon_C}^2 = \sigma_{\varepsilon_S}^2 = \sigma_{\varepsilon_P}^2 = 0$, and uncertainty satisfying $\sigma_C^2 > \sigma_S^2$. Ownership by M is efficient if and only if $(\sigma_C^2, \sigma_S^2, \sigma_P^2)$ is such that $\lambda_S \geq \frac{1}{2}$, or, equivalently,*

$$\frac{(2-\alpha)\,\sigma_S^2 - 2\alpha\sigma_C^2}{(4-\alpha)\,\sigma_S^2 + (4-2\alpha)\,\sigma_C^2}\sigma_S^2 + \frac{2-\alpha}{4-\alpha}\sigma_P^2 \geq 0.$$

Proof of Theorem 5. For $\sigma_C^2 > \sigma_S^2$, some form of integration is efficient. The question is whether S or M should be in charge of decision-making. For $\lambda = \frac{1}{2}$, M's objective at the stage of deciding who decides corresponds to social surplus. Hence, ownership by M is efficient if and only if M finds it better to make the decision rather than to delegate it to S at $\lambda = \frac{1}{2}$. Simplifying $\pi_M^{delS} \leq \pi_M^{con}$ (see Lemma 1) for $\lambda = \frac{1}{2}$

yields the inequality in the statement that expresses this requirement algebraically. □

6.3 When integration is not possible

Integration is not always the answer. If $\sigma_C^2 < \sigma_S^2$, then the only possible way to integrate is with the help of M. However, if S and P benefit from adaptation to very different extents, integration is not feasible. It is easiest to illustrate this insight in a pure private value environment. The point obviously does not depend on this.

We have demonstrated above that in the pure private value environment with symmetric uncertainty, integration is always feasible and optimal. We now show the converse.

Theorem 6 *Consider a pure private-value environment with fully revealing signals,* $\sigma_{\varepsilon_C}^2 = \sigma_{\varepsilon_S}^2 = \sigma_{\varepsilon_P}^2 = 0$, *and a fifty-fifty revenue sharing arrangement between S and P,* $\omega_S = \omega_P$. *Fix* $\sigma_P^2 > 0$ $(\sigma_S^2 > 0)$. *Integration is not feasible for* σ_S^2 (σ_P^2) *sufficiently close to zero.*

Intuitively, consider the extreme case in which S already knows his costs perfectly, while P is uncertain about the own costs. If M comes in, M adapts the quantity to some extent to P, which clearly benefits P. From the ex ante perspective of S, this adaptation decision is unfavorable. Instead of a fixed and certain production choice, S faces a lottery with an outcome that depends on the realization of P's costs. Due to the concavity of the loss function, S dislikes such an increase in risk. Hence, S will veto and the resulting organizational form is stand-alone.

In sum, we observe that M acquires ownership rights only if M expects to make use of them. While formal authority need not confer real authority, in our model, the two occur together in equilibrium. Moreover, whenever some form of authority arises, then the reason is to enable adaptation that would otherwise not occur.

7 Extensions: endogenous choice of information

Information conveys authority – and thereby surplus – to the one who has it. Clearly, the information will be chosen with a view to individual gains from the ensuing decision-making process. We now consider this natural extension of our model. Imagine that M has already acquired ownership rights – formally, take λ again as given – but suppose that information still needs to be acquired. In particular, suppose that M can decide what information to acquire.[24] Since we are interested in the consequences of information for authority, we abstract from costs of information acquisition in what follows. It is easy to show that M does best by keeping authority over information acquisition; if discretion over information acquisition is delegated to S and P, then they acquire perfect information.[25] Clearly, M can replicate this outcome. However, M sometimes benefits from leaving some noise in the observations, because this can reduce the conflicts in information transmission.

Consider now the information choices in detail. Common value information is free of noise at the optimum, $c^* = \sigma_C^2$, regardless of which institution is chosen. The reason is that everybody wishes to use the information ideally in the same way. On top, more common value information reduces conflicts in communication with S. Private value information needs to be analyzed for each institution separately. The gain from delegating to S is maximized for $s^* = \sigma_S^2$ for $\lambda > \frac{1}{2}$, resulting in an overall gain of

$$\Delta\pi_M^{delS,endo} = \sigma_C^2 + (2\lambda - 1)\sigma_S^2. \tag{8}$$

[24] In the communication subgames, we focus on equilibria with the most informative communication on and off path. This rules out threats not to listen or not to talk if the information provided is not ideal from a player's perspective. We believe that threats not to use available information are particularly difficult to enforce within an organization.

[25] This is intuitive but not obvious. Choosing perfect private value information comes at the cost of a large bias in communication. However, better information trumps the desire to eliminate the bias in communication. This is a consequence of logconcavity of the density which corresponds to $\alpha \in \left[\frac{1}{2}, 1\right]$. The converse to this argument is given in Deimen and Szalay (2019). There, in environments with fat tails ($\alpha > 1$), the sender does not benefit from observing better information from his perspective, because this introduces conflicts; any improvement in information is lost in biased communication.

For $\lambda \leq \frac{1}{2}$, no private value information is provided to S conditional on delegating to S, $s^* = 0$, and moreover, delegating to S is dominated by delegating to P; the maximum payoff from delegating to P is achieved for $p^* = \sigma_P^2$ and equals

$$\Delta\pi_M^{delP,endo} = \sigma_C^2 + (1 - 2\lambda)\,\sigma_P^2. \tag{9}$$

To see this is true, observe that S can communicate truthfully to P if S does not have information about S's private value, $s^* = 0$. Hence, P can make perfect use of common value information too. It can be shown that the maximum gain from delegating to S *or* P is given by the maximum over (8) and (9), so delegating to P is optimal for $\lambda \leq \frac{1}{2}$ and delegating to S is optimal for $\lambda > \frac{1}{2}$.[26]

Consultative decision-making achieves a gain of

$$\Delta\pi_M^{con,endo} = \max\left\{\sigma_C^2, \frac{(2-\alpha)\frac{\sigma_C^2+\lambda\sigma_S^2}{\sigma_C^2+\sigma_S^2}}{2-\alpha\frac{\sigma_C^2+\lambda\sigma_S^2}{\sigma_C^2+\sigma_S^2}}\left(\sigma_C^2+\lambda\sigma_S^2\right)\right\} + (1-\lambda)^2\,\frac{2-\alpha}{2-\alpha\,(1-\lambda)}\sigma_P^2. \tag{10}$$

The gain from communicating with S is convex in S's private value information, so that only the extreme cases of $s^* = 0$ or $s^* = \sigma_S^2$ are candidates for an optimum. M faces the trade-off of allowing S to observe private value information, at the cost of receiving common value information garbled by private value observations. The weight attached to S's payoff needs to be sufficiently high to outweigh the loss from strategic communication with S, to make it optimal to endow S with private value information. No such trade-off arises with respect to P's information, so this is observed without noise.

[26] The reason is that delegating to P and giving information $s = \sigma_S^2$ to S is dominated by delegating directly to S, for $\lambda > \frac{1}{2}$. This argument recognizes that conditional on delegating to S for $\lambda > \frac{1}{2}$, it is optimal to have $p^* = 0$, $s^* = 0$ for $\lambda \in \left(\frac{1}{2}, \frac{1}{2-\alpha}\right)$, and $s^* = \sigma_S^2$ for $\lambda \geq \frac{1}{2-\alpha}$. Conditional on delegating to P, if M assigns a high weight to S's payoff, then it becomes optimal to endow S with private value information too. The cost is that communication works less effectively, but this is the only way to make sure, information about S's costs finds its way into decision-making. However, since a lot of information is lost this way, this way of delegating is dominated.

7.1 Almost common values: shifting effective control from S to M and from M to P

Endogenous information choice puts M back at the helm, even in situations that would otherwise deprive M of real authority. To demonstrate this, we focus on the case where M's real authority relative to S is minimized: the case of almost common value information. To solve for the optimal mode of decision-making, we, thus, compare equations (8), (9), and (10), in the limit of $\sigma_C^2 \to \infty$. For consultative decision-making, by l'Hôpital's rule, we find that the maximum in (10) is achieved at $s^* = 0$ for $\lambda < \lambda^* := \frac{2}{4-\alpha}$, and at $s^* = \sigma_S^2$ for $\lambda \geq \lambda^*$.

Theorem 7 *In the almost common value environment, the power to choose P's and S's information shifts real authority from S to M and from M to P. We have* $\lambda_P^{endo} = \frac{\alpha}{2+\alpha}$ *and*

$$\lambda_S^{endo} = \frac{1}{\frac{\sigma_S^2}{\sigma_P^2}\frac{\alpha^2}{(2-\alpha)^2}+1}\left(1-\frac{\sigma_S^2}{\sigma_P^2}\frac{\alpha}{2-\alpha}\right) \quad \textit{for} \;\; \frac{\sigma_S^2}{\sigma_P^2} \leq \frac{1}{\alpha}\frac{(2-\alpha)^3}{\alpha^2-4\alpha+8};$$

otherwise, λ_S^{endo} *solves* $(1-\lambda)^2 \frac{2-\alpha}{2-\alpha(1-\lambda)}\sigma_P^2 = (2\lambda-1)\,\sigma_S^2$.

For $\lambda \leq \frac{1}{2}$, delegation to P is better than delegation to S. Moreover, since $\lambda^* = \frac{2}{4-\alpha} > \frac{1}{2}$, S has no private value information under consultative decision-making. Hence, the trade-off that M faces between communicating with P or delegating to P is exactly as in a pure private value setting in which P has private value information and S is absent. Therefore, it is optimal to delegate to P for all $\lambda \leq \frac{\alpha}{2+\alpha}$.

Consider next the trade-off between delegating to S and consultative decision-making. If $\frac{\sigma_S^2}{\sigma_P^2}$ is relatively low, the switching point λ_S that we obtain for the analysis of exogenous information structures remains valid. In contrast, if $\frac{\sigma_S^2}{\sigma_P^2}$ is relatively high, then M needs to compare the payoff from consultative decision-making when S has no private value information and P does have private value information with the payoff from delegating to S who does have private value information. The indifference point, in this scenario, is the solution to the last equation in the theorem.

The power to withhold information from the units, in particular from S, shifts authority from S to M: λ_S is lower if $\frac{\sigma_S^2}{\sigma_P^2}$ is higher. However, even in the most extreme case in which the variation in S's costs is much more pronounced than the variation in P's costs, λ_S is bounded below by $\frac{1}{2}$. Hence, the power to choose what the organization learns, ensures that M has effective control over decisions for payoff distributions around $\lambda = \frac{1}{2}$.

7.2 Endogenous information and risk

How does the allocation of real authority depend on the riskiness of the environment? Intuitively, making mistakes is more costly in more risky environments. Hence, we would expect that institutions dominate that make the best use of information and thereby avoid making mistakes. In general, it is not obvious how to make this insight formal. However, it turns out that our stochastic environment can reveal some insights.

A priori, there are two ways in which one can think about riskiness, in our model. Let $\boldsymbol{\sigma}^2 = (\sigma_C^2, \sigma_S^2, \sigma_P^2)$ denote the vector of uncertainty, and let $\kappa > 1$ denote a scalar. Evidently, an environment with uncertainty $\kappa\boldsymbol{\sigma}^2$ is – in some sense – more risky than an environment with uncertainty $\boldsymbol{\sigma}^2$. Another measure of riskiness is the parameter α; a higher value of α reflects a higher tail risk in the environment. It turns out that variance is irrelevant but that tail risk is exactly what matters.

Theorem 8 *Let $\sigma_C^2, \sigma_S^2, \sigma_P^2 > 0$, $\kappa > 0$, and suppose an optimal information choice. The optimal mode of decision-making is determined by $0 < \lambda_P^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha) < \lambda_S^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha) < 1$. We find that, $\lambda_P^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ and $\lambda_S^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ are independent of κ. Moreover, a higher tail risk crowds out consultative decision-making: $\lambda_P^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ is increasing in α and $\lambda_S^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ is decreasing in α.*

Proof. Equation (10) is equal to equation (9) for $\lambda = 0$, and equation (10) is equal to equation (8) for $\lambda = 1$. Equation (10) is convex and takes a higher value for $\lambda = \frac{1}{2}$ than equations (9) and (8), since $\sigma_C^2 + \frac{1}{4}\frac{2-\alpha}{2-\alpha\frac{1}{2}}\sigma_P^2 > \sigma_C^2$. This implies that $0 < \lambda_P^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha) < \lambda_S^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha) < 1$.

The expressions $\Delta\pi_M^{delS,endo}$, $\Delta\pi_M^{delP,endo}$, and $\Delta\pi_M^{con,endo}$ given in equations (8), (9), and (10) are all homogenous of degree one in κ. Hence, the points of indifference $\lambda_S^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ and $\lambda_P^{endo}(\kappa\boldsymbol{\sigma}^2, \alpha)$ are homogenous of degree zero in κ.

Equation (10) is strictly decreasing in α, while equations (8) and (9) are independent of α. $\square$

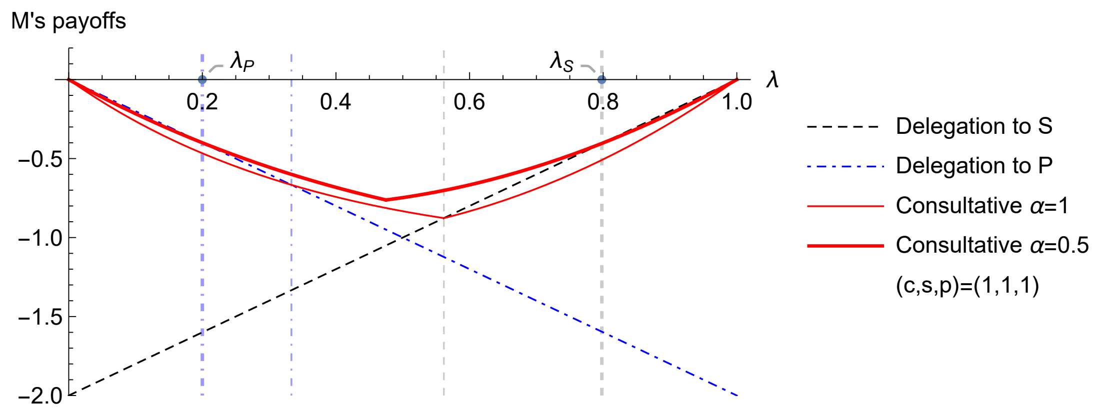

Figure 5: Payoff comparison for different levels of tail risk; thick lines $\alpha = 0.5$ (Uniform) and thin lines $\alpha = 1$ (Laplace).

A uniform scaling of all variances – arguably a change in the riskiness of the environment – has no effect at all on the relative merits of the institutions. The reason is that only relative magnitudes matter for the quality of communication. Evidently, this does not prove that risk has no effect. Rather, it demonstrates that all the ingredients in our model are necessary to get the full picture. Risk matters if it affects the tail risk of the distribution. The tail risk parameter α magnifies any losses that arise from strategic communication (for more details, recall Figure 2 and the text before). Through the optimal choice of information, the delegation modes rely purely on nonstrategic communication and, hence, achieve perfect transmission of common value information. In contrast, consultative decision-making always involves losses from strategic communication. Even if M chooses to withhold private value information from S – so that there is fully revealing communication with S – M still needs to communicate with P, who has private value information exclusively. As a result, consultative decision-making becomes relatively less attractive and M has real authority only in a smaller set, if the environment is more risky.

8 Conclusions

We propose a model to think about firm boundaries based on decision-making inside the firm. Ownership as an institution is explained through the formal authority it confers. An integrated structure is one with an established authority structure: someone has the right to make adaptation decisions on behalf of everyone. In contrast, a market based interaction involves decentralized trade at will.

The crucial insight in our theory is that acquiring formal authority through ownership rights is efficient only if the owner expects to obtain real authority this way: the owner must find it optimal to make adaptation decisions by himself rather than delegate them to someone else. Otherwise, a structure without the potential owner is more desirable.

We explain the allocation of real authority as a function of the amount and kind of information available to the firm, the distribution of surplus in the firm, and the riskiness of the environment. For a given ownership structure, we find that real authority resides at the top of the organization when information and surplus are distributed evenly in the organization. By contrast, if one unit within the organization holds an informational monopoly, then authority tends to gravitate towards that unit. Likewise, if the gains from decision-making are distributed very unevenly within the organization, then real authority tends to gravitate towards that prime beneficiary of adaptation. Risk, as measured by the likelihood of extreme events, magnifies these effects. Higher risk favors solutions that rely less on biased communication.

Taking one step back, we analyze the implications for the acquisition of ownership rights. Based on the nature and magnitudes of uncertainty, the model predicts when an integrated and when a stand alone structure arises. The model can rationalize the whole spectrum from stand alone decisions to integration between a producer and a retailer and integration involving in addition an owner with financial interests in both the producer and the retailer. This last structure is our main interest here and shown to serve two economic roles. First, it can make integration possible that could not occur without the involvement of a third party. Second, it can make an integrated arrangement more efficient. In both cases, the third party enables adaptation decisions that the producer and the retailer cannot replicate on their own.

We leave many extensions for future investigation. We have looked at the smallest possible form of an integrated firm. An exciting generalization will be to study multidivisional firms with a coordination motive. By studying a general model featuring both common and private value information components, we provide a toolbox that should be useful to study strategic information transmission more broadly. Finally, we identify a measure of riskiness that affects the performance of strategic information transmission crucially.

A Appendix

Lemma A.1 *Let $\boldsymbol{Z}$ follow an elliptically contoured symmetric distribution, or simply elliptical distribution. We write $\boldsymbol{Z} \sim EC_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, where we assume that d is the dimension of $\boldsymbol{Z}$, $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix with $rank(\boldsymbol{\Sigma}) = k$, and ϕ is the characteristic generator. Further let*

$$\boldsymbol{Z} = (\boldsymbol{Z_1}, \boldsymbol{Z_2}), \quad \boldsymbol{\mu} = (\boldsymbol{\mu_1}, \boldsymbol{\mu_2}), \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix},$$

where the dimensions of $\boldsymbol{Z_1}$, $\boldsymbol{\mu_1}$, and $\boldsymbol{\Sigma_{11}}$ are m, m, and $m \times m$ for $m \leq d$, respectively.

i) The distribution is symmetric about $\boldsymbol{\mu}$.

ii) The moments of the conditional distribution $(\boldsymbol{Z_1} | \boldsymbol{Z_2} = \boldsymbol{z_2})$ are given by the mean vector

$$\mathbb{E}\left[\boldsymbol{Z_1} | \boldsymbol{Z_2} = \boldsymbol{z_2}\right] = \boldsymbol{\mu_1} + (\boldsymbol{z_2} - \boldsymbol{\mu_2})\, \boldsymbol{\Sigma_{22}^{-1}} \boldsymbol{\Sigma_{21}}$$

and the conditional covariance matrix satisfying

$$\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma_{11}} - \boldsymbol{\Sigma_{12}} \boldsymbol{\Sigma_{22}^{-1}} \boldsymbol{\Sigma_{21}}.$$

iii) Let $\boldsymbol{A}$ be an $d \times l$ matrix and $\boldsymbol{b}$ be an $l \times 1$ vector. Then

$$\boldsymbol{b} + \boldsymbol{A}'\boldsymbol{Z} \sim EC_l\left(\boldsymbol{b} + \boldsymbol{A}'\boldsymbol{Z}, \boldsymbol{A}'\boldsymbol{Z}\boldsymbol{A}, \phi\right).$$

Proof of Lemma A.1. i) by definition, ii) Fang et al. (1990) Theorem 2.18, iii) Fang et al. (1990) Theorem 2.16. □

Proof of Proposition 1. The following proof generalizes the proof of Proposition 1 in Deimen and Szalay (2019), which uses the functional form of the Laplace distribution. The steps of the proof are exactly the same, except for the fact that we do not use any functional form here, but rather assume the general class of logconcave densities.

The proof of the proposition consists of three lemmas. Lemma A.2 proves uniqueness of finite equilibria, Lemma A.3 proves existence, and Lemma A.4 the existence of a limit equilibrium. □

Lemma A.2 *For any finite number N, if there exists an equilibrium with N distinct actions, then the equilibrium is unique.*

Proof of Lemma A.2. Fix N. For notational simplicity we oppress the dependence on N and write for example a_i instead of a_i^N. Define a forward equation as follows. Start with an arbitrary value $a_1 = t$ and compute the solution $a_2(t)$ as the value of a_2 that satisfies

$$t - \beta\mathbb{E}[\theta|\,\theta \leq t] = \beta\mathbb{E}[\theta|\,\theta \in [t, a_2(t)]] - t.$$

The right-hand side is increasing in a_2, so if a solution exists, it is unique. Differentiating totally, we find that

$$\frac{da_2}{dt} = \frac{\left(2 - \beta\frac{\partial}{\partial t}\mathbb{E}[\theta|\,\theta \leq t] - \beta\frac{\partial}{\partial t}\mathbb{E}[\theta|\,\theta \in [t, a_2(t)]]\right)}{\beta\frac{\partial}{\partial a_2}\mathbb{E}[\theta|\,\theta \in [t, a_2(t)]]}.$$

We have $\frac{da_2}{dt} > 1$ if and only if

$$2 > \beta\frac{\partial}{\partial t}\mathbb{E}[\theta|\,\theta \leq t] + \beta\frac{\partial}{\partial t}\mathbb{E}[\theta|\,\theta \in [t, a_2(t)]] + \beta\frac{\partial}{\partial a_2}\mathbb{E}[\theta|\,\theta \in [t, a_2(t)]]. \qquad (11)$$

As shown in Szalay (2012), for a logconcave distribution

$$\frac{\partial}{\partial \underline{a}}\mathbb{E}[\theta|\,\theta \in [\underline{a}, \overline{a}]] + \frac{\partial}{\partial \overline{a}}\mathbb{E}[\theta|\,\theta \in [\underline{a}, \overline{a}]] \leq 1 \quad \text{for } \underline{a} < \overline{a}.$$

Hence, condition (11) is satisfied due to the fact that $f_\theta(\theta)$ is logconcave and $\beta < 1$.

Now consider

$$a_i - \beta\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right] = \beta\mathbb{E}\left[\theta|\,\theta \in [a_i, a_{i+1}]\right] - a_i,$$

and

$$\frac{da_{i+1}}{da_i} = \frac{\left(2 - \beta\frac{\partial}{\partial a_i}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right] - \beta\frac{\partial}{\partial a_i}\mathbb{E}\left[\theta|\,\theta \in [a_i, a_{i+1}]\right]\right) - \beta\frac{\partial}{\partial a_{i-1}}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right]\frac{da_{i-1}}{da_i}}{\beta\frac{\partial}{\partial a_{i+1}}\mathbb{E}\left[\theta|\,\theta \in [a_i, a_{i+1}]\right]}.$$

If $\frac{da_i}{da_{i-1}} > 1$, then $\frac{da_{i-1}}{da_i} < 1$. Moreover, $\frac{da_{i+1}}{da_i} > 1$ if and only if

$$\begin{aligned} 2 >&\beta\frac{\partial}{\partial a_i}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right] + \beta\frac{\partial}{\partial a_i}\mathbb{E}\left[\theta|\,\theta \in [a_i, a_{i+1}]\right] + \beta\frac{\partial}{\partial a_{i-1}}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right]\frac{da_{i-1}}{da_i} \\ &+ \beta\frac{\partial}{\partial a_{i+1}}\mathbb{E}\left[\theta|\,\theta \in [a_i, a_{i+1}]\right]. \end{aligned} \tag{12}$$

Noting that $\beta\frac{\partial}{\partial a_{i-1}}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right]\frac{da_{i-1}}{da_i} \leq \beta\frac{\partial}{\partial a_{i-1}}\mathbb{E}\left[\theta|\,\theta \in [a_{i-1}, a_i]\right]$, we observe that the right-hand side of (12) is bounded above by 2β, so that inequality (12) is satisfied and we have indeed $\frac{da_{i+1}}{da_i} > 1$.

Take now $a_2(t), \ldots, a_N(t)$ as determined by the forward equations up to and including $a_N(t)$ and consider the difference

$$\Delta(t) \equiv 2a_N(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [a_{N-1}(t), a_N(t)]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_N(t)\right].$$

By the now familiar reasoning, this difference is a strictly monotonic function of t, as

$$\begin{aligned} &\left(2 - \beta\mathbb{E}\frac{\partial}{\partial a_N}\left[\theta|\,\theta \in [a_{N-1}, a_N]\right] - \beta\frac{\partial}{\partial a_N}\mathbb{E}\left[\theta|\,\theta \geq a_N\right]\right)da_N \\ &- \beta\frac{\partial}{\partial a_{N-1}}\mathbb{E}\left[\theta|\,\theta \in [a_{N-1}, a_N]\right]\frac{da_{N-1}}{da_N}da_N > 0. \end{aligned}$$

Therefore, there is at most one value of t, say $\tilde{t}_N$, such that the sequence $\{\tilde{a}_1, \ldots, \tilde{a}_N\}$ with $\tilde{a}_1 := \tilde{t}_N$ and $\tilde{a}_i := a_i\left(\tilde{t}_N\right)$ solves the system of indifference conditions. Hence, the equilibrium is unique. □

Lemma A.3 *For any n, there exists an equilibrium inducing $N = 2(n+1)$ actions and there exists an equilibrium inducing $N = (2n+1)$ actions.*

Proof of Lemma A.3. Lemma A.2 shows uniqueness of equilibria that do exist. By symmetry of payoffs and the density, the model has symmetric equilibria. Together this implies that all finite equilibria must be symmetric around 0.

We, here, focus on the equilibria with an even number of induced actions. All the results extend to the equilibria with an odd number of induced actions.

Consider the truncated distribution, where the truncation is at zero and to the positive side. An equilibrium, if it exists, satisfies for $i = 2, \ldots, n-1$

$$\begin{aligned} a_1 - \beta \mathbb{E}[\theta | \theta \in [0, a_1]] &= \beta \mathbb{E}[\theta | \theta \in [a_1, a_2]] - a_1 \\ a_i - \beta \mathbb{E}[\theta | \theta \in [a_{i-1}, a_i]] &= \beta \mathbb{E}[\theta | \theta \in [a_i, a_{i+1}]] - a_i \\ a_n - \beta \mathbb{E}[\theta | \theta \in [a_{n-1}, a_n]] &= \beta \mathbb{E}[\theta | \theta \geq a_n] - a_n. \end{aligned}$$

We construct an equilibrium as follows. We first consider the forward solution for arbitrary $a_1 = t$ and show that for any n, the forward equation is guaranteed to have solutions up to a_n as long as $t \leq \tilde{t}_n$. Then, we show that an equilibrium of the communication game (which satisfies the forward equation and the closure condition) exists for a value of t consistent with that condition.

The forward equation for $a_2(t)$ is the value of a_2 such that

$$t - \beta \mathbb{E}[\theta | \theta \in [0, t]] = \beta \mathbb{E}[\theta | \theta \in [t, a_2]] - t. \tag{13}$$

The difference between left and right side is

$$2t - \beta \mathbb{E}[\theta | \theta \in [0, t]] - \beta \mathbb{E}[\theta | \theta \in [t, a_2]].$$

In the limit as $a_2 \to t$, the difference is strictly positive as $(2-\beta)t - \beta \mathbb{E}[\theta | \theta \in [0, t]] > 0$. The difference is decreasing in a_2. In the limit as $a_2 \to \infty$, we get

$$2t - \beta \mathbb{E}[\theta | \theta \in [0, t]] - \beta \mathbb{E}[\theta | \theta \geq t].$$

By logconcavity, $\mathbb{E}[\theta | \theta \in [0, t]]$ and $\mathbb{E}[\theta | \theta \geq t]$ increase with t each at rate smaller than or equal to one. Hence, there exists a finite solution $a_2(t)$ if and only if $t < \tilde{t}_2$, where $\tilde{t}_2$ is defined as the unique value of t that solves

$$2\tilde{t}_2 - \beta \mathbb{E}\left[\theta | \theta \in \left[0, \tilde{t}_2\right]\right] - \beta \mathbb{E}\left[\theta | \theta \geq \tilde{t}_2\right] = 0. \tag{14}$$

For future reference, note that by logconcavity the forward equation satisfies

$$\frac{da_2}{dt} = \frac{2 - \beta\frac{\partial}{\partial t}\mathbb{E}\left[\theta|\,\theta \in [0,t]\right] - \beta\frac{\partial}{t}\mathbb{E}\left[\theta|\,\theta \in [t,a_2]\right]}{\beta\frac{\partial}{\partial a_2}\mathbb{E}\left[\theta|\,\theta \in [t,a_2]\right]} > 1. \tag{15}$$

Moreover, for $t \to 0$ we have $a_2(t) \to 0$, and $a_2(t) - t$ is increasing in t.

Consider next the forward solution for $a_3(t)$, which is the value of a_3 that solves

$$2a_2(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] - \beta\mathbb{E}\left[\theta|\,\theta \in [a_2(t), a_3]\right] = 0. \tag{16}$$

For $a_3 \to a_2(t)$, the difference takes value

$$2a_2(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] - \beta a_2(t) > 0.$$

The difference is decreasing in a_3. Hence, there exists a finite solution if and only if

$$2a_2(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_2(t)\right] < 0.$$

Differentiating totally, we observe

$$\left(2 - \beta\frac{\partial}{\partial a_2}\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] - \beta\frac{\partial}{\partial a_2}\mathbb{E}\left[\theta|\,\theta \geq a_2(t)\right]\right) da_2 - \beta\frac{\partial}{\partial t}\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] dt.$$

As $\frac{dt}{da_2} < 1$ by (15), the difference is increasing in t. Hence, there exists a unique value $\tilde{t}_3$ such that a finite solution $a_3(t)$ exists for $t < \tilde{t}_3$. The value $\tilde{t}_3$ satisfies

$$2a_2\left(\tilde{t}_3\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[\tilde{t}_3, a_2\left(\tilde{t}_3\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_2\left(\tilde{t}_3\right)\right] = 0. \tag{17}$$

At $\tilde{t}_3$, the forward equation for $a_2\left(\tilde{t}_3\right)$, equation (13), implies that

$$2\tilde{t}_3 - \beta\mathbb{E}\left[\theta|\,\theta \in \left[0, \tilde{t}_3\right]\right] = \beta\mathbb{E}\left[\theta|\,\theta \in \left[\tilde{t}_3, a_2\left(\tilde{t}_3\right)\right]\right]$$

Substituting back into (17) gives

$$2a_2\left(\tilde{t}_3\right) - \beta\mathbb{E}\left[\theta|\,\theta \geq a_2\left(\tilde{t}_3\right)\right] = 2\tilde{t}_3 - \beta\mathbb{E}\left[\theta|\,\theta \in \left[0, \tilde{t}_3\right]\right].$$

Subtracting $\beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_3\right]$ from each side, we get

$$2\tilde{t}_3 - \beta\mathbb{E}\left[\theta|\,\theta \in \left[0, \tilde{t}_3\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_3\right] = 2a_2\left(\tilde{t}_3\right) - \beta\mathbb{E}\left[\theta|\,\theta \geq a_2\left(\tilde{t}_3\right)\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_3\right].$$

Since

$$2a_2\left(\tilde{t}_3\right) - \beta\mathbb{E}\left[\theta|\,\theta \geq a_2\left(\tilde{t}_3\right)\right] = \beta\mathbb{E}\left[\theta|\,\theta \in \left[\tilde{t}_3, a_2\left(\tilde{t}_3\right)\right]\right],$$

by (17), the right side takes value

$$\beta\mathbb{E}\left[\theta|\,\theta \in \left[\tilde{t}_3, a_2\left(\tilde{t}_3\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_3\right] < 0,$$

and hence

$$2\tilde{t}_3 - \beta\mathbb{E}\left[\theta|\,\theta \in \left[0, \tilde{t}_3\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_3\right] < 0.$$

Now recall equation (14):

$$2\tilde{t}_2 - \beta\mathbb{E}\left[\theta|\,\theta \in \left[0, \tilde{t}_2\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq \tilde{t}_2\right] = 0.$$

Since $2t - \beta\mathbb{E}\left[\theta|\,\theta \in [0,t]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t\right]$ is increasing in t by logconcavity, we have shown that $\tilde{t}_3 < \tilde{t}_2$.

Totally differentiating (16) gives

$$\frac{da_3}{da_2} = \frac{2 - \beta\frac{\partial}{\partial a_2}\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right] - \beta\frac{\partial}{\partial a_2}\mathbb{E}\left[\theta|\,\theta \in [a_2(t), a_3]\right] - \beta\frac{\partial}{\partial t}\mathbb{E}\left[\theta|\,\theta \in [t, a_2(t)]\right]\frac{dt}{da_2}}{\beta\frac{\partial}{\partial a_3}\mathbb{E}\left[\theta|\,\theta \in [a_2(t), a_3]\right]}.$$

Hence, $\frac{da_3}{da_2} > 1$ given that $\frac{da_2}{dt} > 1$. It follows that $a_3(t) - a_2(t)$ is increasing in t. Likewise, it is obvious that $a_3(t)$ goes to zero as $t \to 0$.

Suppose that the forward solutions exist up to $a_{n-1}(t)$ and all have the above properties. Consider the forward solution for $a_n(t)$, defined as the value that satisfies

$$a_{n-1}(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [a_{n-2}(t), a_{n-1}(t)]\right] = \beta\mathbb{E}\left[\theta|\,\theta \in [a_{n-1}(t), a_n]\right] - a_{n-1}(t). \quad (18)$$

At $a_n = a_{n-1}(t)$ the right side is negative, while the left side is positive. The right side is increasing in a_n, so there exists a unique finite solution if and only if

$$2a_{n-1}(t) - \beta\mathbb{E}\left[\theta|\,\theta \in [a_{n-2}(t), a_{n-1}(t)]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n-1}(t)\right] < 0.$$

Totally differentiating, we note that the difference is increasing in t by the fact that $\frac{da_{n-1}}{da_{n-2}} > 1$. Hence, there is a unique value $\tilde{t}_n$ such that a forward solution $a_n(t)$ exists for any $t < \tilde{t}_n$, where $\tilde{t}_n$ is defined by the condition

$$2a_{n-1}\left(\tilde{t}_n\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-2}\left(\tilde{t}_n\right), a_{n-1}\left(\tilde{t}_n\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n-1}\left(\tilde{t}_n\right)\right] = 0.$$

We now argue that $\tilde{t}_n < \tilde{t}_{n-1}$. Consider

$$2a_{n-2}\left(\tilde{t}_{n-1}\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-3}\left(\tilde{t}_{n-1}\right), a_{n-2}\left(\tilde{t}_{n-1}\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n-2}\left(\tilde{t}_{n-1}\right)\right] = 0. \tag{19}$$

At $\tilde{t}_n$, the forward equation for $a_{n-1}(t)$ implies

$$2a_{n-2}\left(\tilde{t}_n\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-3}\left(\tilde{t}_n\right), a_{n-2}\left(\tilde{t}_n\right)\right]\right] = \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-2}\left(\tilde{t}_n\right), a_{n-1}\left(\tilde{t}_n\right)\right]\right].$$

Hence, at $\tilde{t}_n$,

$$\begin{aligned} & 2a_{n-2}\left(\tilde{t}_n\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-3}\left(\tilde{t}_n\right), a_{n-2}\left(\tilde{t}_n\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n-2}\left(\tilde{t}_n\right)\right] \\ = \; & \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-2}\left(\tilde{t}_n\right), a_{n-1}\left(\tilde{t}_n\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n-2}\left(\tilde{t}_n\right)\right] \\ < \; & 0. \end{aligned}$$

Since the left side of (19) is increasing in t, it follows that $\tilde{t}_{n-1} > \tilde{t}_n$ is necessary to restore equality with zero.

Consider now the closure condition. A sequence of thresholds $t, a_2(t), \ldots, a_n(t)$ forms an equilibrium if and only if the thresholds $a_{n-1}(t)$ and $a_n(t)$ are such that

$$a_n(t) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-1}(t), a_n(t)\right]\right] = \beta\mathbb{E}\left[\theta|\,\theta \geq a_n(t)\right] - a_n(t).$$

Define the difference

$$\Delta_n(t) \equiv 2a_n(t) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_{n-1}(t), a_n(t)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_n(t)\right].$$

By the now familiar argument, $\Delta_n(t)$ is strictly increasing in t, so there is a unique value of t, say t_n^*, that solves the equation. We note that the value of t_n^* is exactly $\tilde{t}_{n+1}$, the value such that the next forward solution just goes out of the support.

It follows that for any n, we can construct an equilibrium. Moreover, in such an equilibrium, the value of the first threshold a_1^n is $\tilde{t}_{n+1}$, a decreasing function of n. □

Lemma A.4 *There exists an infinite equilibrium.*

Proof of Lemma A.4. We know from Lemma A.3 that the value of the first threshold $\tilde{t}_{n+1}$ is a monotone decreasing function of n. Since the sequence is bounded by zero it must converge. Likewise, $a_n(t_n^*)$ is bounded above: suppose for contradiction that $a_n(t)$ goes out of bounds as n goes out of bounds, and consider the closure condition, $D_n(t) = 0$, with

$$D_n(t) = 2a_n(t) - \beta\mathbb{E}[\theta|\, \theta \in [a_{n-1}(t), a_n(t)]] - \beta\mathbb{E}[\theta|\, \theta \geq a_n(t)].$$

Note that

$$-\beta\mathbb{E}[\theta|\, \theta \in [a_{n-1}(t), a_n(t)]] \geq -\beta\mathbb{E}[\theta|\, \theta \geq a_n(t)]$$

and that $a - \beta\mathbb{E}[\theta|\, \theta \geq a]$ is increasing in a for a logconcave distribution. This implies a contradiction. Therefore, we must have $\lim_{n\to\infty} a_n(t_n^*) < \infty$ and the sequence $a_n(t_n^*)$ is bounded above.

Claim 1) The equilibrium features increasing intervals,

$$a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n \quad \forall n \text{ and } \forall i < n.$$

Proof: Consider the equilibrium indifference condition for a_1,

$$a_1 - \beta\mathbb{E}[\theta|\, \theta \in [0, a_1]] = \beta\mathbb{E}[\theta|\, \theta \in [a_1, a_2]] - a_1.$$

Logconcave densities are unimodal. By symmetry, the mode is at 0 and hence the density truncated at zero is non-increasing. This implies that for an interval of given length Δ,

$$\mathbb{E}[\theta|\, \theta \in [a_1, a_1 + \Delta]] \leq a_1 + \frac{\Delta}{2}.$$

Consider $a_1 = \Delta$ and $a_2 = 2\Delta$. Then, $\Delta - \beta\mathbb{E}[\theta|\, \theta \in [0, \Delta]] \geq \Delta - \beta\frac{\Delta}{2}$ and $\beta\mathbb{E}[\theta|\, \theta \in [\Delta, 2\Delta]] - \Delta \leq \beta\frac{3}{2}\Delta - \Delta$, where the inequalities are strict if the density is strictly decreasing. Since $\Delta - \beta\frac{\Delta}{2} > \beta\frac{3}{2}\Delta - \Delta$, a_2 must increase to satisfy the equilibrium condition.

Likewise, consider

$$a_i - \beta\mathbb{E}[\theta|\, \theta \in [a_{i-1}, a_i]] = \beta\mathbb{E}[\theta|\, \theta \in [a_i, a_{i+1}]] - a_i$$

and suppose $a_i - a_{i-1} = \Delta = a_{i+1} - a_i$. Then $\beta\mathbb{E}[\theta|\, \theta \in [a_i - \Delta, a_i]] \leq \beta\left(a_i - \frac{\Delta}{2}\right)$ and $\beta\mathbb{E}[\theta|\, \theta \in [a_i, a_i + \Delta]] \leq \beta\left(a_i + \frac{\Delta}{2}\right)$ (with strict inequalities for a strictly decreasing density) imply that $a_i - \beta\mathbb{E}[\theta|\, \theta \in [a_{i-1}, a_i]] \geq a_i - \beta\left(a_i - \frac{\Delta}{2}\right)$ and $\beta\mathbb{E}[\theta|\, \theta \in [a_i, a_{i+1}]] -$

$a_i \leq \beta\left(a_i + \frac{\Delta}{2}\right) - a_i$. Since $a_i - \beta\left(a_i - \frac{\Delta}{2}\right) > \beta\left(a_i + \frac{\Delta}{2}\right) - a_i$ for all a_i, we must again have that $a_{i+1} - a_i > a_i - a_{i-1} = \Delta$ to restore equilibrium. $\square$

Claim 2) The sequence $\left(a_1^n\right)_n$ is monotone decreasing, while the sequence $\left(a_n^n\right)_n$ is monotone increasing. Moreover, equilibrium thresholds are nested,

$$a_1^{n+1} < a_1^n < a_2^{n+1} < \cdots a_n^{n+1} < a_n^n < a_{n+1}^{n+1} \qquad \forall n. \tag{20}$$

Proof: Recall the notation $a_1^n = \tilde{t}_n = t_{n+1}^*$ and $a_1^{n+1} = \tilde{t}_{n+1} = t_{n+2}^*$ from Lemma A.3. Since by Lemma A.3 the solution of the forward equation is monotonic in the initial condition, t, we have that $a_i^{n+1} < a_i^n$ for $i = 1, \ldots, n$. Hence, it suffices to prove that $a_i^n < a_{i+1}^{n+1}$ for $i = 1, \ldots, n$.

We start with two preliminary observations. First, the "next" solution of the forward equation, $a_{i+1}^k(t)$ for $i = 1, \ldots, k-1$, and $k = n, n+1$ is monotonic in $a_i^k(t)$, and the length of the previous interval, $a_i^k(t) - a_{i-1}^k(t)$. To see this, note that the forward equations for a_2^k, a_3^k, and a_{i+1}^k, for $i = 3, \ldots, k-1$ and $k = n, n+1$ satisfy:

$$t - \beta \mathbb{E}\left[\theta \mid \theta \in [0, t]\right] = \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_2^k, t\right]\right] - t,$$

$$a_2^k(t) - \beta \mathbb{E}\left[\theta \mid \theta \in \left[t, a_2^k(t)\right]\right] = \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_2^k(t), a_3^k\right]\right] - a_2^k(t),$$

and

$$a_i^k(t) - \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_{i-1}^k(t), a_i^k(t)\right]\right] = \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_i^k(t), a_{i+1}^k\right]\right] - a_i^k(t).$$

Let $a_{i-1}^k(t) = a_i^k(t) - \Delta$ and substitute into the forward equation for a_{i+1}^k :

$$a_i^k(t) - \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_i^k(t) - \Delta, a_i^k(t)\right]\right] = \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_i^k(t), a_{i+1}^k\right]\right] - a_i^k(t).$$

Monotonicity follows from the fact that $a_i^k(t)$ decreases the value of the right-hand side by logconcavity of the density and increases the value of the left-hand side again by that property. Moreover, an increase in Δ increases the left-hand side further, implying that $a_{i+1}^k(t)$ has to increase to restore the equality.

Second, it is impossible that $a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_n^n\left(\tilde{t}_n\right)$ and $a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right) - a_n^{n+1}\left(\tilde{t}_{n+1}\right) < a_n^n\left(\tilde{t}_n\right) - a_{n-1}^n\left(\tilde{t}_n\right)$. If these conditions would hold, then one of the closure conditions,

$$0 = 2a_n^n\left(\tilde{t}_n\right) - \beta \mathbb{E}\left[\theta \mid \theta \in \left[a_{n-1}^n\left(\tilde{t}_n\right), a_n^n\left(\tilde{t}_n\right)\right]\right] - \beta \mathbb{E}\left[\theta \mid \theta \geq a_n^n\left(\tilde{t}_n\right)\right]$$

and

$$0 = 2a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_n^{n+1}\left(\tilde{t}_{n+1}\right), a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right)\right]$$

would necessarily be violated. To see this, take $\delta_1, \delta_2 > 0$ and suppose that $a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right) = a_n^n\left(\tilde{t}_n\right) - \delta_1$, $a_n^n\left(\tilde{t}_n\right) - a_{n-1}^n\left(\tilde{t}_n\right) = \Delta$, and $a_{n+1}^{n+1}\left(\tilde{t}_{n+1}\right) - a_n^{n+1}\left(\tilde{t}_{n+1}\right) = \Delta - \delta_2$. Now consider the closure conditions

$$0 = 2a_n^n\left(\tilde{t}_n\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_n^n\left(\tilde{t}_n\right) - \Delta, a_n^n\left(\tilde{t}_n\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_n^n\left(\tilde{t}_n\right)\right]$$

and

$$0 = 2\left(a_n^n\left(\tilde{t}_n\right) - \delta_1\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[a_n^n\left(\tilde{t}_n\right) - \delta_1 - \left(\Delta - \delta_2\right), a_n^n\left(\tilde{t}_n\right) - \delta_1\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq a_n^n\left(\tilde{t}_n\right) - \delta_1\right].$$

By logconcavity, $\delta_1 > 0$ reduces the right-hand side of the second condition. Moreover, $\delta_2 > 0$ increases the lower bound $a_n^n\left(\tilde{t}_n\right) - \delta_1 - \Delta + \delta_2$, so decreases the right-hand side further. Hence, one of the closure conditions must necessarily be violated.

We now show that $a_{j+1}^{n+1} > a_j^n$ for all $j \leq n$. Suppose for contradiction that the property is violated for the first time at $j = l$. Suppose $a_{j+1}^{n+1}\left(\tilde{t}_{n+1}\right) > a_j^n\left(\tilde{t}_n\right)$ for all $j = 1, \ldots, l-1$ and $a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_l^n\left(\tilde{t}_n\right)$. Taken together, these inequalities immediately imply that $a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) - a_l^{n+1}\left(\tilde{t}_{n+1}\right) < a_l^n\left(\tilde{t}_n\right) - a_{l-1}^n\left(\tilde{t}_n\right)$. In turn, the monotonicity property of the next forward solution implies that $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+1}^n\left(\tilde{t}_n\right)$.

It also follows then that $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) - a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+1}^n\left(\tilde{t}_n\right) - a_l^n\left(\tilde{t}_n\right)$. To see this, suppose instead that $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) - a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) \geq a_{l+1}^n\left(\tilde{t}_n\right) - a_l^n\left(\tilde{t}_n\right)$ or equivalently that $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) \geq a_{l+1}^n\left(\tilde{t}_n\right) + \left(a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) - a_l^n\left(\tilde{t}_n\right)\right)$. However, this is impossible since both $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+1}^n\left(\tilde{t}_n\right)$ and $a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_l^n\left(\tilde{t}_n\right)$. Hence, the claim follows.

However, if $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+1}^n\left(\tilde{t}_n\right)$ and $a_{l+2}^{n+1}\left(\tilde{t}_{n+1}\right) - a_{l+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+1}^n\left(\tilde{t}_n\right) - a_l^n\left(\tilde{t}_n\right)$, then $a_{l+3}^{n+1}\left(\tilde{t}_{n+1}\right) < a_{l+2}^n\left(\tilde{t}_n\right)$ and so forth. Hence, we would have $a_{j+1}^{n+1}\left(\tilde{t}_{n+1}\right) < a_j^n\left(\tilde{t}_n\right)$ and $a_{j+1}^{n+1}\left(\tilde{t}_{n+1}\right) - a_j^{n+1}\left(\tilde{t}_{n+1}\right) < a_j^n\left(\tilde{t}_n\right) - a_{j-1}^n\left(\tilde{t}_n\right)$ for all $j \geq l$ and in particular for $j = n$, leading to a violation of one of the closure conditions.

The same argument can be given for a Class II equilibrium. This is omitted. □

Claim 3) The limit of the sequences of thresholds and actions is an equilibrium.

Proof: The limit is an equilibrium if $\lim_{n\to\infty}\beta\mu_i^n \leq \lim_{n\to\infty} a_i^n \leq \lim_{n\to\infty}\beta\mu_{i+1}^n$. Therefore, we have to show that equilibrium thresholds remain ordered in the limit, $\lim_{n\to\infty} a_i^n < \lim_{n\to\infty} a_{i+1}^n$. For all finite n, thresholds are ordered in equilibrium, $a_i^n < a_{i+1}^n$, since they are ordered for any forward equation. By Claim 2) equilibrium thresholds converge. Denote the limits by $\overline{a}_i = \lim_{n\to\infty} a_i^n$ for all i. By convergence, for any ε there is N such that for all $n > N$: $a_i^n \geq \overline{a}_i - \frac{\varepsilon}{2}$ and $a_{i+1}^n \leq \overline{a}_{i+1} + \frac{\varepsilon}{2}$. Suppose for contradiction that $\overline{a}_i \geq \overline{a}_{i+1} + \delta$ for some $\delta > 0$; this implies

$$a_i^n \geq \overline{a}_i - \frac{\varepsilon}{2} \geq \overline{a}_{i+1} + \delta - \frac{\varepsilon}{2} \geq a_{i+1}^n - \frac{\varepsilon}{2} + \delta - \frac{\varepsilon}{2} > a_{i+1}^n,$$

for all $\varepsilon < \delta$. Hence thresholds remain ordered in the limit and the limit is an equilibrium. □

□

Proof of Lemma 1. Consultative decision-making.

M's optimal action is $\Delta y^{con} = \beta_{MS} \cdot \mu_S + \beta_{MP} \cdot \mu_P$, with $\beta_{MS} = \frac{c+\lambda s}{c+s}$ and $\beta_{MP} = 1 - \lambda$.

Note that $\mathbb{E}\left[\mu_\theta \mu_P\right] = 0 = \mathbb{E}\left[\mu_P\left(X_C + X_S\right)\right]$, since these variables are uncorrelated. Moreover, we have $\mathbb{E}\left[\mu_S\left(X_C + X_S\right)\right] = \mathbb{E}\left[\mu_S^2\right] = var(\mu_S)$ and $\mathbb{E}\left[\mu_P X_P\right] = \mathbb{E}\left[\mu_P^2\right] = var(\mu_P)$.

We can calculate S's expected loss as

$$\begin{aligned} & \mathbb{E}\left[\beta_{MS} \cdot \mu_S + \beta_{MP} \cdot \mu_P - \left(X_C + X_S\right)\right]^2 \\ = \; & \left(\beta_{MS}\right)^2 \mathbb{E}\left[\mu_S^2\right] + \left(\beta_{MP}\right)^2 \mathbb{E}\left[\mu_P^2\right] + \sigma_C^2 + \sigma_S^2 + 2\beta_{MS}\beta_{MP}\mathbb{E}\left[\mu_S \mu_P\right] \\ & -2\beta_{MP}\mathbb{E}\left[\mu_P\left(X_C + X_S\right)\right] - 2\beta_{MS}\mathbb{E}\left[\mu_S\left(X_C + X_S\right)\right] \\ = \; & \beta_{MS}\left(\beta_{MS} - 2\right) var(\mu_S) + \left(\beta_{MP}\right)^2 var(\mu_P) + \sigma_C^2 + \sigma_S^2. \end{aligned} \tag{21}$$

Similarly, we can calculate P's expected loss as

$$\begin{aligned} & \mathbb{E}\left(\beta_{MS} \cdot \mu_S + \beta_{MP} \cdot \mu_P - \left(X_C + X_P\right)\right)^2 \\ = \; & \left(\beta_{MS}\right)^2 \mathbb{E}\left[\mu_S^2\right] + \left(\beta_{MP}\right)^2 \mathbb{E}\left[\mu_P^2\right] + \sigma_C^2 + \sigma_P^2 - 2\beta_{MS}\mathbb{E}\left[\mu_S X_C\right] - 2\beta_{MP}\mathbb{E}\left[\mu_P X_P\right]. \end{aligned}$$

To compute $\mathbb{E}\left[\mu_S X_C\right]$, observe that the joint distribution of μ_S and X_C can be computed from the joint distribution of Θ_S and X_C. For $\theta_S \in (\theta_{S,i-1}, \theta_{S,i}]$, let $\mu_S =$

$\mu_{\theta_{S,i}}$, and let P denote the corresponding random variable with typical realization $P_i = (\theta_{S,i-1}, \theta_{S,i}]$.

$$\begin{aligned}
\mathbb{E}\left[\mu_S X_C\right] &= \mathbb{E}_P\left[\mathbb{E}\left[\mu_S X_C\right]\middle| P = P_i\right] = \mathbb{E}_P\left[\mu_S \mathbb{E}\left[X_C\right]\middle| P = P_i\right] \\
&= \mathbb{E}_P\left[\mu_S \mathbb{E}_{\Theta_S \in P_i}\left[\mathbb{E}\left[X_C\right]\middle| \Theta_S = \theta_S\right]\right] = \mathbb{E}_P\left[\mu_S \mathbb{E}_{\Theta_S \in P_i}\left[\frac{Cov\left(X_C, \Theta_S\right)}{Var\left(\Theta_S\right)}\Theta_S\right]\right] \\
&= \frac{c}{c+s}\mathbb{E}\left[\mu_S^2\right] = \beta_{PS}\, var(\mu_S).
\end{aligned}$$

Hence P's expected loss can be written as

$$\beta_{MS}\left(\beta_{MS} - 2\beta_{PS}\right) var(\mu_S) + \beta_{MP}\left(\beta_{MP} - 2\right) var(\mu_P) + \sigma_C^2 + \sigma_P^2. \tag{22}$$

To calculate M's payoff, we rewrite the variances according to equation (5). After rearranging, M's expected payoff can be written as

$$\begin{aligned}
&-\lambda \mathbb{E}\left[\left(\Delta y^{con} - (X_C + X_S)\right)^2\right] - (1-\lambda)\,\mathbb{E}\left[\left(\Delta y^{con} - (X_C + X_P)\right)^2\right] \\
=&\frac{(2-\alpha)\,\beta_{MS}}{2-\alpha\cdot\beta_{MS}}\left(c + \lambda s\right) + (1-\lambda)^2\,\frac{2-\alpha}{2-\alpha\,(1-\lambda)}p - \sigma_C^2 - \lambda\sigma_S^2 - (1-\lambda)\,\sigma_P^2.
\end{aligned}$$

Delegated decision-making to P. The decision-rule of A is $\Delta y^{delP} = \beta_{PS}\cdot\mu_S + \frac{\sigma_P^2}{\sigma_P^2+\sigma_{\varepsilon_P}^2}\cdot s_P$, with $\beta_{PS} = \frac{c}{c+s}$. As before, we can calculate S's expected loss as

$$\mathbb{E}\left(\beta_{PS}\cdot\mu_S + \frac{\sigma_P^2}{\sigma_P^2 + \sigma_{\varepsilon_P}^2}\cdot s_P - (X_C + X_S)\right)^2 = \beta_{PS}\left(\beta_{PS} - 2\right) var(\mu_S) + p + \sigma_C^2 + \sigma_S^2.$$

Similarly, we can calculate P's expected loss as

$$\mathbb{E}\left(\beta_{PS}\cdot\mu_S + \frac{\sigma_P^2}{\sigma_P^2 + \sigma_{\varepsilon_P}^2}\cdot s_P - (X_C + X_P)\right)^2 = -\left(\beta_{PS}\right)^2 var(\mu_S) - p + \sigma_C^2 + \sigma_P^2.$$

To calculate M's payoff, we rewrite the variances according to equation (5). After rearranging, M's expected payoff can be written as

$$\begin{aligned}
&-\lambda \mathbb{E}\left[\left(\Delta y^{delP} - (X_C + X_S)\right)^2\right] - (1-\lambda)\,\mathbb{E}\left[\left(\Delta y^{delP} - (X_C + X_P)\right)^2\right] \\
=&\left(2\lambda + (1-2\lambda)\,\beta_{PS}\right)\frac{2-\alpha}{2-\alpha\beta_{PS}}c + (1-2\lambda)\,p - \sigma_C^2 - \lambda\sigma_S^2 - (1-\lambda)\,\sigma_P^2.
\end{aligned}$$

Delegated decision-making to S. The decision-rule of S is $\Delta y^{delS} = \theta_S = \frac{\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2} s_C + \frac{\sigma_S^2}{\sigma_S^2+\sigma_{\varepsilon_S}^2} s_S$. S's expected loss is

$$\mathbb{E}\left(\frac{\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2} s_C + \frac{\sigma_S^2}{\sigma_S^2+\sigma_{\varepsilon_S}^2} s_S - (X_C+X_S)\right)^2 = (\frac{-\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2}+1)\sigma_C^2 + (\frac{-\sigma_S^2}{\sigma_S^2+\sigma_{\varepsilon_S}^2}+1)\sigma_S^2.$$

Similarly, we can calculate P's expected loss as

$$\mathbb{E}\left(\frac{\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2} s_C + \frac{\sigma_S^2}{\sigma_S^2+\sigma_{\varepsilon_S}^2} s_S - (X_C+X_P)\right)^2 = s + \sigma_P^2 + (\frac{-\sigma_C^2}{\sigma_C^2+\sigma_{\varepsilon_C}^2}+1)\sigma_C^2 \tag{23}$$

Thus, M's expected payoff can be written as

$$\begin{aligned}
&-\lambda\mathbb{E}\left[\left(\Delta y^{delS} - (X_C+X_S)\right)^2\right] - (1-\lambda)\,\mathbb{E}\left[\left(\Delta y^{delS} - (X_C+X_P)\right)^2\right] \\
=&c + (2\lambda-1)s - \sigma_C^2 - \lambda\sigma_S^2 - (1-\lambda)\,\sigma_P^2.
\end{aligned}$$

□

Proof of Proposition 3. The proof of the proposition consists of two lemmas. Lemma A.5 characterizes the winners of the pairwise comparisons of the three modes of decision-making. In Lemma A.6, we construct the overall optimum of the pairwise winners identified in Lemma A.5. Verifying that the characterization in the proposition fits the general characterization completes the proof. □

Lemma A.5 *There exists three unique cutoffs $\lambda', \lambda'', \lambda''' \in [0,1)$ with the following properties:*
M weakly prefers delegating to S over delegating to P if and only if $\lambda \geq \lambda'''$; M weakly prefers consultative decision-making over delegating to P if and only if $\lambda \geq \lambda'$; and M weakly prefers delegating to S over consultative decision-making if and only if $\lambda \geq \lambda''$.

Proof of Lemma A.5. The values from delegation to P and S are linear functions with the property that $\frac{\partial}{\partial\lambda}\left(\Delta\pi_M^{delS}(\lambda) - \Delta\pi_M^{delP}(\lambda)\right) \geq 0$. Inserting the values from Lemma 1, we have

$$\lambda''' = \max\left\{0, \frac{1}{2} - \frac{1}{2}\frac{cs\alpha}{2s^2+2ps+(2-\alpha)\,pc}\right\}.$$

Note that if $\lambda''' = 0$, then the region where M prefers delegating to P is empty.

Consider next the values from consultative decision-making and delegating to P. The gain $\Delta\pi_M^{con}(\lambda)$ is convex in λ, while $\Delta\pi_M^{delP}(\lambda)$ is linear in λ. Moreover, $\Delta\pi_M^{con}(0) = \Delta\pi_M^{delP}(0)$ and $\Delta\pi_M^{con}(1) > \Delta\pi_M^{delP}(1)$. Hence, there exists $\lambda' \in [0,1)$ defined as the largest solution of $\Delta\pi_M^{con}(\lambda) = \Delta\pi_M^{delP}(\lambda)$.

Finally, the value from delegating to S is linear and satisfies $\Delta\pi_M^{con}(1) = \Delta\pi_M^{delS}(1)$ and $\frac{\partial}{\partial\lambda}\left(\Delta\pi_M^{con}(\lambda) - \Delta\pi_M^{delS}(\lambda)\right)\big|_{\lambda=1} > 0$. Hence, there exists $\lambda \in [0,1)$ such that $\Delta\pi_M^{con}(\lambda) - \Delta\pi_M^{delS}(\lambda) = 0$ if and only if $\Delta\pi_M^{con}(0) \geq \Delta\pi_M^{delS}(0)$. We define λ'' as the maximum of zero and the unique value $\lambda < 1$ satisfying $\Delta\pi_M^{con}(\lambda) - \Delta\pi_M^{delQ}(\lambda) = 0$. For $\lambda'' = 0$, the region in which M prefers consultative decision-making over delegation to S is empty. $\square$

Lemma A.6 *A mode of decision-making is optimal if and only if it takes the following form.*

i) For $\lambda = 1$, consultative decision-making or delegation to S are optimal.

ii) If $\lambda' = \lambda'' = \lambda''' = 0$, then delegation to S is optimal for all $\lambda \in [0,1]$.

iii) If $\lambda''', \lambda'' > 0$, and

(a) $\lambda' \leq \lambda''' \leq \lambda''$, then delegation to S is optimal if and only if $\lambda \in [\lambda'', 1]$, consultative decision-making is optimal if and only if $\lambda \in [\lambda', \lambda''] \cup \{0,1\}$, and delegation to P is optimal if and only if $\lambda \in [0, \lambda']$,

(b) $\lambda' > \lambda''' > \lambda''$, then delegation to P is optimal if and only if $\lambda \in [0, \lambda''']$, delegation to S is optimal if and only if $\lambda \in [\lambda''', 1]$, and consultative decision-making is optimal if and only if $\lambda \in \{0,1\}$.

Proof of Lemma A.6. i) Consultative decision-making and delegation to S achieve the same value for $\lambda = 1$, and the achieved value is higher than the one from delegating to P, $\Delta\pi_M^{delS}\big|_{\lambda=1} = \Delta\pi_M^{con}\big|_{\lambda=1} = c + s > c - p = \Delta\pi_M^{delP}\big|_{\lambda=1}$.

ii) By definition of λ''.

iii) For $\lambda = 0$, we have $\beta_{PS} = \beta_{MS}$, and delegation to P and consultative decision-making achieve the same value, $\Delta\pi_M^{delP}\big|_{\lambda=0} = \beta_{PS}\frac{2-\alpha}{2-\alpha\beta_{PS}}c + p = \beta_{MS}\frac{2-\alpha}{2-\alpha\beta_{MS}}c + p = \Delta\pi_M^{con}\big|_{\lambda=0}$. By i), delegation to S and consultative decision-making achieve the same value for $\lambda = 1$. The value from consultative decision-making is convex in λ; the delegation values are linear in λ with $\frac{\partial}{\partial\lambda}\left(\Delta\pi_M^{delS}(\lambda) - \Delta\pi_M^{delP}(\lambda)\right) \geq 0$. Thus, the upper envelope of the three functions is attained by consultative decision-making at $\lambda = \lambda'''$ if and only if $\lambda' \leq \lambda''' \leq \lambda''$. Otherwise the upper envelope is formed by the delegation functions only. $\square$

Proof of Theorem 1. Consider, first, the difference between the gains from delegating to S and consultative decision-making:

$$\left(\Delta\pi_M^{delS} - \Delta\pi_M^{con}\right)\Big|_{c=0} = (1-\lambda)\left(\frac{(2+\alpha)\lambda - 2}{2-\alpha\lambda}s - (1-\lambda)\frac{2-\alpha}{2-\alpha(1-\lambda)}p\right).$$

There are two points of indifference: $\lambda = 1$ and the unique value of λ_S setting the term in brackets equal to zero. Uniqueness follows from monotonicity in λ. Existence follows from continuity and the fact that the term in brackets is negative for $\lambda = 0$ and positive for $\lambda = 1$. By the implicit function theorem, λ_S is decreasing in s and increasing in p. For $p = 0$, the solution is $\lambda_S = \frac{2}{2+\alpha}$.

Likewise, consider the difference between the gains from delegating to P and consultative decision-making:

$$\left(\Delta\pi_M^{delP} - \Delta\pi_M^{con}\right)\Big|_{c=0} = \lambda\left(-\lambda\frac{2-\alpha}{2-\alpha\lambda}s + \frac{\alpha - (2+\alpha)\lambda}{2-\alpha(1-\lambda)}\right)p.$$

There are two points of indifference: $\lambda = 0$ and the unique value of λ_P setting the term in brackets equal to zero. Uniqueness follows from monotonicity in λ. Existence follows from continuity and the fact that the term in brackets is negative for $\lambda = 1$ and positive for $\lambda = 0$. By the implicit function theorem, λ_P is decreasing in s and increasing in p. For $s = 0$, we have $\lambda_P = \frac{\alpha}{2+\alpha}$. $\square$

Proof of Theorem 2. Note that for $0 < \lambda_P < \lambda_S < 1$, from the proof of Proposition 3, λ_P is the highest value of λ such that $\Delta\pi_M^{con}(\lambda; c) - \Delta\pi_M^{delP}(\lambda; c) = 0$. Likewise, λ_S is the smallest value of λ solving $\Delta\pi_M^{con}(\lambda; c) - \Delta\pi_M^{delS}(\lambda; c) = 0$.

For simplicity, we suppress everything that is kept constant. By the implicit function theorem,

$$\frac{d\lambda_P}{dc} = -\frac{\left(\frac{\partial}{\partial C}\Delta\pi_M^{con}(\lambda_P; c) - \frac{\partial}{\partial C}\Delta\pi_M^{delP}(\lambda_P; c)\right)}{\left(\frac{\partial}{\partial\lambda}\Delta\pi_M^{con}(\lambda_P; c) - \frac{\partial}{\partial\lambda}\Delta\pi_M^{delP}(\lambda_P; c)\right)}.$$

Note that $\frac{\partial}{\partial\lambda}\Delta\pi_M^{con}(\lambda_P;c) > \frac{\partial}{\partial\lambda}\Delta\pi_M^{delP}(\lambda_P;c)$, because M's payoff under consultative decision-making crosses the payoff of delegation to P at λ_P from below. Thus the denominator is positive. The derivative with respect to c of the numerator can be signed as follows. Due to the fact that $\Delta\pi_M^{con}(0;c) = \Delta\pi_M^{delP}(0;c)$, we can write

$$\Delta\pi_M^{con}(\lambda;c) - \Delta\pi_M^{delP}(\lambda;c) = \int_0^{\lambda_P} \frac{\partial}{\partial z}\left(\Delta\pi_M^{con}(z;c) - \Delta\pi_M^{delP}(z;c)\right) dz,$$

so that

$$\frac{\partial}{\partial c}\Delta\pi_M^{con}(\lambda_P;c) - \frac{\partial}{\partial c}\Delta\pi_M^{delP}(\lambda_P;c) = \int_0^{\lambda_P} \frac{\partial^2}{\partial z\partial c}\left(\Delta\pi_M^{con}(z;c) - \Delta\pi_M^{delP}(z;c)\right) dz.$$

Using the specific functional form, we find

$$\frac{\partial^2}{\partial z\partial c}\left(\Delta\pi_M^{con}(z;c) - \Delta\pi_M^{delP}(z;c)\right) = \frac{8s^2(2-\alpha)(1-\lambda)(c+s)}{(c(2-\alpha)+s(2-\lambda\alpha))^3} - \frac{4s^2(2-\alpha)}{(2s+c(2-\alpha))^2}.$$

Evaluated at $\lambda = 0$, the difference is positive. The first term is a concave function of λ. Evaluated at $\lambda = \frac{\alpha}{2+\alpha}$, the difference is positive if and only if

$$\frac{4}{2+\alpha}(c+s)((2-\alpha)c+2s)^2 > ((2-\alpha)c - s\alpha)^3,$$

which is easily verified to be true. Since λ_p is bounded above by $\frac{\alpha}{2+\alpha}$, λ_P is decreasing in the amount of common value information c.

Second, recall that λ_S is defined as the smallest value that solves $\Delta\pi_M^{con}(\lambda;c) - \Delta\pi_M^{delS}(\lambda;c) = 0$. By the implicit function theorem

$$\frac{d\lambda_S}{dc} = -\frac{\left(\frac{\partial}{\partial c}\Delta\pi_M^{con}(\lambda_S;c) - \frac{\partial}{\partial c}\Delta\pi_M^{delS}(\lambda_S;c)\right)}{\left(\frac{\partial}{\partial\lambda}\Delta\pi_M^{con}(\lambda_S;c) - \frac{\partial}{\partial\lambda}\Delta\pi_M^{delS}(\lambda_S;c)\right)},$$

where $\frac{\partial}{\partial\lambda}\Delta\pi_M^{con}(\lambda_S;c) - \frac{\partial}{\partial\lambda}\Delta\pi_M^{delS}(\lambda_S;c) < 0$ because the value of consultative crosses the value of delegation to S from above at λ_S. Straightforward differentiation shows that

$$\frac{\partial}{\partial c}\Delta\pi_M^{con}(\lambda_S;c) - \frac{\partial}{\partial c}\Delta\pi_M^{delS}(\lambda_S;c) = \frac{-4s^2(\lambda_S-1)^2}{(c(2-\alpha)+s(2-\lambda_S\alpha))^2} < 0.$$

Hence, λ_S is decreasing in the common value c. $\square$

Proof of Theorem 3. i) Since $\Delta\pi_M^{con}(0) = \Delta\pi_M^{delP}(0)$, we have that $\lambda_S > 0$ if and only if $\Delta\pi_M^{con}(0) > \Delta\pi_M^{delS}(0)$. Given equation (6), for $\lim_{c\to\infty}$, the comparison at $\lambda = 0$ reduces to

$$\frac{\alpha}{2-\alpha}s - p > 0.$$

ii) We have $\lambda_P > 0$ if and only if $\frac{\partial}{\partial\lambda}\left(\Delta\pi_M^{con}(\lambda) - \Delta\pi_M^{delP}(\lambda)\right)\big|_{\lambda=0} < 0$. Given equation (7), this is equivalent to

$$-\frac{\alpha}{2-\alpha}(p-s) < 0.$$

Given equation (6), λ_S is the solution to $\frac{\alpha}{2-\alpha}s = \frac{2(1-\lambda)-\alpha(1-\lambda)}{2-\alpha(1-\lambda)}p$.

iii) We have that $p > s$ implies $p > \frac{\alpha}{2-\alpha}s$, thus, for $\lambda_S > 0$ case ii) applies. Given equation (7), λ_P is the solution to $\frac{\alpha}{2-\alpha}s = -\frac{2\lambda-\alpha(1-\lambda)}{2-\alpha(1-\lambda)}p$, which is strictly positive under the stated conditions. $\square$

Proof of Theorem 4. We need to show i) that S and P are willing to accept, ii) that M benefits from proposing a deal that S and P accept, and iii) that such a deal effectively puts M at the helm.

i) The participation constraint of S is

$$\delta_S\frac{1}{2}\left(\mathbb{E}\left[\pi_S^*\right] + \mathbb{E}\left[\pi_S\right] - \mathbb{E}\left[\pi_S^*\right]\right) \geq \frac{1}{2}\left(\mathbb{E}\left[\pi_S^*\right] - \sigma_S^2\right),$$

while P participates if

$$\delta_P\frac{1}{2}\left(\mathbb{E}\left[\pi_P^*\right] + \mathbb{E}\left[\pi_P\right] - \mathbb{E}\left[\pi_P^*\right]\right) \geq \frac{1}{2}\left(\mathbb{E}\left[\pi_P^*\right] - \sigma_P^2\right).$$

It is easy to verfiy that $\mathbb{E}\left[\pi_S^*\right] = \mathbb{E}\left[\pi_P^*\right] = \Pi + \sigma^2$, where $\Pi = \left(\frac{a-k_S-k_P}{2}\right)^2$ is firm profit without uncertainty and $\sigma^2 = Var\left(X_P\right) = Var\left(X_S\right)$. Moreover, using the explicit expressions in the proof of Lemma 1, equation (21), we get

$$\delta_S\frac{1}{2}\left(\Pi + \lambda\left(2-\lambda\right)\frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2 - \left(1-\lambda\right)^2\frac{2-\alpha}{2-\alpha\left(1-\lambda\right)}\sigma_P^2\right) \geq \frac{1}{2}\Pi, \qquad (24)$$

and using equation (22), we get

$$\delta_P \frac{1}{2}\left(\Pi + (1-\lambda)\left(2-(1-\lambda)\right)\frac{2-\alpha}{2-\alpha\left(1-\lambda\right)}\sigma_P^2 - \lambda^2 \frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2\right) \geq \frac{1}{2}\Pi. \tag{25}$$

With $\sigma_S^2 = \sigma_P^2 = \sigma^2$, the terms in both brackets simplify for $\lambda = \frac{1}{2}$ to

$$\begin{aligned} & \lambda\left(2-\lambda\right)\frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2 - (1-\lambda)^2 \frac{2-\alpha}{2-\alpha\left(1-\lambda\right)}\sigma_P^2 \\ = \; & (1-\lambda)\left(2-(1-\lambda)\right)\frac{2-\alpha}{2-\alpha\left(1-\lambda\right)}\sigma_P^2 - \lambda^2\frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2 \\ = \; & \frac{1}{2}\frac{2-\alpha}{2-\alpha\frac{1}{2}}\sigma^2. \end{aligned}$$

Thus, the participation constraints reduce to

$$\frac{\delta_i}{1-\delta_i}\frac{1}{2}\frac{2-\alpha}{2-\alpha\frac{1}{2}}\sigma^2 \geq \Pi, \qquad \text{for } i = S, P.$$

Since the participation constraints of S and P are identical and $\frac{\delta}{1-\delta}$ can take any non-negative value, there exists some δ^* that makes the contraints hold with equality.

ii) M receives a fraction of the profits, which were positive already at the outset. Moreover, the surplus increases from adapting to the changes in costs. Hence, M benefits from the deal.

iii) Finally, since $\delta_S = \delta_P$ and $\omega_S = \omega_P = \frac{1}{2}$,

$$\lambda^* = \frac{\left(1-\delta_S\right)\omega_S}{\left(1-\delta_S\right)\omega_S + \left(1-\delta_P\right)\omega_P} = \frac{1}{2},$$

implying that M prefers consultative decision-making over delegating to S or P. □

Proof of Theorem 6. Recall the participations constraints of S (equation (24)) and P (equation (25)) from the proof of Theorem 4. The constraints have the same structure. Without loss of generality, fix σ_P^2 and focus on S's participation decision

$$\delta_S \frac{1}{2}\left(\Pi + \lambda\left(2-\lambda\right)\frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2 - (1-\lambda)^2 \frac{2-\alpha}{2-\alpha\left(1-\lambda\right)}\sigma_P^2\right) \geq \frac{1}{2}\Pi.$$

Since $\omega_S = \omega_P$, we have
$$\lambda = \frac{1-\delta_S}{1-\delta_S+1-\delta_P}.$$
Note that for any $\delta_P < 1$, λ is well defined for all $\delta_S < 1$. For $\sigma_S^2 = 0$,
$$\delta_S \frac{1}{2}\left(\Pi - (1-\lambda)^2 \frac{2-\alpha}{2-\alpha(1-\lambda)}\sigma_P^2\right) < \frac{1}{2}\Pi \quad \text{for all } \delta_S.$$
Moreover,
$$\max_{\delta_S,\delta_P} \delta_S \frac{1}{2}\left(\Pi + \lambda(2-\lambda)\frac{2-\alpha}{2-\alpha\lambda}\sigma_S^2 - (1-\lambda)^2 \frac{2-\alpha}{2-\alpha(1-\lambda)}\sigma_P^2\right)$$
is increasing in σ_S^2. Hence, for σ_S^2 close enough to zero, S's participation constraint is not met. □

Proof of Theorem 7. Recall equation (10). Under consultative decision-making, information about S's private value is provided if and only if $\sigma_C^2 \leq \frac{(2-\alpha)\frac{\sigma_C^2+\lambda\sigma_S^2}{\sigma_C^2+\sigma_S^2}}{2-\alpha\frac{\sigma_C^2+\lambda\sigma_S^2}{\sigma_C^2+\sigma_S^2}}\left(\sigma_C^2+\lambda\sigma_S^2\right)$, which can be reduced to $\lambda \geq \lambda^* = \frac{\sqrt{(4-\alpha)^2+8(2-\alpha)\frac{\sigma_S^2}{\sigma_C^2}}-(4-\alpha)}{2(2-\alpha)\frac{\sigma_S^2}{\sigma_C^2}}$, a decreasing function of $\frac{\sigma_S^2}{\sigma_C^2}$.

In the almost common value case, by L'Hospital's rule, we obtain
$$\lim_{\frac{\sigma_S^2}{\sigma_C^2}\to 0} \frac{\sqrt{(4-\alpha)^2+8(2-\alpha)\frac{\sigma_S^2}{\sigma_C^2}}-(4-\alpha)}{2(2-\alpha)\frac{\sigma_S^2}{\sigma_C^2}} = \frac{2}{4-\alpha}.$$

Note first that $\frac{2}{4-\alpha} > \frac{1}{2}$. Hence, for $\lambda \leq \frac{1}{2}$, under consultative decision-making, S has no private value information. Therefore, the point of indifference λ_P between (9) and (10) is computed for $s = 0$. This implies $\lambda_P = \frac{\alpha}{2+\alpha}$, the threshold for the case where the only information to extract is private value information of P.

Consider now λ_S. Note that $\Delta\pi_M^{con}(\lambda)$ is continuous and convex in λ. Moreover, $\Delta\pi_M^{con}(1) = \Delta\pi_M^{delS}(1)$ and $\Delta\pi_M^{con}$ is steeper at $\lambda = 1$ than $\Delta\pi_M^{delS}$. This implies that if $\lambda^* \geq \lambda_S$, then consultative decision-making remains dominated by delegation to S, even for $\lambda \geq \lambda^*$Hence, there are two cases to consider:

i) $\lambda^* = \frac{2}{4-\alpha} \leq \lambda_S$ (S has private value information under consultative decision-making at the point of indifference). In this case, it is straightforward to see that the comparison is as if information were exogenously given and perfect. Using the explicit expression in the theorem, this arises if and only if $\frac{\sigma_S^2}{\sigma_P^2} \leq \frac{1}{\alpha}\frac{(2-\alpha)^3}{\alpha^2-4\alpha+8}$.

ii) $\lambda^* = \frac{2}{4-\alpha} > \lambda_S$ (S has no private value information under consultative decision-making at the point of indifference). Since σ_C^2 drops out of the comparison in this case, the point of indifference is as indicated in the theorem. □

References

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

Alonso, R., Dessein, W. and Matouschek, N. (2008). When does coordination require centralization?, *American Economic Review* **98**(1): 145–179.

Alonso, R. and Matouschek, N. (2008). Optimal delegation, *The Review of Economic Studies* **75**(1): 259–293.

Ambrus, A. and Takahashi, S. (2008). Multi-sender cheap talk with restricted state spaces, *Theoretical Economics* **3**(1): 1–27.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk, *Mathematical finance* **9**(3): 203–228.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk, *Econometrica* **70**(4): 1379–1401.

Blume, A., Board, O. J. and Kawamura, K. (2007). Noisy talk, *Theoretical Economics* **2**(4): 395–440.

Blume, A., Deimen, I. and Inoue, S. (2020). Information processing: Contracts versus communication. Working paper.

Cachon, G. P. and Lariviere, M. A. (2005). Supply chain coordination with revenue-sharing contracts: strengths and limitations, *Management science* **51**(1): 30–44.

Chakraborty, A. and Harbaugh, R. (2007). Comparative cheap talk, *Journal of Economic Theory* **132**(1): 70–94.

Coase, R. H. (1937). The nature of the firm, *Economica* **4**(16): 386–405.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission, *Econometrica* **50**(6): 1431–1451.

Deimen, I. and Szalay, D. (2019). Delegated expertise, authority, and communication, *American Economic Review* **109**(4): 1349–74.

Dessein, W. (2002). Authority and communication in organizations, *The Review of Economic Studies* **69**(4): 811–838.

Dessein, W. (2013). Incomplete contracts and firm boundaries: New directions, *Journal of Law, Economics, and Organization* **30**(suppl 1): i13–i36.

Fang, K.-T., Kotz, S. and Ng, K. W. (1990). *Symmetric multivariate and related distributions*, Chapman and Hall.

Gibbons, R., Matouschek, N. and Roberts, J. (2012). *Decisions in Organizations*, Princeton University Press, pp. 373–431.

Gordon, S. (2010). On infinite cheap talk equilibria. Working paper.

Grossman, S. J. and Hart, O. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* **94**(4): 691–719.

Hagenbach, J. and Koessler, F. (2010). Strategic Communication Networks, *The Review of Economic Studies* **77**(3): 1072–1099.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm, *Journal of Political Economy* **98**(6): 1119–1158.

Hidir, S. and Migrow, D. (2019). Credibility, efficiency, and the structure of authority. Working paper.

Holmström, B. and Tirole, J. (1991). Transfer pricing and organizational form, *Journal of Law, Economics, and Organization* **7**(2): 201–228.

Kotz, S., Kozubowski, T. and Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhäuser.

Lafontaine, F. and Slade, M. (2007). Vertical integration and firm boundaries: The evidence, *Journal of Economic Literature* **45**(3): 629–685.

Levy, G. and Razin, R. (2007). On the limits of communication in multidimensional cheap talk: a comment, *Econometrica* **75**(3): 885–893.

Li, M. (2010). Advice from multiple experts: A comparison of simultaneous, sequential, and hierarchical communication, *The B.E. Journal of Theoretical Economics* **10**: 18–18.

Li, M. and Madarász, K. (2008). When mandatory disclosure hurts: Expert advice and conflicting interests, *Journal of Economic Theory* **139**(1): 47–74.

Liu, S. and Migrow, D. (2019). Designing organizations in volatile markets. Working paper.

McGee, A. and Yang, H. (2013). Cheap talk with two senders and complementary information, *Games and Economic Behavior* **79**: 181 – 191.

Meyer, M., de Barreda, I. M. and Nafziger, J. (forthcoming). Robustness of full revelation in multisender cheap talk, *Theoretical Economics* .

Meyer-ter Vehn, M., Smith, L. and Bognar, K. (2018). A Conversational War of Attrition, *The Review of Economic Studies* **85**(3): 1897–1935.

Rantakari, H. (2008). Governing adaptation, *The Review of Economic Studies* **75**(4): 1257–1285.

Simon, H. A. (1951). A formal theory of the employment relationship, *Econometrica* **19**(3): 293–305.

Szalay, D. (2012). Strategic information transmission and stochastic orders. Working paper.

Wolinsky, A. (2002). Eliciting information from multiple experts, *Games and Economic Behavior* **41**(1): 141 – 160.