Organized Information Transmission

Laurent Mathevet and Ina Taneva[*]

January 24, 2022

Abstract

We formalize the concepts of horizontal and vertical information transmission and introduce two families of information structures, namely single-meeting schemes and delegated hierarchies, that specialize along these dimensions. We characterize the strategic outcomes that they implement in general finite incomplete information games and illustrate the resulting linear programming approach in a linear network example. We build on the characterizations to show that these families are unconstrained-optimal in binary-action games with strategic complementarities. Finally, we generalize these families to multiple meetings and random hierarchies and characterize the corresponding strategic outcomes.

(*JEL codes*: C72, D82, D83.)

Keywords: Incomplete information, information hierarchy, delegated transmission, meeting scheme, Bayes correlated equilibrium, information design.

[*]Department of Economics, European University Institute, laurent.mathevet@eui.eu. Department of Economics, University of Edinburgh, ina.taneva@ed.ac.uk. We thank Sylvain Chassang, Russell Cooper, Olivier Gossner (discussant), Elliot Lipnowski, Meg Meyer, Stephen Morris, Alessandro Pavan, David Pearce, József Sákovics, Joel Sobel, Guylene Spaziani, Jakub Steiner and Dong Wei for very valuable discussions and suggestions. We would also like to thank conference participants at the CRETA conference at Warwick, Frontiers in Design at UCL, Information Design and Splitting Games in Paris, the HSE-Vienna Workshop at Higher School of Economics in Moscow, SAET 2019 in Ischia, the quarterly CS and Economics Workshop at Northwestern, the Kent-Bristol-City Workshop in Economic Theory, ESWC 2020, CETC 2021, GAMES 2021, Workshop on Strategic Communication and Learning at Stony Brook, as well as seminar audiences at Bilkent, Bonn, Cambridge, Edinburgh, EUI, Glasgow, Graz, Leicester, NY Microsoft Research, Oxford, Saint-Etienne, Southampton, Toulouse, UCL, Venice and Zurich. Earlier versions of this paper circulated under the titles "Information Hierarchies" and "Optimal Information Hierarchies".

1 Introduction

Organizations operate under uncertainty about consumer demand, regulations, raw materials, equipment performance, production quality, technological change, financing costs, and so on. Protocols are put in place, either formally or informally, to transmit relevant information to the right members of the organization. In practice, horizontal and vertical transmission protocols are ubiquitous in organizations.

Horizontal transmission refers to informing a group of individuals symmetrically and simultaneously. Meetings are the ultimate example, where a selectively invited audience jointly receives information about exogenous events. For instance, a sudden shortage of raw materials typically prompts a meeting between the general director, supply chain management and the production team, in which related information is shared. Reported noncompliance with quality standards results in a meeting between the general director and the quality department, while a meeting between the marketing and the engineering teams is organized to share information about new regulations in view of product adaptation.

Vertical transmission, instead, refers to information passed down sequentially, and perhaps partially, from one individual to another. Downstream and upstream communications in organizations are obvious examples. Restructuring in response to competitive pressures and decreasing demand is typically implemented by downstream communication: the headquarters becomes aware of the economic situation and communicates with the manufacturing director, who in turn conveys the necessary actions to the local head of human resources. Conversely, equipment failure typically initiates an upstream communication sequence: the problem appears to a worker on the factory floor, who reports it to his shift supervisor, who forwards it to the production manager, who forwards it to the chief production officer.

In economic theory, these protocols are modeled by information structures, which specify what information agents learn about the realization of a payoff-relevant state. In this paper, we formally describe all information structures according to their capacity to transmit information horizontally or vertically, based on a new informativeness order, and within that formalism analyze two special families of information structures as proof of concept. These families of information structures formalize simple organizational protocols, inspired by the practical constraints of real-world transmission.

The first family of information structures we consider are *single-meeting schemes*. They represent a natural limit case of horizontal transmission and, yet, a generalization of public information, as they require that each message profile be communicated publicly to a *subset* of agents, while all other agents remain uninformed. In organizations, for example, this is akin to a meeting during which an announcement will be made to all people who have been invited to attend.[1] These information structures are intuitively appealing, because only one meeting ever takes place or one email is ever sent to selected recipients. This can save on physical communication costs, since the same information is transmitted to many people at once instead of many times to each of them individually.

The second family of information structures we consider are *delegated hierarchies*. They represent a natural limit case of vertical transmission, as they require that each message profile be communicated to one agent only, after which it gets transmitted, possibly partially but truthfully, down a fixed hierarchy of agents in a decentralized way. This requires agents to be totally ordered in terms of informedness, so that they are able to pass down the relevant information to the next agent in the hierarchy. It also requires that transmission incentives be satisfied, so that agents are willing to pass down information truthfully. Delegated transmission is widely used in many contexts as it allows for efficient communication with only one individual at the top of the hierarchy, while ensuring that the relevant information reaches all levels of the organization. As such, delegating information transmission to the receivers themselves can also economize on physical communication costs.

We characterize the strategic outcomes that can emerge in general finite incomplete information games when information is restricted to those families of information structures. Horizontal and vertical transmission naturally come to mind when thinking about how a group of agents could self-organize, or be organized by a third party, to receive information. Our characterizations delineate the outcome distributions that emerge under pure Bayes Nash equilibrium when the information protocol takes the form of single-meeting schemes (Theorem 1) or delegated hierarchies (Theorem 2).[2] These theorems connect the organizational constraints on information to the resulting strategic constraints by means of linear inequalities, which are stronger than the Bayes correlated equilibrium constraints

[1] American businesses hold millions of meetings a day, a billion meetings a year, and the average employee spends hours in meetings every week. See Rogelberg, Scott, and Kello (2007) for numbers.

[2] While these characterizations are in terms of distributions over actions and states, direct implementation à la Bergemann and Morris (2016) and Taneva (2019) — by using action recommendations as messages — will generally not succeed in implementing the desired outcome distributions by a single-meeting scheme or a delegated hierarchy.

of Bergemann and Morris (2016).

The information design literature is often challenged for giving unlimited freedom of choice of information structures. Our formalization of horizontal and vertical transmission is a novel proposal for how to conceptually address this concern which also leads to tractable analysis. The characterizing inequalities can be applied to constrained information design, where they impose organizational restrictions on the set of admissible information structures. In that perspective, we apply our characterizations to a linear network example, where we demonstrate the linear programming approach and draw some basic insights from the optimal single-meeting scheme and delegated hierarchy.

In comparison, direct information structures, which invoke the revelation principle (Myerson (1991)) and make incentive-compatible action recommendations, do not constrain information to be commonly observed by some agents or to be transmittable from one agent to another in an incentive compatible manner. This can make them very difficult to implement in reality. Indeed, in many environments,[3] optimal direct information structures are not single-meeting schemes or delegated hierarchies, as the action recommendations they involve are private information to each agent, which others are uncertain about. To implement these privacy requirements, there is hardly any other way but to communicate with every single agent individually. In large organizations or markets, this describes an unrealistic picture of information transmission.[4]

We build on the characterizations to show that our families of information structures are unconstrained-optimal in binary-action games with strategic complementarities for a wide class of objectives. This encompasses important economic environments, such as global games of regime change, team effort problems and purchase decisions with network effects, where maximizing the probability of regime change, total expected efforts, or total expected purchases, are natural objectives. These results answer the question of when optimal outcomes can be implemented by simple, realistic protocols, which is analogous to questions in mechanism design concerning the optimality of, for example, posted-price mechanisms.

[3]These include monotone environments, in which agents play an incomplete information game with strategic complementarities and the sender wants to foster large actions. For example, consider a manager who wants to maximize total efforts by disclosing information about the profitability of a project to a team of workers.

[4]Concerns about direct mechanisms have been raised previously by Van Zandt (2007): "The Revelation Principle in mechanism design is both a blessing and a curse [. . .] It is a curse because direct mechanisms provide such an unrealistic picture of decision-making in organizations." We express similar concerns about direct information structures and information transmission in organizations.

The first optimality result (Proposition 1) establishes the optimality of single-meeting schemes in binary-action environments where each agent's payoff is increasing in the actions of the other agents. Intuitively, if agents' actions are strategic complements, then some degree of shared information should help coordination. Yet, we know that pure public information can be strictly suboptimal in such contexts, for example when the objective is to maximize the total probability of the high action among heterogenous agents. The result sheds light on this matter by showing that single-meeting schemes provide the optimal degree of shared information. The second optimality result (Proposition 2) demonstrates that delegated hierarchies are optimal under the stronger complementarity condition that each agent's utility be supermodular in action profiles and states. This result implies that an informational line network, where agents are totally ordered by informedness, is optimal, even though agents' utilities may be only partially ordered according to the strengths of their dependencies on the state and each other's actions.

We apply our optimality results to a classical global game of regime change à la Sakovics and Steiner (2012), where we show how the optimal single-meeting scheme and delegated hierarchy change with the objective function. In a different application, we illustrate that optimization may require agents to be treated equally, because they should receive the same information, while, at the same time, delegation requires them to be ordered in one specific way.

Finally, we generalize our concepts of organized information transmission to *multiple-meeting schemes*, where more than one meeting can be organized at each message profile, and to *random delegated hierarchies*, where the hierarchy of agents can vary with the message profile. The corresponding definitions and outcome characterizations are discussed in Section 6 and Online Appendix B.3.

Related Literature. Various definitions of correlated equilibrium in games with incomplete information have been proposed (Forges (1993, 2006), Bergemann and Morris (2016)), depending on what variables the mediator or the correlating device can condition on to correlate agents' actions (e.g., the state, private types). Our characterizations propose new forms of correlated equilibrium, where the ability to correlate behavior is not limited by the conditioning variables, but by the organizational structure of information. In doing so, we bring an organizational perspective to the formulation of incomplete information and study the resulting strategic implications.

This paper contributes to the information design literature, surveyed in Bergemann and Morris (2019) and Kamenica (2019), by importing organizational considerations into the designer's problem. In particular, our optimality results contribute to the recent interest in binary-action supermodular games. Arieli and Babichenko (2019), Candogan and Drakopoulos (2020) and Candogan (2020) study the optimal design of *direct* information structures in binary-action supermodular games (with binary states or in linear networks). Our motivation is different and highlights the role of *indirect* information structures. In binary-action supermodular games, Oyama and Takahashi (2020) focus on equilibrium robustness and Morris, Oyama, and Takahashi (2020) on implementation in the smallest equilibrium through information design. Under adversarial equilibrium selection, Inostroza and Pavan (2020) derive an optimal public information structure, while Li, Song, and Zhao (2019) and Mathevet, Perego, and Taneva (2020) show the importance of indirect private information structures. In contrast to these papers, our focus on indirect information structures is not driven by equilibrium multiplicity or selection, but by organizational concerns that cannot generall be satisfied by direct information structures.

In a multi-player email game, Morris (2002) introduces the concept of locally public communication where subgroups of players meet sequentially and share information that becomes common knowledge within each meeting. The size of the meetings is shown to play a pivotal role in making coordination possible. Two recent papers examine how information is structured or shared among many agents. Brooks, Frankel, and Kamenica (2020) define an information hierarchy as a partially ordered set of agents, where the order describes who is better informed about the state in the Blackwell sense. They characterize the information hierarchies that are compatible with the strong "informedness" order, in which higher ranked agents know all the information of less informed ones. Our paper is concerned with the strategic implications and the optimality of related information structures.[5] Galperti and Perego (2020) study the impact of information spillovers on the outcomes of incomplete information games. Their notion of an information system assumes that messages are automatically shared between linked agents on a network.

Hierarchical transmission relates to strategic communication. Ambrus, Azevedo, and Kamada (2013) and Laclau, Renou, and Venel (2020) study rich cheap talk intermediation networks between a sender and a single receiver, where the intermediators do not interact with each other beyond the transmission of messages. Our agents both receive and send information, and strategically choose an action; their ordering and strategic interdepen-

[5]Our definition of delegated hierarchy completely ranks the agents under the strong informedness order and adds incentive compatibility to information transmission.

dence play a pivotal role in determining what outcomes are implementable with delegated hierarchies (a special case of a line network). Hagenbach and Koessler (2010) and Galeotti, Ghiglino, and Squintani (2013) examine strategic (and simultaneous) pre-play communication and characterize the communication networks that emerge in equilibrium under quadratic payoffs. Our delegated hierarchies represent an equilibrium communication network in pre-play communication over multiple rounds, when only the highest-ranked agent starts off with private information. Rivera (2018) characterizes the communication networks between the designer and the agents, for which one can implement the entire set of (Bayes) correlated equilibria for any game. Instead, we impose natural restrictions on the communication network, which, in general, do not allow for implementability of the whole set of (Bayes) correlated equilibria.

Finally, our work is also related to information transmission within organizations (Radner (1993), Van Zandt (1999), Rantakari (2008), Alonso, Dessein, and Matouschek (2008), Hori (2006), Dessein (2002), Crémer, Garicano, and Prat (2007) among others). In our framework, however, the principal is not trying to elicit information from better informed agents, but instead to disseminate it effectively throughout the organization. Mookherjee and Tsumagari (2014) consider mechanism design under restricted communication protocols motivated by the presence of communication costs associated with the length of messages sent, the capacity of the communication channel or time delays.

2 Model

A set of agents $\mathscr{I} = \{1,\ldots,n\}$ interact in an environment where an unknown state ω is drawn from a finite set Ω according to prior $\mu \in \Delta(\Omega)$, which is common knowledge. Agents simultaneously choose actions $a_i \in A_i$, where A_i is finite for each $i \in \mathscr{I}$. Payoffs are given by $u_i : A \times \Omega \to \mathbb{R}$ for each $i \in \mathscr{I}$, where $A = \times_i A_i$.

2.1 Information and Outcomes

An information structure formalizes the protocol by which information about the state is distributed to the agents. Formally, an *information structure* is a pair (S,P), where $S = \prod_i S_i$ is a finite message space and $P = \{P(\cdot|\omega)\}_{\omega\in\Omega}$ is a family of conditional distributions over S. In any state ω, message profile $s = (s_i)_i$ is drawn with probability $P(s|\omega)$ and i observes $s_i \in S_i$. Abusing notation, we also use P to denote the marginal distribution of any subset of messages, such as $P(s_i)$ or $P(s_i,s_j)$. Without loss, assume $P(s_i) > 0$ for all $s_i \in S_i$ and $i \in \mathscr{I}$.

Given prior μ, information structure (S,P), and upon receiving message $s_i \in S_i$, agent i formulates beliefs

$$\mu_i\left(s'_{-i},\omega|s_i\right) = \frac{\mu(\omega)P(s'_{-i},s_i|\omega)}{P(s_i)} \tag{1}$$

that the state is ω and that the other agents have received s'_{-i}. For reasons that will become clear later, we are also interested in the beliefs agent i would form if he were to observe agent j's message in addition to his own. For each $s_i \in S_i$ and $s_j \in S_j$ such that $P(s_i,s_j) > 0$, those beliefs are given by:

$$\mu_i\left(s'_{-i},\omega|s_i,s_j\right) = \frac{\mu(\omega)P(s'_{-i},s_i|\omega)}{P(s_i,s_j)}\,\mathbb{1}_{\{s_j=s'_j\}}.$$

Given an information structure (S,P), agents play a *pure Bayes Nash equilibrium (BNE)* and the set of those is defined as

$$\mathscr{E}(S,P) := \Big\{a^* = (a_i^*)_i : a_i^* : S_i \to A_i \text{ and}$$
$$a_i^*(s_i) \in \operatorname*{argmax}_{a_i \in A_i} \sum_{\omega,s_{-i}} u_i(a_i,a^*_{-i}(s_{-i});\omega)\mu_i(s_{-i},\omega|s_i) \quad \forall i \in \mathscr{I}, s_i \in S_i\Big\}. \tag{2}$$

Play across all states can be described by an outcome distribution $p \in \Delta(A \times \Omega)$. An outcome distribution p is (weakly) *implemented* by (S,P) if it results from equilibrium play, that is, if there is $a^* \in \mathscr{E}(S,P)$ such that $p(a,\omega) = \sum_{s\in S}\mu(\omega)P(\{s : a^*(s) = a\}|\omega)$ for all $a \in A$ and $\omega \in \Omega$.

From Bergemann and Morris (2016, Theorem 1), we know that p is implementable by some (S,P) if and only if $p(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$, and for all $i \in \mathscr{I}$ and $a_i \in A_i$,

$$\sum_{\omega\in\Omega}\sum_{a_{-i}\in A_{-i}} p(a,\omega)\big(u_i(a;\omega) - u_i(a'_i,a_{-i};\omega)\big) \geq 0 \quad \forall a'_i \in A_i. \tag{3}$$

These conditions define the set of *Bayes correlated equilibria (BCE)* for a given prior μ, which we denote by $BCE(\mu)$.

2.2 Organized Information

Given an information structure (S,P) and $s \in S$ such that $P(s) > 0$, agent i is (weakly) *more informed* than j at s, denoted $i \succeq^s_{\text{Inf}} j$, if $\mu_i(\omega, s'_{-i}|s_i, s_j) = \mu_i(\omega, s'_{-i}|s_i)$ for all $\omega \in \Omega$ and $s'_{-i} \in S_{-i}$.[6] This is the strongest definition of 'being more informed,' as i knows everything that j knows, including j's message. We say that i and j are *equally informed* at s, denoted $i =^s_{\text{Inf}} j$, if $i \succeq^s_{\text{Inf}} j$ and $j \succeq^s_{\text{Inf}} i$.

The informedness order $\succeq_{\text{Inf}} = \left(\succeq^s_{\text{Inf}}\right)_s$ provides a useful language to formalize how an information structure organizes agents based on their interactive knowledge. An information structure (S,P) allows *horizontal transmission* to i and j at s if $i =^s_{\text{Inf}} j$. Since both agents know each other's message, they both know that they both know it and so on, so that this fact is common knowledge among them; it is as if (s_i, s_j) were communicated simultaneously and overtly to both i and j. The ultimate example of horizontal transmission is public information, which requires that $i =^s_{\text{Inf}} j$ for all $s \in S$ and $i, j \in \mathscr{I}$. An information structure (S,P) allows *vertical transmission* from i to j at s if $i \succeq^s_{\text{Inf}} j$ and i satisfies communication incentives toward j at s, as formalized in Section 3.2. Thus, while vertical transmission is more permissive than horizontal transmission from an informedness perspective, it requires strategic incentives that horizontal transmission does not.

Many information structures are such that $\left\{(i,j) \in \mathscr{I}^2 : i \succeq^s_{\text{Inf}} j\right\} = \emptyset$ for all s and hence do not allow horizontal or vertical transmission between any i and j at any s. Some may allow it only between a few agents and at a few messages. Yet others may allow horizontal transmission at some messages and vertical transmission at other messages. The information structures in this paper are special because they specialize in either horizontal or vertical transmission across all messages in their support.

3 Characterizations of Implementable Outcomes

We focus on two particular classes of information structures, which represent natural limit cases of organized transmission and serve as proof of concept. The first class, *single-meeting schemes*, captures horizontal transmission through the requirement that each message profile be communicated via a single meeting. The second class, *delegated hierarchies*, captures vertical transmission by requiring that information flow along the

[6] While the informedness ranking between i and j at s is independent of $s_{-(ij)}$, we write $i \succeq^s_{\text{Inf}} j$ instead of $i \succeq^{(s_i,s_j)}_{\text{Inf}} j$, to simplify notation.

same hierarchical order for each message profile. We formally define each class of information structures and characterize the set of outcomes that they implement in pure BNE. The characterizations can also be viewed as solutions concepts that capture common certainty of rationality and that agents have self-organized or been organized to receive their information in a single meeting or through a delegated hierarchy. In Section 6.1 and Online Appendix B.3.1 we generalize these classes of information structures by allowing for multiple-meetings per message profile and for the hierarchical informedness order to vary with the message profile.

3.1 Horizontal Transmission: Single-Meeting Schemes

The transmission of private messages can come at significant costs due to the necessity of creating and using separate communication channels to maintain the required privacy. In such cases, meetings present a cost-effective alternative by communicating content, simultaneously and overtly (i.e., on one and the same channel) to subsets of agents. Meetings embody horizontal transmission to those in attendance, as it is common knowledge amongst them that they receive the same information. At the same time, unlike purely public information, they do allow for informational asymmetries: those who are not invited to a meeting are less informed than those who attend it and not necessarily equally informed to each other. While meetings can take many forms and serve many purposes, in this section we introduce a family of information structures that require a single communication channel, that is, only one meeting, to communicate the content of any message profile. Our definition formalizes these stylized points by designating only one message per agent, $\tilde{s}_i$, to represent i's non-participation in any meeting, and by building common knowledge between the participants in a meeting.

Definition 1. *An information structure* (S,P) *is a single-meeting scheme if there exist a collection* $\{M(s) \subseteq \mathscr{I} : s \in S \text{ s.t. } P(s) > 0\}$ *and at most one* $\tilde{s}_i \in S_i$ *for each* $i \in \mathscr{I}$ *such that* $i \in M(s)$ *implies* $i \succeq^s_{\text{Inf}} j$ *for all* $j \in \mathscr{I}$ *and* $i \notin M(s)$ *implies* $s_i = \tilde{s}_i$.

For each message profile s, the subset of agents $M(s)$ is invited to a meeting at which s is communicated to them. Although many different meetings may be possible ex-ante, as described by the collection in the definition, at most one meeting is ever organized given a realized message profile s. The ability to decide which meeting to call depending on the nature of information that needs to be communicated is an important feature of organizations, as outlined in the introduction. When i and j are in $M(s)$, $i =^s_{\text{Inf}} j$, that is, they are equally informed. In addition, participation in a meeting perfectly reveals, to those invited, the subset of non-participants $\mathscr{I} \setminus M(s)$ and, therefore, their respective beliefs. Note that an agent can have different beliefs depending on the meeting he participates in and the s that gets announced at that meeting, while there is only one way of not participating in any meeting and, hence, every i has only one belief associated with message $\tilde{s}_i$. This does not mean $\mu_i(\cdot|\tilde{s}_i) = \mu_j(\cdot|\tilde{s}_j)$ for distinct i and j, as non-participation may carry different information for different agents. The reason why non-participation in a meeting needs to be associated with only one message per agent is due to the requirement that communication be done on a single channel, only with the agents invited to a meeting. If we allowed an agent to have multiple ways of not participating in any meeting, that would require communication with him about the particular way in which he is not participating, and hence, an additional communication channel, besides the meeting, is needed.

Note that single-meeting schemes are, by definition, more general than public information. Moreover, they need not be thought of as physical meetings. For example, they could also represent group emails such that in every contingency at most one message is ever sent to a subgroup of agents listed in the "To:" field. This avoids having to send that same email individually to each recipient in order to satisfy privacy requirements. Finally, while at most one meeting ever takes place in Definition 1, we generalize this concept to multiple simultaneous meetings in Section 6 and Online Appendix B.3.1.

Our first result characterizes the outcome distributions that can be implemented by single-meeting schemes.

Theorem 1. *A distribution $p \in BCE(\mu)$ can be implemented by a single-meeting scheme, if and only if, for all $i \in \mathscr{I}$, there is $\tilde{a}_i \in A_i$ such that for all $a_i \in A_i \setminus \{\tilde{a}_i\}$*

$$\sum_{\omega \in \Omega} p(a_i, a_{-i}, \omega)\big(u_i(a_i, a_{-i}; \omega) - u_i(a_i', a_{-i}; \omega)\big) \geq 0 \tag{4}$$

for all $a_i' \in A_i$ and $a_{-i} \in A_{-i}$.

For any prior μ, denote by $SMS(\mu)$ the set of BCE distributions that satisfy these necessary and sufficient conditions. Note that the characterization provides a system of linear inequalities for each $\tilde{a} \in A$, the solutions of which, $\mathscr{C}(\tilde{a}, \mu)$, represent a class of single-meeting schemes. Thus, $SMS(\mu) = \cup_{\tilde{a} \in A} \mathscr{C}(\tilde{a}, \mu)$.

The theorem makes clear the connection between the informational constraints of single-meeting schemes and the resulting strategic constraints: Each agent i has one action $\tilde{a}_i$

that satisfies the BCE obedience constraint (3), while every other action that i plays should be a best response to any a_{-i} against which it is played with positive probability. In comparison, the BCE constraints are summations over all a_{-i} of the constraints specified separately for each a_{-i} in (4). Thus, the above incentive constraints are stronger than the obedience constraints of a BCE, but weaker than (pure strategy) equilibrium play under public information, where all constraints are of the form (4) without any exception.[7]

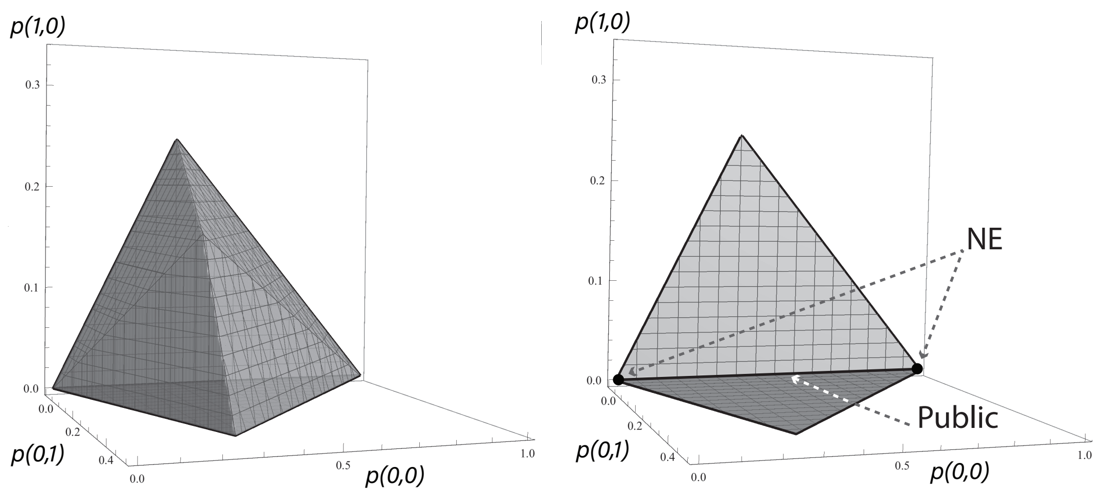
For each distribution in $\mathscr{C}(\tilde{a},\mu)$, a realization $a \in A$ can be interpreted as a meeting amongst $\{i \in \mathscr{I} : a_i \neq \tilde{a}_i\}$, because by (4) each of these agents would want to follow his action recommendation even if he knew the actions of all other agents. Indeed, the proof of the theorem shows that any single-meeting scheme $p \in \mathscr{C}(\tilde{a},\mu)$ can be implemented by a canonical information structure that organizes meetings in this way and sends incentive compatible, *augmented* action recommendations. For every a with $p(a) > 0$, the canonical information structure sends the message (a_i, a_{-i}) to every agent in a meeting; that is, to every $i \in \mathscr{I}$ such that $a_i \neq \tilde{a}_i$, it sends a message with his own action recommendation a_i augmented by the action recommendations of all other agents a_{-i}. Furthermore, to every agent outside the meeting, that is to every $i \in \mathscr{I}$ such that $a_i = \tilde{a}_i$, the canonical information structure sends only their own action recommendation $\tilde{a}_i$. Note that an agent can play the same action in different meetings, that is, under different augmentations of the same own action recommendation, as well as in and outside a meeting (if $\tilde{a}_i$ satisfies (4) for some a_{-i}). In contrast, direct information structures, based on the Revelation Principle, associate a different action to each message an agent receives.

In the rest of this section, we emphasize two properties of the set of single-meeting schemes. To allow graphical representation, we focus on the special case of complete information, where Ω is singleton. However, the properties generalize to incomplete information games.

	0	1
0	3,2	0,0
1	0,0	2,3

Table 1: Battle of the Sexes

[7] Relative to the framework of Galperti and Perego (2020), our single-meeting schemes can be interpreted as stochastic networks where the network itself conveys information about the state.



(a) Correlated Equilibria (b) SMS

Figure 1: Correlated Equilibria and SMS

Consider the Battle of the Sexes game in Table 1. Figure 1 (a) depicts $BCE(\mu)$, which is simply the set of (complete information) correlated equilibria, and part (b) depicts $SMS(\mu)$, which consists of four classes of single-meeting schemes:

$$
\begin{aligned}
\mathscr{C}((0,1),\mu) &= \{p: \quad p(0,0), p(0,1), p(1,1) \geq 0 \text{ and } p(1,0)=0\} \\
\mathscr{C}((1,0),\mu) &= \{p: \quad p(0,0), p(1,0), p(1,1) \geq 0 \text{ and } p(0,1)=0\} \\
\mathscr{C}((0,0),\mu) = \mathscr{C}((1,1),\mu) &= \{p: \quad p(0,1)=p(1,0)=0\}.
\end{aligned}
$$

In $\mathscr{C}((0,1),\mu)$, the row agent plays 1 when invited to a meeting and 0 when not. The opposite is true for the other agent. This class corresponds to the bottom triangle in Figure 1 (b). $\mathscr{C}((1,0),\mu)$ is the mirror image of $\mathscr{C}((0,1),\mu)$, given by the top triangle. In $\mathscr{C}((0,0),\mu)$ and $\mathscr{C}((1,1),\mu)$, the agents are always together in a meeting and coordinate their actions.

By comparing the two panels of Figure 1, we can see that $SMS(\mu)$ consists of faces of $BCE(\mu)$. This is a general property in coordination games with strict Nash equilibria.

Claim 1. *(Face structure) Suppose there exist distinct $a^*, a^{**} \in A$ which are strict Nash equilibria of $(\mathscr{I}, \{A_i, u_i(\cdot,\omega)\})$ for all $\omega \in \Omega$. Then, $SMS(\mu)$ is a union of faces of $BCE(\mu)$.*

Let

$$\text{Public}(\mu)=\left\{p\in BCE(\mu):\sum_{\omega\in\Omega}p(a_i,a_{-i},\omega)\big(u_i(a_i,a_{-i};\omega)-u_i(a_i',a_{-i};\omega)\big)\geq 0\right.$$
$$\left.\forall i\in\mathscr{I},\,a_i,a_i'\in A_i,\;a_{-i}\in A_{-i}\right\}$$

denote the set of outcome distributions that can emerge from pure BNE under public information. Note that Public(μ) is just the convex combination of the two pure Nash equilibria in Battle of the Sexes, visually at the intersection of the triangles $\mathscr{C}((1,0),\mu)$ and $\mathscr{C}((0,1),\mu)$, which also coincides with the sets $\mathscr{C}((0,0),\mu)=\mathscr{C}((0,1),\mu)$. As stated in our next claim, this property generalizes to all games: not only does public information produce outcomes in $SMS(\mu)$, as it is a special kind of single-meeting schemes, but its outcomes lie in the intersection of all classes of single-meeting schemes.

Claim 2. *(Public intersection)* $\text{Public}(\mu)\subseteq\bigcap_{\tilde{a}\in A}\mathscr{C}(\tilde{a},\mu)\subseteq SMS(\mu)$.

3.2 Vertical Transmission: Delegated Hierarchies

Vertical transmission, whereby information flows from one individual to another, is also frequently observed in reality. Delegated hierarchies are a mode of vertical transmission in which information is delivered directly to a single agent, de facto the most informed one, and subsequently gets (partially) transmitted from agent to agent down the hierarchy in an incentive compatible way. Therefore, communication happens on one channel at a time along a sequence of one-to-one transmissions. Delegated hierarchies are cost-effective when the communication costs are convex in the size of the audience, that is, when the cost of communicating directly with n individuals (privately or publicly) is larger than n times the cost of communicating with one person.[8]

First, in order for agents to be able to vertically transfer information to one another, they must be ordered with respect to how informed they are. This suggests the notion of an information hierarchy. Second, in order for agents to be willing to transfer information to one another, an information hierarchy must satisfy certain incentive constraints.

Definition 2. *An information structure (S,P) is an information hierarchy if there exists a total order $\succeq$ on $\mathscr{I}$ such that $\succeq^s_{\text{Inf}}$ coincides with $\succeq$ for all s such that $P(s)>0$.*

[8] In this case, the sender would prefer to delegate transmission and compensate each agent for the cost of transmission, rather than transmit the messages herself.

Given this definition, we denote an information hierarchy by $((S,P),\succeq)$, using the associated total order $\succeq$. We use $\succ$ to refer to the corresponding strict order (i.e., the irreflexive version of $\succeq$). Notice that, by definition, $\succeq$ is independent of s, so that the hierarchy does not change with the information to be transmitted.

In an information hierarchy $((S,P),\succeq)$, although agent i such that $i \succeq j$ is able to deliver s_j to j, i may have an incentive to report something else. Therefore, truthful information transmission down the hierarchy needs to be incentive compatible: Not only must all agents have an incentive to play their equilibrium action, but also to pass down the relevant information, for other agents to do the same.

Given an information hierarchy $((S,P),\succeq)$, an agent $i \in \mathscr{I}$ and $s \in S$, let $s_{\prec i} = \left(s_j : j \prec i\right)$ and $s_{\succ i} = \left(s_j : j \succ i\right)$ and let the corresponding sets of profiles be denoted $S_{\prec i}$ and $S_{\succ i}$, respectively. Similarly, we use notation $a_{\prec i} \in A_{\prec i}$ and $a_{\succ i} \in A_{\succ i}$ to denote the action profiles of i's predecessors and i's successors.

Definition 3. *A distribution $p \in \Delta(A \times \Omega)$ can be implemented by a delegated hierarchy if there exist an information hierarchy $((S,P),\succeq)$ and an equilibrium $a^* \in \mathscr{E}(S,P)$ such that*

$$p(a,\omega) = \sum_{s \in S} \mu(\omega) P(\{s : a^*(s) = a\}|\omega) \quad \forall a \in A, \omega \in \Omega$$

and for all $i \in \mathscr{I}$, $s_i \in S_i$ and $s_{\prec i} \in S_{\prec i}$ such that $P(s_i, s_{\prec i}) > 0$,

$$\mathbb{E}\left[u_i\left(a_i^*(s_i), a_{\prec i}^*(s_{\prec i}), a_{\succ i}^*(s_{\succ i}); \omega\right) \middle| s_i\right] \geq \mathbb{E}\left[u_i\left(a_i', a_{\prec i}^*(s_{\prec i}'), a_{\succ i}^*(s_{\succ i}); \omega\right) \middle| s_i\right] \tag{5}$$

for all $a_i' \in A_i$ and $s_{\prec i}'$ such that $P(s_{\prec i}') > 0$.

Delegated hierarchies allow for direct communication with only the most informed agent, $i^* = \max_{\succeq} \mathscr{I}$, who gets informed according to (S,P). The definition describes a truthful equilibrium of a sequential cheap talk game in which, starting with i^*, each agent is both a receiver choosing an action strategically and an (strategic) informed sender to his immediate $\succeq$-predecessor. Each agent can deviate both from truthful information transmission and from his equilibrium action. Condition (5) requires that these deviations not be strictly profitable in the desired equilibrium a^*.[9] In this delegated process, each agent must be more informed in a strong sense than all of his $\succeq$-predecessors, as he is their only source of information. Hence, this process must build on an information hierarchy.

[9] We assume that any off-path message (such that $P(s_{\prec i}') = 0$) is interpreted by receiver $i-1$ as belief-equivalent to some particular on-path message.

Additionally, since $((S,P),\succeq)$ is common knowledge among the agents and chosen with commitment, no one would believe a message $s'_{\prec i}$ passed down by i that has zero probability under P.[10] Finally, while in the definition the same fixed transmission order holds for every message profile s, in Section 6.2 and Online Appendix B.3.2 we generalize this concept to random delegated hierarchies, where the transmission order can vary with s.

The definition also builds in robustness to communication as a by-product. In the standard information design framework, agents are assumed to receive all their information from (S,P), and strategic communication between the agents is assumed away. However, once agents are in possession of their messages, they could potentially want to share some of their information with each other. If agents have this ability, then by (5), they will be happy with the actions played by less informed agents and, therefore, will not have an incentive to induce them to change their actions by disclosing different information to them. At the same time, no agent can reveal anything to more informed agents that the latter do not already know and, therefore, cannot impact their actions. Overall, no communication would be an equilibrium in this extended game and a^* would still be played, thus being robust to inter-agent communication.

The next result characterizes the outcome distributions that can be implemented by delegated hierarchies.

Theorem 2. *A distribution $p \in \Delta(A \times \Omega)$ can be implemented by a delegated hierarchy, if and only if, $p(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$ and there exists a total order $\succeq$ on $\mathscr{I}$ such that for all $a_i \in A_i$, $a_{\prec i} \in A_{\prec i}$ and $i \in \mathscr{I}$,*

$$\sum_{\omega\in\Omega}\sum_{a_{\succ i}} p(a_i, a_{\prec i}, a_{\succ i}, \omega)\big(u_i(a_i, a_{\prec i}, a_{\succ i}; \omega) - u_i(a'_i, a'_{\prec i}, a_{\succ i}; \omega)\big) \geq 0 \tag{6}$$

for all $a'_i \in A_i$ and $a'_{\prec i}$ such that $p(a'_{\prec i}) > 0$.[11]

For any μ, denote by $DH(\mu)$ the set of outcome distributions that satisfy the above necessary and sufficient conditions. Like Theorem 1, this theorem also makes clear the con-

[10] For the special case of a line network where information is seeded at i^*, the above definition can be related to the framework of Galperti and Perego (2020), where information transmission between linked agents is full and mechanical. Here, in contrast, information transmission is partial and accounts for agents' incentives.

[11] Given this theorem, it is straightforward to characterize the outcome distributions that can be implemented by information hierarchies: A distribution $p \in \Delta(A \times \Omega)$ can be implemented by an information hierarchy, if and only if, $p(A \times \{\omega\}) = \mu(\omega)$ for all ω and there exists a total order $\succeq$ on $\mathscr{I}$ such that for all $a_i \in A_i$, $a_{\prec i} \in A_{\prec i}$ and $i \in \mathscr{I}$, $\sum_{\omega\in\Omega}\sum_{a_{\succ i}} p(a_i, a_{\prec i}, a_{\succ i}, \omega)\big(u_i(a_i, a_{\prec i}, a_{\succ i}; \omega) - u_i(a'_i, a_{\prec i}, a_{\succ i}; \omega)\big) \geq 0$ for all $a'_i \in A_i$.

nection between the informational constraints of delegated hierarchies and the resulting strategic constraints: condition (6) requires that if agent i knew the actions of his predecessors $a_{\prec i}$, he would not want to deviate from his recommended action or transmit any other recommendation $a'_{\prec i}$ that has positive probability under the outcome distribution.

Note that these incentive constraints are stronger than the BCE obedience constraints, which sum (6) over all $a_{\prec i} = a'_{\prec i}$.[12] They take the form of linear inequality systems for each total order $\succeq$ on $\mathscr{I}$, the solutions to which form a class of delegated hierarchies built on the same ordering of agents, denoted by $\mathscr{C}(\succeq, \mu)$. Thus, $DH(\mu) = \cup_{\succeq} \mathscr{C}(\succeq, \mu)$. [13]

The proof shows that a delegated hierarchy $p \in \mathscr{C}(\succeq, \mu)$ can be implemented by a canonical information structure, which sends every i an incentive compatible action recommendation augmented by the action recommendations of his predecessors. That is, for every a with $p(a) > 0$, the canonical information structure sends the message $(a_i, a_{\prec i})$ to every $i \in \mathscr{I}$ and, upon receiving this message, it is incentive compatible for agent i to play a_i and to transmit $a_{\prec i}$ to his immediate predecessor in the order $\succeq$. Therefore, the same action a_i can be played under different messages, (i.e., under different augmentations of the same a_i), which distinguishes delegated hierarchies from direct information structures, where distinct messages always lead to distinct actions.

In the Battle of the Sexes of Section 3.1, the two pure Nash equilibrium distributions p and p', defined as $p(1,1) = 1$ and $p'(0,0) = 1$, are the only outcome distributions in $DH(\mu)$. In particular, (strict) public randomizations between p and p' are not elements of $DH(\mu)$, because the higher ranked agent would have an incentive to deviate both in own action and information transmission so as to always induce his most preferred equilibrium. This is reflected in the next claim (to be compared with Claim 2):

Claim 3. $\text{ext}(\text{Public}(\mu)) \subseteq \bigcap_{\succeq} \mathscr{C}(\succeq, \mu) \subseteq DH(\mu)$.[14]

[12] They are also stronger than the obedience condition of Doval and Ely (2020), which requires that deviations from the recommended path of play are observed, due to the sequential nature of the game, and subsequent agents respond to the deviation with their own contingent recommendations. Hence, their obedience constraint is weaker than the standard BCE obedience. In comparison, our condition is stronger, because it assumes that deviating agents can choose the play of subsequent agents and also deviate in own action without being detected.

[13] If some underlying hierarchy of agents $\succeq$ needs to be respected, as in the military, for example, the respective class $\mathscr{C}(\succeq, \mu)$ gives the relevant outcome distributions.

[14] The extreme points of Public(μ) are simply the pure strategy Nash outcomes in the ex-ante normal form game in which it is common knowledge that all agents have belief μ

3.3 Horizontal vs. Vertical Transmission Outcomes

Comparison. In light of Theorems 1 and 2, note that $SMS(\mu)$ and $DH(\mu)$ are not, in general, related by inclusion. In some cases, as in the Battle of the Sexes of Section 3.1, we do have $DH(\mu) \subsetneq SMS(\mu)$. In Online Appendix B.1, we provide an example in which the set $DH(\mu)$ contains the optimal BCE distribution, while $SMS(\mu)$ does not. In that example, there is a uniformly-distributed binary state $\{a,b\}$. Agent 1's action set is $\{A,B\}$ and he has a different dominant action in each state: action A is dominant in state a and action B is dominant in state b. Agent 2 has the opposite dominant actions – B in state a and A in state b. However, agent 2 also has a third action, M which is dominant when he is left completely uninformed and acts under the prior. Consider the BCE that maximizes coordination between the actions and the state; that is, that maximizes the probability that (A,A) is played in state a and (B,B) is played in state b. It turns out this distribution cannot be implemented by a single-meeting scheme, for agent 2 would want to deviate in the meeting where both he and agent 1 are supposed to play the same action. However, p^* can be implemented by a delegated hierarchy under the order $1 \succ 2$.

Existence. Well-known families of Bayesian games have a pure BNE for all information structures. If the ex-post game $(\mathscr{I}, \{A_i,\ u_i(\cdot,\omega)\})$ is supermodular for all $\omega \in \Omega$ (Milgrom and Roberts (1990)),[15] then the ex-ante Bayesian game is also supermodular for all priors and information structures. The same is true for potential games (Monderer and Shapley (1996)): if the ex-post game admits a potential $\phi_\omega : A \to \mathbb{R}$ for all $\omega \in \Omega$, then the ex-ante Bayesian game is also a potential game for all priors and information structures (Heumen et al. (1996)). Therefore, in both of these families of games, existence of a pure equilibrium is guaranteed for all information structures, which includes all single-meeting schemes and information hierarchies. Notice, however, that the non-emptiness of $SMS(\mu)$, for example, is weaker than existence of a pure BNE for *every* single-meeting scheme.

3.4 Constrained Information Design: A Linear Programming Approach

This section illustrates the linear programming (LP) approach enabled by our characterizations and the type of economic questions it can be used to address. Since the characterizations deliver inequality constraints that are linear in the outcome probabilities, and

[15]Assuming the partial orders on the action sets are the same across ω.

the expected value of any objective function over actions and states is also linear in those, the problem of finding an optimal single-meeting scheme or delegated hierarchy can be solved using the tools of LP.

Consider a linear network environment with finite state space Ω, prior μ, and n workers, indexed by i, who each choose a positive effort level from a finite set A_i. Their payoffs are given by

$$u_i(a,\omega) = a_i\Big(\gamma_{ii}\cdot\omega + \sum_{j\neq i}\gamma_{ij}a_j\Big) - c_i a_i^2$$

where $\gamma_{ii} > 0$ and $\gamma_{ij} > 0$ are benefits from aligning own effort with the state and with agent j's effort, respectively, and $c_i > 0$ is a cost-of-effort parameter. The organization wants to maximize the expected value of a given objective $v : A\times\Omega \to \mathbb{R}$ by setting up a single-meeting scheme or a delegated hierarchy as its information transmission protocol.

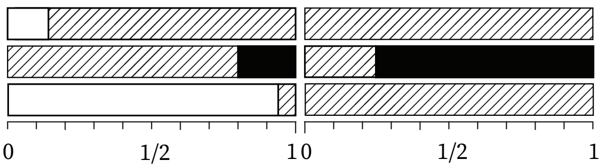
Single-Meeting Schemes: For each $\tilde{a} \in A$, Theorem 1 describes $\mathscr{C}(\tilde{a},\mu)$ as a system of linear inequalities. Thus, $V^*(\tilde{a},\mu) = \max_{p\in\mathscr{C}(\tilde{a},\mu)}\mathbb{E}_p[v]$ is a standard LP problem for each $\tilde{a} \in A$. Hence, the optimal expected value of the objective attainable through a single-meeting scheme, $\max\limits_{p\in SMS(\mu)}\mathbb{E}_p[v] = \max\limits_{\tilde{a}\in A} V^*(\tilde{a},\mu)$, is a maximum of $|A|$ LP problems.

Delegated Hierarchies: Optimization over delegated hierarchies adds a layer of complexity, because the linear inequalities in (6) are required to hold only at action profiles that are played with strictly positive probability. Nevertheless, in this environment, the strict monotonicity of $u_i(a_i,a_{-i};\omega)$ in a_{-i} when $a_i > 0$ for all $\omega\in\Omega$ and $i \in \mathscr{I}$, simplifies the problem computationally: For any given order $\succeq$ on $\mathscr{I}$, each i will recommend the largest $a_{\prec i}$ such that $p(a_{\prec i}) > 0$ to his immediate predecessor.[16] For each $\varepsilon > 0$, total order $\succeq$ on $\mathscr{I}$ and $\bar{a} \in$ and $\bar{a} \in A$, we solve the following LP problem:

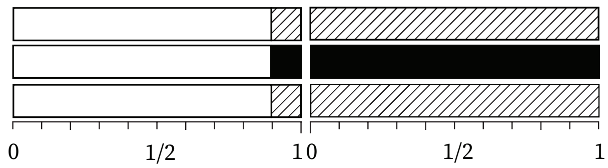
$$\begin{aligned} V^*_\varepsilon(\succeq,\bar{a},\mu) \quad = \quad & \max_{p\in\Delta(\Omega\times A)} \mathbb{E}_p[v] \\ & \text{s.t.} \sum_a p(a,\omega) = \mu(\omega) \text{ for all } \omega \\ & \quad (7) \text{ holds for all } i\in\mathscr{I} \text{ and } a_i, a_i' \in A_i \\ & \quad p(\bar{a}) \geq \varepsilon \text{ and } p(\hat{a}) = 0 \text{ for all } \hat{a} \not\leq \bar{a}, \end{aligned}$$

$$\text{where} \qquad \sum_{\omega\in\Omega}\sum_{a_{\succ i}} p(a_i,a_{\prec i},a_{\succ i},\omega)\big(u_i(a_i,a_{\prec i},a_{\succ i};\omega) - u_i(a_i',\bar{a}_{\prec i},a_{\succ i};\omega)\big) \geq 0. \tag{7}$$

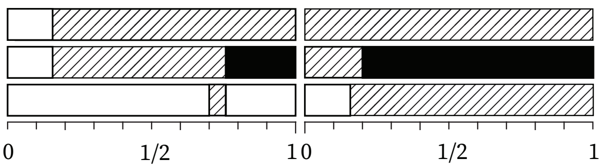
[16] This LP approach can be easily extended to the general linear network model, where γ_{ij}'s are arbitrary. When $\gamma_{ij} < 0$ and j is i's immediate predecessor, i will instead recommend the smallest a_j such that $p(a_j) > 0$.



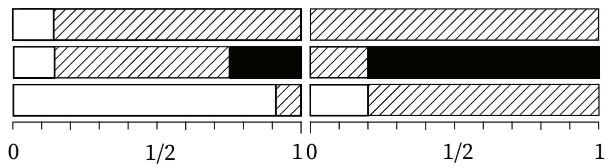
(a) Optimal Information Structure



(b) Optimal Public Information



(c) Optimal Single-Meeting Scheme
(optimal $\tilde{a} = (1/2, 1, 1/2)$)



(d) Optimal Delegated Hierarchy
(optimal order $3 \succ 2 \succ 1$)

Figure 2: Optimal Information Structures

Notes: Each panel from (a) to (d) consists of two sets of three bars. The bars on the left depict $p(\cdot|0)$ and the bars on the right $p(\cdot|1)$. The top bar represents manager 1's action, the middle bar manager 2's action, and the bottom bar manager 3's action, using the following color code: action 1 = solid black, action 1/2 = lined, and action 0 = white. See Tables 2 to 5 in the online appendix for the detailed numerical results.

The optimal expected value of the objective attainable through a delegated hierarchy is $\max_{p \in DH(\mu)} \mathbb{E}_p[v] = \lim_{\varepsilon \to 0} V_\varepsilon^*(\mu)$ where $V_\varepsilon^*(\mu) = \max_{\succeq, \bar{a}} V_\varepsilon^*(\succeq, \bar{a}, \mu)$, is a maximum of $(n!)|A|$ LP problems.

Results: For concreteness consider a binary state $\Omega = \{0,1\}$ distributed according to a uniform prior μ. There are $n = 3$ workers with $\gamma_{ij} = 1/10$ and $c_i = 3/4$ for all i and $j \neq i$, and $\gamma_{11} = 1/2$, $\gamma_{22} = 4/3$ and $\gamma_{33} = 1/4$. Each worker can choose low, intermediate or high effort, that is $A_i = \{0, 1/2, 1\}$. Assume that the organization aims to maximize total efforts $v(a) = \sum_i a_i$.[17]

Figure 2 displays the outputs of the LP approach.[18] Below we highlight qualitative features of the optimal solutions that are broadly generalizable.

(**i**) Fundamental tradeoff when $|A_i| \geq 3$: It is unconstrained-optimal (see Figure 2(a)) to induce worker 2 to exert both intermediate and high efforts in both states, even though

[17] The qualitative features of the LP solutions to this problem extend to objective functions v that take wages into account as an expenditure for the organization.

[18] The LP approach was implemented in R using the 'lp' function of package 'lpSolve.' The code is available on the authors' personal websites. The detailed computational results can be found in Online Appendix B.2. We thank Orestis Vravosinos for his programming expertise and excellent RA work.

he could be incentivized to exert high effort with probability 1 in $\omega = 1$. This is because intermediate effort cannot be incentive compatible if it is played only in $\omega = 0$, and thus, to maximize total efforts, it is preferable to give up some of the probability of recommending $a_2 = 1$ in $\omega = 1$ in order to incentivize $a_2 = 1/2$ in $\omega = 0$.

(ii) Optimal single-meeting scheme and meeting composition: It is optimal in (c) to choose $\tilde{a} = (1/2, 1, 1/2)$ and the optimal collection of meetings is $\{\emptyset, \{2\}, \{2,3\}, \{3\}, \{1,2,3\}\}$. Some meetings, such as $\{2\}$ and $\{2,3\}$, do incentivize strictly positive total effort: in those meetings worker 2 chooses effort 1/2. Relative to optimal public information, worker 3 exerts less effort in the optimal single-meeting scheme, yet that helps the overall objective by increasing the efforts of workers 1 and 2, because $a_3 = 0$ does not reveal the low state to worker 2.

(iii) Optimal delegated hierarchy and order: The optimal order of workers $3 \succ 2 \succ 1$ (Figure 2(d)) is different from simply ordering them according to their dependence on the state, which is the only asymmetry between them and would imply the order $3 \succ 1 \succ 2$. Under the optimal order, the highest effort profile recommended with positive probability is $\bar{a} = (1/2, 1, 1/2)$, and worker 2 also exerts effort 1/2 with positive probability at other profiles. Under $3 \succ 1 \succ 2$, we could not induce the same outcome because worker 1 would always transmit $a_2 = 1$ to worker 2 and thus violate truthful transmission at profile $(a_1, a_2, a_3) = (1/2, 1/2, 0)$. The optimal order $3 \succ 2 \succ 1$ ensures truthful transmission from worker 2 to worker 1: While worker 2 would like to tell worker 1 to play 1 when she herself follows the recommendation 1/2, that would always be detected as a lie, as the highest effort worker 1 is ever recommended to exert is 1/2.

(iv) Relative performance: Given the following (constrained and unconstrained) optimal expected values of the objective:

$$
\begin{array}{lclclcl}
\max\limits_{p \in BCE(\mu)} \mathbb{E}_p[v] & = & 1.47 & ; & \max\limits_{p \in DH(\mu)} \mathbb{E}_p[v] & = & 1.4025 \\
\max\limits_{p \in SMS(\mu)} \mathbb{E}_p[v] & = & 1.404 & ; & \max\limits_{p \in PUBLIC(\mu)} \mathbb{E}_p[v] & = & 1.11
\end{array}
$$

we conclude that single-meeting schemes and delegated hierarchies achieve at least 95% of the maximum possible value against 75% for public information. A simple look at Figure 2(b) shows that public information is significantly more restrictive when compared to panels (c) and (d).

4 Optimality in Binary-Action Supermodular Environments

In this section, we show that in binary-action environments with strategic complementarities single-meeting schemes and delegated hierarchies are optimal when the objective is to promote welfare or activity.

Assumption 1. (Binary Actions) For all $i \in \mathscr{I}$, $A_i = \{0, 1\}$.

Assumption 2. (Outside Option) For all $i \in \mathscr{I}$, $u_i = 0$ whenever $a_i = 0$.

Assumption 3. (Complementarities) For all $i \in \mathscr{I}$ and each $\omega \in \Omega$, $u_i(1, a_{-i}; \omega)$ is weakly increasing in a_{-i}.

These assumptions describe a binary-action framework, where the payoff from the low action is normalized to zero and agents' actions are strategic complements. This structure accommodates well-known economic applications such as the investment game of Carlsson and van Damme (1993), the team-production model of Moriya and Yamashita (2020), as well as "beauty contest" descriptions of social phenomena, such as location choice of city versus suburb or entry into versus exit from the labor force (see for example Brock and Durlauf (2001)). It also captures regime-change models, where a status quo is abandoned when a sufficiently large number of agents choose the high action. Examples of such models abound in the global games literature (see Morris and Shin (2003)).

We define partial orders on the set of outcome distributions $\Delta(A \times \Omega)$ to capture the activity and welfare enhancement featured in our results. Let $v : A \times \Omega \to \mathbb{R}$ be weakly increasing if $a' \geq a$ implies $v(a'; \omega) \geq v(a; \omega)$ for all $\omega \in \Omega$.

Definition 4. *Distribution p' dominates distribution p, denoted $p' \succeq_d p$, if $\mathbb{E}_{p'}[v] \geq \mathbb{E}_p[v]$ for all weakly increasing v and $\mathbb{E}_{p'}[u_i] \geq \mathbb{E}_p[u_i]$ for all $i \in \mathscr{I}$.*

A distribution dominates another one if the former first-order stochastically dominates the latter and also weakly improves every agent's expected utility.[19] Notice that the second requirement is not implied by the first, because u_i is not necessarily weakly increasing, since it is not weakly increasing in a_i.

[19] We follow the characterization of stochastic dominance provided by Shaked and Shanthikumar (2007). Let $\hat{A} \subseteq A$ be an upper set if $a \in \hat{A}$ and $a' \geq a$ imply $a' \in \hat{A}$.[20] Then, $\mathbb{E}_{p'}[v] \geq \mathbb{E}_p[v]$ for all weakly increasing v if and only if $p'\left(\hat{A} \times \{\omega\}\right) \geq p\left(\hat{A} \times \{\omega\}\right)$ for all upper sets $\hat{A}$ and $\omega \in \Omega$.

Proposition 1. *Under Assumptions 1-3, if $p \in BCE(\mu)$, then there exists $p' \geq_d p$ such that p' can be implemented by a single-meeting scheme.*

Given $\mathcal{V} \subseteq \{v : A \times \Omega \to \mathbb{R}\}$, cone($\mathcal{V}$) is the convex cone of $\mathcal{V}$. Let $\mathcal{V}^{\mathrm{M}} = \{v : A \times \Omega \to \mathbb{R} : a' \geq a \Rightarrow v(a';\omega) \geq v(a;\omega)\ \forall \omega \in \Omega\}$ be the family of action-wise weakly increasing functions.

Corollary 1. *If $\{u_i\}$ satisfy Assumptions 1-3 and $v \in cone\left(\mathcal{V}^M \cup \{u_i\}\right)$, then there exists $p^* \in \operatorname{argmax}_{p \in BCE(\mu)} \mathbb{E}_p[v]$ that can be implemented by a single-meeting scheme.*

Dominance of the outcome distribution in the proposition translates into optimality in the corollary for many objective functions. Those functions include all action-wise weakly increasing functions, weighted welfare functions, and their positive linear combinations. The latter capture, for instance, the objective to maximize the probability that some agents play action 1 while simultaneously maximizing the welfare of the others.

The argument behind the results is as follows. Let $p^* \in \operatorname{argmax}_{p \in BCE(\mu)} \mathbb{E}_p[v]$. If p^* is not implementable by a single-meeting scheme, then, by Theorem 1, there exist i and $\hat{a}_{-i}$ such that i has a strictly profitable deviation from action 0 to 1 when he knows that everyone else is playing $\hat{a}_{-i}$. This implies that another distribution, p', which would recommend action 1 instead of 0 to i against $\hat{a}_{-i}$, would also be a BCE and satisfy (4) for the given i and $\hat{a}_{-i}$. Since it stochastically dominates p^* and v is increasing, p' would also be optimal. By repeating this procedure, we can guarantee that some optimal distribution can be implemented by a single-meeting scheme.

These results hold regardless of the relationship between each agent's action and the state. For example, if two agents have opposite relationships with the state, a higher state may incentivize agent 1 to choose the high action, but deter agent 2. Hence, all else equal, a high action by agent 2 is interpreted as bad news about the state by agent 1, who may then choose the low action in response, even if actions are strategic complements. Single-meeting schemes can induce some agents to choose the high action without depressing the beliefs about the state of others, by not inviting them to the same meetings and thus creating different beliefs about the state that incentivize the same (high) action.

We now present an example of an optimal BCE distribution which is implementable by a single-meeting scheme, but not by a delegated hierarchy. This will help us motivate an additional assumption used in the optimality results for delegated hierarchies.

Example 1. (Optimal SMS and lack of delegation). Consider a team effort problem with three agents $\mathcal{I} = \{1,2,3\}$, $\Omega = \{0,1\}$, prior $\mu(\omega = 0) = 4/5$ and binary actions $A_i = \{0,1\}$.

Let $u_1(a;\omega) = a_1(2\omega - 1)$, $u_2(a;\omega) = a_2(2\omega - 2 + a_1)$ and

$$u_3(a;\omega) = a_3 \times \begin{cases} -2 & \text{if } a_1 = a_2 = \omega = 0 \\ a_1 + a_2 & \text{otherwise.} \end{cases}$$

Agent 1 wants to exert effort only if the state is high; agent 2 only if the state is high and agent 1 does so as well; and agent 3's utility from effort is given by 1 and 2's total efforts, except in the low state where exerting effort alone is detrimental.

Consider maximizing $\text{Prob}(a_3 = 1)$, that is, let $v(a) := a_3$. The following $p^* \in BCE(\mu)$ uniquely maximizes $\mathbb{E}_p[v]$ (rows correspond to agent 1's actions, and columns to action profiles of agents 2 and 3):

$p^*(\cdot;0)$	0,0	1,0	0,1	1,1
0	3/20	0	7/20	1/10
1	0	0	1/5	0

$p^*(\cdot;1)$	0,0	1,0	0,1	1,1
0	0	0	0	0
1	0	0	0	1/5

By Corollary 1, p^* is implementable by a single-meeting scheme. Nevertheless, the corresponding direct information structure $(A, \{p^*(\cdot|\omega)\})$, which implements p^*, is not a single-meeting scheme.[21]

We now illustrate the limitations of Assumptions 1-3 in the perspective of delegation. If p^* could be implemented by a delegated hierarchy, one agent would be the most informed and at the top of the hierarchy. If that agent were agent 1, then he would not be willing to play 1 at $a = (1,0,1)$, as that would reveal to him that $\omega = 0$ and he would prefer playing 0. If that were agent 2, he would not be willing to play 1 at (0,1,1). Finally, if that were agent 3, then he would always have an incentive to recommend action 1 to the others and switch to playing 1 himself. Thus, no total ordering that is constant across all messages can satisfy (5). To guarantee optimality of a delegated hierarchy, we need a stronger supermodularity assumption than Assumption 3, which we introduce next.

In what follows we assume that $\Omega \subseteq \mathbb{R}$ and define

$$\underline{\omega} = \min\left\{\omega \in \Omega : \quad \omega' \geq \omega \Rightarrow u_i(1,\ldots,1;\omega') \geq 0 \ \ \forall i \in \mathscr{I}\right\}.$$

[21]It can be seen from p^* itself that at action profiles $(1,0,1)$ and $(0,0,1)$, agents are pairwise unordered with respect to $\succeq^s_{Inf}$. Indeed, at $(1,0,1)$ under $(A, \{p^*(\cdot|\omega)\})$, agent 1 will not know that $a_2 = 0$; agent 2 will not know that $a_1 = 1$ or that $a_3 = 1$, and agent 3 will not know that $a_1 = 1$ or that $a_2 = 0$. Hence, each agent must receive his recommendation privately at both profiles.

Assumption 4. *(Supermodularity) For all $i \in \mathcal{I}$,* (a) *$u_i(1,a_{-i};\omega)$ is weakly increasing in a_{-i} for each $\omega \geq \underline{\omega}$ and* (b) *u_i is supermodular on $A \times \{\omega < \underline{\omega}\}$.*

Compared to Assumption 3, this assumption adds the requirement that for all u_i and states below $\underline{\omega}$, any pair of variables from $(a_1,\ldots,a_n,\omega)$ be weak complements. In particular, every agent is incentivized to choose action 1 in larger states and regards the other agents' actions as complementary to each other in his own payoff. Note also that the choice of the domain in this assumption is not arbitrary: there are important applications, such as global games of regime change, in which agents' payoffs fail to be supermodular on $A \times \Omega$ and yet are supermodular on $A \times \{\omega < \underline{\omega}\}$.

Before stating the next proposition, we introduce our second ordering on outcome distributions.

Definition 5. *Distribution p' supermodular-dominates distribution p, denoted $p' \succeq_{sd} p$, if $\mathbb{E}_{p'}[v] \geq \mathbb{E}_p[v]$ for all weakly increasing v that are supermodular on $A \times \{\omega < \underline{\omega}\}$ and $\mathbb{E}_{p'}[u_i] \geq \mathbb{E}_p[u_i]$ for all $i \in \mathcal{I}$.*[22]

Supermodular dominance does not guarantee a larger expected value for all weakly increasing functions, but only for those that *also* value coordination between actions and the state in the sense of supermodularity.

Proposition 2. *Under Assumptions 1-2 and 4, if $p \in BCE(\mu)$, then there exists $p' \succeq_{sd} p$ such that p' can be implemented by a delegated hierarchy.*

Let $\mathcal{V}^{\mathrm{SM}} = \{v : A \times \Omega \to \mathbb{R} : v \text{ is supermodular on } A \times \{\omega < \underline{\omega}\}\} \bigcap \mathcal{V}^{\mathrm{M}}$ be the family of weakly increasing supermodular functions.

Corollary 2. *If $\{u_i\}$ satisfy Assumptions 1-2 and 4 and $v \in cone\left(\mathcal{V}^{SM} \cup \{u_i\}\right)$, then there is $p^* \in \underset{p \in BCE(\mu)}{\operatorname{argmax}} \mathbb{E}_p[v]$ that can be implemented by a delegated hierarchy.*

In supermodular environments, delegated hierarchies are a welfare and activity enhancing mode of information transmission. It is also interesting to note that, although the environment may display no hierarchical ordering of agents in terms of their payoff dependencies on the state and each other's actions, the results imply that optimization always totally orders them in a way that enables delegated transmission. While the results

[22]As standard, given a lattice $(Y,\geq)$ and a sublattice $(X,\geq)$, $f : Y \to \mathbb{R}$ is supermodular on X if for all $x', x'' \in X$, $f(x' \vee x'') + f(x' \wedge x'') \geq f(x') + f(x'')$.

are agnostic about the exact optimal delegation order, Section 5.1 illustrates how the order may change with the objective v in a regime-change game. Moreover, note that the implementing information structure in the theorem and its corollary is also a single-meeting scheme.

The proof of Proposition 2 establishes that for any $p \in BCE(\mu)$ (i) there is a distribution $p^* \in BCE(\mu)$ that supermodular-dominates p and satisfies an inclusion property (Lemma 2) and (ii) given p^* from (i), there exists p' that supermodular-dominates p^* and can be implemented by a delegated hierarchy (Lemma 3). Part (i) generalizes Lemma 5 of Arieli and Babichenko (2019) to many states, state-dependent v such as weighted welfare, and to a weaker supermodularity condition which, for instance, captures regime-change games. Part (ii) provides an algorithm that elicits a total order $\succeq$ on $\mathscr{I}$ that ranks agents by informedness and for which the incentive requirements of delegation are met. The algorithm constructs this order by eliminating profitable coalitional deviations of subgroups of agents. Unless there is room for welfare and activity enhancement, there cannot be room for such coalitional deviations.

5 Applications

5.1 Regime Change

An organization consists of finitely many agents deciding whether to exert effort in a common project of uncertain quality $\omega \in \Omega$. Let $\kappa_i > 0$ be i's contribution to success, $c_i > 0$ his effort cost, and $b_i > c_i$ his benefit from a successful project. If i exerts effort, his payoff is:

$$u_i(a_{-i};\omega) = \begin{cases} b_i - c_i & \text{if} \quad \kappa_i + \sum_{j \neq i} \kappa_j a_j > 1 - \omega \\ -c_i & \text{otherwise.} \end{cases}$$

A manager wants to maximize the expected value of $v : A \times \Omega \to \mathbb{R}$ by optimally choosing an information structure, so that

$$\begin{aligned} &\max_{(S,P)} \mathbb{E}v(S,P) \\ &\text{where } \mathbb{E}v(S,P) = \max_{a^* \in \mathscr{E}(S,P)} \sum_{\omega \in \Omega} \sum_{s \in S} v(a^*(s);\omega) P(s|\omega)\mu(\omega) \end{aligned} \tag{8}$$

describes the design problem. In case there are multiple equilibria, (8) assumes favorable selection, as a^* is the equilibrium that yields the largest expected value of the objective. Recall that $\underline{\omega} = \min\{\omega \in \Omega : \sum_i \kappa_i > 1-\omega\}$ is the lowest state at which success is possible.

This payoff structure is typical of global games of regime change (see Sakovics and Steiner (2012)), and will illustrate nicely the variety of optimal single-meeting schemes and delegated hierarchies as a function of the objective.

Example 2. Suppose the manager wants to maximize the probability that the project succeeds:

$$v(a;\omega) = \begin{cases} 1 & \text{if } \sum_{i\in\mathscr{I}} \kappa_i a_i \geq 1-\omega \\ 0 & \text{otherwise.} \end{cases}$$

Note v is supermodular on $A \times \{\omega < \underline{\omega}\}$ (since $v = 0$ on $A \times \{\omega < \underline{\omega}\}$) and weakly increasing in a.

By Propositions 1-2 and Corollaries 1-2, there is a solution to (8) that is a single-meeting scheme or a delegated hierarchy.

It is trivial to see $\mathbb{E}v$ is maximized by having all agents play action 1 when $\omega \geq \underline{\omega}$ and to play action 0 otherwise. That is, $p^*(1,\ldots,1,\omega) = \mu(\omega)$ if $\omega \geq \underline{\omega}$ and $p^*(0,\ldots,0,\omega) = \mu(\omega)$ otherwise. From the point of view of horizontal transmission, p^* can be induced by inviting all agents to one meeting when the state is smaller than $\underline{\omega}$ (and announcing that "the state is less than $\underline{\omega}$"), and by not inviting anyone to a meeting otherwise. From the point of view of vertical transmission, p^* can be implemented by delegated hierarchy $((S^*,P^*),\succeq)$ for any total order $\succeq$ on $\mathscr{I}$, with $S^*_i = A_i \times A_{\preceq i}$ and $P^*(a_i,a_{\preceq i}|\omega) = \sum_{a_{\succeq i}} P(a_{\succeq i},a_i,a_{\preceq i})|\omega)$, for all $a_i \in A_i$, $a_{\preceq i} \in A_{\preceq i}$, $\omega \in \Omega$ and $i \in \mathscr{I}$. Agent i will either receive the message $(a_i,a_{\preceq i}) = (1,(1,\ldots,1))$, in which case he knows the project can succeed, will play 1, and want to forward the truthful recommendation to his successor, or he will receive the message $(a_i,a_{\preceq i}) = (0,(0,\ldots,0))$, in which case he knows that the project will fail, will play 0 and also be happy to forward the truthful recommendation to his successor.

Example 3. A (weighted) welfare-maximizing / utilitarian manager

$$v = \sum_{i\in\mathscr{I}} \lambda_i u_i \quad (\text{for any } \lambda_i \geq 0)$$

is an example of an objective that may not be weakly increasing, because an agent may have a strictly negative utility from playing the high action. Nevertheless, Propositions 1-2 and Corollaries 1-2 apply to this objective. The same single-meeting scheme and delegated hierarchy as in the previous example are optimal.

Example 4. Monotone objectives are natural in this context, such as maximizing the total probability of the high action or insisting on the participation of a key agent i^*

$$v(a,\omega) = \sum_{i \in \mathscr{I}} a_i \quad \text{or} \quad v(a,\omega) = a_{i^*} \sum_{i \in \mathscr{I}} a_i.$$

The reasoning below applies to all v that satisfy Assumption 4.

Fixing $\omega_i \in (0, \underline{\omega})$ and $\beta_i \in (0,1]$, define p^* implicitly as follows: $p^*(\{a_i = 1\}, \omega) = \mu(\omega)$ whenever $\omega > \omega_i$, $p^*(\{a_i = 1\}, \omega_i) = \beta_i \mu(\omega_i)$, and $p^*(\{a_i = 0\}, \omega) = \mu(\omega)$ whenever $\omega < \omega_i$. In words, this outcome distribution has every agent i play $a_i = 1$ with certainty in each state above ω_i, with probability β_i at state ω_i, and play 0 with certainty in each state below ω_i. Under p^*, hence,

$$\text{Prob}\left(\left\{\sum_{i \in \mathscr{I}} \kappa_i a_i \geq 1 - \omega\right\} \bigcap \{a_i = 1\}\right) = p^*\left(\{\omega \geq \underline{\omega}\} \cap \{a_i = 1\}\right) = \sum_{\omega : \omega \geq \underline{\omega}} \mu(\omega)$$

is the joint probability that the project succeeds and $a_i = 1$, and $p^*(\{a_i = 1\}) = \sum_{\omega:\omega > \omega_i} \mu(\omega) + \beta_i \mu(\omega_i)$ is the total probability that i exerts effort. Thus, the obedience constraint for $a_i = 1$ is given by

$$p^*\left(\{\omega \geq \underline{\omega}\} \cap \{a_i = 1\}\right) b_i - p^*(\{a_i = 1\}) c_i = \sum_{\omega:\omega \geq \underline{\omega}} \mu(\omega) b_i - \left(\sum_{\omega:\omega > \omega_i} \mu(\omega) + \beta_i \mu(\omega_i)\right) c_i \geq 0. \quad (9)$$

We choose ω_i and β_i such that (9) holds with equality for all i (recall $b_i > c_i$), that is,

$$p^*\left(\{\omega \geq \underline{\omega}\} | a_i = 1\right) = \frac{p^*\left(\{\omega \geq \underline{\omega}\} \cap \{a_i = 1\}\right)}{p^*(\{a_i = 1\})} = \frac{\sum_{\omega:\omega \geq \underline{\omega}} \mu(\omega)}{\sum_{\omega:\omega > \omega_i} \mu(\omega) + \mu(\omega_i) \beta_i} = \frac{c_i}{b_i}. \quad (10)$$

The obedience constraint for $a_i = 0$ is also satisfied, since $p^*\left(\{\omega \geq \underline{\omega}\} \cap \{a_i = 0\}\right) = 0$. We conclude that p^* is a BCE.

From the point of view of horizontal transmission, p^* can be implemented by a single-meeting scheme which, for each $a \in A$, organizes a meeting amongst $\{i \in \mathscr{I} : a_i = 0\}$ with probability $p^*(a) > 0$.

From the point of view of vertical transmission, p^* can be implemented by a delegated hierarchy $((S,P), \succeq)$ such that

1. $i \succeq j$ if and only if $\frac{c_i}{b_i} > \frac{c_j}{b_j}$
2. $S_i = A_i \times A_{\prec i}$ for all $i \in \mathscr{I}$

3. $P((a_i, a_{<i})|\omega) = \sum_{a_{>i}} p^*(a_{>i}, a_i, a_{<i}), \omega)/\mu(\omega)$ for all $i \in \mathscr{I}$, $a_i \in A_i$, $a_{<i} \in A_{<i}$ and $\omega \in \Omega$.

Agents who have a lower cost-to-benefit ratio from the project are given less information (both about the state and about the other agents' messages); play the high action more often; and occupy a lower position in the delegated hierarchy. Next, consider the incentives for obedient play and truthful transmission. First, consider $s_i^* = (a_i, a_{<i})$ such that $a_i = 1$: by construction, it must be that $a_i = 1$ is obedient and $a_{<i} = (1, \ldots, 1)$; agent i wants all agents below him in the hierarchy to also play 1, which is achieved by truthfully transmitting $a_{<i}$. Next consider $s_i^* = (a_i, a_{<i})$ such that $a_i = 0$: by construction, $a_i = 0$ is obedient and since i's payoff is zero independently of anyone else's action, truthful transmission of any $a_{<i}$ is optimal.

Example 5. There is a rich class of objectives that incorporate (additively and separably) the well-being of some agents and the effort choices of the other agents. For example, assume $n \geq 3$ and let

$$v(a;\omega) = u_1(a;\omega) + u_2(a;\omega) + \sum_{i \neq 1,2} a_i.$$

This describes benevolence toward 1 and 2 together with a desire to induce the rest of the population to adopt the high action. Under the optimal BCE p^*, 1 and 2 should play in the same way as in Example 2, while all other agents should play as in Example 4. That is, for agents $i = 1,2$, $p^*(\{a_i = 1\}, \omega) = \mu(\omega)$ if $\omega \geq \underline{\omega}$ and $p^*(\{a_i = 1\}, \omega) = 0$ otherwise, while for all agents $i \geq 3$, $p^*(\{a_i = 1\}, \omega) = \mu(\omega)$ if $\omega > \omega_i$, $p^*(\{a_i = 1\}, \omega) = \beta_i \mu(\omega)$ if $\omega = \omega_i$, and $p^*(\{a_i = 0\}, \omega) = \mu(\omega)$ if $\omega < \omega_i$.

From the point of view of horizontal transmission, p^* can be implemented by a single-meeting scheme which, for each $a \in A$, organizes a meeting amongst $\{i \in \mathscr{I} : a_i = 0\}$ with probability $p^*(a) > 0$.

From the point of view of vertical transmission, p^* can be implemented by a delegated hierarchy $((S, P), \succeq)$ such that

1. $1 \succeq 2 \succeq i$ (or $2 \succeq 1 \succeq i$) for all $i \in \mathscr{I} \setminus \{1,2\}$
2. $i \succeq j \iff c_i/b_i > c_j/b_j$ for all $\{i, j\} \subseteq \mathscr{I} \setminus \{1,2\}$
3. $S_i^* = A_i \times A_{<i}$ for all $i \in \mathscr{I}$
4. $P((a_i, a_{<i})|\omega) = \sum_{a_{>i}} p^*(a_{>i}, a_i, a_{<i}), \omega)/\mu(\omega)$ for all $i \in \mathscr{I}$, $a_i \in A_i$, $a_{<i} \in A_{<i}$ and $\omega \in \Omega$.

The incentives for obedient play and truthful transmission are met for all agents based on the same arguments as in Examples 2 and 4.

5.2 Delegation Ordering

In some environments, optimization requires agents to be treated equally in terms of informedness, yet delegation requires them to be treated differently, because only one specific total order implements the delegated hierarchy. The next example illustrates this possibility.

Example 6. Consider the following environment with $\mathscr{I} = \{1,2,3\}$, $\Omega = \{-1/2, 1/2\}$, $\mu(\omega = 1/2) = \mu$, $u_1(a;\omega) = a_1\omega$, $u_2(a;\omega) = a_2(\omega + a_3)$, and $u_3(a;\omega) = a_3(\omega + a_1)$.

Given a utilitarian objective $v = \lambda_1 u_1 + \lambda_2(u_2 + u_3)$ with $\lambda_1 > 2\lambda_2 > 0$, the optimal BCE distribution p^* is $p^*(1,1,1,1/2) = \mu$ and $p^*(0,0,0,-1/2) = 1-\mu$. The information structure that implements p^* gives all agents full information about the state, so that all agents are equally informed: $i \succeq_{\text{Inf}} j$ and $j \succeq_{\text{Inf}} i$ for all i and j.

There are six possible total orders $\succeq$ on $\mathscr{I}$ compatible with $\succeq_{\text{Inf}}$, one for each permutation of the agents. Nevertheless, only one of those, namely $1 \succeq 3 \succeq 2$, allows delegated transmission. Indeed, any total order with $3 \succeq 1$ fails the incentive for truthful transmission: upon receiving message 0, agent 3 would not want to forward 0 to agent 1 but instead prefer to play 1 and tell agent 1 to play 1, even though he knows that the state is -1/2. For an analogous reason, any total order with $2 \succeq 3$ fails to promote truthful transmission. agent 2 would want to misreport to agent 3 and simultaneously deviate in action given message 0.

6 Extensions

The concepts presented in this paper can be extended in various ways. Here we discuss some of them, but leave their detailed exploration to future research.

6.1 Multiple Meetings

Organizing more than one parallel meetings allows for a greater diversity of incentives, which is especially useful beyond binary actions. In Definition 1, there is *only one* way of keeping an agent imperfectly informed about others' information, which is to not invite him to any meeting. With many simultaneous meetings, there are many ways of keeping agents imperfectly informed, as participation to a meeting does not give perfect information about what is said in another or who has not been invited to any meeting.

In Online Appendix B.3.1, we generalize Definition 1 to m-meeting schemes, by allowing at most m simultaneous meetings to be organized at any message profile. We also characterize the outcome distributions that are implementable by m-meeting schemes in Proposition 3. Indeed, for each $m > 1$, the set of implementable outcomes of $(m-1)$-meeting schemes is included in that of m-meeting schemes, so a higher m provides greater flexibility in terms of implementable outcomes.

6.2 Random Delegated Hierarchies

In organizations, although it is simpler to have a fixed hierarchy, it is not unreasonable to assume that an executive would be able to choose the hierarchical order of transmission as function of the message she wants to transmit. In Online Appendix B.3.2, we generalize Definition 3 to random delegated hierarchies, by allowing the order of delegated transmission to change with the message profile, and characterize in Proposition 4 the distributions that can be implemented by such information structures. The extra flexibility allows some distributions, which cannot be implemented by a delegated hierarchy, to be implemented by a *random* delegated hierarchy.

A Appendix: Proofs

A.1 Characterizations

Proof of Theorem 1. (**Necessity**). Suppose $p \in BCE(\mu)$ can be implemented by a single-meeting scheme (S,P). Then, there is $a^* \in \mathscr{E}(S,P)$ such that $p(a,\omega) = \sum_{s\in S}\mu(\omega)P(\{s : a^*(s) = a\}|\omega)$ for all $a \in A$ and $\omega \in \Omega$. By definition of a single-meeting scheme, for all i and all s such that $P(s) > 0$ and $s_i \in S_i \setminus \{\tilde{s}_i\}$, $i \in M(s)$ and hence $\mu_i(s_{-i}|s_i) = 1$, implying

$$P(s|\omega)\mu(\omega) = \mu_i(\omega|s_i)P(s_i). \tag{11}$$

Let $\tilde{a}_i = a_i^*(\tilde{s}_i)$. Then, for all $a_i \in A_i \setminus \{\tilde{a}_i\}$ and all s such that $P(s) > 0$ and $a_i^*(s_i) = a_i$ (thus, $s_i \neq \tilde{s}_i$), we have

$$\sum_{\omega\in\Omega}\mu_i(\omega|s_i)\big(u_i(a_i,a^*_{-i}(s_{-i});\omega) - u_i(a'_i,a^*_{-i}(s_{-i});\omega)\big) \geq 0, \tag{12}$$

for all $a'_i \in A_i$ in virtue of equilibrium. By (11) and (12), we obtain

$$\sum_{\omega\in\Omega}\mu(\omega)P(s|\omega)\big(u_i(a_i,a^*_{-i}(s_{-i});\omega) - u_i(a'_i,a^*_{-i}(s_{-i});\omega)\big) \geq 0,$$

for all $s_i \in S_i$ such that $a_i^*(s_i) = a_i$ and all $s_{-i} \in S_{-i}$. Now, given arbitrary $a_{-i} \in A_{-i}$ and $a_i \in A_i \setminus \{\tilde{a}_i\}$, let

$$S(a_i, a_{-i}) = \left\{ s \in S : \quad a_i^*(s_i) = a_i \text{ and } a_{-i}^*(s_{-i}) = a_{-i} \right\}.$$

For all $a_{-i} \in A_{-i}$ and $a_i \in A_i \setminus \{\tilde{a}_i\}$,

$$\begin{aligned}
&\sum_{s \in S(a_i, a_{-i})} \sum_{\omega \in \Omega} \mu(\omega) P(s|\omega) \left(u_i(a_i, a_{-i}^*(s_{-i}); \omega) - u_i(a_i', a_{-i}^*(s_{-i}); \omega) \right) \\
= &\sum_{\omega \in \Omega} \mu(\omega) P\left(\{s : a_i^*(s_i) = a_i, a_{-i}^*(s_{-i}) = a_{-i}\} | \omega \right) \left(u_i(a_i, a_{-i}; \omega) - u_i(a_i', a_{-i}; \omega) \right) \\
= &\sum_{\omega \in \Omega} p(a_i, a_{-i}, \omega) \left(u_i(a_i, a_{-i}; \omega) - u_i(a_i', a_{-i}; \omega) \right) \geq 0.
\end{aligned}$$

(**Sufficiency**). Suppose now that $p \in BCE(\mu)$ and for all $i \in \mathscr{I}$, there is $\tilde{a}_i \in A_i$ such that for all $a_i \in A_i \setminus \{\tilde{a}_i\}$, $\sum_\omega p(a_i, a_{-i}, \omega) \left(u_i(a_i, a_{-i}; \omega) - u_i(a_i', a_{-i}; \omega) \right) \geq 0$ for all $a_i' \in A_i$ and $a_{-i} \in A_{-i}$. Then, define (S, P) as follows:

1. for each $i \in \mathscr{I}$, let $S_i = \{\tilde{a}_i\} \bigcup A$ and $s_i : A \to S_i$ be $s_i(a) = \begin{cases} a_i & \text{if} \quad a_i = \tilde{a}_i \\ a & \text{if} \quad a_i \neq \tilde{a}_i \end{cases}$, and
2. let $S = \prod_i S_i$ and $P\left((s_i(a))_i | \omega \right) = p(a|\omega)$ for all $a \in A$ and $\omega \in \Omega$.

Consider strategy profile a^* such that $a_i^*(s_i(a)) = a_i$ for all $a \in A$ and $i \in \mathscr{I}$. Since $p \in BCE(\mu)$, for all $i \in \mathscr{I}$ and $a_i \in A_i$,

$$\begin{aligned}
&\sum_{\omega \in \Omega} \sum_{a_{-i} \in A_{-i}} p(a_i, a_{-i}, \omega) \left(u_i(a_i, a_{-i}, \omega) - u_i(a_i', a_{-i}, \omega) \right) \\
= &\sum_{\omega \in \Omega} \sum_{a_{-i} \in A_{-i}} \mu(\omega) P\left((s_i(a))_i | \omega \right) \left(u_i(a_i^*(s_i(a)), a_{-i}^*(s_{-i}(a)), \omega) - u_i(a_i', a_{-i}^*(s_{-i}(a)), \omega) \right) \geq 0,
\end{aligned}$$

which shows that a^* is a BNE. Clearly, given that $a_i^*(s_i(a)) = a_i$ for all $a \in A$, $\sum_{s \in S} \mu(\omega) P(\{s : a^*(s) = a\} | \omega) = \mu(\omega) p(a|\omega) = p(a, \omega)$ for all $a \in A$ and $\omega \in \Omega$, so that p is implemented by (S, P). Finally, for each $i \in \mathscr{I}$, let $\tilde{s}_i = \tilde{a}_i$ and note that $\mu_i(s_{-i}(a)|s_i) = 1$ for all $s_i \neq \tilde{s}_i$. Therefore, (S, P) is a single-meeting scheme. ∎

Proof of Theorem 2. (**Necessity**). Suppose $p \in \Delta(A \times \Omega)$ can be implemented by a delegated hierarchy $((S, P), \succeq)$. Then, there exists an equilibrium $a^* \in \mathscr{E}(S, P)$ such that

$$p(a, \omega) = \sum_{s \in S} \mu(\omega) P(\{s : a^*(s) = a\} | \omega) \quad \forall a \in A, \omega \in \Omega \tag{13}$$

and, by condition (5), for all $i \in \mathscr{I}$, $s_i \in S_i$ and $s_{\preceq i} \in S_{\preceq i}$ such that $P(s_i, s_{\preceq i}) > 0$,

$$\sum_{\omega \in \Omega} \sum_{s_{\succeq i}} u_i \left(a_i^*(s_i), a_{\preceq i}^*(s_{\preceq i}), a_{\succeq i}^*(s_{\succeq i}); \omega\right) \mu_i(\omega, s_{\succeq i} | s_i) \geq$$
$$\sum_{\omega \in \Omega} \sum_{s_{\succeq i}} u_i \left(a_i', a_{\preceq i}^*(s_{\preceq i}'), a_{\succeq i}^*(s_{\succeq i}); \omega\right) \mu_i(\omega, s_{\succeq i} | s_i)$$

for all $a_i' \in A_i$ and $s_{\preceq i}' \in S_{\preceq i}$ such that $P(s_{\preceq i}') > 0$. Equivalently, for each $s_i \in S_i$ and $s_{\preceq i} \in S_{\preceq i}$ such that $P(s_i, s_{\preceq i}) > 0$, we have

$$\sum_{\omega \in \Omega} \sum_{a_{\succeq i}} \sum_{s_{\succeq i} : a_{\succeq i}^*(s_{\succeq i}) = a_{\succeq i}} u_i \left(a_i^*(s_i), a_{\preceq i}^*(s_{\preceq i}), a_{\succeq i}; \omega\right) \mu_i(\omega, s_{\succeq i} | s_i) \geq$$
$$\sum_{\omega \in \Omega} \sum_{a_{\succeq i}} \sum_{s_{\succeq i} : a_{\succeq i}^*(s_{\succeq i}) = a_{\succeq i}} u_i \left(a_i', a_{\preceq i}^*(s_{\preceq i}'), a_{\succeq i}; \omega\right) \mu_i(\omega, s_{\succeq i} | s_i) \quad (14)$$

for all $a_i' \in A_i$ and $s_{\preceq i}' \in S_{\preceq i}$ such that $P(s_{\preceq i}') > 0$. Since $((S,P), \succeq)$ is an information hierarchy, for each $s_i \in S_i$ there is at most one $s_{\preceq i} \in S_{\preceq i}$ such that $P(s_i, s_{\preceq i}) > 0$, which implies $\mu_i(s_{\preceq i} | s_i) = 1$. Hence, multiplying each side of (14) by $P(s_i)$ and summing over all $s_i : a_i^*(s_i) = a_i$ yields the following inequalities for each $s_{\preceq i} \in S_{\preceq i}$:

$$\sum_{\omega \in \Omega} \sum_{s_i : a_i^*(s_i) = a_i} \sum_{a_{\succeq i}} \sum_{s_{\succeq i} : a_{\succeq i}^*(s_{\succeq i}) = a_{\succeq i}} P(s_i, s_{\preceq i}, s_{\succeq i} | \omega) \mu(\omega) u_i \left(a_i^*(s_i), a_{\preceq i}^*(s_{\preceq i}), a_{\succeq i}; \omega\right) \geq$$
$$\sum_{\omega \in \Omega} \sum_{s_i : a_i^*(s_i) = a_i} \sum_{a_{\succeq i}} \sum_{s_{\succeq i} : a_{\succeq i}^*(s_{\succeq i}) = a_{\succeq i}} P(s_i, s_{\preceq i}, s_{\succeq i} | \omega) \mu(\omega) u_i \left(a_i', a_{\preceq i}^*(s_{\preceq i}'), a_{\succeq i}; \omega\right) \quad (15)$$

for all $a_i' \in A_i$ and $s_{\preceq i}' \in S_{\preceq i}$ such that $P(s_{\preceq i}') > 0$. Summing each side of (15) over all $s_{\prec i} : a_i^*(s_{\preceq i}) = a_{\preceq i}$, and using (13), we get for all $i \in \mathscr{I}$, $a_i \in A_i$ and $a_{\preceq i} \in A_{\preceq i}$ such that $p(a_i, a_{\preceq i}) > 0$,

$$\sum_{\omega \in \Omega} \sum_{a_{\succeq i}} p(a_i, a_{\preceq i}, a_{\succeq i}, \omega) \left(u_i(a_i, a_{\prec i}, a_{\succeq i}; \omega) - u_i(a_i', a_{\preceq i}', a_{\succeq i}; \omega)\right) \geq 0$$

for all $a_i' \in A_i$ and $a_{\preceq i}'$ such that $p(a_{\preceq i}') > 0$.

(**Sufficiency**). Suppose now that there exists a total order $\succeq$ on $\mathscr{I}$ such that (6) holds. Then, define (S,P) as follows:

1. for each i, define $S_i = A_i \bigcup A_{\prec i}$ and $s_i : A \to S_i$ such that $s_i(a) = (a_i, a_{\preceq i})$.

2. let $S = \prod_i S_i$ and P be such that $P\left((s_i(a))_i | \omega\right) = p(a|\omega)$ for all $a \in A$ and $\omega \in \Omega$.

We first prove that (S,P) is a delegated hierarchy. By construction, $s_i = (a_i, s_j)$ for all $i \in \mathscr{I}$, $j = \max\{j' : i \succeq j'\}$ and s such that $P(s) > 0$. Therefore, i's message contains the messages of all of $j \preceq i$. Hence, $i \succeq j$ implies $i \succeq_{\text{Inf}} j$ and so (S,P) is an information hierarchy. Furthermore, (6) implies that for all $a_i \in A_i$ and $a_{\preceq i} \in A_{\preceq i}$ such that $p(a_i, a_{\preceq i}) > 0$,

$$\sum_{\omega\in\Omega}\sum_{a_{\succeq i}} p(\omega, a_{\succeq i}|a_i, a_{\preceq i})\big(u_i(a_i, a_{\preceq i}, a_{\succeq i};\omega) - u_i(a'_i, a'_{\preceq i}, a_{\succeq i};\omega)\big) \geq 0 \tag{16}$$

for all $a'_i \in A_i$ and $a'_{\preceq i}$ such that $p(a'_{\preceq i}) > 0$. Given strategy profile a^* defined as $a^*_i(s_i(a)) = a_i$ for all $a \in A$ and $i \in \mathscr{I}$, (16) implies that for all $s_i \in S_i$ and $s_{\preceq i} \in S_{\preceq i}$ such that $P(s_i, s_{\preceq i}) > 0$

$$\sum_{\omega\in\Omega}\sum_{s_{\succeq i}} \mu_i(\omega, s_{\succeq i}|s_i)\big(u_i(a^*_i(s_i), a^*_{\preceq i}(s_{\preceq i}), a^*_{\succeq i}(s_{\succeq i});\omega) - u_i(a'_i, a'_{\preceq i}, a^*_{\succeq i}(s_{\succeq i});\omega)\big) \geq 0$$

for all $a'_i \in A_i$ and $a'_{\preceq i}$ such that $p(a'_{\preceq i}) > 0$. Hence, for all $i \in \mathscr{I}$, $s_i \in S_i$ and $s_{\preceq i} \in S_{\preceq i}$ such that $P(s_i, s_{\preceq i}) > 0$,

$$\begin{aligned}
& \mathbb{E}\left[u_i\left(a^*(s);\omega\right)\middle|s_i\right] - \mathbb{E}\left[u_i\left(a'_i, a^*_{\preceq i}(s'_{\preceq i}), a^*_{\succeq i}(s_{\succeq i});\omega\right)\middle|s_i\right] \\
= \quad & \sum_{\omega\in\Omega}\sum_{s_{\succeq i}} \mu_i(\omega, s_{\succeq i}|s_i)\big(u_i(a^*_i(s_i), a^*_{\preceq i}(s_{\preceq i}), a^*_{\succeq i}(s_{\succeq i});\omega) \\
& \qquad\qquad - u_i(a'_i, a^*_{\preceq i}(s'_{\preceq i}), a^*_{\succeq i}(s_{\succeq i});\omega)\big) \geq 0,
\end{aligned}$$

for all $a'_i \in A_i$ and $s'_{\preceq i} \in S_{\preceq i}$ such that $P(s'_{\preceq i}) > 0$. This establishes the delegation property and also that a^* is a BNE. Clearly, by definition of a^*,

$$\sum_{s\in S} \mu(\omega)P(\{s : a^*(s) = a\}|\omega) = \mu(\omega)p(a|\omega) = p(a,\omega)$$

for all $a \in A$ and $\omega \in \Omega$, so that p is implemented by (S,P). ∎

A.2 Claims

Proof of Claim 1. Since a^* and a^{**} are distinct, for any $\tilde{a} \in A$ there is $i \in \mathscr{I}$ such that either $\tilde{a}_i \neq a^*_i$ or $\tilde{a}_i \neq a^{**}_i$. Without loss, pick $\tilde{a}$ such that $\tilde{a}_i \neq a^*_i$. Then, invoking Theorem 1, pick $p^{\text{sms}}_1 \in \mathscr{C}(\tilde{a}, \mu)$. Since a^{**} is a strict equilibrium for all $\omega \in \Omega$, it must be that $u_i(a^*_i, a^{**}_{-i};\omega) - u_i(a^{**}_i, a^{**}_{-i};\omega) < 0$ for all $\omega \in \Omega$. Thus, the only way of satisfying (4) for $a_i = a^*_i$ and for all $a_{-i} \in A_{-i}$ is by setting $p^{\text{sms}}_1(a^*_i, a^{**}_{-i}, \omega) = 0$ for all $\omega \in \Omega$.

To show p^{sms}_1 lies on a face, we need to show that there exists $p \in \text{BCE}(\mu)$ such that, for all $\omega \in \Omega$, $p^{\text{sms}}_1(a,\omega) > 0$ implies $p(a,\omega) > 0$, and, for some $\omega' \in \Omega$, $p(a^*_i, a^{**}_{-i}, \omega') > 0$. We next construct $p^{\text{sms}}_2 \in \mathscr{C}(a^*_i, a^{**}_{-i}, \mu)$ such that $\text{supp}\sum_\omega p^{\text{sms}}_2(\cdot,\omega) = \{a^*, a^{**}, (a^*_i, a^{**}_{-i})\}$.

Since a^* and a^{**} are strict Nash equilibria, (4) holds for all agents and $p_2^{\text{sms}}(a^*,\cdot) \neq 0$ and $p_2^{\text{sms}}(a^{**},\cdot) \neq 0$. Next, by definition of $\mathscr{C}$, we must check that

$$\sum_{\omega\in\Omega} p_2^{\text{sms}}(a_i^*, a_{-i}^*, \omega)\big(u_i(a_i^*, a_{-i}^*, \omega) - u_i(a_i', a_{-i}^*, \omega)\big) \\ + \sum_{\omega\in\Omega} p_2^{\text{sms}}(a_i^*, a_{-i}^{**}, \omega)\big(u_i(a_i^*, a_{-i}^{**}, \omega) - u_i(a_i', a_{-i}^{**}, \omega)\big) \geq 0 \quad (17)$$

for all $a_i' \in A_i$. Since a^* is a strict Nash equilibrium for all $\omega \in \Omega$, the first expression in (17) is strictly positive for all $a_i' \in A_i$. Hence, there exists $\omega' \in \Omega$ and $\varepsilon > 0$ small enough such that (17) holds for all $a_i' \in A_i$ with $p_2^{\text{sms}}(a_i^*, a_{-i}^{**}, \omega') = \varepsilon$. By an analogous argument, all $j \neq i$ satisfy the corresponding version of (17) at a_j^{**} with $p_2^{\text{sms}}(a_i^*, a_{-i}^{**}, \omega') = \varepsilon$. Next, define $p = (1-\alpha)p_1^{\text{sms}} + \alpha p_2^{\text{sms}}$ for any $\alpha \in (0,1)$. Given that $p_1^{\text{sms}}, p_2^{\text{sms}} \in \text{BCE}(\mu)$ and $\text{BCE}(\mu)$ is convex, $p \in \text{BCE}(\mu)$. Moreover, by construction, $p(a,\omega) > 0$ whenever $p_1^{\text{sms}}(a,\omega) > 0$ and, at the same time, $p(a_i^*, a_{-i}^{**}, \omega') > 0$ since $p_2^{\text{sms}}(a_i^*, a_{-i}^{**}, \omega') > 0$. ∎

Proof of Claim 2. The proof is straightforward because, by definition, for any $p \in \text{Public}(\mu)$ and any $\tilde{a} \in A$, p must satisfy (4) for all $a_i \in A_i \setminus \{\tilde{a}_i\}$, $a_i' \in A_i$, $a_{-i} \in A_{-i}$ and $i \in \mathscr{I}$. Therefore, $p \in \mathscr{C}(\tilde{a}, \mu)$ for all $\tilde{a} \in A$, that is, $p \in \bigcap_{\tilde{a}\in A} \mathscr{C}(\tilde{a}, \mu)$. ∎

Proof of Claim 3. Given $\mu \in \Delta(\Omega)$, let

$$\text{NE}(\mu) = \Big\{ p \in \Delta(A \times \Omega) : \quad \exists\, a^* \in A \text{ s.t. } p(a^*, \cdot) = \mu \text{ and} \\ \sum_{\omega\in\Omega} \mu(\omega) u_i(a^*;\omega) \geq \sum_{\omega\in\Omega} \mu(\omega) u_i(a_i, a_{-i}^*;\omega) \quad \forall i \in \mathscr{I}, a_i \in A_i \Big\}$$

be the set of pure strategy Nash outcomes in the ex-ante normal form game in which it is common knowledge that all agents have belief μ, and note that $\text{NE}(\mu) = \text{ext}(\text{Public}(\mu))$. If $p \in \text{NE}(\mu)$, there exists $a^* \in A$ such that $p(a^*,\omega) = \mu(\omega)$ for all $\omega \in \Omega$ and

$$\sum_{\omega\in\Omega} \mu(\omega)(u_i(a^*;\omega) - u_i(a_i, a_{-i}^*;\omega)) \geq 0 \quad (18)$$

for all $i \in \mathscr{I}$ and $a_i \in A_i$. Take any total order $\succeq$ on $\mathscr{I}$. Since $p(a^*,\cdot) = \mu$, (18) is equivalent to

$$\sum_{\omega\in\Omega} p(a_i^*, a_{\preceq i}^*, a_{\succeq i}^*, \omega)\big(u_i(a_i^*, a_{\preceq i}^*, a_{\succeq i}^*;\omega) - u_i(a_i', a_{\preceq i}^*, a_{\succeq i}^*;\omega)\big) \geq 0 \quad (19)$$

for all $a_i' \in A_i$. Since $a_{\preceq i}^*$ is the only action profile $a_{\preceq i} \in A_{\preceq i}$ such that $p(a_{\preceq i}) > 0$, (19) holds for all $a_{\preceq i} \in A_{\prec i}$ such that $p(a_{\preceq i}) > 0$. Since $a_{\succeq i}^*$ is the only action profile $a_{\succeq i} \in A_{\succeq i}$ such that $p(a_{\geq i})>0$, summing up (19) across all $a_{\geq i}$ maintains the inequality. ■

A.3 Optimality Results

Proof of Proposition 1. Take any $p\in BCE(\mu)$. Let $\tilde{a}_i=1$ for all $i\in\mathscr{I}$. If for all $i\in\mathscr{I}$ and $a_{-i}\in A_{-i}$ it holds that

$$\sum_{\omega\in\Omega} p(0,a_{-i},\omega)(u_i(0,a_{-i},\omega)-u_i(1,a_{-i},\omega))>0,$$

then (4) holds and $p\in SMS(\mu)$ by Theorem 1, so we are done.

Suppose instead that there exist $i_0\in\mathscr{I}$ and nonempty $A'_{-i_0}\subseteq A_{-i_0}$ such that

$$\sum_{\omega\in\Omega} p(0,a_{-i_0},\omega)\big(u_{i_0}(0,a_{-i_0},\omega)-u_{i_0}(1,a_{-i_0},\omega)\big)\leq 0 \text{ iff } a_{-i_0}\in A'_{-i_0}. \tag{20}$$

Then, define $p^0\in\Delta(A\times\Omega)$ such that $p^0(0,a_{-i_0},\omega)=0$ and $p^0(1,a_{-i_0},\omega)=p(1,a_{-i_0},\omega)+p(0,a_{-i_0},\omega)$ for all $a_{-i_0}\in A'_{-i_0}$ and $\omega\in\Omega$, and $p^0(a,\omega)=p(a,\omega)$ for all a_{i_0}, $a_{-i_0}\in A_{i_0}\backslash A'_{-i_0}$ and $\omega\in\Omega$. Notice that this transformation does not impact the consistency with the prior, i.e. $p^0(A\times\{\omega\})=p(A\times\{\omega\})=\mu(\omega)$ for all $\omega\in\Omega$. Given $p\in BCE(\mu)$ and (20),

$$\begin{aligned}
\sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A_{-i_0}} p^0(1,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &= \\
\sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A_{-i_0}} p(1,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &+ \\
\sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A'_{-i_0}} p(0,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &\geq 0,
\end{aligned}\tag{21}$$

so that $a_{i_0}=1$ is obedient under p^0. Given (20),

$$\begin{aligned}
\sum_{\omega\in\Omega} p^0(0,a_{-i_0},\omega)\big(u_{i_0}(0,a_{-i_0};\omega)-u_{i_0}(1,a_{-i_0};\omega)\big) &= \\
\sum_{\omega\in\Omega} p(0,a_{-i_0},\omega)\big(u_{i_0}(0,a_{-i_0};\omega)-u_{i_0}(1,a_{-i_0};\omega)\big) &> 0,
\end{aligned}$$

for all $a_{-i_0}\in A_{-i_0}\backslash A'_{-i_0}$. Since $p^0(0,a_{-i_0},\omega)=0$ for all $a_{-i_0}\in A'_{-i_0}$, this implies that (4) is satisfied for agent i_0 under p^0. Together with (21), this implies that the BCE obedience constraints (3) are satisfied for i_0 under p^0. Given $p\in BCE(\mu)$ and by Assumption 3, the BCE obedience constraints for all $i\neq i_0$ and $a_i=1$ are immediately satisfied under p^0 (the recommended probability of play is the same, but the profile a_{-i} is now weakly higher due to the switch of agent i_0 from action 0 to action 1 under some profiles). If, in addition, for all $i \neq i_0$ and $a_{-i} \in A_i$ it holds that

$$\sum_{\omega \in \Omega} p^0(0, a_{-i}, \omega)(u_i(0, a_{-i}, \omega) - u_i(1, a_{-i}, \omega)) > 0,$$

then (4) is satisfied for all $i \neq i_0$ under p^0. Hence, $p^0 \in BCE(\mu)$ and by Theorem 1, $p^0 \in SMS(\mu)$. By construction, $p^0\left(\hat{A} \times \{\omega\}\right) \geq p\left(\hat{A} \times \{\omega\}\right)$ for all upper sets $\hat{A}$ and $\omega \in \Omega$. By (21) and Assumptions 2-3, $\mathbb{E}_{p^0}[u_i] \geq \mathbb{E}_p[u_i]$ for all $i \in \mathscr{I}$. Hence, $p^0 \succeq_d p$ and we are done.

Instead, suppose that there exists $i_1 \neq i_0$ and nonempty $A'_{-i_1} \subseteq A_{-i_1}$ such that

$$\sum_{\omega \in \Omega} p^0(0, a_{-i_1}, \omega)\left(u_{i_1}(0, a_{-i_1}; \omega) - u_{i_1}(1, a_{-i_1}; \omega)\right) \leq 0.$$

if and only if $a_{-i_1} \in A'_{-i_1}$. Then, repeat the above construction, obtaining $p^1 \in SMS(\mu)$ and $p^1 \succeq_d p^0$. Due to the finiteness of $\mathscr{I}$ and A, this process must terminate at some $p' \in SMS(\mu)$ and $p' \succeq_d p$.[23]

■

Proof of Proposition 2.

Lemma 1. *Suppose $p \in BCE(\mu)$ and $\hat{p} \succeq_{sd} p$ where $\hat{p}(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$. Then there exists $p^* \in BCE(\mu)$ such that $p^* \succeq_{sd} p$.*

Proof. Since $\mathbb{E}_{\hat{p}}[u_i] \geq \mathbb{E}_p[u_i] \geq 0$ for all $i \in \mathscr{I}$ and by Assumption 2, $a_i = 1$ is obedient for all $i \in \mathscr{I}$ under $\hat{p}$. If for all $i \in \mathscr{I}$

$$\sum_{\omega \in \Omega} \sum_{a_{-i} \in A_{-i}} \hat{p}(0, a_{-i}, \omega)[u_i(0, a_{-i}; \omega) - u_i(1, a_{-i}; \omega)] > 0, \tag{22}$$

then $a_i = 0$ is obedient for all $i \in \mathscr{I}$ under $\hat{p}$. Hence, $\hat{p} \in BCE(\mu)$ and by setting $p^* = \hat{p} \succeq_{sd} p$ we are done.

If, instead, (22) does not hold for some $i_0 \in \mathscr{I}$, define $p^0 \in \Delta(A \times \Omega)$ such that $p^0(0, a_{-i_0}, \omega) = 0$ and $p^0(1, a_{-i_0}, \omega) = \hat{p}(1, a_{-i_0}, \omega) + \hat{p}(0, a_{-i_0}, \omega)$ for all $a_{-i_0} \in A_{-i_0}$ and $\omega \in \Omega$. Note that $p^0(A \times \{\omega\}) = \hat{p}(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$. Since $a_{i_0} = 1$ is obedient under $\hat{p}$ and (22) is

[23]Notice that the same agent may appear at different steps in the process, that is more than once. For example, it could be that $i_2 = i_0$.

violated for i_0,

$$\begin{aligned}\sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A_{-i_0}} p^0(1,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &= \\ \sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A_{-i_0}} \hat{p}(1,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &+ \\ \sum_{\omega\in\Omega}\sum_{a_{-i_0}\in A'_{-i_0}} \hat{p}(0,a_{-i_0},\omega)\big(u_{i_0}(1,a_{-i_0};\omega)-u_{i_0}(0,a_{-i_0};\omega)\big) &\geq 0.\end{aligned}$$

Hence, $a_{i_0}=1$ is obedient under p^0 and since $\sum_{a_{-i_0}\in A_{-i_0}} p^0(1,a_{-i_0},\omega)=1$, this is the only action i_0 plays under p^0. Since the obedience constraints for all $i\neq i_0$ and $a_i=1$ were satisfied under $\hat{p}$, by Assumption 4 and by construction of p^0, they continue to hold under p^0 (the probabilities of play are the same, but all a_{-i} are now weakly higher since i_0 has switched to always playing 1 under p^0 from sometimes playing 0 under $\hat{p}$).

If, in addition

$$\sum_{\omega\in\Omega}\sum_{a_{-i}\in A_{-i}} p^0(0,a_{-i},\omega)[u_i(0,a_{-i};\omega)-u_i(1,a_{-i};\omega)]>0 \tag{23}$$

for all $i\neq i_0$, then $a_i=0$ is obedient for all $i\neq i_0$ under p^0. Hence, $p^0\in BCE(\mu)$. As argued above, $\sum\limits_{\omega\in\Omega}\sum\limits_{a_{-i}\in A_{-i}} u_i(1,a_{-i};\omega)[p^0(1,a_{-i},\omega)-\hat{p}(1,a_{-i},\omega)]\geq 0$ (due to i_0 always playing 1 under p^0). Combined with Assumption 2, this implies $\mathbb{E}_{p^0}[u_i]\geq\mathbb{E}_{\hat{p}}[u_i]$ for all $i\in\mathscr{I}$. By construction, $p^0\left(\hat{A}\times\{\omega\}\right)\geq\hat{p}\left(\hat{A}\times\{\omega\}\right)$ for all upper sets $\hat{A}$ and $\omega\in\Omega$. Hence, $p^0\succeq_d\hat{p}$, and thus $p^0\succeq_{sd}\hat{p}\succeq_{sd}p$. By setting $p^*=p^0$, we are done.

Instead, suppose that for some $i_1\neq i_0$ (23) does not hold. Then, repeat the above construction, obtaining $p^1\in BCE(\mu)$ and $p^1\succeq_d p^0$, and thus $p^1\succeq_{sd}p^0\succeq_{sd}p$. Due to the finiteness of $\mathscr{I}$ and A, this process must terminate at some $p^*\in BCE(\mu)$ and $p^*\succeq_{sd}p$.

■

For any $a\in A$, let $I(a)=\{i\in\mathscr{I}:a_i=1\}$.

Lemma 2. *For any $p\in BCE(\mu)$, there exists $p^*\in BCE(\mu)$ such that $p^*\succeq_{sd}p$ and*

1. *For all $\omega\in\Omega$ and $a',a''\in A$, if $p^*(a',\omega)>0$ and $p^*(a'',\omega)>0$, then $I(a')\subseteq I(a'')$ or $I(a'')\subseteq I(a')$.*

2. *For all $\omega',\omega''\in\Omega$ such that $\omega'<\omega''$ and $a',a''\in A$, if $p^*(a',\omega')>0$ and $p^*(a'',\omega'')>0$, then $I(a')\subseteq I(a'')$.*

Proof.

(**Part 1**). Suppose that for some ω', a' and a'', $p(a',\omega')>0$ and $p(a'',\omega')>0$, yet $I(a'')\subsetneq I(a')$ and $I(a')\subsetneq I(a'')$. Assume without loss that $p(a'',\omega')\geq p(a',\omega')$.

Case 1: If $\omega'\geq\underline{\omega}$, then define $\hat{p}$ as

$$\hat{p}(a,\omega)=\begin{cases}0 & \text{if} \quad a=a' \text{ and } \omega=\omega' \\ p(a,\omega)+p(a',\omega') & \text{if} \quad a=(1,\ldots,1) \text{ and } \omega=\omega' \\ p(a,\omega) & \text{otherwise.}\end{cases}$$

Notice that $\hat{p}(A\times\{\omega\})=p(A\times\{\omega\})=\mu(\omega)$ for all $\omega\in\Omega$.

If $i\in I(a')$, then $\mathbb{E}_{\hat{p}}[u_i]-\mathbb{E}_p[u_i]=p(a',\omega')(u_i(1,\ldots,1;\omega')-u_i(a';\omega'))\geq 0$, by Assumption 4(a). If $i\notin I(a')$, then $\mathbb{E}_{\hat{p}}[u_i]-\mathbb{E}_p[u_i]=p(a',\omega')u_i(1,\ldots,1;\omega')\geq 0$, where the equality follows from Assumption 2 and the inequality by $\omega'\geq\underline{\omega}$.

Case 2: If $\omega'<\underline{\omega}$, then define $\hat{p}$ as

$$\hat{p}(a,\omega)=\begin{cases}0 & \text{if} \quad a=a' \text{ and } \omega=\omega' \\ p(a,\omega)-p(a',\omega') & \text{if} \quad a=a'' \text{ and } \omega=\omega' \\ p(a,\omega)+p(a',\omega') & \text{if} \quad a=a'\vee a'' \text{ and } \omega=\omega' \\ p(a,\omega)+p(a',\omega') & \text{if} \quad a=a'\wedge a'' \text{ and } \omega=\omega' \\ p(a,\omega) & \text{otherwise.}\end{cases}$$

Notice that $\hat{p}(A\times\{\omega\})=p(A\times\{\omega\})=\mu(\omega)$ for all $\omega\in\Omega$.

If $i\in I(a')$ but $i\notin I(a'')$, then $\mathbb{E}_{\hat{p}}[u_i]-\mathbb{E}_p[u_i]=p(a',\omega')(u_i(a'\vee a'';\omega')-u_i(a';\omega'))\geq 0$, by Assumptions 2 and 4(b).

If $i\in I(a'')$ but $i\notin I(a')$, $\mathbb{E}_{\hat{p}}[u_i]-\mathbb{E}_p[u_i]=p(a',\omega')(u_i(a'\vee a'';\omega')-u_i(a'';\omega'))\geq 0$, which follows from Assumptions 2 and 4(b).

If $i\in I(a')\cap I(a'')$, then

$$\mathbb{E}_{\hat{p}}[u_i]-\mathbb{E}_p[u_i]=p(a',\omega')\Big(\big[u_i(a'\vee a'';\omega')-u_i(a';\omega')\big]-\big[u_i(a'';\omega')-u_i(a'\wedge a'';\omega')\big]\Big)\geq 0,$$

which follows from Assumption 4(b). If $i\notin I(a')\cup I(a'')$, then $\mathbb{E}_{\hat{p}}[u_i]=\mathbb{E}_p[u_i]$. Furthermore, for all weakly increasing v that are supemodular on $A\times\{\omega<\underline{\omega}\}$, in Case 1 we have

$$\mathbb{E}_{\hat{p}}[v]-\mathbb{E}_p[v]=p(a',\omega')\big[v(1,\ldots,1;\omega')-v(a';\omega')\big]\geq 0,$$

which follows from v being weakly-increasing, and in Case 2 we have

$$\mathbb{E}_{\hat{p}}[v]-\mathbb{E}_p[v]=p(a',\omega')\Big(\big[v(a'\vee a'';\omega')-v(a';\omega')\big]-\big[v(a'';\omega')-v(a'\wedge a'';\omega')\big]\Big)\geq 0,$$

which follows from supermodularity of v on $A \times \{\omega < \underline{\omega}\}$. Therefore, $\hat{p} \succeq_{sd} p$.

By repeating the above procedure for all a', $a'' \in A$ and $\omega' \in \Omega$ such that $p(a',\omega') > 0$, $p(a'',\omega') > 0$, $I(a') \subsetneq I(a'')$ and $I(a'') \subsetneq I(a')$, we eventually obtain $\hat{p} \succeq_{sd} p$ for which $\hat{p}(a',\omega) > 0$ and $\hat{p}(a'',\omega) > 0$ at any $\omega \in \Omega$ imply $I(a') \subseteq I(a'')$ or $I(a'') \subseteq I(a')$. By Lemma 1, there exists $p^* \in BCE(\mu)$ such that $p^* \succeq_{sd} \hat{p} \succeq_{sd} p$ and p^* preserves the inclusion properties of $\hat{p}$. Indeed, if (22) does not hold, then $a_i = 1$ for all $a \in A$ and $\omega \in \Omega$ such that $p^*(a,\omega) > 0$.

(**Part 2**). Now suppose that for some $\omega' < \omega''$, p is such that $p(a',\omega') > 0$, $p(a'',\omega'') > 0$ and yet $I(a') \subsetneq I(a'')$.

Case 1. If $\omega'' \geq \underline{\omega}$, then define $\hat{p}$ as

$$\hat{p}(a,\omega) = \begin{cases} 0 & \text{if} \quad a = a'' \text{ and } \omega = \omega'' \\ p(a,\omega) + p(a'',\omega'') & \text{if} \quad a = (1,\ldots,1) \text{ and } \omega = \omega'' \\ p(a,\omega) & \text{otherwise.} \end{cases}$$

Notice that $\hat{p}(A \times \{\omega\}) = p(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$.
If $i \in I(a'')$, then $\mathbb{E}_{\hat{p}}[u_i] - \mathbb{E}_p[u_i] = p(a'',\omega'')\big[u_i(1,\ldots,1;\omega'') - u_i(a'';\omega'')\big] \geq 0$, by Assumption 4(a).
If $i \notin I(a'')$, then, $\mathbb{E}_{\hat{p}}[u_i] - \mathbb{E}_p[u_i] = p(a'',\omega'')u_i(1,\ldots,1;\omega'') \geq 0$ where the equality follows from Assumption 2 and the inequality by $\omega'' \geq \underline{\omega}$.

Case 2. Suppose $\omega'' < \underline{\omega}$. Let $\underline{p} = \min\{p(a',\omega'), p(a'',\omega'')\}$ and define $\hat{p}$ as

$$\hat{p}(a,\omega) = \begin{cases} p(a,\omega) - \underline{p} & \text{if} \quad a = a'' \text{ and } \omega = \omega'' \\ p(a,\omega) - \underline{p} & \text{if} \quad a = a' \text{ and } \omega = \omega' \\ p(a,\omega) + \underline{p} & \text{if} \quad a = a' \vee a'' \text{ and } \omega = \omega'' \\ p(a,\omega) + \underline{p} & \text{if} \quad a = a' \wedge a'' \text{ and } \omega = \omega' \\ p(a,\omega) & \text{otherwise.} \end{cases}$$

Notice that $\hat{p}(A \times \{\omega\}) = p(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$.
If $i \in I(a')$ and $i \notin I(a'')$, then $\mathbb{E}_{\hat{p}}[u_i] - \mathbb{E}_p[u_i] = \underline{p}\big[u_i(a' \vee a'';\omega'') - u_i(a';\omega')\big] \geq 0$, which follows from Assumptions 2 and 4(b).
If $i \in I(a'')$ and $i \notin I(a')$, then $\mathbb{E}_{\hat{p}}[u_i] - \mathbb{E}_p[u_i] = \underline{p}\big[u_i(a' \vee a'';\omega'') - u_i(a'';\omega'')\big] \geq 0$, which follows from Assumptions 2 and 4(b).
If $i \in I(a') \cap I(a'')$, then

$$\mathbb{E}_{\hat{p}}[u_i] - \mathbb{E}_p[u_i] = \underline{p}\big[u_i(a' \vee a'';\omega'') - u_i(a'';\omega'') + u_i(a' \wedge a'';\omega') - u_i(a';\omega')\big] \geq 0,$$

which follows from Assumption 4(b). If $i \notin I(a') \cup I(a'')$, then $\mathbb{E}_{\hat{p}}[u_i] = \mathbb{E}_p[u_i]$.

Furthermore, for all weakly increasing v that are supemodular on $A \times \{\omega < \underline{\omega}\}$, in Case 1 we have

$$\mathbb{E}_{\hat{p}}[v] - \mathbb{E}_p[v] = p(a'', \omega'')\left[v(1, \ldots, 1; \omega'') - v(a''; \omega'')\right] \geq 0,$$

which follows from v being weakly increasing; in Case 2 we have,

$$\mathbb{E}_{\hat{p}}[v] - \mathbb{E}_p[v] = \underline{p}\Big(v(a' \vee a''; \omega'') - v(a''; \omega'') + v(a' \wedge a''; \omega') - v(a'; \omega')\Big) \geq 0,$$

which follows from v being supermodular on $A \times \{\omega < \underline{\omega}\}$. Therefore, $\hat{p} \succeq_{sd} p$.

By repeating the above procedure for all $\omega' < \omega''$ such that $p(a', \omega') > 0$, $p(a'', \omega') > 0$ and $I(a') \subsetneq I(a'')$, we eventually obtain $\hat{p} \succeq_{sd} p$ for which for any $\omega' < \omega''$, $\hat{p}(a', \omega') > 0$ and $\hat{p}(a'', \omega'') > 0$ imply $I(a') \subseteq I(a'')$. By Lemma 1, there exists $p^* \in BCE(\mu)$ such that $p^* \succeq_{sd} \hat{p} \succeq_{sd} p$ and p^* preserves the inclusion properties of $\hat{p}$. Indeed, if (22) does not hold, then $a_i = 1$ for all $a \in A$ and $\omega \in \Omega$ such that $p^*(a, \omega) > 0$. ■

Lemma 3. *Given p^* from Lemma 2, there exists $p' \succeq_{sd} p^*$ that can be implemented by a delegated hierarchy.*

Proof. The proof starts with an order and sequentially transforms it to ensure (6) is satisfied. Step 1 ensures that there are no profitable deviations in both action and transmission for all $i \in \mathscr{I}$ and $a_i = 1$, while Steps 2 and 3 ensure that there are no profitable deviations in both action and transmission for all $i \in \mathscr{I}$ and $a_i = 0$.

Step 1: Take p^* from Lemma 2. Let $\mathscr{N} = \bigcup_{a,\omega: p^*(a,\omega)>0} I(a)$ be the collection of sets of agents playing 1 in the action profiles occurring with positive probability under p^*. By Lemma 2, $(\mathscr{N}, \subseteq)$ is a totally ordered set, the elements of which can be denoted $\{I_k\}_{k=0}^K$ such that $I_0 = \emptyset$ and $I_{k'} \subseteq I_{k''}$ iff $k'' > k'$. Now, define $G_k = I_k \backslash I_{k-1}$ for all $k \geq 1$. Define $\succeq$ such that $j \succeq i$ iff $[i \in G_{k''},\ j \in G_{k'}$ and $k'' > k']$ or $[\{i, j\} \subseteq G_k,\ j > i]$. By the inclusion property of p^* and by construction of $\succeq$, for all $i \in \mathscr{I}$ and $\omega \in \Omega$, $p^*(1, a_{<i}, a_{>i}, \omega) > 0$ implies $a_{<i} = (1, \cdots, 1)$. Thus, for all $i \in \mathscr{I}$

$$\begin{aligned}\sum_{\omega \in \Omega} \sum_{a_{>i}} p^*(1, a_{<i}, a_{>i}, \omega)(u_i(1, a_{<i}, a_{>i}; \omega) - u_i(0, a_{<i}, a_{>i}; \omega)) \quad &= \\ \sum_{\omega \in \Omega} \sum_{a_{-i}} p^*(1, a_{-i}, \omega)(u_i(1, a_{-i}; \omega) - u_i(0, a_{-i}; \omega)) \quad &\geq \quad 0\end{aligned} \tag{24}$$

where the inequality follows from Lemma 2 (since $p^* \in BCE(\mu)$). By Assumption 2, $u_i(0, \cdot) = 0$ and thus, (24) is equivalent to

$$\sum_{\omega \in \Omega} \sum_{a_{>i}} p^*(1, a_{<i}, a_{>i}, \omega)\big(u_i(1, a_{<i}, a_{>i}; \omega) - u_i(0, a'_{<i}, a_{>i}; \omega)\big) \geq 0 \tag{25}$$

for all $i \in \mathscr{I}$ and $a_{<i}, a'_{<i} \in A_{>i}$. This implies that any deviation in action from $a_i = 1$ to $a'_i = 0$, irrespective of the deviation in transmission, is not strictly profitable for any $i \in \mathscr{I}$. Additionally, since $(1,\ldots,1) \geq a'_{<i}$ for all $a'_{<i} \in A_{<i}$, Assumption 4 guarantees that

$$\begin{aligned}\sum_{\omega\in\Omega}\sum_{a_{>i}} p'(1,a_{<i},a_{>i},\omega)\big(u_i(1,a_{<i},a_{>i};\omega)-u_i(1,a'_{<i},a_{>i};\omega)\big) &= \\ \sum_{\omega\in\Omega}\sum_{a_{>i}} p'(1,1,\ldots,1,a_{>i},\omega)\big(u_i(1,1,\ldots,1,a_{>i};\omega)-u_i(1,a'_{<i},a_{>i};\omega)\big) &\geq 0\end{aligned} \tag{26}$$

for all $a'_{<i} \in A_{<i}$ and $i \in \mathscr{I}$. This shows that there is no profitable deviation in transmission only, upon receiving action recommendation $a_i = 1$. By (25) and (26), we conclude that there is no profitable deviation in both action and transmission for all $i \in \mathscr{I}$ and $a_i = 1$, that is:

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p^*(1,a_{<i},a_{>i},\omega)\big(u_i(1,a_{<i},a_{>i};\omega)-u_i(a'_i,a'_{<i},a_{>i};\omega)\big) \geq 0 \tag{27}$$

for all $i \in \mathscr{I}$, $a'_i \in A_i$, and $a'_{<i} \in A_{<i}$. By Assumption 2,

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p^*(0,a_{<i},a_{>i},\omega)\big(u_i(0,a_{<i},a_{>i};\omega)-u_i(0,a'_{<i},a_{>i};\omega)\big) = 0$$

for all $i \in \mathscr{I}$ and $a_{<i}, a'_{<i} \in A_{<i}$, which holds for any p^*. If, in addition,

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p^*(0,a_{<i},a_{>i},\omega)\big(u_i(0,a_{<i},a_{>i};\omega)-u_i(1,a'_{<i},a_{>i};\omega)\big) > 0 \tag{28}$$

for all $i \in \mathscr{I}$ and $a_{<i}, a'_{<i} \in A_{<i}$, then (27)-(28) imply (6). Hence, by Theorem 2, p^* can be implemented by a delegated hierarchy and setting $p' = p^*$, we are done.

Suppose instead that there exist $i_0 \in \mathscr{I}$ and nonempty $A'_{<i_0} \subseteq A_{<i_0}$ such that

$$\sum_{\omega\in\Omega}\sum_{a_{>i_0}} p^*(0,a_{<i_0},a_{>i_0},\omega)\big(u_{i_0}(0,a_{<i_0},a_{>i_0};\omega)-u_{i_0}(1,a_{<i_0},a_{>i_0};\omega)\big) \leq 0 \iff a_{<i_0} \in A'_{<i_0}. \tag{29}$$

That is, for some agent i_0 it is weakly profitable to deviate in action from 0 to 1 while truthfully reporting $a_{<i_0} \in A'_{<i_0}$. Then, proceed to Step 2.

Alternatively, if for all $i \in \mathscr{I}$ and $a_{<i} \in A_{<i}$

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p^*(0,a_{<i},a_{>i},\omega)(u_i(0,a_{<i},a_{>i};\omega)-u_i(1,a_{<i},a_{>i};\omega)) > 0, \tag{30}$$

then proceed to Step 3.

Step 2: Define $p^0 \in \Delta(A \times \Omega)$ such that $p^0(0, a_{-i_0}, \omega) = 0$ and $p^0(1, a_{-i_0}, \omega) = p^*(1, a_{-i_0}, \omega) + p^*(0, a_{-i_0}, \omega)$ for all $a_{-i_0} \in A_{-i_0}$ such that $a_{<i_0} \in A'_{<i_0}$ and $\omega \in \Omega$, and $p^0(a, \omega) = p^*(a, \omega)$ otherwise. Notice that $p^0(A \times \{\omega\}) = p^*(A \times \{\omega\}) = \mu(\omega)$ for all $\omega \in \Omega$.

Hence, for all $a_{<i_0} \in A'_{<i_0}$

$$\begin{aligned}
&\sum_{\omega\in\Omega}\sum_{a_{>i_0}} p^0(1, a_{<i_0}, a_{>i_0}, \omega)\big(u_{i_0}(1, a_{<i_0}, a_{>i_0}; \omega) - u_{i_0}(0, a_{<i_0}, a_{>i_0}; \omega)\big) \quad = \\
&\sum_{\omega\in\Omega}\sum_{a_{>i_0}} p^*(1, a_{<i_0}, a_{>i_0}, \omega)\big(u_{i_0}(1, a_{<i_0}, a_{>i_0}; \omega) - u_{i_0}(0, a_{<i_0}, a_{>i_0}; \omega)\big) \quad + \\
&\sum_{\omega\in\Omega}\sum_{a_{>i_0}} p^*(0, a_{<i_0}, a_{>i_0}, \omega)\big(u_{i_0}(1, a_{<i_0}, a_{>i_0}; \omega) - u_{i_0}(0, a_{<i_0}, a_{>i_0}; \omega)\big) \quad \geq \quad 0
\end{aligned} \tag{31}$$

which follows from (24) and (29), and the same inequality holds for all $a_{<i_0} \in A_{<i_0} \setminus A'_{<i_0}$ by construction of p^0 and (24). Since $p^0(0, a_{<i_0}, a_{>i_0}, \omega) = 0$ for all $a_{<i_0} \in A'_{<i_0}$,

$$\sum_{\omega\in\Omega}\sum_{a_{>i_0}} p^0(0, a_{<i_0}, a_{>i_0}, \omega)\big(u_{i_0}(0, a_{<i_0}, a_{>i_0}; \omega) - u_{i_0}(1, a_{<i_0}, a_{>i_0}; \omega)\big) \geq 0$$

trivially holds for all $a_{<i_0} \in A_{<i_0}$.

For all agents $i > i_0$, (24) is unaffected under p^0. For all agents $i < i_0$, the lhs of (24) is weakly larger under p^0 by Assumption 4 (because agent i_0 switches from playing 0 to 1 in some profiles). Hence, for all $i \neq i_0$, $a_i = 1$, and $a_{<i} \in A_{<i}$

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p^0(1, a_{<i}, a_{>i}, \omega)(u_i(1, a_{<i}, a_{>i}; \omega) - u_i(0, a_{<i}, a_{>i}; \omega)) \geq 0.$$

By construction, $p^0\left(\hat{A} \times \{\omega\}\right) \geq p^*\left(\hat{A} \times \{\omega\}\right)$ for all upper sets $\hat{A}$ and $\omega \in \Omega$. By Assumption 2 and because the lhs of (24) under p^0 is unchanged for $i > i_0$, weakly larger for $i < i_0$, and weakly larger for i_0 (by (31)), we conclude that $\mathbb{E}_{p^0}[u_i] \geq \mathbb{E}_{p^*}[u_i]$ for all $i \in \mathscr{I}$. Hence, $p^0 \succeq_d p^*$, which implies $p^0 \succeq_{sd} p^*$. Next, apply Lemma 2 to p^0 to obtain $p^{**} \in BCE(\mu)$ such that $p^{**} \succeq_{sd} p^0 \succeq_{sd} p^*$, and apply Step 1 to p^{**}.

Step 3: We start with p' from Step 1 which satisfies (27) and (30) under the order $\succeq$. It remains to show that there is no profitable deviation in both action and transmission upon receiving an action recommendation $a_i = 0$ and $a_{<i}$.

Denote by $\hat{a}_{<i} = \max\left\{a_{<i} \in A_{<i} : p'(a_i = 0, a_{<i}) > 0\right\}$ and $\bar{a}_{<i} = \max\left\{a_{<i} \in A_{<i} : p'(a_{<i}) > 0\right\}$, which are both well defined by the inclusion property of p' and by construction of $\succeq$. This, together with Assumption 4, also implies that the most profitable deviation is from $a_i = 0$

and $a_{<i} = \hat{a}_{<i}$ to $a'_i = 1$ and $a'_{<i} = \bar{a}_{<i}$, since the belief about the state is highest at $a_{<i} = \hat{a}_{<i}$ and, therefore, so is the incentive to switch to action 1.

Case 1: $\hat{a}_{<i} = \bar{a}_{<i}$. By (30) we know that

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p'(0,\hat{a}_{<i},a_{>i},\omega)(u_i(0,\hat{a}_{<i},a_{>i};\omega) - u_i(1,\hat{a}_{<i},a_{>i};\omega)) > 0 \tag{32}$$

for all $i \in \mathscr{I}$ and thus the deviation is not weakly profitable for any $i \in \mathscr{I}$ with $\hat{a}_{<i} = \bar{a}_{<i}$.

Case 2: $\hat{a}_{<i} < \bar{a}_{<i}$. Hence, by construction of p' we have that $\bar{a}_{<i} = (1,\dots,1)$. If

$$\sum_{\omega\in\Omega}\sum_{a_{>i}} p'(0,\hat{a}_{<i},a_{>i},\omega)(u_i(0,\hat{a}_{<i},a_{>i},;\omega) - u_i(1,1,\dots,1,a_{>i};\omega)) > 0 \tag{33}$$

for all $i \in \mathscr{I}$ with $\hat{a}_{<i} < \bar{a}_{<i}$, then the deviation is not weakly profitable. Thus, p' can be implemented by a delegated hierarchy and we are done.

Suppose instead that for some $i \in \mathscr{I}$ with $\hat{a}_{<i} < \bar{a}_{<i}$ (33) is violated. Then, consider the set $J^i = I(\bar{a}_{<i}) \setminus I(\hat{a}_{<i})$ and let $\{\succeq^j\}_{j\in J^i}$ be the set of orders which coincide with $\succeq$ except that the positions of i and j are switched. Hence, by definition, $\hat{a}_{\prec^j j} = \hat{a}_{<i}$ and $a_{>i} = a_{\succ^j j}$.

If for all $j \in J^i$ and $\succeq^j$ it holds that

$$\sum_{\omega\in\Omega}\sum_{a_{\succ^j j}} p'(0,\hat{a}_{\prec^j j},a_{\succ^j j},\omega)\Big(u_j(0,\hat{a}_{\prec^j j},a_{\succ^j j};\omega) - u_i(1,1,\dots,1,a_{\succ^j j};\omega)\Big) \le 0 \tag{34}$$

then define $p'' \in \Delta(A\times\Omega)$ such that $p''(0,\hat{a}_{<i},a_{>i},\omega) = 0$ and $p''(1,\bar{a}_{<i},a_{>i},\omega) =$ $p'(1,\bar{a}_{<i},a_{>i},\omega) + p'(0,\hat{a}_{<i},a_{>i},\omega)$ for all $a_{>i} \in A_{>i}$ and $\omega\in\Omega$, and $p''(a,\omega) = p'(a,\omega)$ otherwise. Notice that $p''(A\times\{\omega\}) = p'(A\times\{\omega\}) = \mu(\omega)$ for all $\omega\in\Omega$.

Hence, for all $j \in J^i$ it holds that

$$\begin{aligned}
\sum_{\omega\in\Omega}\sum_{a_{>j}} p''(1,\bar{a}_{<j},a_{>j},\omega)\big(u_j(1,\bar{a}_{<j},a_{>j};\omega) - u_j(0,\bar{a}_{<j},a_{>j};\omega)\big) &= \\
\sum_{\omega\in\Omega}\sum_{a_{>j}} p'(1,\bar{a}_{<j},a_{>j},\omega)\big(u_j(1,\bar{a}_{<j},a_{>j};\omega) - u_j(0,\bar{a}_{<j},a_{>j};\omega)\big) &+ \\
\sum_{\omega\in\Omega}\sum_{a_{>j}} p'(0,\hat{a}_{<j},a_{>j},\omega)\big(u_j(1,\bar{a}_{<j},a_{>j};\omega) - u_j(0,\bar{a}_{<j},a_{>j};\omega)\big) &\ge 0
\end{aligned} \tag{35}$$

which follows from (24) and (34). Note also that for all $j \notin J^i$, (27) continues to hold under p'' and $\succeq$.

By construction, $p''\big(\hat{A}\times\{\omega\}\big) \ge p'\big(\hat{A}\times\{\omega\}\big)$ for all upper sets $\hat{A}$ and $\omega\in\Omega$. By (35), the fact that (27) continues to hold under p'' for all $j \notin J^i$, and Assumptions 2 and 4, we conclude that $\mathbb{E}_{p''}[u_i] \geq \mathbb{E}_{p'}[u_i]$ for all $i \in \mathscr{I}$. Hence, $p'' \succeq_d p'$, which implies $p'' \succeq_{sd} p'$. Next, apply Lemma 2 to p'' to obtain $p^{***} \in BCE(\mu)$ such that $p^{***} \succeq_{sd} p' \succeq_{sd} p^*$, and apply Step 1 to p^{***}.

Suppose instead that there exists $j \in J^i$ and $\succeq^j$ such that

$$\sum_{\omega \in \Omega} \sum_{a_{\succ^j j}} p'(0, \hat{a}_{\prec^j j}, a_{\succ^j j}, \omega) \Big(u_j(0, \hat{a}_{\prec^j j}, a_{\succ^j j}; \omega) - u_i(1, 1, \ldots, 1, a_{\succ^j j},; \omega) \Big) > 0. \tag{36}$$

Then choose the order $\succeq^j$ instead of $\succeq$. Notice that (27) and (30) continue to hold for all $i \in \mathscr{I}$ under p' and the order $\succeq^j$. Go back to the beginning of Step 3 with p' and $\succeq^j$.

Eventually, we obtain p' and an order $\succeq'$ such that

$$\sum_{\omega \in \Omega} \sum_{a_{\succ' i}} p'(0, \hat{a}_{\prec' i}, a_{\succ' i} \omega)(u_i(0, \hat{a}_{\prec' i}, a_{\succ' i}; \omega) - u_i(1, 1, \ldots, 1, a_{\succ' i}; \omega)) > 0.$$

for all $i \in \mathscr{I}$ with $\hat{a}_{\prec' i} < \bar{a}_{\prec' i}$. Thus, the deviation is not weakly profitable and we conclude that (28) holds. Hence, (6) is satisfied, and by Theorem 2, p' can be implemented by a delegated hierarchy.

■

Lemmas 2-3 prove Proposition 2. ■

References

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When does Coordination Require Centralization?" *American Economic Review* 98 (1):145–179.

Ambrus, Attila, Eduardo Azevedo, and Yuichiro Kamada. 2013. "Hierarchical Cheap Talk." *Theoretical Economics* 8 (1).

Arieli, Itai and Yakov Babichenko. 2019. "Private Bayesian Persuasion." *Journal of Economic Theory* :forthcoming.

Bergemann, Dirk and Stephen Morris. 2016. "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games." *Theoretical Economics* 11:487–522.

———. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature* 57 (1):44–95.

Brock, William and Steven Durlauf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68 (2):235–260.

Brooks, Benjamin, Alexander Frankel, and Emir Kamenica. 2020. "Information Hierarchies." Working paper, University of Chicago.

Candogan, Ozan. 2020. "Reduced Form Information Design: Persuading a Privately Informed Receiver." Working paper.

Candogan, Ozan and Kimon Drakopoulos. 2020. "Optimal Signaling of Content Accuracy: Engagement vs. Misinformation." *Operations Research* 68 (2).

Carlsson, H. and E. van Damme. 1993. "Global games and equilibrium selection." *Econometrica* 61 (5):989–1018.

Crémer, Jacques, Luis Garicano, and Andrea Prat. 2007. "Language and the Theory of the Firm." *The Quarterly Journal of Economics* 122 (1):373–407.

Dessein, Wouter. 2002. "Authority and Communication in Organizations." *Review of Economic Studies* 69:811–838.

Doval, Laura and Jeffrey Ely. 2020. "Sequential Information Design." *Econometrica* 86 (6):2575–2608.

Forges, Francoise. 1993. "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information." *Theory and Decision* 35:277.

———. 2006. "Correlated Equilibrium in Games with Incomplete Information Revisited." *Theory and Decision* 61:329.

Galeotti, Andrea, Christian Ghiglino, and Francesco Squintani. 2013. "Strategic information transmission networks." *Journal of Economic Theory* 148 (5):1751–1769.

Galperti, Simone and Jacopo Perego. 2020. "Information Systems." Working paper.

Hagenbach, Jeanne and Frederic Koessler. 2010. "Strategic Communication Networks." *Review of Economic Studies* 77 (3):1072–1099.

Halac, Marina, Elliot Lipnowski, and Daniel Rappoport. 2020. "Rank Uncertainty in Organizations." Working paper.

Heumen, Robert, Bezalel Peleg, Stef Tijs, and Peter Borm. 1996. "Axiomatic characterizations of solutions for Bayesian games." *Theory and Decision* 40:103–129.

Hori, Kazumi. 2006. "Organizing Information Transmission with Multiple Agents." Working paper.

Inostroza, Nicolas and Alessandro Pavan. 2020. "Persuasion in Global Games with Application to Stress Testing." Working Paper.

Kamenica, Emir. 2019. “Bayesian Persuasion and Information Design.” *Annual Review of Economics* 11:249–272.

Laclau, Marie, Ludovic Renou, and Xavier Venel. 2020. “Robust Communication on Networks.” Working Paper.

Li, Fei, Yangbo Song, and Mofei Zhao. 2019. “Global Manipulation by Local Obfuscation.” Working Paper.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2020. “On Information Design in Games.” *Journal of Political Economy* 128 (4):1370–1404.

Milgrom, Paul and John Roberts. 1990. “Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities.” *Econometrica* 58 (6):1255–1277.

Monderer, Dov and Lloyd Shapley. 1996. “Potential Games.” *Games and Economic Behavior* 14 (1):124–143.

Mookherjee, Dilip and Masatoshi Tsumagari. 2014. “Mechanism Design with Communication Constraints.” *Journal of Political Economy* 122 (5):1094–1129.

Moriya, Fumitoshi and Takuro Yamashita. 2020. “Asymmetric Information Allocation to Avoid Coordination Failure.” *Journal of Economics and Management Strategy* 29 (175):201–247.

Morris, S. and H. S. Shin. 2003. “Global Games: Theory and Applications.” In *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge University Press.

Morris, Stephen. 2002. “Faulty Communication: Some Variations on the Electronic Mail Game.” *Advances in Theoretical Economics* 1 (1):Article 5.

Morris, Stephen, Daisuke Oyama, and Satoru Takahashi. 2020. “Implementation via Information Design in Binary-Action Supermodular Games.” Working Paper.

Myerson, Roger B. 1991. *Game theory: Analysis of conflict*. Cambridge, Massachussets: Harvard University Press, 5th ed.

Oyama, Daisuke and Satoru Takahashi. 2020. “Generalized Belief Operator and Robustness in Binary Action Supermodular Games.” *Econometrica* 88 (2):693–726.

Radner, Roy. 1993. “The Organization of Decentralized Information Processing.” *Econometrica* 61:1109–1146.

Rantakari, Heikki. 2008. “Governing Adaptation.” *Review of Economic Studies* 75:1257–1285.

Rivera, Thomas J. 2018. “Incentives and the Structure of Communication.” *Journal of Economic Theory* 8 (1).

Rogelberg, Steven G., Cliff Scott, and John Kello. 2007. "The Science and Fiction of Meetings." *MIT Sloan Management Review*, Opinions & Analysis.

Sakovics, Jozsef and Jakub Steiner. 2012. "Who Matters in Coordination Problems?" *American Economic Review* 102 (7):3439–61.

Shaked, Moshe and J. George Shanthikumar. 2007. *Stochastic Orders*. New York, NY: Springer.

Taneva, Ina. 2019. "Information Design." *American Economic Journal: Microeconomics* :Forthcoming.

Van Zandt, Timothy. 1999. "Real-Time Decentralized Information Processing as a Model of Organizations with Boundedly Rational Agents." *Review of Economic Studies* 66 (3):633–658.

———. 2007. "Communication Complexity and Mechanism Design." *Journal of the European Economic Association* 5 (2-3):543–553.

B ONLINE APPENDIX

B.1 Example of an optimal BCE that is a DH but not an SMS

Consider the following state-dependent payoffs, where agent 1 is the row player and agent 2 is the column player:

$\omega = a$	A	M	B
A	$3,0$	$2,2$	$3,5$
B	$-1,-5$	$0,0$	$-1,1$

$\omega = b$	A	M	B
A	$-1,1$	$0,0$	$-1,-5$
B	$3,5$	$2,2$	$3,0$

If the state is common knowledge, then it is strictly dominant for agent 1 to match the state – play A in state a and B in state B, while it is dominant for agent 2 to mismatch the state and play B in state a and A in state b. Importantly, if agent 2 acts only under the prior, then M becomes strictly dominant for him. Hence, leaving agent 2 completely uninformed would result in 0 expected payoff for the designer.

Consider maximizing the probability of coordination between both agents' actions and the state, that is $p(A,A,a)$ and $p(B,B,b)$. The optimal BCE p^* is given by:[24]

$\omega = a$	A	M	B
A	0.26	0	0.14
B	0	0	0.1

$\omega = b$	A	M	B
A	0.1	0	0
B	0.14	0	0.26

Figure 3: Optimal BCE outcome distribution p^*

This distribution cannot be implemented by a single-meeting scheme: If $\tilde{a}_2 = A$, then condition (4) fails to hold for $a_2 = B$ against $a_1 = B$, because this meeting is associated with a posterior of 0.72 on $\omega = b$ and, hence, agent 2 would want to deviate to his dominant action A in that state. Similarly, if $\tilde{a}_2 = B$, then condition (4) fails to hold for $a_2 = A$ against $a_1 = A$, since this meeting is associated with a posterior of 0.72 on $\omega = a$ and, hence, agent 2 would want to deviate to his dominant action B in that state.

[24]If we removed action M for agent 2 but kept everything else the same, then the optimal BCE would be a public information structure and thus implementable by a single-meeting scheme.

However, p^* can be implemented by a delegated hierarchy under the order $1 \succ 2$. When Agent 1 is recommended to play action A (B), his posterior belief is 0.8 on the state in which that action is dominant, i.e. $\omega = a$ ($\omega = b$), so following his own action recommendation is incentive compatible. Additionally, agent 1's payoff is the same across all messages he is asked to pass down to agent 2, so transmission incentives are satisfied. For agent 2, the incentives are the same as the BCE obedience constraints, which are also satisfied.

B.2 Computational Results from Section 3.4

a_1	a_2	a_3	ω	$p(a,\omega)$	a_1	a_2	a_3	ω	$p(a,\omega)$
0.0	0.0	0.0	0	0.000	0.0	0.0	0.0	1	0.000
0.0	0.0	0.5	0	0.000	0.0	0.0	0.5	1	0.000
0.0	0.0	1.0	0	0.000	0.0	0.0	1.0	1	0.000
0.0	0.5	0.0	0	0.075	0.0	0.5	0.0	1	0.000
0.0	0.5	0.5	0	0.000	0.0	0.5	0.5	1	0.000
0.0	0.5	1.0	0	0.000	0.0	0.5	1.0	1	0.000
0.0	1.0	0.0	0	0.000	0.0	1.0	0.0	1	0.000
0.0	1.0	0.5	0	0.000	0.0	1.0	0.5	1	0.000
0.0	1.0	1.0	0	0.000	0.0	1.0	1.0	1	0.000
0.5	0.0	0.0	0	0.000	0.5	0.0	0.0	1	0.000
0.5	0.0	0.5	0	0.000	0.5	0.0	0.5	1	0.000
0.5	0.0	1.0	0	0.000	0.5	0.0	1.0	1	0.000
0.5	0.5	0.0	0	0.316	0.5	0.5	0.0	1	0.000
0.5	0.5	0.5	0	0.000	0.5	0.5	0.5	1	0.124
0.5	0.5	1.0	0	0.000	0.5	0.5	1.0	1	0.000
0.5	1.0	0.0	0	0.081	0.5	1.0	0.0	1	0.000
0.5	1.0	0.5	0	0.028	0.5	1.0	0.5	1	0.376
0.5	1.0	1.0	0	0.000	0.5	1.0	1.0	1	0.000
1.0	0.0	0.0	0	0.000	1.0	0.0	0.0	1	0.000
1.0	0.0	0.5	0	0.000	1.0	0.0	0.5	1	0.000
1.0	0.0	1.0	0	0.000	1.0	0.0	1.0	1	0.000
1.0	0.5	0.0	0	0.000	1.0	0.5	0.0	1	0.000
1.0	0.5	0.5	0	0.000	1.0	0.5	0.5	1	0.000
1.0	0.5	1.0	0	0.000	1.0	0.5	1.0	1	0.000
1.0	1.0	0.0	0	0.000	1.0	1.0	0.0	1	0.000
1.0	1.0	0.5	0	0.000	1.0	1.0	0.5	1	0.000
1.0	1.0	1.0	0	0.000	1.0	1.0	1.0	1	0.000

Table 2: Optimal BCE

a_1	a_2	a_3	ω	$p(a,\omega)$	a_1	a_2	a_3	ω	$p(a,\omega)$
0.0	0.0	0.0	0	0.081	0.0	0.0	0.0	1	0.000
0.0	0.0	0.5	0	0.000	0.0	0.0	0.5	1	0.000
0.0	0.0	1.0	0	0.000	0.0	0.0	1.0	1	0.000
0.0	0.5	0.0	0	0.000	0.0	0.5	0.0	1	0.000
0.0	0.5	0.5	0	0.000	0.0	0.5	0.5	1	0.000
0.0	0.5	1.0	0	0.000	0.0	0.5	1.0	1	0.000
0.0	1.0	0.0	0	0.000	0.0	1.0	0.0	1	0.000
0.0	1.0	0.5	0	0.000	0.0	1.0	0.5	1	0.000
0.0	1.0	1.0	0	0.000	0.0	1.0	1.0	1	0.000
0.5	0.0	0.0	0	0.000	0.5	0.0	0.0	1	0.000
0.5	0.0	0.5	0	0.000	0.5	0.0	0.5	1	0.000
0.5	0.0	1.0	0	0.000	0.5	0.0	1.0	1	0.000
0.5	0.5	0.0	0	0.267	0.5	0.5	0.0	1	0.086
0.5	0.5	0.5	0	0.036	0.5	0.5	0.5	1	0.009
0.5	0.5	1.0	0	0.000	0.5	0.5	1.0	1	0.000
0.5	1.0	0.0	0	0.116	0.5	1.0	0.0	1	0.000
0.5	1.0	0.5	0	0.000	0.5	1.0	0.5	1	0.405
0.5	1.0	1.0	0	0.000	0.5	1.0	1.0	1	0.000
1.0	0.0	0.0	0	0.000	1.0	0.0	0.0	1	0.000
1.0	0.0	0.5	0	0.000	1.0	0.0	0.5	1	0.000
1.0	0.0	1.0	0	0.000	1.0	0.0	1.0	1	0.000
1.0	0.5	0.0	0	0.000	1.0	0.5	0.0	1	0.000
1.0	0.5	0.5	0	0.000	1.0	0.5	0.5	1	0.000
1.0	0.5	1.0	0	0.000	1.0	0.5	1.0	1	0.000
1.0	1.0	0.0	0	0.000	1.0	1.0	0.0	1	0.000
1.0	1.0	0.5	0	0.000	1.0	1.0	0.5	1	0.000
1.0	1.0	1.0	0	0.000	1.0	1.0	1.0	1	0.000

Table 3: Optimal Single-Meeting Scheme

a_1	a_2	a_3	ω	$p(a,\omega)$	a_1	a_2	a_3	ω	$p(a,\omega)$
0.0	0.0	0.0	0	0.444	0.0	0.0	0.0	1	0.000
0.0	0.0	0.5	0	0.000	0.0	0.0	0.5	1	0.000
0.0	0.0	1.0	0	0.000	0.0	0.0	1.0	1	0.000
0.0	0.5	0.0	0	0.000	0.0	0.5	0.0	1	0.000
0.0	0.5	0.5	0	0.000	0.0	0.5	0.5	1	0.000
0.0	0.5	1.0	0	0.000	0.0	0.5	1.0	1	0.000
0.0	1.0	0.0	0	0.000	0.0	1.0	0.0	1	0.000
0.0	1.0	0.5	0	0.000	0.0	1.0	0.5	1	0.000
0.0	1.0	1.0	0	0.000	0.0	1.0	1.0	1	0.000
0.5	0.0	0.0	0	0.000	0.5	0.0	0.0	1	0.000
0.5	0.0	0.5	0	0.000	0.5	0.0	0.5	1	0.000
0.5	0.0	1.0	0	0.000	0.5	0.0	1.0	1	0.000
0.5	0.5	0.0	0	0.000	0.5	0.5	0.0	1	0.000
0.5	0.5	0.5	0	0.000	0.5	0.5	0.5	1	0.000
0.5	0.5	1.0	0	0.000	0.5	0.5	1.0	1	0.000
0.5	1.0	0.0	0	0.000	0.5	1.0	0.0	1	0.000
0.5	1.0	0.5	0	0.056	0.5	1.0	0.5	1	0.500
0.5	1.0	1.0	0	0.000	0.5	1.0	1.0	1	0.000
1.0	0.0	0.0	0	0.000	1.0	0.0	0.0	1	0.000
1.0	0.0	0.5	0	0.000	1.0	0.0	0.5	1	0.000
1.0	0.0	1.0	0	0.000	1.0	0.0	1.0	1	0.000
1.0	0.5	0.0	0	0.000	1.0	0.5	0.0	1	0.000
1.0	0.5	0.5	0	0.000	1.0	0.5	0.5	1	0.000
1.0	0.5	1.0	0	0.000	1.0	0.5	1.0	1	0.000
1.0	1.0	0.0	0	0.000	1.0	1.0	0.0	1	0.000
1.0	1.0	0.5	0	0.000	1.0	1.0	0.5	1	0.000
1.0	1.0	1.0	0	0.000	1.0	1.0	1.0	1	0.000

Table 4: Optimal Public Outcome Distribution

a_1	a_2	a_3	ω	$p(a,\omega)$	a_1	a_2	a_3	ω	$p(a,\omega)$
0.0	0.0	0.0	0	0.082	0.0	0.0	0.0	1	0.000
0.0	0.0	0.5	0	0.000	0.0	0.0	0.5	1	0.000
0.0	0.0	1.0	0	0.000	0.0	0.0	1.0	1	0.000
0.0	0.5	0.0	0	0.000	0.0	0.5	0.0	1	0.000
0.0	0.5	0.5	0	0.000	0.0	0.5	0.5	1	0.000
0.0	0.5	1.0	0	0.000	0.0	0.5	1.0	1	0.000
0.0	1.0	0.0	0	0.000	0.0	1.0	0.0	1	0.000
0.0	1.0	0.5	0	0.000	0.0	1.0	0.5	1	0.000
0.0	1.0	1.0	0	0.000	0.0	1.0	1.0	1	0.000
0.5	0.0	0.0	0	0.000	0.5	0.0	0.0	1	0.000
0.5	0.0	0.5	0	0.000	0.5	0.0	0.5	1	0.000
0.5	0.0	1.0	0	0.000	0.5	0.0	1.0	1	0.000
0.5	0.5	0.0	0	0.301	0.5	0.5	0.0	1	0.097
0.5	0.5	0.5	0	0.000	0.5	0.5	0.5	1	0.000
0.5	0.5	1.0	0	0.000	0.5	0.5	1.0	1	0.000
0.5	1.0	0.0	0	0.073	0.5	1.0	0.0	1	0.000
0.5	1.0	0.5	0	0.045	0.5	1.0	0.5	1	0.403
0.5	1.0	1.0	0	0.000	0.5	1.0	1.0	1	0.000
1.0	0.0	0.0	0	0.000	1.0	0.0	0.0	1	0.000
1.0	0.0	0.5	0	0.000	1.0	0.0	0.5	1	0.000
1.0	0.0	1.0	0	0.000	1.0	0.0	1.0	1	0.000
1.0	0.5	0.0	0	0.000	1.0	0.5	0.0	1	0.000
1.0	0.5	0.5	0	0.000	1.0	0.5	0.5	1	0.000
1.0	0.5	1.0	0	0.000	1.0	0.5	1.0	1	0.000
1.0	1.0	0.0	0	0.000	1.0	1.0	0.0	1	0.000
1.0	1.0	0.5	0	0.000	1.0	1.0	0.5	1	0.000
1.0	1.0	1.0	0	0.000	1.0	1.0	1.0	1	0.000

Table 5: Optimal Delegated Hierarchy

B.3 Extensions

In this section of the online appendix, we define the notion of multiple-meeting scheme, provide a simple example of it, define the notion of a random delegated hierarchy, where the order of delegation can depend on the message realization, and characterize the corresponding implementable distributions.

B.3.1 m-Meeting Schemes

Organizing multiple meetings in parallel offers extra flexibility.

Definition 6. *An information structure (S,P) is an m-meeting scheme if, for all s with $P(s)>0$, $\mathscr{I}$ is partitioned in at most $m+1$ groups $\{G_1(s),\ldots,G_{m+1}(s)\}$ such that $i =^s_{Inf} j$ for all $i,j \in G_k(s)$ and $k \le m$, and $s_i = s'_i$ for all $s,s' \in S$ such that $i \in G_{m+1}(s) \cap G_{m+1}(s')$.*

An m-meeting scheme organizes at most m simultaneous meetings, $\{G_1(s),\ldots, G_m(s)\}$, each of which makes its content common knowledge among the participants. The agents who are not invited to any meeting at message profile s are contained in $G_{m+1}(s)$. There is an important distinction between the uncertainty from participation in a meeting (about what is said in other meetings) and the uncertainty from participation in no meeting. The latter must be the same across all messages, because if i is not present in any of the meetings $\{G_1(s),\ldots,G_m(s)\}$ or $\{G_1(s'),\ldots,G_m(s')\}$, then i must have the same belief given s_i as he does given s'_i.

For any $I \subseteq \mathscr{I}$, let $a_I = (a_i)_{i\in I}$. Let $\pi = \{\pi_k\}_{k=1}^K$ be a collection of up to m disjoint subsets of $\mathscr{I}$ (that is, $K \le m$). Let Π^m be the set of all such collections. Let $\tilde{m}(\pi) := \mathscr{I} \setminus \cup_k \pi_k$. Define $\pi(i)$ as follows: $\pi(i)$ is the element of π that contains i if such an element exists, and $\pi(i) = \tilde{m}(\pi)$ otherwise. The next result characterizes the outcome distributions that can be implemented by m-meeting schemes.

Proposition 3. *A distribution $p \in BCE(\mu)$ can be implemented by an m-meeting scheme, if and only if, there exist $\beta \in \Delta(\Pi^m \times A \times \Omega)$ such that $p = marg_{A\times\Omega}\beta$ and $\tilde{a}_i \in A_i$ for all $i \in \mathscr{I}$ such that:*

1. *for all π and $i \in \tilde{m}(\pi)$, $\beta(\pi,a,\omega)=0$ for all $a \in A$ such that $a_i \neq \tilde{a}_i$;*

2. *for all $i \in \mathscr{I}$, $a_i \in A_i \setminus \tilde{a}_i$ and $I \subseteq \mathscr{I}$,*

$$\sum_{\pi:\pi(i)=I} \sum_{a_{\mathscr{I}\setminus I}} \sum_{\omega} \beta(\pi,a_i,a_{I\setminus i},a_{\mathscr{I}\setminus I},\omega)\left(u_i(a_i,a_{I\setminus i},a_{\mathscr{I}\setminus I};\omega)-u_i(a'_i,a_{I\setminus i},a_{\mathscr{I}\setminus I};\omega)\right)\ge 0 \tag{37}$$

for all $a_i' \in A_i$ *and* $a_{I\setminus i} \in A_{I\setminus i}$.

Given the definition and characterization presented above, there are a number of relevant questions that can be asked for a given strategic interaction and designer objective. For example, we could ask: What is the minimal m such that an optimal outcome distribution $p^* \in \Delta(A \times \Omega)$ can be implemented by an m-meeting scheme? Of course, an n-meeting scheme imposes no restriction at all, because any message profile can be transmitted through individual meetings with each agent. Therefore, the answer to this question is at most n, making it especially interesting to identify problems for which the answer is strictly less than n. Additionally, we could ask a related question: For which environments is restricting attention to m-meeting schemes, where $m < n$, without loss?

B.3.2 Random Delegated Hierarchies

We generalize Definition 3 and Theorem 2 by allowing the order of delegation to depend on the message profile.

Definition 7. *A distribution* $p \in \Delta(A \times \Omega)$ *can be implemented by a random delegated hierarchy if there is an information structure* (S, P) *and an equilibrium* $a^* \in \mathscr{E}(S, P)$ *such that*

$$p(a, \omega) = \sum_{s \in S} \mu(\omega) P(\{s : a^*(s) = a\} | \omega) \quad \forall a \in A, \omega \in \Omega$$

and if, for every s *such that* $P(s) > 0$, *there exists a total order* $\succeq^s$ *on* $\mathscr{I}$ *such that* $\succeq^s_{\text{Inf}}$ *coincides with* $\succeq^s$ *and*

$$\mathbb{E}\left[u_i\left(a_i^*(s_i), a^*_{\prec^s i}(s_{\prec^s i}), a^*_{\succ^s i}(s_{\succ^s i}); \omega\right) \Big| s_i\right] \geq \mathbb{E}\left[u_i\left(a_i', a^*_{\prec^{s'} i}(s'_{\prec^{s'} i}), a^*_{\succ^s i}(s_{\succ^s i}); \omega\right) \Big| s_i\right] \tag{38}$$

for all $a_i' \in A_i$ *and* $s' \in S$ *such that* $P(s') > 0$ *and* $\{j : j \prec^{s'} i\} = \{j : j \prec^s i\}$.

The definition assigns to each message profile realization s a total order $\succeq^s$ such that two conditions hold. First, each s corresponds to a "local information hierarchy" in which every i knows the messages of his $\succeq^s$-predecessors in the information structure (S, P). This condition is satisfied by all single-meeting schemes with a total order that ranks the agents invited to a meeting at s arbitrarily amongst each other but above the non-invited agents at s, who are also ranked arbitrarily amongst each other. Single-meeting schemes, however, may not satisfy (38).

In the expectations in (38), taken over $(s_{\succ^s i}, \omega)$, agent i can deviate to any a'_i and also misreport to all of his $\succeq^s$-predecessors by switching to any positive probability message profile $s'_{\prec^{s'} i} = (s'_j : j \prec^{s'} i)$ such that the set of i's predecessors at s' is the same as at s.[25] One subtlety in (38) is that upon observing s_i, i learns his rank in the total order $\prec^s$, because he can infer it from $\{j : j \prec^s i\}$, whose messages he is asked to forward and can manipulate. Hence, unlike in Halac, Lipnowski, and Rappoport (2020) and Morris, Oyama, and Takahashi (2020), there is no own-rank uncertainty in our definition, as the only rank uncertainty pertains to agents higher-up in the hierarchy.

In addition, we assume that for each message profile s, agent i knows the identity of his immediate successor, $i^+_s = \min\{j : j \succ^s i\}$, from whom he receives his message. This information is used in (38) to compute the expectation over all $s_{\succ^s i}$ such that i^+_s is i's immediate successor. We could alternatively assume that i does not know the identity of i^+, in which case the expectation in (38) would be over all possible $s_{\succ^s i}$, rather than just those that have i^+_s as i's immediate predecessor.

The next proposition generalizes the characterization in Theorem 2 by letting the total order depend on the action profile.

Proposition 4. *A distribution $p \in \Delta(A \times \Omega)$ can be implemented by a random delegated hierarchy, if and only if, for each $a \in A$ there exists a total order $\succeq^a$ on $\mathscr{I}$ such that for all $i \in \mathscr{I}$,*

$$\sum_{\omega \in \Omega} \sum_{a_{\succ^a i}} p(a_i, a_{\prec^a i}, a_{\succ^a i}, \omega)\big(u_i(a_i, a_{\prec^a i}, a_{\succ^a i}; \omega) - u_i(a'_i, a'_{\prec^{a'} i}, a_{\succ^a i}; \omega)\big) \geq 0 \tag{39}$$

for all $a'_i \in A_i$ and $a' \in A$ such that $p(a') > 0$ and $\{j : j \prec^{a'} i\} = \{j : j \prec^a i\}$.

For any μ, denote by $RDH(\mu)$ the set of outcome distributions that satisfy the above necessary and sufficient conditions.

What distributions are in $RDH(\mu)$ but *not* in $DH(\mu)$? Before giving a partial answer, let us see if p^* from Example 1 can be implemented by a random delegated hierarchy. Start with $a = (0,1,1)$. As discussed in Example 1, agents 2 and 3 cannot be first in the ordering and so the possible orderings are $1 \succ^a 2 \succ^a 3$ or $1 \succ^a 3 \succ^a 2$. Whichever one is chosen should also apply to $a' = (1,1,1)$, for otherwise agent 2 would infer from his rank that the state is 0 at a and hence refuse to play 1. This, however, creates a problem at $a'' = (1,0,1)$, where 2 has to be first and the possible orderings are $2 \succ^{a''} 1 \succ^{a''} 3$ or $2 \succ^{a''} 3 \succ^{a''} 1$. In either case,

[25] We abstract from deviations to message profiles where $\{j : j \prec^s i\} \subseteq \{j : j \prec^{s'} i\}$ as those may be detected as misreports at some point in the hierarchy.

agent 1 will learn that $\omega = 0$ at a'' because, unlike at a', he knows he is not first in rank. Hence, 1 will refuse to play 1 at a''. We conclude that p^* is *not* implementable by a random delegated hierarchy either.

From Claim 3 in Section 3.2 and the discussion that precedes it, we know that strict randomizations between profiles in NE(μ) are in general not included in $DH(\mu)$. Yet randomizations between strong Nash equilibria are included in $RDH(\mu)$. Let

$$\mathrm{SNE}(\mu) = \Big\{ p \in \Delta(A \times \Omega) : \quad \exists\, a^* \in A \text{ s.t. } p(a^*, \cdot) = \mu \quad \text{and}$$
$$\text{for all } J \subseteq \mathscr{I} \text{ and } a_J \in \times_{j \in J} A_j \text{ there exists } i \in J \text{ s.t.}$$
$$\sum_{\omega \in \Omega} \mu(\omega) u_i(a^*; \omega) \geq \sum_{\omega \in \Omega} \mu(\omega) u_i(a_J, a^*_{-J}; \omega) \Big\}$$

be the set of pure strategy *strong*-Nash outcomes in the ex-ante normal form game in which it is common knowledge that all agents share belief μ. Let

$$\mathrm{SPublic}(\mu) = \bigcup \left\{ \sum_{\hat{\mu}} \alpha(\hat{\mu}) \mathrm{Co}(\mathrm{SNE}(\hat{\mu})) : \quad \alpha \in \Delta(\Delta(\Omega)) \text{ s.t. } \sum_{\hat{\mu}} \alpha(\hat{\mu}) \hat{\mu} = \mu \right\}.$$

Claim 4. $\mathrm{SPublic}(\mu) \subseteq RDH(\mu)$.