Optimal Information Hierarchies*

LAURENT MATHEVET† INA TANEVA‡

New York University University of Edinburgh

January 2020

[Preliminary & Incomplete]

ABSTRACT. The way information allocates knowledge in a group affects its implementation. Are there information structures with desirable implementation properties that are optimal in a wide range of environments? We study this question as a standard information design problem, but emphasize the organizational characteristics of optimal disclosure. We propose the notions of a single-meeting scheme and of a decentralizable information hierarchy to capture notions of "easier", or less costly, implementation. Decentralization relies on partitioning agents into groups which are ranked by how informed they are in a strong sense, so that information can flow down the hierarchy after communicating once with the group at the top. Our results establish a connection between implementation and strategic complementarities, showing conditions under which single-meeting schemes and decentralizable information hierarchies are optimal in binary-action environments. We apply our results to a regime-change framework and outline the effects which determine the agents' ranks in the information hierarchy.

Keywords: information design, belief hierarchy, signal structure, organizations, networks.

JEL Codes: D82, D83, D91.

*We would like to thank conference participants at the CRETA conference at Warwick, Frontiers in Design at UCL, Information Design and Splitting Games in Paris, HSE-Vienna Workshop at Higher School of Economics in Moscow, SAET 2019 in Ischia, as well as seminar audiences at Cambridge, Glasgow, Graz, Edinburgh, Oxford, Saint Etienne, Toulouse, and Zurich.

†lmath@nyu.edu

‡ina.taneva@ed.ac.uk

1 INTRODUCTION

The way information allocates knowledge in a group affects its implementation. For example, if someone wants to invite two people to an event, then she can send two separate (private) invitations or send one (public) invitation to both. In the first option, neither of them may know that the other is invited, while it is common knowledge in the second. Since communicating content once is cheaper than communicating the same content twice, public communication should be preferred if it induces the right incentives. While these considerations may be anecdotal in small groups, they are very important for information sharing or advertising in large organizations or markets. Even if the physical costs of communication have fallen, the costs of implementing certain knowledge distributions remain non-trivial. In the current email systems, for example, it is impossible for a director to send one email to three managers such that one manager knows that the other two have received the email while one manager is kept uncertain about any other recipient.[1] Such prescriptions might seem strange when taken out of context, but they are commonplace in theory, regardless of whether and how they can actually be implemented. Are there information structures with desirable implementation properties that are optimal in a wide range of scenarios?

We study this question in the standard information design framework; however, unlike the existing literature on information design (Kamenica and Gentzkow (2011), Bergemann and Morris (2016), Taneva (2019), Mathevet, Perego, and Taneva (2019), and others), we emphasize the organizational characteristics of the optimal information structure. In our model, a set of agents make a binary decision under incomplete information about a payoff-relevant state of the world. For example, think of it as the decision of whether or not to exert effort in a team working on a project with uncertain prospects; whether or not to buy a product of uncertain quality;

1. The director would need two *Bcc* fields such that the manager in Bcc_2 knows the identity of the recipients in Bcc_1 and *To* fields; the manager in Bcc_1 knows the identity of the recipient in *To* and the manager in *To* does not know if anyone else has received the email.

whether or not to invest in a venture with uncertain viability; and so on. The information provided to the agents about the state affects the equilibrium outcomes of their interaction not only through their beliefs about the state, but also through their beliefs about each other's beliefs about the state. Information can be thus used for various objectives, e.g., to maximize expected total effort or total sales, or to maximize expected welfare.

We introduce two notions of what it might mean for information to be easily implemented. A natural place to start is to consider information communicated publicly but to a restricted audience. This is what we call a single-meeting scheme. In organizations, this is equivalent to a director calling a meeting during which she will make a public announcement to all who are present. However, she has leeway in choosing (perhaps, with some randomness) whom to invite to the meeting. Such information structures are intuitively appealing, because only one meeting ever takes place during which information is communicated publicly to the participants — or, equivalently, one (email) message is ever sent to the selected recipients. Our second notion is decentralizable information hieararchies, where information transmission can be delegated after communicating only with the top group in the hierarchy. In organizations, a director may not have the time to communicate directly, on a regular basis, with all her subordinates, from the managers to the workers. Being able to delegate information transmission, so that it flows down the hierarchy of employees in a decentralized way, is therefore a desirable property of an information structure. In marketing, high-profile individuals are often targeted for their endorsement of a product or an event, in the hope that their followers will do the same, and so will the followers of their followers, and so on. Such viral marketing relies on delegating information transmission.

In our formal definition of an information hierarchy, agents are partitioned into groups such that all members of a group are equally informed, and more informed than all members of lower groups. To be "more informed" is meant in a strong sense here, as an agent is more informed than another if there is nothing the former would learn by knowing the information of the latter. These organizational forms of distributed knowledge, which are shaped like pyramids and totally order agents by the degree of their informedness, make it feasible for more informed agents to pass down information to those less informed than them. But would they have the strategic incentives to truthfully pass down information? That constitutes the second requirement, which makes an information hierarchy decentralizable: agents have an incentive to forward the prescribed information to the group of agents immediately below them in the hierarchy, understanding that those agents in turn will do the same to the group below them and so on.

Note that in our environment the Revelation Principle holds. Hence, there is always an optimal direct information structure whereby an agent is simply sent an action recommendation, which is how he would have optimally responded to his message. However, direct information structures are typically not single-meeting schemes or information hierarchies which can be decentralized. We would like to emphasise that the latter indirect structures come with desirable properties in terms of ease of implementation, which the former do not have. Although direct information structures simplify the analyst's optimization problem, they are deceptively simple, because it is often difficult to implement their privacy requirements. When only two messages can be used (as is the case with only two actions per player), the knowledge relationships that ensue are often packed with epistemic requirements across agents, so that epistemic complexity compensates for the simplicity of the message space. In practice, using email as an illustration, this may require as many *Bcc* fields as there are agents and to carefully choose which recipient is put in each field depending on the message to be sent.

Our results establish a natural connection between implementation and strategic complementarities. Let us return to the invitation example. Intuition suggests that if the invitees' decisions to attend the event depended positively on the other's attendance, then a public invitation to a subset of agents would be optimal. This is confirmed by our first proposition, which proves that single-meeting schemes are optimal with considerable generality in binary-action environments, provided each agent finds it more desirable to exert effort, buy a product and so on, if others do so as well. In particular, this does not require any complementarity with the state of the world, and many different objectives are covered by the result. It also holds whether an agent regards others' decisions as complements to each other. If we invoke this stronger form of complementarity, we obtain our second result on decentralizable information hierarchies. If an agent prefers working with different subgroups of other agents better than on his own or with everyone jointly, then maximizing his likelihood of effort should involve disparate partial coordination. Clearly, this make obtaining an informational ranking amongst agents, and therefore decentralization of information flows, difficult. Our second result shows that a stronger complementarity condition, which involves complementarities with the state as well, is sufficient for information hierarchies to be both optimal and decentralizable.

One interesting aspect of information hierarchies and their ability to implement delegation has to do with the composition of the different groups. Who should be given most information and who should be given relatively less information? We answer this question in the context of the regime-change model of Sakovics and Steiner (2012). Regime-change games from the global game literature (e.g., Morris and Shin (2003)) are one of the main applications of our framework.[2]

The paper is organised as follows. In Section 2 we present the model and the assumptions we impose on the primitives. Section 3 contains our main definitions and results. In Section 4 we present our application to a regime-change environment, and Section 5 concludes. All proofs are relegated to the Appendix.

2. Assuming an arbitrary but finite number of agents and states.

2 The Model

A set of agents $\mathscr{I} = \{1,\dots,n\}$ interact in an uncertain environment where the state of the world, ω, can take on finitely many values in $\Omega \subseteq \mathbb{R}$ according to the prior $\mu \in \Delta(\Omega)$. Every agent $i \in \mathscr{I}$ chooses $a_i \in \{0,1\}$, without observing the actions of any other player, and receives a payoff $u_i : A \to \mathbb{R}$. We have normalized payoffs so that $u_i(0, a_{-i}; \omega) = 0$ for all $a_{-i} \in \{0,1\}^{n-1}$ and $\omega \in \Omega$. Let $\mathscr{G} = ((u_i, \{0,1\})_{i \in \mathscr{I}}, \mu)$ denote the underlying strategic environment.

A priori, the agents only know that ω is distributed according to the prior $\mu \in \Delta\Omega$. They receive additional information, which is modelled as an information structure (S,P), where $S = \prod_i S_i$ is a message space and $P = \{P(\cdot|\omega)\}_{\omega \in \Omega}$ is a set of conditional distributions, one for each possible realization of the state. That is, in any state ω, the message profile $s = (s_i)_i$ is drawn according to $P(s|\omega)$ and player i privately observes s_i. For all $i \in \mathscr{I}$, assume each $s_i \in S_i$ is sent with strictly positive probability in some state, which is without loss.

For a Bayesian game $(\mathscr{G}, (S,P))$, denote a pure strategy of player i by $a_i : S_i \to A_i$ and let a^* denote a (pure) Bayes Nash equilibrium (BNE) strategy profile of the game.[3] An information scheme is chosen to maximize the expected value of an objective $v : A \times \Theta \to \mathbb{R}$, so that

$$\sup_{(S,P)} \mathbb{E}v \quad \text{where } \mathbb{E}v = \sum_{\omega,s} v\left((a_i^*(s_i))_i; \omega\right) P(s|\omega)\mu(\omega)$$

describes the design problem. In case of equilibrium multiplicity for a given information structure (S,P), we focus on favorable selection, i.e. we assume players coordinate on the equilibrium which maximizes the expected value of the objective.

3. In our environment, it is without loss to focus on pure strategies as the designer-preferred equilibrium is always the largest one, which is a pure strategy equilibrium.

2.1 Interactions

We study situations where actions are strategic complements and also where there is state complementarity. This means that choosing the higher actions is more desirable to an agent when others choose it as well and also when the state is higher. This covers many different relevant situations. Choosing whether to exert effort in a team, whether to buy a product or to adopt a technology, whether to invest in a project, or whether to participate in a revolution can all be modeled as environments with strategic and state complementarities.

We make the following assumptions on the players' payoffs.

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

Let

$$\underline{\omega} = \min\{\omega \in \Omega : u_i(1, \ldots, 1; \omega) \geq 0 \text{ for all} i \in \mathscr{I}\}$$

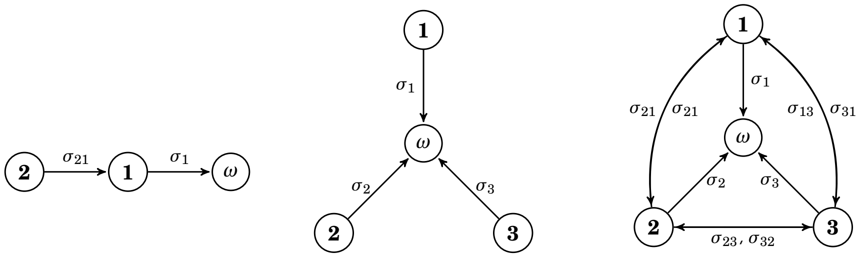
and define $\Psi = A \times \{\omega < \underline{\omega}\}$.

Assumption 1'. (Supermodularity) For all $i \in \mathscr{I}$, u_i is supermodular on Ψ.[4]

According to the first assumption, a_i and a_{-i} are strategic complements. This captures the idea that choosing 1 is more desirable to an agent when others choose 1 as well. The second assumption adds the requirement that a_i and ω are complementary (larger states encourage choosing 1) and that each i regards the other agents' actions as strategic complements. Moreover, there are important applications, which we would like to capture in our framework, but in which the agents' utilities fail to be supermodular on $A \times \Omega$ and yet are supermodular on Ψ. So, the choice of supermodularity on Ψ is not arbitrary.

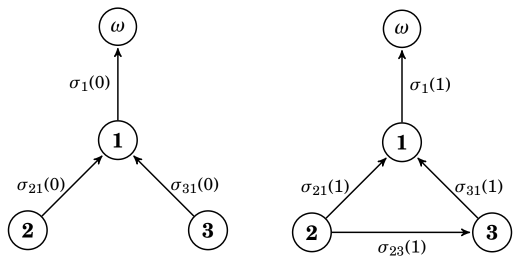
There are two main applications that we would like to consider.

4. As standard, given a lattice $(X, \geq)$, supermodularity of $f : (X, \geq) \to \mathbb{R}$ means that for all $x', x'' \in X$, $f(x' \vee x'') + f(x' \wedge x'') \geq f(x') + f(x'')$.



(a) Line (b) Star (c) Complete network

FIGURE I: FIXED NETWORKS



$\omega = 0$ $\omega = 1$

FIGURE II: RANDOM NETWORKS

Network Interactions. Agents make binary decisions while on a network à la Ballester, Calvo-Armengol, and Zenou (2006):

$$u_i(a_i, a_{-i}; \omega) = \sigma_i(\omega) a_i + \sum_{j \neq i} \sigma_{ij}(\omega) a_i a_j$$

where σ_i measures i's intrinsic preference for the high action and σ_{ij} expresses i's dependence on j. Assume that for all i and j, $\sigma_{ij}(\cdot) \geq 0$ and that $\sigma_i(\cdot)$ and $\sigma_{ij}(\cdot)$ are weakly increasing. In this incomplete-information version of the model, the state may affect the intrinsic preferences for action 1 of each agent, but also the network interactions between agents. Therefore, this covers fixed networks, as in Figure I, but also random networks, as in Figure II, where the network structure is unknown. While the state and the neighbors always have a positive influence, the relative importance of each one might change with the state.

This formulation includes the investment game of Carlsson and van Damme (1993) and Morris and Shin (2003) with $n = 2$, $\sigma_i(\omega) = \omega - 1$ and

$\sigma_{ij}(\omega) = 1$ and other celebrated applications (for example, "beauty contest" descriptions of social phenomena such as location in a city or suburb, and entry or withdrawal from the labor force (see Brock and Durlauf (2001) and Glaeser and Scheinkman (1999)).[5]

Regime-Change Models. These models are coordination games in which a status quo is abandoned when a sufficiently large number of agents choose action 1. Examples abound in the global game literature (see Morris and Shin (2003)), such as currency attacks (Morris and Shin (1998b)), bank runs (Goldstein and Pauzner (2005)), debt crisis (Morris and Shin (1998a)) and many others. To illustrate our findings, we take the regime-change model of Sakovics and Steiner (2012). A finite set of investors make binary investment decisions into a common project. Let $\kappa_i > 0$ be i's contribution to the success of the project, $c_i > 0$ be his cost of investment, and $b_i > 0$ his benefit from a successful project. If i invests, her payoff is:

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

where $b_i - c_i > 0$. Although u_i is *not* supermodular on $A \times \Omega$, because the inf of two action profiles at which the project succeeds can make the project fail, u_i is supermodular on Ψ for all i.

2.2 Objectives

We focus on objectives in which higher actions are desirable and so is coordination between agents and between actions and states. We make the following assumptions on the objective function.

Assumption 2. (Mutually Beneficial Monotonicity) If $P(\{s \in S : a^*(s) \in A\}|\omega) \geq P'(\{s' \in S' : a^*(s') \in A\}|\omega)$ for all $\omega \in \Omega$ and upper sets A and $\mathbb{E}u_i(S,P) \geq \mathbb{E}u_i(S',P')$ for all $i \in \mathscr{I}$, then $\mathbb{E}v(S,P) \geq \mathbb{E}v(S',P')$.

5. In Brock and Durlauf (2001), the action choice is also binary, but they study a variety of interactions. Basic payoffs take the form: $\sigma_i(\omega) = 2\beta_i\omega$ and $\sigma_{ij}(\omega) = (1-\beta_i)\frac{2\sum_{j\neq i}a_j}{n-1} - 1$ in the normalized version (which ensures 0 payoff when $a_i = 0$).

Assumption 2'. (Supermodularity) v is supermodular on Ψ.

3 Implementation

In this section we present our main definitions and results. We introduce two different notions which relate to ease of implementability of an information structure in natural ways. For example, it is reasonable to argue that there are costs increasing in the number of one-to-one meetings that a sender has to hold or the number of individual messages that she has to write and send. This would create a wedge between direct information structures, which have message simplicity but come with belief-complexity that is costly to implement, and the indirect information structures we consider, which come with a larger message space, but allow for more straightforward distribution of informations and its decentralized implementation. Once these costs are taken into considerations, working only with direct information structures is no longer without loss, and their optimality is not guaranteed.

Definition 1. (S,P) is a single-meeting scheme if for all $i \in \mathscr{I}$ there is at most one $\tilde{s}_i \in S_i$ such that $\mu_i(s_{-i}|\tilde{s}_i) \neq 1$ for some $s_{-i} \in S_{-i}$.

A single-meeting scheme can be implemented by the designer through meetings with subgroups of the players. Each signal profile realisation calls for the designer to summon a different subset of the players to a meeting. As a result, only a single meeting is ever organised. However, ex ante, there are different meetings possible with different players present at each one. Being invited to a meeting itself carries the information embedded in the signal profile and that information becomes common knowledge among those present at the meeting. Importantly, a player who is not called to a meeting, always has the same first- and higher-order beliefs, as represented by the signal $\tilde{s}_i$. Therefore, the players who are present at a meeting know both what signal the other players at the meeting have observed, as well as what signals those who are not at the meeting have observed; that is, $\mu_i(s_{-i}|s_i) = 1$ for all player i and $s_i \neq \tilde{s}_i$. Notice that a player can have different first- and higher-order beliefs depending on which meeting he is invited to, while not being invited to a meeting always implies the same first- and higher-order beliefs.

An alternative implementation of a single-meeting scheme is possible in the context of electronic mail, where only one message is sent to a subgroup of the players and the recipient list is visible to all who receive the email, i.e. all recipients can be identified in the "To" or "Cc" field. In this setting, not receiving an email is associated with the signal $\tilde{s}_i$, while receiving an email is perfectly informative about both the beliefs of the other players in the recipient list and the beliefs of the players who are not in the recipient list.

PROPOSITION 1. Under Assumptions 1 and 2, a single-meeting scheme is optimal.

To discuss the features of different information structures and their implementation properties, it is useful to define some concepts of informedness. In particular, we introduce a concept of being more informed in a very strong sense — both about the state of the world and about the signals of other players. Given an information structure (S,P), player i is (weakly) **more informed** than player j if $\mu_i(\omega, s_{-i}|s_i, s_j) = \mu_i(\omega, s_{-i}|s_i)$ for all s and ω. This is, indeed, a very strong form of being more informed: there is nothing j knows, not even his own message, that i does not. When two agents are weakly more informed than one another, they are **equally informed**. For example, all agents are equally informed when information is public. Moreover, if i is weakly more informed than j but j is not weakly more informed than i, then i is **strictly more informed** than j.

In a single meeting scheme, there is ex ante no ranking of players regarding their level of informedness about the state and about the signals of others. However, at the interim level, each signal realisation results in a binary ranking: the players who are invited to the meeting are equally informed amongst themselves and are strictly more informed than all players who are not invited to the meeting. Importantly, the players who are not invited to the meeting need not be ordered in terms of their informedness, nor are they necessarily equally informed amongst themselves.

We next introduce a definition of a class of information structures which provide an ex-ante ordering of all players, so that players are more or less informed relative to other players for all possible signal realisations. In other words, for these information structures, the ex ante informedness ordering is also preserved at the interim stage.

DEFINITION 2. (S,P) is an information hierarchy if all players can be partitioned into groups $\{G_k\}_{k=1}^{K}$ such that for any $k = 1,\ldots,K-1$ everyone in group G_k is equally informed, and strictly more informed than everyone in $\cup_{\ell>k} G_\ell$.

An information hierarchy is a total ordering of agents according to the strong informedness criterion introduced above. Notice that the agents who are in the same group are always equally informed and so it is without loss to assume that they observe the same signal. A public information structure is the simplest example of an information hierarchy, where all agents are equally informed and therefore there is only one group. More generally, when there are some agents who are strictly more informed than others, there are multiple groups and agents in lower-indexed groups always know the signals of agents in higher-indexed groups. If we denote the signal observed by all agents in group G_k by s_k, the previous sentence implies that

$$\mu_k(s_{k+1},\ldots,s_K|s_k) = 1,$$

where μ_k is the belief of any agent in group k. This turns out to be an important feature of information hierarchies, as it is a necessary requirement for a decentralized implementation of the information structure, where the designer communicates only once with the most informed group, G_1, and lets information flow down the hierarchy in a decentralised way. For this to work, there is a second condition that needs to be satisfied — the truthful transfer of information needs to be incentive compatible. That is, more informed players need to have an incentive to pass down the information they have been told to report to less informed players, without distorting it. Whether these incentives are satisfied or not depends on the nature of the underlying strategic environment.

DEFINITION 3. An information hierarchy (S, P) is decentralizable in an environment $\mathscr{G}$ if for all $i \in G_k$ and all $k = 1, \ldots, K$,

$$\mathbb{E}\left[u_i\left(a^*(s);\omega\right)\Big|s_k\right] \geq$$
$$\mathbb{E}\left[u_i\left((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},\ldots,(a_j^*(s_k))_{j\in G_k, j\neq i}, a_i', (a_j'(s_m))_{j\in\bigcup_{m>k}G_m};\omega\right)\Big|s_k\right] \quad (1)$$

for all $(a_i', (a_j'(s_m))_{j\in\bigcup_{m>k}G_m})$.

PROPOSITION 2. Under Assumptions 1-1' and 2-2', a decentralizable information hierarchy is optimal. It is also a single-meeting scheme.

The intuition behind the proof of this proposition is as follows. In a first step (Lemma 1) we show that there exists an optimal *direct* information structure which satisfies an inclusion property. This property allows us to rank the sets of agents who play the high action from largest to smallest in every state, and also to show that every set of agents who is recommended to play the high action in a low state is included in every set of agents who is recommended to play the high action in any higher state. This lemma is a generalization to multiple states and a less restrictive supermodularity requirement of a result by Arieli and Babichenko (2019). These generalizations allow us to capture important economic environments not covered by previous results. A pivotal step in the proof relies on using transfers of probability obeying the supermodular ordering of Meyer and Strulovici (2012).

The second step in the proof (Lemma 2) involves augmenting the messages of the optimal direct information structure from the first step to inform certain groups of agents about the signal realizations of others. In this way, we create groups of agents which are totally ranked in terms of informativeness across groups and all agents within a group are equally informed. We also show that this augmentation of messages can never result in a lower value of the objective. The resulting indirect information structure is both an information hierarchy and a single-meeting scheme. The latter allows information to be disclosed by communicating once with a subgroup of all agents. The former allows for the possibility of decentralization.

The last step in the proof (Lemma 3) speaks to the decentralizability of the optimal (indirect) information structure and shows that in the environments we consider it is indeed decentralizable. Here, we show that agents that are more informed do not have an incentive to misreport the relevant signals to the agents in groups below them, which are less informed. We show that there are no profitable deviation while allowing for the possibility to deviate in both own action and in the report to the agents in the group below. Therefore, there is an optimal information structure which is an information hierarchy and in which the transmission of information down the chain can be decentralized after communicating once with the most informed group, which is highest up in the hierarchy.

We next consider an example where the weak complementarities of Assumption 1 and the weak monotonicity of Assumption 2 ensure optimality of a single-meeting scheme, but are not sufficient to allow for a hierarchical decentralization.

EXAMPLE. Consider a team effort problem where $\Omega = \{0,1\}$, $\mu(\omega = 0) = 4/5$ and

$$u_1(1, a_{-1}; \omega) = \begin{cases} 1 & \text{if } \omega = 1 \\ -1 & \text{if } \omega = 0 \end{cases}, \quad u_2(1, a_{-2}; \omega) = \begin{cases} -1 + 2 \cdot \mathbb{1}_{\{\omega=1\}} & \text{if } a_1 = 1 \\ -2 + 2 \cdot \mathbb{1}_{\{\omega=1\}} & \text{if } a_1 = 0 \end{cases},$$

and

$$u_3(1, a_{-3}; \omega) = \begin{cases} -2 & \text{if } a_1 = a_2 = \omega = 0 \\ a_1 + a_2 & \text{otherwise.} \end{cases}$$

Agent 1 wants to work if the state is high; agent 2 wants to work only if 1 works too and he intrinsically likes state 1; and agent 3's utility for effort is driven by the total efforts of 1 and 2, except in the low state where exerting effort alone carries a penalty. Clearly, all $u_i(1, a_{-i}; \omega)$ are weakly increasing in a_{-i} for each ω so that Assumption 1 holds. For the sake of simplicity, choose $v = a_3$, which satisfies Assumption 2. Then, the optimal direct information structure is given by:

$\omega = 0$	0,0	1,0	0,1	1,1
0	3/16	0	7/16	1/8
1	0	0	1/4	0

$\omega = 1$	0,0	1,0	0,1	1,1
0	0	0	0	0
1	0	0	0	1

This optimal direct information structure can be converted into a single-meeting scheme. For each signal realization this is done by summoning only the players who are recommended to play action 0 to a meeting. However, it cannot be converted into an information hierarchy and therefore does not allow for decentralized implementation. The reason is that agent 1 cannot be higher in the hierarchy than agent 2, for otherwise 1 would sometimes (at some message profile s) have to play 1 and recommend 2 to play 0, which agent 1 would not do (since he would prefer recommending 1). Agent 2 cannot be higher in the hierarchy than agent 1, for otherwise 2 would sometimes (at some message profile s') have to play 1 and recommend agent 1 to play 0 which agent 2 would not do. This implies that agents 1 and 2 must be at the same level in the hierarchy, and hence receive their recommendations from agent 3. Yet, agent 3 will sometimes play 1 while agents 1 and 2 play 0, which 3 will refuse to do. Therefore, no ranking of informativeness is possible, which also precludes the possibility for decentralized implementation.

4 Applications

In the context of linear networks, note that although the environment might display no hierarchical ordering of agents, optimal design treats them hierarchically with totally ordered information. Proposition 2 also invites the obvious question of group composition. In the regime-change model of Sakovics and Steiner (2012), laid out in Section 2.1, the answer is simple and illuminating. To begin with, suppose the objective is to maximize the probability that the project succeeds:

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

Note v is supermodular on Ψ (since $v=0$ on Ψ) and v is weakly monotone in a so that it satisfies Assumption 2. Since $\underline{\omega}$ is the lowest state at which the project can ever succeed, which assumes everyone invests, it is trivial to see $\mathbb{E}v$ is maximized by telling all agents that $\omega \geq \underline{\omega}$ when this is true, and telling them $\omega < \underline{\omega}$ when this is true.[6] Therefore, there is only one group, G_1, in which everyone is equally and perfectly informed.

Now consider another objective, $v = \sum_i a_i$. As before, truthful revelation induces everyone to invest at $\omega \geq \underline{\omega}$ and no one to invest at $\omega < \underline{\omega}$. Surely, it may be possible to generate more investment by pooling states below $\underline{\omega}$ together with all states above. For every $\omega' \leq \underline{\omega}$, consider

$$\text{Prob}(\{\omega \geq \underline{\omega}\}|\{\omega \geq \omega'\}) = \frac{\sum_{\omega:\omega\geq\underline{\omega}} \mu(\omega)}{\sum_{\omega:\omega\geq\omega'} \mu(\omega)}$$

which is the probability that the project succeeds despite knowing only that $\omega \geq \omega'$. Because each agent invests iff

$$\text{Prob}(\{\omega \geq \underline{\omega}\}|\{\omega \geq \omega'\})\, b_i - c_i \geq 0, \tag{2}$$

the goal is to extract as much investment as possible from each i by choosing $\omega' = \omega_i^*$ such that (2) holds with equality. Such agents i would no longer find it worthwhile to invest for $\omega' < \omega_i^*$, but others might. We conclude that there are as many groups as there are distinct subset who believe that the probability of success is not worth investing, that is,

$$\left|\left\{\frac{c_i}{b_i} : \exists\, i \in \mathscr{I} \text{ and } \omega' \in \Omega \text{ s.t. } \frac{c_i}{b_i} > \text{Prob}(\{\omega \geq \underline{\omega}\}|\{\omega \geq \omega'\})\right\}\right|$$

6. There are other optimal information structures, such as revealing the state exactly.

and

$$\begin{aligned} G_1 &= \left\{ i \in \mathscr{I} : \frac{c_i}{b_i} = \max_{i \in \mathscr{I}} \frac{c_i}{b_i} \right\} \\ G_2 &= \left\{ i \in \mathscr{I} : \frac{c_i}{b_i} = \max_{i \in \mathscr{I} \setminus G_1} \frac{c_i}{b_i} \right\} \\ &\vdots \end{aligned}$$

Agents at the top are more sensitive to the state than to their neighbors, compared to these relative sensitivities in the remaining agents.

5 Conclusion

In this paper, we study the optimality of a special class of information structures, which we call information hierarchies. In an information hierarchy, there is no information that agents in lower groups have that agents in higher groups do not. We show that information hierarchies are optimal in a general class of problems with complementarities under favorable equilibrium selection. How the optimal hierarchies are structured depends on agents' sensitivity to the state and on the strength of the strategic complementarities.

In the context of organizational design, the flow of information inherent in an information hierarchy is particularly attractive. It has useful properties in terms of implementation (e.g., delegation) and robustness to certain forms of communication between agents. These advantages demonstrate that even when two information structures deliver the same expected payoff, there may be other important considerations to take into account that might make one preferable to the other.

A Appendix: Proofs

A.1 Proposition 1

Proof.

Take any IC direct information structure (A, P). For all i, define

$$S_i = \{1\} \bigcup \{(0, a_{-i})\}$$

and $s_i : A \to S_i$ such that

$$s_i(a) = \begin{cases} a_i & \text{if} \quad a_i = 1, \\ (a_i, a_{-i}) & \text{if} \quad a_i = 0. \end{cases}$$

Consider the indirect information structure $(S, \hat{P})$ such that $S = \prod_i S_i$ and $\hat{P}\big((s_i(a))_i | \omega\big) = P(a|\omega)$ for all a and ω. Take the following interim strategy

$$\beta_i(a_i|s_i) = \begin{cases} 1 & \text{if} \quad s_i = a_i \text{ or } s_i = (a_i, a_{-i})\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Given $(S, \hat{P})$, the interim payoff of agent i observing signal $s_i = 1$ and choosing action $a_i = 1$ when his opponents follow β_{-i} is:

$$\begin{aligned} \sum_{\omega} \sum_{a_{-i}, a'_{-i}} \mu(\omega) \hat{P}(1, s_{-i}(a)|\omega) \left(\prod_{j \neq i} \beta_j(a'_j | s_j(a)) \right) u_i(1, a'_{-i}; \omega) \\ = \sum_{\omega} \sum_{a_{-i}} \mu(\omega) \hat{P}(1, s_{-i}(a)|\omega) u_i(1, a_{-i}; \omega) \\ = \sum_{\omega} \sum_{a_{-i}} \mu(\omega) P(1, a_{-i}|\omega) u_i(1, a_{-i}; \omega) \geq 0, \quad (3) \end{aligned}$$

where the last equality follows since $\hat{P}(1, s_{-i}(a)|\omega) = P(1, a_{-i}|\omega)$ by definition of $(S, \hat{P})$, and the last inequality follows by IC of (A, P). Therefore, for all i, playing $a_i = 1$ is optimal conditional upon observing signal $s_i = 1$.

The interim payoff of agent i observing signal $s_i = (0, a_{-i})$ and choosing action $a_i = 1$ when his opponents follow β_{-i} is:

$$\sum_{\omega} \mu(\omega)\hat{P}(0, s_{-i}(a)|\omega) \left(\prod_{j \neq i} \beta_j(a'_j | s_j(a)) \right) u_i(1, a'_{-i}; \omega)$$
$$= \sum_{\omega} \mu(\omega)\hat{P}(0, s_{-i}(a)|\omega) u_i(1, a_{-i}; \omega) = \sum_{\omega} \mu(\omega) P(0, a_{-i}|\omega) u_i(1, a_{-i}; \omega). \quad (4)$$

If $\sum_{\omega} \mu(\omega) P(0, a_{-i}|\omega) u_i(1, a_{-i}; \omega) < 0$ for all i and $a_{-i} \in T_{-i}$, then action $a_i = 0$ is incentive compatible conditional upon observing signal $s_i = (0, a_{-i})$ for any $a_{-i} \in A_{-i}$. Therefore, the strategy profile β is an equilibrium and

$$\sum_{a' \in A} \hat{P}\big((s_i(a'))_i | \omega\big) \prod_{i=1}^{N} \beta_i\big(a_i | s(a')\big) = P(a|\omega),$$

which ensures that $\mathbb{E}v(S, \hat{P}) = \mathbb{E}v(A, P)$.

Alternatively, if there is a non-empty set $J \in \mathscr{I}$ and $A'_{-j} \subseteq A_{-j}$ for each $j \in J$ such that $\sum_{\omega} \mu(\omega) P(0, a_{-j}|\omega) u_i(1, a_{-j}; \omega) \geq 0$ if and only if $(j, a_{-j}) \in J \times A'_{-j}$, then for any i such that $J \setminus \{i\} \neq \emptyset$, upon receiving $s_i = 1$ and choosing $a_i = 1$, the difference in expected utility relative to (3) is given by

$$\sum_{\omega} \sum_{a: a_{-j} \in A'_j \forall j \in J} \mu(\omega)\hat{P}(1, s_{-i}(a)|\omega)[u_i(1, a_{J \setminus \{i\}} = 1, a_{-iJ}; \omega)$$
$$- u_i(1, a_{J \setminus \{i\}} = 0, a_{-iJ}; \omega)] \geq 0 \quad (5)$$

where the inequality follows by Assumption 1. Therefore, the incentives to play $a_i = 1$ upon observing $s_i = 1$ are only strengthened if some player j decides to deviate from β_j upon observing $s_j = (0, a_{-j})$ and play $a_j = 1$ instead of $a_j = 0$. Thus, there must exist a BNE strategy profile β^* such that

$$\sum_{a \in \hat{A}} \sum_{a' \in A} \hat{P}\big((s_i(a'))_i | \omega\big) \prod_{i=1}^{N} \beta_i^*\big(a_i | s(a')\big) \geq P(a \geq \hat{A}|\omega),$$

for all upper sets $\hat{A}$ and for all $\omega \in \Omega$. Hence, the distribution over equilibrium actions for each $\omega \in \Omega$ induced by $(S, \hat{P})$ strongly stochastically dominates that induced by (A, P), i.e. $P(a|\omega)$, for each $\omega \in \Omega$.

By equation **??** we know that the expected utility under $(S, \hat{P})$ from observing $s_i = 1$ and playing $a_i = 1$ is weakly higher than that under (A,P) from following action recommendation $a_i = 1$. Moreover, for any $i \in J$ and any $A'_{-i} \subseteq A_{-i}$, the expected utility from observing $s_i = (0, s_{-i})$ and playing action $a_i = 1$ is

$$\sum_{\omega} \mu(\omega) P(0, a_{-i}|\omega) u_i(1, a_{-i}; \omega) \geq 0,$$

which is by definition of J and A'_{-i} weakly greater than their payoff of 0 under (A,P), where they observed $a_i = 0$ and followed the recommendation. Also, the utility of any player i, for whom it is optimal to play $a_i = 0$ upon observing $s_i = (0, a_{-i})$, is equal to 0 and thus also equal to their utility under (A,P), where they observed $a_i = 0$ and followed the recommendation. Therefore, for all $i \in \mathscr{I}$ it holds that $\mathbb{E}u_i(S,\hat{P}) \geq \mathbb{E}u_i(A,P)$.

Hence, by Assumption 2 we obtain $\mathbb{E}v(S,\hat{P}) \geq \mathbb{E}v(A,P)$. ■

A.2 *Proposition 2*

For any $a \in A$, let

$$I(a) = \{i \in \mathscr{I} : a_i = 1\}.$$

LEMMA 1. There is an optimal information scheme (A, P^*) such that:

1. For all ω, 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 $\omega' < \omega''$, if $P(a'|\omega') > 0$ and $P(a''|\omega'') > 0$, then $I(a') \subseteq I(a'')$.

Proof. For any (A,P) and $i \in \mathscr{I}$, let

$$\mathbb{E}u_i(A,P) = \sum_{\omega \in \Omega} \sum_{a \in A} \mu(\omega) P(a|\omega) u_i(a;\omega).$$

(**Part 1**). Suppose that for some ω', a' and a'', (A,P^*) is such that $P^*(a'|\omega') > 0$ and $P^*(a''|\omega') > 0$, yet $I(a'') \subsetneq I(a')$ and $I(a') \subsetneq I(a'')$. Assume without loss that $P^*(a''|\omega') > P^*(a'|\omega')$.

Case 1: If $\omega' \geq \underline{\omega}$, then define $(A,\hat{P})$ as

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

If $i \in I(a')$, then[7]

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = \mu(\omega')P^*(a'|\omega')(u_i(1,\ldots,1;\omega') - u_i(a';\omega')) \geq 0,$$

which follows by Assumption 1. If $i \in I(a'')$ but $i \notin I(a')$, then

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = \mu(\omega')P^*(a'|\omega')u_i(1,\ldots,1;\omega') \geq 0,$$

which follows by the assumption of this case. For all other $i \in \mathscr{I}$, the incentive constraints remain unaltered.

Case 2: If $\omega' < \underline{\omega}$, then define $(A,\hat{P})$ as

$$\hat{P}(a|\omega) = \begin{cases} 0 & \text{if} \quad I(a) = I(a') \text{ and } \omega = \omega' \\ P^*(a|\omega) - P^*(a'|\omega') & \text{if} \quad I(a) = I(a'') \text{ and } \omega = \omega' \\ P^*(a|\omega) + P^*(a'|\omega') & \text{if} \quad I(a) = I(a') \cup I(a'') \text{ and } \omega = \omega' \\ P^*(a|\omega) + P^*(a'|\omega') & \text{if} \quad I(a) = I(a') \cap I(a'') \text{ and } \omega = \omega' \\ P^*(a|\omega) & \text{otherwise.} \end{cases}$$

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

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = \mu(\omega')P^*(a'|\omega')(u_i(a' \vee a'';\omega') - u_i(a';\omega')) \geq 0,$$

which follows by Assumption 1. If $i \in I(a'')$ but $i \notin I(a')$, then

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = \mu(\omega')P^*(a'|\omega')(u_i(a' \vee a'';\omega') - u_i(a'';\omega')) \geq 0,$$

which follows by Assumption 1. If $i \in I(a') \cap I(a'')$, then

$$\begin{aligned} \mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = \mu(\omega')P^*(a'|\omega')\Big(&\big[u_i(a' \vee a'';\omega') - u_i(a';\omega')\big] \\ &- \big[u_i(a'';\omega') - u_i(a' \wedge a'';\omega')\big]\Big) \geq 0, \end{aligned}$$

which follows by Assumption 1'. For all other $i \in \mathscr{I}$, the incentive constraints remain unaltered.

In both Cases 1 and 2, since the incentive compatibility condition to

7. Since i receives 0 when playing 0, this describes the difference in i's expected payoff between following recommendation $a_i = 1$ under $(A,\hat{P})$ and under (A,P^*).

follow action recommendation $a_i = 1$ was satisfied under (A,P^*) and is (weakly) greater under $(A,\hat{P})$, following the recommendations remains incentive compatible under $(A,\hat{P})$. In Case 1,

$$\mathbb{E}v(A,\hat{P}) - \mathbb{E}v(A,P^*) \;=\; \mu(\omega')P^*(a'|\omega')\big[v(1,\ldots,1;\omega') - v(a';\omega')\big] \;\geq\; 0, \quad (6)$$

which holds by Assumption 2. In Case 2,

$$\mathbb{E}v(A,\hat{P}) - \mathbb{E}v(A,P^*) = \mu(\omega')P^*(a'|\omega')\Big(\big[v(a' \vee a'';\omega') - v(a';\omega')\big] - \big[v(a'';\omega') - v(a' \wedge a'';\omega')\big]\Big) \geq 0, \quad (7)$$

which holds by Assumption 2'. Hence, the designer is (weakly) better off under $(A,\hat{P})$ than under (A,P^*).

By doing this procedure iteratively, for all a, a' and ω' such that $P^*(a|\omega') > 0$, $P^*(a'|\omega') > 0$, $I(a') \subsetneq I(a)$ and $I(a) \subsetneq I(a')$, we obtain (A,P^{**}) such that (i) action 1 is IC for all i; (ii) for any ω, if $P^{**}(a|\omega) > 0$ and $P^{**}(a'|\omega) > 0$, then $I(a') \subseteq I(a)$ or $I(a) \subseteq I(a')$; and (iii) $\mathbb{E}v$ is weakly larger under (A,P^{**}). Hence, (A,P^{**}) must also be optimal.

(**Part 2**). Now suppose that for some $\omega' < \omega''$, (A,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 $(A,\hat{P})$ as

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

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

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = P^*(a''|\omega'')\mu(\omega'')\big[u_i(1,\ldots,1;\omega'') - u_i(a'';\omega'')\big] \geq 0,$$

which follows by Assumption 1. If $i \in I(a')$ and $i \notin I(a'')$, then

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = P^*(a''|\omega'')\mu(\omega'')u_i(1,\ldots,1;\omega'') \geq 0,$$

which follows by the assumption of this case. For all other $i \in \mathscr{I}$, the incentive constraints remain unaltered.

Case 2. Suppose $\omega'' < \underline{\omega}$.

Case 2.1. If $P^*(a'|\omega')\mu(\omega') \geq P^*(a''|\omega'')\mu(\omega'')$, then define $(A, \hat{P})$ as

$$\hat{P}(a|\omega) = \begin{cases} 0 & \text{if} \quad I(a) = I(a'') \text{ and } \omega = \omega'' \\ P^*(a|\omega) - \frac{\mu(\omega'')}{\mu(\omega')} P^*(a''|\omega'') & \text{if} \quad I(a) = I(a') \text{ and } \omega = \omega' \\ P^*(a|\omega) + P^*(a''|\omega'') & \text{if} \quad I(a) = I(a') \cup I(a'') \text{ and } \omega = \omega'' \\ P^*(a|\omega) + \frac{\mu(\omega'')}{\mu(\omega')} P^*(a''|\omega'') & \text{if} \quad I(a) = I(a') \cap I(a'') \text{ and } \omega = \omega' \\ P^*(a|\omega) & \text{otherwise.} \end{cases}$$

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

$$\mathbb{E}u_i(A, \hat{P}) - \mathbb{E}u_i(A, P^*) = P^*(a''|\omega'')\mu(\omega'')\left[u_i(a' \vee a''; \omega'') - u_i(a'; \omega')\right] \geq 0,$$

which follows by Assumption 1', since $u_i(a' \vee a''; \omega'') - u_i(a'; \omega') \geq u_i(a''; \omega'') - u_i(a' \wedge a''; \omega') = 0$. If $i \in I(a'')$ and $i \notin I(a')$, then

$$\mathbb{E}u_i(A, \hat{P}) - \mathbb{E}u_i(A, P^*) = P^*(a''|\omega'')\mu(\omega'')\left[u_i(a' \vee a''; \omega'') - u_i(a''; \omega'')\right] \geq 0,$$

which follows by Assumption 1. If $i \in I(a') \cap I(a'')$, then

$$\begin{aligned} \mathbb{E}u_i(A, \hat{P}) - \mathbb{E}u_i(A, P^*) = P^*(a''|\omega'')\mu(\omega'')\big[&u_i(a' \vee a''; \omega'') - u_i(a''; \omega'') \\ &+ u_i(a' \wedge a''; \omega') - u_i(a'; \omega')\big] \geq 0, \end{aligned}$$

which follows by Assumption 1'. For all other $i \in \mathscr{I}$, the incentive constraints remain unaltered.

Case 2.2. If instead $P^*(a'|\omega')\mu(\omega') < P^*(a''|\omega'')\mu(\omega'')$, then define $(A, \hat{P})$ as

$$\hat{P}(a|\omega) = \begin{cases} 0 & \text{if} \quad I(a) = I(a') \text{ and } \omega = \omega' \\ P^*(a|\omega) - \frac{\mu(\omega')}{\mu(\omega'')} P^*(a'|\omega') & \text{if} \quad I(a) = I(a'') \text{ and } \omega = \omega'' \\ P^*(a|\omega) + \frac{\mu(\omega')}{\mu(\omega'')} P^*(a'|\omega') & \text{if} \quad I(a) = I(a') \cup I(a'') \text{ and } \omega = \omega'' \\ P^*(a|\omega) + P^*(a'|\omega') & \text{if} \quad I(a) = I(a') \cap I(a'') \text{ and } \omega = \omega' \\ P^*(a|\omega) & \text{otherwise.} \end{cases}$$

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

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = P^*(a'|\omega')\mu(\omega')\big[u_i(a' \vee a'';\omega'') - u_i(a';\omega')\big] \geq 0,$$

which follows by Assumption 1' since $u_i(a' \vee a'';\omega'') - u_i(a';\omega') \geq u_i(a'';\omega'') - u_i(a' \wedge a'';\omega') = 0$. If $i \in I(a'')$ and $i \notin I(a')$, then

$$\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = P^*(a'|\omega')\mu(\omega')\big[u_i(a' \vee a'';\omega'') - u_i(a'';\omega'')\big] \geq 0,$$

which follows by Assumption 1. Finally, if $i \in I(a') \cap I(a'')$, then

$$\begin{aligned}\mathbb{E}u_i(A,\hat{P}) - \mathbb{E}u_i(A,P^*) = P^*(a'|\omega')\mu(\omega')\big[&u_i(a' \vee a'';\omega'') - u_i(a'';\omega'') \\ &+ u_i(a' \wedge a'';\omega') - u_i(a';\omega')\big] \geq 0,\end{aligned}$$

which follows by Assumption 1'. For all other $i \in \mathscr{I}$, the incentive constraints remain unaltered.

In both Cases 1 and 2, since the incentive compatibility condition to follow action recommendation $a_i = 1$ was satisfied under (A,P^*) and is (weakly) greater under $(A,\hat{P})$, following the recommendations remains incentive compatible under $(A,\hat{P})$. In Case 1,

$$\mathbb{E}v(A,\hat{P}) - \mathbb{E}v(A,P^*) = \mu(\omega'')P^*(a''|\omega'')\big[v(1,\dots,1;\omega'') - v(a'';\omega'')\big] \geq 0, \quad (8)$$

which holds by Assumption 2. In Case 2.1 we have

$$\begin{aligned}\mathbb{E}v(A,\hat{P}) - \mathbb{E}v(A,P^*) = \mu(\omega'')P^*(a''|\omega'')\Big(&v(a' \vee a'';\omega'') - v(a'';\omega'') \\ &+ v(a' \wedge a'';\omega') - v(a';\omega')\Big) \geq 0\end{aligned}$$

and in Case 2.2,

$$\begin{aligned}\mathbb{E}v(A,\hat{P}) - \mathbb{E}v(A,P^*) = \mu(\omega')P^*(a'|\omega')\Big(&v(a' \vee a'';\omega'') - v(a'';\omega'') \\ &+ v(a' \wedge a'';\omega') - v(a';\omega')\Big) \geq 0,\end{aligned}$$

which hold by Assumption 2'.

By doing this iteratively, for all $\omega' < \omega''$ such that $P^*(a'|\omega') > 0$, $P^*(a''|\omega'') > 0$ and $I(a') \subsetneq I(a'')$, we obtain (A, P^{**}) such that (i) action 1 is IC for all i; (ii) for any $\omega' < \omega''$, if $P^{**}(a'|\omega') > 0$ and $P^{**}(a''|\omega'') > 0$, then $I(a') \subseteq I(a'')$; and (iii) $\mathbb{E}[v]$ is weakly larger under (A, P^{**}). Hence, (A, P^{**}) must also be optimal. ■

The next lemma shows how to build an information hierarchy from the optimal information structure derived in Lemma 1.

LEMMA 2. By Lemma 1 we know that there is an optimal direct information structure (A, P^{**}) that satisfies the inclusion properties 1 and 2. For all i, define

$$S_i = \{1\} \cup A$$

and $s_i : A \to S_i$ such that

$$s_i(a) = \begin{cases} a_i & \text{if} \quad a_i = 1, \\ a & \text{if} \quad a_i = 0. \end{cases}$$

Consider the indirect information structure $(S, \hat{P})$ such that $S = \prod_i S_i$ and $\hat{P}\big((s_i(a))_i|\omega\big) = P^{**}(a|\omega)$ for all $a \in A$ and $\omega \in \Omega$. $(S, \hat{P})$ is an information hierarchy, a single-meeting scheme, and $\mathbb{E}v(S, \hat{P}) = \mathbb{E}v(A, P^{**})$.

Proof. Consider $(S, \hat{P})$ as defined in the lemma. First we show that $(S, \hat{P})$ is an information hierarchy. Let $\mathbb{S} = \bigcup_{a,\omega: P^{**}(a|\omega)>0} I(a)$. The properties of (A, P^{**}) allows us to order the sets in $\mathbb{S}$ by inclusion, from largest I_1 to smallest $I_{K+1} = \emptyset$, where $K \le n$. Then, let

$$G_k = I_k \setminus I_{k+1}$$

for all $k = 1, \ldots, K$, which implies $G_K = I_K$.

For any k, consider all $i \in G_k$. By the inclusion property of (A, P^{**}) from Lemma 1 and by construction of $(S, \hat{P})$, we have that $s_i(a) = 1$ if and only if $a_j = 1$ for all $j \in \bigcup_{m \ge k} G_m$ and $s_i(a) = a$ if and only if $a_j = 0$ for all $j \in \bigcup_{m \le k} G_m$ Hence, all $i \in G_k$ observe the same signal s_k and are therefore equally informed. The above also implies that:

(i) $\mu_i(a_j = 1|s_k = 1) = 1 = \mu_i(s_m = 1|s_k = 1)$ for all $j \in \bigcup_{m \geq k} G_m$ and $m \geq k$,

(ii) $\mu_i(a'_j|s_k = a) = 1 = \mu_i(s_m(a)|s_k = a)$ for all $a'_j = a_j$, $j \in \bigcup_{m \geq k} G_m$, and $m \geq k$, and

(iii) $\mu_i(a_j = 0|s_k = a) = 1 = \mu_i(s_m = a|s_k = a)$ for all $j \in \bigcup_{m \leq k} G_m$ and $m \leq k$.

By (i) and (ii) we obtain that agent $i \in G_k$ has beliefs such that

$$\mu_i(\omega, s_{-i}|s_k) = \mu_i(\omega, s_{-i}|s_k, s_j) \tag{9}$$

for any s_k and for all $j \in \bigcup_{m \geq k} G_m$. Therefore, every $i \in G_k$ is weakly more informed than every $j \in \cup_{m \geq k} G_m$. However, for any $j \in G_m$ and $m > k$ and for $s_m = 1$ it holds that $\mu_j(s_k|s_m = 1) < 1$ for any s_k. Thus, $j \in \cup_{m > k} G_m$ is not weakly more informed than $i \in G_k$, implying that $i \in G_k$ is strictly more informed than any $j \in \cup_{m > k} G_m$. Hence, $(S, \hat{P})$ is an information hierarchy.

Further, for all $i \in G_k$, all $k = 1, \ldots, K$, and $s_k = a$ for any $a \in A$, it holds that $\mu_i(s_{-i}|s_k) = 1$. Therefore, for each $i \in G_k$ and all $k = 1, \ldots, K$, $\mu_i(s_{-i}|s_k) < 1$ is possible for some s_{-i} only when $s_k = 1$. Hence, $(S, \hat{P})$ is a single-meeting scheme.

Next, it follows by Proposition 1 that since $(S, \hat{P})$ is a single-meeting scheme derived by augmenting (A, P^{**}) in the way specified in the statement of the lemma, it must be that $\mathbb{E}v(S, \hat{P}) \geq \mathbb{E}v(A, P^{**})$. However, since (A, P^{**}) is optimal, it can only be that $\mathbb{E}v(S, \hat{P}) = \mathbb{E}v(A, P^{**})$. ■

LEMMA 3. Take environment $\mathcal{G}$, which satisfies Assumptions 1 and 1', and $(S, \hat{P})$ as defined in Lemma 2. $(S, \hat{P})$ is decentralizable for environment $\mathcal{G}$.

Proof. Take player $i \in G_k$ for some $k = 1, \ldots, K$. Consider first $s_k = 1$. By (i) we know that $\mu_i(s_m = 1|s_k = 1) = 1$ for all $m > k$, so player i would have to transmit $s_m = 1$ for all agents $j \in \bigcup_{m > k} G_m$. Two deviation are possible — deviation in transmission and a simultaneous deviation in both action and transmission. Deviations only in action are not profitable due to $(S, \hat{P})$ being IC.

Take agent $i \in G_k$ with $s_k = 1$ and suppose he deviates in transmission only. His expected utility from truthful transmission is

$$\mathbb{E}\left[u_i\left(((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=1)_{j\in\bigcup_{m\geq k}G_m};\omega\right)\middle|s_k\right]\geq 0 \tag{10}$$

where the inequality follows by IC of $(S,\hat{P})$. By Assumption 1

$$u_i\left(((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=1)_{j\in\bigcup_{m\geq k}G_m};\omega\right)\geq$$
$$u_i\left(((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=1)_{j\in G_k},((a_j'(s_m))_{j\in\bigcup_{m>k}G_m};\omega\right) \tag{11}$$

for any $((a_j'(s_m))_{j\in G_m,m=k+1,\ldots,K}$ and ω. Hence, it holds that

$$\mathbb{E}\left[u_i\left(((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=1)_{j\in\bigcup_{m\geq k}G_m};\omega\right)\middle|s_k\right]\geq$$
$$\mathbb{E}\left[u_i\left(((a_j^*(s_\ell))_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=1)_{j\in G_k},((a_j'(s_m))_{j\in\bigcup_{m>k}G_m};\omega\right)\middle|s_k\right] \tag{12}$$

for any $((a_j'(s_m))_{j\in G_m,m=k+1,\ldots,K}$. Next, consider a deviation in both action and transmission. The only possible deviation to action is to $a_i'=0$ which results in an expected payoff of 0 irrespective of the deviation in transmission. By (10) this is not profitable.

Take agent $i \in G_k$ with $s_k = (a^*)$ and suppose he deviates in transmission only. Then, his expected payoff is 0 in both cases, and so this is not profitable. Next, consider a deviation in both action and transmission. His expected utility from the most profitable deviation is

$$\mathbb{E}\left[u_i\left((a_j^*=0)_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=0)_{j\neq i,j\in G_k},1,(a_j=1)_{j\in\bigcup_{m>k}G_m};\omega\right)\middle|s_k\right]$$

where he tells all players $j \in G_m, m = k+1,\ldots,K$ to play 1 and witches himself to action 1. Since $(S,\hat{P})$ is an optimal information structure, we know that all player $i \in G_k$ who observe signal $s_k = a$ for any $a \in A$ have a strict incentive to player zero. Therefore, it must be that even when the observed signal is $\tilde{s}_k = ((a_j=0)_{j\in\bigcup_{\ell<k}G_\ell},(a_j=0)_{j\neq i,j\in G_k},0,(a_j=1)_{j\in\bigcup_{m>k}G_m})$, which is the signal associated with the most optimistic beliefs about play of opponents in $\bigcup_{m>k}G_m$ and about the state of the world ω (the latter follows by property 2 of Lemma 1), agent $i \in G_k$ has a strict incentive to play 0. That is,

$$\mathbb{E}\left[u_i\left((a_j^*=0)_{j\in\bigcup_{\ell<k}G_\ell},(a_j^*=0)_{j\neq i,j\in G_k},1,(a_j=1)_{j\in\bigcup_{m>k}G_m};\omega\right)\Big|s_k\right]\leq$$
$$\mathbb{E}\left[u_i\left((a_j=0)_{j\in\bigcup_{\ell<k}G_\ell},(a_j=0)_{j\neq i,j\in G_k},1,(a_j=1)_{j\in\bigcup_{m>k}G_m};\omega\right)\Big|\tilde{s}_k\right]<0. \quad (13)$$

Therefore, this deviation is not profitable.

■

Lemmas 1-3 prove Proposition 2. ■

References

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

Ballester, Coralio, Antoni Calvo-Armengol, and Yves Zenou. 2006. "Who's Who in Networks. Wanted: The Key Player." *Econometrica* 74 (5):1403–1417.

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

Best, James and Daniel Quigley. 2018. "Persuasion for the Long-Run." Oxford University WP.

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

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

Glaeser, Edward and Jose Scheinkman. 1999. "Measuring Social Interactions." In *Social Dynamics*, edited by S. Durlauf and P. Young. Cambridge: MIT Press.

Goldstein, I and A. Pauzner. 2005. "Demand-Deposit Contracts and the Probability of Bank Runs." *Journal of Finance* 60 (3):1293–1327.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6):2590–2615.

Mathevet, Laurent, David Pearce, and Ennio Stacchetti. 2018. "Reputation and Information Design." WP, New York University.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2019. "On Information Design in Games." *Journal of Political Economy* :forthcoming.

Meyer, Margaret and Bruno Strulovici. 2012. “Increasing interdependence of multivariate distributions.” *Journal of Economic Theory* 147 (4):1460–1489.

Morris, S. and H. S. Shin. 1998a. “Coordination Risk and the Price of Debt.” *American Economic Review* 88 (3):587–97.

———. 1998b. “Unique equilibrium in a model of self-fulfilling currency attacks.” *The American Economic Review* 88 (3):587–97.

———. 2003. “Global games: theory and applications.” In *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge University Press.

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

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