On Information Design in Games*

Laurent Mathevet
New York University

Jacopo Perego
Yale University

Ina Taneva†
University of Edinburgh

August 29, 2017

Abstract

Information disclosure in games influences behavior by affecting the players' beliefs about the state, as well as their higher-order beliefs. We first characterize the extent to which a designer can manipulate players' beliefs by disclosing information. Building on this, our next results show that all optimal solutions to information design problems are made of an optimal private and of an optimal public component, where the latter comes from concavification. This representation subsumes Kamenica and Gentzkow (2011)'s single-agent result. In an environment where the Revelation Principle fails, and hence direct manipulation of players' beliefs is indispensable, we use our results to compute the optimal solution. In a second example, we illustrate how the private–public decomposition leads to a particularly simple and intuitive resolution of the problem.

Keywords: information design, disclosure, belief manipulation, belief distributions, extremal decomposition, concavification.

JEL Classification: C72, D82, D83.

*This work is a different version of a paper that circulated under the title "Information Design: The Epistemic Approach," first posted on September 27, 2016.

†Special thanks are due to Anne-Katrin Roesler for her insightful comments as a discussant at the 2016 Cowles Foundation conference, and to Philippe Jehiel, Elliot Lipnowski, Stephen Morris, David Pearce, József Sákovics, and Siyang Xiong for their helpful comments and conversations. We are also grateful to Andrew Clausen, Olivier Compte, Jon Eguia, Willemien Kets, Matt Jackson, Qingmin Liu, Efe Ok, Ennio Stacchetti, and Max Stinchcombe, as well as to seminar audiences at Columbia University, Institut d'Anàlisi Econòmica, New York University, Michigan State University, Paris School of Economics, Stony Brook University, University of Cambridge, University of Edinburgh, UT Austin, the SIRE BIC Workshop, the 2016 Decentralization conference, the 2016 Canadian Economic Theory conference, the 2016 Cowles Foundation conference, and the 2016 North American Summer Meeting of the Econometric Society. Taneva acknowledges financial support from ESRC grant ES/N00776X/1.
E-mail: lmath@nyu.edu and jacopo.perego@yale.edu and ina.taneva@ed.ac.uk

1. Introduction

In incomplete-information environments, a designer can affect an agent's behavior by manipulating his payoffs, such as by taxing bad behavior, rewarding efforts, offering insurance, and so on. Alternatively, when the setting allows it, she can affect his behavior by manipulating his beliefs. Information design studies the latter: a designer commits to disclosing information to a group of players so as to induce them to behave in a desired way. The analysis of information design problems, especially in games, has been a focal point of the recent literature.

Under incomplete information, a player chooses an action based, in part, on his beliefs about the uncertain state of the world. Since his choice also depends on the other players' actions, his beliefs about their beliefs about the state also affect his decision, as do his beliefs about their beliefs about his beliefs about the state, and so on. These higher-order beliefs are absent from the one-agent problem, but they are an inevitable part of strategic interactions under incomplete information.

This paper contributes to the foundations of information design in three ways:

– We characterize the feasible distributions of player's beliefs that a designer can induce through her choice of information structure. Information in games influences players' behavior by affecting their beliefs and higher-order beliefs. Thus, information design is ultimately an exercise in belief manipulation, whether it is explicitly modeled as such or solved—if at all possible—in terms of distributions over actions and states. However, not every combination of beliefs can be induced by an information designer with commitment power.

– We characterize the theoretical structure of optimal solutions. Our main theorem and its two-step decomposition show how all optimal solutions can be seen as a combination of basic communication schemes. We show that the minimal common-belief components, defined later, are the basic communication schemes from which *all* others are constructed. This leads to a representation of any optimal information structure as a combination of *optimal* private and *optimal* (purely) public signals, where the latter comes from concavification of a value function.

– We demonstrate the *necessity* of explicit belief manipulation in environments where the Revelation Principle fails. In these problems, harnessing beliefs is essential because action recommendations are insufficient for optimal design. Section 5 gives an example of a class of information design problems that can only be solved by working *directly* with players' higher-order beliefs, and solves it by using the results in this paper.

First, there is a *structural* value to these results, independent of their methodological value. Our approach can be used to decompose an optimal information structure irrespective of the method that was used to compute it. This is important as it allows for all optimal solutions to information design problems to be interpreted in light of a fundamental distinction in information economics — that between private and public information.

Second, there is a methodological value to our results, as they can be used to compute an optimal information structure. The appeal of this methodological application varies with the environment. This is particularly clear when no other approach is available to compute a solution, as in Section 5. Even when other approaches are available, our results may still be appealing as they can lead to a particularly simple and intuitive resolution of the problem. Section 6 introduces an example, called the *Manager's Problem*, that illustrates both the structural and the methodological contribution.[1]

The one-agent problem has been a rich subject of study since Kamenica and Gentzkow (2011) (e.g., Ely, Frankel, and Kamenica (2015), Lipnowski and Mathevet (2015), Kolotilin et al. (2015), etc.). By contrast, the theory of information design in games is not as well understood. Bergemann and Morris (2016) and Taneva (2014) formulate the Myersonian approach to Bayes Nash information design as a linear program.[2] As will become clear, our results develop the "belief-based" approach to information design, which is the multi-player analogue to Kamenica and Gentzkow (2011)'s single-agent formulation. Optimal solutions have been derived in specific environments, as in Vives (1988), Morris and Shin (2002) and Angeletos and Pavan (2007). More recent works study information design in voting games (Alonso and Câmara (2015), Chan et al. (2016)); dynamic bank runs (Ely (2017)); stress testing (Inostroza and Pavan (2017)); auctions (Bergemann, Brooks, and Morris (2017)); contests (Zhang and Zhou (2016)); or focus on public information in games (Laclau and Renou (2016)).

2. The Information Design Problem

Let Θ be a finite set. A (**base**) **game** $G = ((A_i, u_i)_{i\in N}, \mu_0)$ describes a set of players, $N = \{1, \ldots, n\}$, interacting in an environment with uncertain state $\theta \in \Theta$, distributed according to $\mu_0 \in \Delta\Theta$. Every $i \in N$ has finite action set A_i and utility function $u_i : A \times \Theta \to \mathbb{R}$, where $A = \prod_i A_i$ is the set of action profiles.

The designer is an external agent who discloses information about the state to the players, but otherwise does not participate in the strategic interaction. The

[1] Our optimal solution to the Manager's Problem is reminiscent of communications in real world organizations, in which increasingly less information is disclosed to lower ranked employees.

[2] The Myersonian approach relies on revelation arguments, and thus cannot be used in environments where the Revelation Principle fails.

designer's utility function is given by $v : A \times \Theta \to \mathbb{R}$. An **information design environment** is a pair $\langle v, G\rangle$ consisting of a designer's preference and a base game, where the state distribution μ_0 is common knowledge to all parties. Information disclosure is modeled by an **information structure** (S, π), where S_i is player i's finite message space; $S = \prod_i S_i$ is the set of message profiles; and $\pi : \Theta \to \Delta S$ is the information map. In any state θ, the message profile $s = (s_i)$ is drawn according to $\pi(s|\theta)$ and player i observes s_i. Let $S_{-i} = \prod_{j\neq i} S_j$ and assume without loss that, for all i and $s_i \in S_i$, there is $s_{-i} \in S_{-i}$ such that $\sum_\theta \pi(s|\theta) > 0$ (otherwise, delete s_i).

The designer chooses the information structure at a time when she does not know the state θ, about which she also holds prior belief μ_0. Therefore, she commits ex-ante to disclosing information to be learned in the future. One can think of an information structure as an experiment concerning the state, such as an audit, a stress test or a medical analysis. The pair $\mathcal{G} = \langle G, (S, \pi)\rangle$ defines a Bayesian game in which players behave according to **solution concept** $\Sigma(\mathcal{G}) \subseteq \{\sigma : S \to \Delta A\}$. The set of distributions over action profiles and states is the important object that determines all payoffs. Thus, define the **outcome correspondence** associated with Σ as

$$O_\Sigma(\mathcal{G}) := \left\{ \gamma \in \Delta(A \times \Theta) : \exists \sigma \in \Sigma(\mathcal{G}) \text{ s.t. } \gamma(a, \theta) = \sum_s \sigma(a|s)\pi(s|\theta)\mu_0(\theta)\, \forall (a, \theta) \right\}. \tag{1}$$

Assume that $O_\Sigma(\mathcal{G})$ is non-empty and compact for any given $\mathcal{G}$. For a fixed base game G, we just write $\Sigma(S, \pi)$ and $O_\Sigma(S, \pi)$.

In games, it is standard to have multiple outcomes under most solution concepts. Multiplicity gives us the opportunity to model the designer's attitude about selection. Define a **selection rule** to be a function $g : D \subseteq \Delta(\Theta \times A) \mapsto g(D) \in D$. Denote $g^{(S,\pi)} := g(O_\Sigma(S, \pi))$. The best and the worst outcomes are natural selection criteria: an optimistic designer uses the max-rule

$$g^{(S,\pi)} \in \underset{\gamma \in O_\Sigma(S,\pi)}{\operatorname{argmax}} \sum_{a,\theta} \gamma(a, \theta) v(a, \theta), \tag{2}$$

while a pessimistic designer uses the min rule (just replace the argmax in (2) with argmin). Other criteria, such as random choice rules, are also sensible. Under selection rule g, the value to the designer of choosing (S, π) is given as

$$V(S, \pi) := \sum_{a,\theta} g^{(S,\pi)}(a, \theta)\, v(a, \theta). \tag{3}$$

Thus, the information design problem is formulated as $\sup_{(S,\pi)} V(S, \pi)$.

3. Information Design as Belief Manipulation

For single-agent problems, Kamenica and Gentzkow (2011) established that choosing an information structure is equivalent to choosing a *Bayes plausible* distribution over posterior beliefs. For multiple-agent problems, to which form of belief manipulation is information disclosure equivalent? To answer, we start off by defining the space of beliefs (now, hierarchies of) and distributions over those, and then prove the equivalence.

3.1. Distributions of Beliefs

A belief hierarchy t_i for player i is an infinite sequence $(t_i^1, t_i^2, \ldots)$ whose components are coherent[3] beliefs of all orders: $t_i^1 \in \Delta\Theta$ is i's first-order belief; $t_i^2 \in \Delta(\Theta \times (\Delta\Theta)^{n-1})$ is i's second-order belief (i.e., a belief about θ and every j's first-order beliefs); and so on. Let T_i be the set of i's belief hierarchies for all i that only assign positive probabilities to coherent beliefs of the other players. Let $T := \prod_i T_i$ and $T_{-i} := \prod_{j \neq i} T_j$. There exists a homeomorphism $\beta_i^* : T_i \to \Delta(\Theta \times T_{-i})$ for all i, which formalizes common belief in coherency (see Brandenburger and Dekel (1993)).[4] This function describes i's beliefs about (θ, t_{-i}) given his hierarchy t_i,[5] and by construction assigns positive probability only to coherent hierarchies of the other players.

Given a prior and an information structure (S, π), every player i formulates posterior beliefs $\mu_i : S_i \to \Delta(\Theta \times S_{-i})$ by Bayes' rule. When player i receives a message s_i from (S, π), he has belief $\mu_i(s_i)$ about the state and others' messages. Since every player $j \neq i$ has a belief $\text{marg}_\Theta \mu_j(s_j)$ about the state given his own message s_j, i's belief about j's messages s_j gives i a belief about j's belief about the state and so on. By induction, every s_i corresponds to a belief hierarchy $h_i(s_i)$ for player i, and every message profile s corresponds to a profile of belief hierarchies $(h_i(s_i))_i$. Let $h : s \mapsto (h_i(s_i))_i$ be the map that associates to every s the corresponding profile of belief hierarchies. Now, say that an information structure (S, π) **induces** a distribution $\tau \in \Delta T$ over profiles of belief hierarchies, called a **belief-hierarchy distribution**, if

$$\tau(t) = \sum_\theta \pi\big(\{s : h(s) = t\}|\theta\big)\mu_0(\theta) \tag{4}$$

[3] A hierarchy t is coherent if, for all k, beliefs of order k, t_i^k, coincide with all beliefs of lower order, $\{t_i^n\}_{n=1}^{k-1}$, on lower order events. For example, $\text{marg}_\Theta t_i^2 = t_i^1$.

[4] Indeed, a player's belief hierarchies could be coherent, but may assign positive probability to other players' beliefs not being coherent. We want to eliminate this.

[5] When there is no confusion, we write $\beta_i^*(t_{-i}|t_i)$ and $\beta_i^*(\theta|t_i)$ to refer to the marginals.

$\pi(\cdot\|0)$	s_1	s_2
s_1	1	0
s_2	0	0

$\pi(\cdot\|1)$	s_1	s_2
s_1	$\frac{1}{2}$	0
s_2	0	$\frac{1}{2}$

Table 1: A (Public) Information Structure

for all t. For example, the information structure in Table 1 induces $\tau = \frac{3}{4}t_{1/3} + \frac{1}{4}t_1$ when $\mu_0 := \mu_0(\theta = 1) = \frac{1}{2}$, where t_μ is the hierarchy profile in which $\mu := \mu(\theta = 1)$ is commonly believed.[6]

We categorize belief distributions into public and private. This distinction is closely linked to the nature of information that induces those distributions. Later on, it will help us disentangle the value of public versus private disclosure.

Definition 1. A belief-hierarchy distribution τ is **public** if $|\text{supp } \tau| \geq 2$ and for all $t \in \text{supp } \tau$, $t_i^1 = t_j^1$ and $\text{marg}_{T_{-i}} \beta_i^*(\theta, t_{-i}|t_i) = \delta_{t_{-i}}$ (where δ is the Dirac measure). We say τ is **private** if it is not public.

The first part says that players share the same first-order beliefs and this is commonly believed among them. This is the natural translation in terms of beliefs of the standard notion of public information. Notice also that we categorize the degenerate case $|\text{supp}\tau| = 1$ as private. When the support is a singleton this distinction is indeed mostly a matter of semantics; yet the current choice makes our characterization later on more transparent.

3.2. Manipulation

In an information design problem with $\theta \in \{0, 1\}$,

$$\begin{aligned} u_i(a_i, \theta) &= -(a_i - \theta)^2 \quad i = 1, 2 \\ v(a, \theta) &= u_1(a_1, \theta) - u_2(a_2, \theta), \end{aligned}$$

where both "players" care only about matching the state and the designer wants to favor 1 while harming 2, the designer could obtain her maximal payoff of 1, if she could somehow tell the truth to 1 while persuading 2 of the opposite. If this were possible, 1 would be certain that the state is θ, 2 would be certain that the state is $1 - \theta$, and this disagreement would be commonly known. Since Aumann (1976), we have known that Bayesian agents cannot agree to disagree if they have a common prior. Say that $p \in \Delta(\Theta \times T)$ is a **common prior** if

$$p(\theta, t) = \beta_i^*(\theta, t_{-i}|t_i)p(t_i) \tag{5}$$

[6] To see why, note that $\Pr(s_1, s_1) = \frac{3}{4}$, $\Pr(s_2, s_2) = \frac{1}{4}$, and a player i receiving message s_ℓ has beliefs $(2\ell - 1)/3$ that $\theta = 1$ and is certain that j also received s_ℓ.

for all θ, t and i. That is, all players i obtain their belief map β_i^* by Bayesian updating of the same distribution p. Denote by Δ^f the probability measures with finite support. From here, define

$$\mathcal{C} := \Big\{\tau \in \Delta^f T : \exists \text{a common prior}\, p \text{ s.t. } \tau = \text{marg}_T p\Big\} \tag{6}$$

to be the space of **consistent (belief-hierarchy) distributions**. In a consistent distribution, all players' beliefs arise from a common prior that draws every t with the same probability as τ, i.e., $\tau = \text{marg}_T p$. Let p_τ be the unique distribution p in (6) (uniqueness follows from Mertens and Zamir (1985, Proposition 4.5))

Proposition 1. *The following are equivalent:*

(i) $\tau \in \Delta^f T$ *is induced by some* (S, π).

(ii) τ *is consistent and* $\sum_{t_i} \text{marg}_\Theta \beta_i^*(\cdot|t_i)\tau_i(t_i) = \mu_0$ *for some* i.

This characterization disciplines the designer's freedom in shaping players' beliefs, but only to the extent that (i) they are consistent and (ii) that some player's (first-order) beliefs satisfy the Bayes plausibility constraint. In the one-agent case, information disclosure is equivalent to choosing a Bayes plausible distribution over (posterior) beliefs. In games, it is equivalent to choosing a Bayes plausible and *consistent* distribution over (hierarchies of) beliefs. Importantly, it does not matter which player i satisfies Bayes plausibility in (ii), because by consistency, if it is true for one player, then it will hold for all.

Returning to the simple example above, if the designer would like to obtain a payoff of 1 with certainty, she must give full information to player 1, for otherwise she will not get an expected payoff of 0 from him. At the same time, she must fool player 2 all the time. Therefore, to reach the upper bound of 1, it would have to be that $\beta_1^*(\theta = 1, t_2|t_1) = 1$ and $\beta_2^*(\theta = 0, t_1|t_2) = 1$ for some (t_1, t_2), so that (5) cannot hold. So, no information structure can deliver an expected payoff of 1. Given the nature of the problem, Proposition 1 rephrases it as saying that the designer wants players to "maximally disagree" on the value of the state, while their beliefs are consistent. From here, it is a small step to conclude that it is optimal for 1 to know the value of the state and for 2 to know as little as possible. In the language of information disclosure, it is optimal to give full information to 1 and no information to 2.

4. The Theoretical Structure Of Optimal Solutions

In this section, we prove that optimal solutions to information design problems in games can be seen as a patchwork of special consistent distributions, independently of the method producing those solutions. An important consequence is that all optimal solutions consists of an optimal private and of an optimal public component, where the latter comes from concavification.

4.1. Assumptions

One advantage of our approach is that it can handle various selection rules and solution concepts, provided the following assumptions hold:

(Linear Selection). Assume g is linear.[7] Linearity of g is a natural assumption that demands that the selection criterion be invariant to the subsets of distributions to which it is applied. The best and the worst outcomes, defined in (2), are linear selection criteria. Appendix B provides further detail about selection in the space of belief-hierarchy distributions.

(Invariant Solution). The results below hold for all solution concepts Σ that satisfy Assumption 1 in Appendix B. For exposition reasons, Σ will be limited to BNE or forms of Interim Correlated Rationalizability (Dekel, Fudenberg, and Morris (2007)) in the main text. The definition of $\Sigma = \text{BNE}$ is standard: every i uses a strategy σ_i such that

$$\operatorname{supp}\sigma_i(s_i) \subseteq \operatorname*{argmax}_{a_i} \sum_{s_{-i},\theta} u_i(a_i, \sigma_{-i}(s_{-i}), \theta)\mu_i(\theta, s_{-i}|s_i)$$

for all s_i and i, where payoffs are extended to mixed actions by linearity.

4.2. Representations

When seen as belief manipulation, information design exhibits a convex structure that allows the designer to induce any belief distribution $\tau = \alpha\tau' + (1-\alpha)\tau''$,[8] provided that τ' and τ'' are consistent and τ is Bayes plausible. In particular, this is true even if τ' and τ'' are themselves *not* Bayes plausible. In technical terms, $\mathcal{C}$ is convex; moreover, it admits extreme points.[9] In the tradition of extremal repre-

[7] For all D', D'' and $0 \le \alpha \le 1$, $g(\alpha D' + (1-\alpha)D'') = \alpha g(D') + (1-\alpha)g(D'')$.

[8] Think of τ'' as a probability distribution on $\{\tau, \tau'\}$.

[9] Recall that an extreme point of $\mathcal{C}$ is a point $\tau \in \mathcal{C}$ with the property that if $\tau = \alpha\tau' + (1-\alpha)\tau''$, given $\tau', \tau'' \in \mathcal{C}$ and $\alpha \in [0,1]$, then $\tau' = \tau$ or $\tau'' = \tau$.

sentation theorems (as in the Minkowski–Caratheodory theorem, the Krein-Milman theorem, Choquet's integral representation theorem, and so on), the designer thus can generate any Bayes plausible distribution of beliefs by "mixing" over extreme points. Importantly, these extreme points have a useful characterization: they are the minimal consistent distributions (Lemma 2). A consistent distribution τ ($\in \mathcal{C}$) is **minimal** if there is no $\tau' \in \mathcal{C}$ such that $\operatorname{supp}\tau' \subsetneq \operatorname{supp}\tau$. The set of all minimal consistent distributions is denoted by $\mathcal{C}^{\mathrm{M}}$,[10] and is nonempty by basic inclusion arguments. Moreover, based on Definition 1 it is easy to show that minimal consistent distributions are always private.

Owing to their mathematical status of extreme points, the minimal consistent distributions correspond to the basic communication schemes at a given distribution of states, from which *all* others are constructed. In the one-agent case, they are (one-to-one to) the agent's posterior beliefs. The results below formalize their special role in optimal design.[11]

Given selection rule g, write the designer's ex ante expected payoff as

$$w : \tau \mapsto \sum_{\theta,t} g^{\tau}(a,\theta)\, v(a,\theta), \tag{7}$$

where $g^{\tau} \in \Delta(A \times \Theta)$ is the selected outcome in the Bayesian game $\langle G, p_{\tau}\rangle$.

Theorem 1 (Representation Theorem)**.** *The designer's maximization problem can be represented as*

$$\begin{aligned} \sup_{(S,\pi)} V(S,\pi) = \quad & \sup_{\lambda \in \Delta^f(\mathcal{C}^M)} \quad \sum_{e \in \mathcal{C}^M} w(e)\lambda(e) \\ & \text{subject to} \quad \sum_{e \in \mathcal{C}^M} marg_{\Theta} p_e\, \lambda(e) = \mu_0. \end{aligned} \tag{8}$$

Corollary 1 (Private–Public Information Decomposition)**.** *Fix an environment* $\langle v, G\rangle$*. For any* $\mu \in \Delta\Theta$*, let*

$$w^*(\mu) := \sup_{e \in \mathcal{C}^M : marg_{\Theta} p_e = \mu} w(e). \tag{9}$$

Then, the designer's maximization problem can be represented as

$$\begin{aligned} \sup_{(S,\pi)} V(S,\pi) = \quad & \sup_{\lambda \in \Delta^f \Delta\Theta} \quad \sum_{\operatorname{supp}\lambda} w^*(\mu)\lambda(\mu) \\ & \text{subject to} \quad \sum_{\operatorname{supp}\lambda} \mu\lambda(\mu) = \mu_0. \end{aligned} \tag{10}$$

[10]Minimal belief subspaces appear in different contexts in Heifetz and Neeman (2006), Barelli (2009), and Yildiz (2015).

[11]We further illustrate the notion of minimal distribution in Appendix F by characterizing minimal distributions for public and conditionally independent information.

From the representation theorem, the designer maximizes her expected utility *as if* she were optimally mixing over minimal consistent distributions, subject to posterior beliefs averaging to μ_0 across those distributions. Every minimal distribution induces a Bayesian game and leads to an outcome for which the designer receives some expected utility. Every minimal distribution also induces a distribution over states, $\mathrm{marg}_\Theta p_e$, and the further it is from μ_0, the "costlier" it is for the designer to use it. In this sense, the constraint in (8) can be seen as a form of budget constraint.

The corollary decomposes the representation theorem into two steps. First, there is a **maximization within**—given by (9)—that takes place among all the minimal distributions with marginal μ and for all μ. All minimal distributions with the same μ contribute equally toward the Bayes plausibility constraint; hence, the designer should choose the best one among them, i.e. the one that gives the highest value of $w(e)$. Most importantly, maximization within delivers the optimal value of private information, which takes the form of a (private) value function $w^* : \mu \mapsto \mathbb{R}$. Given a state distribution μ, $w^*(\mu)$ is the value to the designer if she was constrained to use exclusively private information. The possibility to identify the optimal private information within a solution comes from the fact that all minimal distributions represent private information, as defined previously. Indeed, the process of extremization reveals minimal distributions by breaking up any belief-hierarchy distribution into the smallest common knowledge components (i.e., into distributions e such that, for any belief hierarchy in $\mathrm{supp}\, e$, it is common knowledge among the players that their beliefs are in $\mathrm{supp}\, e$). Therefore, the process of extremization "distills" all the public information out of belief-hierarchy distributions, only leaving private information distributions, namely the minimal ones.

Second, there is a **maximization between** that concavifies the value function, thereby optimally randomizing between the minimal distributions that are solutions to maximizations within. This step is akin to a public signal λ that "sends" all players to different minimal distributions e and thus makes it commonly known. From standard arguments, as in Rockafellar (1970, p.36), the rhs of (10) is a characterization of the **concave envelope** of w^*, defined as $(\mathrm{cav}\, w^*)(\mu) = \inf\{g(\mu) : g \text{ concave and } g \geq w^*\}$. Hence, the corollary delivers a concave-envelope characterization of optimal design. In the one-agent case, $\{e \in \mathcal{C}^{\mathrm{M}} \text{ s.t. } \mathrm{marg}_\Theta p_e = \mu\} = \{\mu\}$, hence $w^* = w$ in (9) and the theorem comes down to maximization between.

The methodological contribution of our results is especially clear when no other approach is available to compute a solution (see Section 5). In general, however, the space of relevant belief hierarchy distributions can be very large and the task of maximizing over it complicated. Nevertheless, the structural contribution of our results remains: any optimal information structure can be viewed as a combination of optimal public and optimal private information—in a sense made precise by the corollary. The tradeoff between private and public information is fundamental in information economics with multiple agents, hence this structural decomposition of optimal solutions belongs to the fundamentals of information design. Indeed, regardless of the method used to derive an optimal information structure, it can be decomposed *ex post* by converting it into a belief distribution and then extremizing it. One could also argue that this decomposed version is more intuitive than the original, more parsimonious information structure, as it employs public and private signals separately. We illustrate these claims at the end of Section 6.

5. The Necessity To Manipulate Beliefs

In single-agent information design, the Revelation Principle applies very generally and makes action recommendations a close substitute for explicit belief manipulation. In information design in games, this is no longer true. As this section illustrates, there are important environments in which the Revelation Principle fails and explicit belief manipulation becomes *essential*. To be clear, the solution concept will *not* be responsible for the failure of the Revelation Principle in this section. Rather, it results from using *min* instead of *max* as a selection rule. *Max* lets the designer choose her favorite solution outcome, while *min* requires all outcomes to provide at least a given payoff for the designer.[12] In these problems, our results can be used to compute an optimal solution.

Since Rubinstein (1989), we have known that strategic behavior in games of incomplete information can depend crucially on the tail of the belief hierarchies. In a design context, we may find it unsettling to induce the desired behavior by relying crucially on information acting on beliefs of very high order. In the example below, the solution concept reflects the designer's intention to disclose information that is robust to misspecification of all beliefs above a given level. We focus on the classic global game of investment decisions of Carlsson and van Damme (1993) and Morris and Shin (2003), but the same analysis would extend to other games, in particular to other global games.

[12]This distinction between one vs. many equilibria is reminiscent of the distinction between mechanism design and implementation theory. In the latter, the Revelation Principle is known to be insufficient. In one-agent information design, there is little difference between using min vs. max, because the multiplicity of outcomes comes from indifferences. Thus, provided the designer is satisfied with ϵ-optimality, her optimal value is nearly unchanged and achievable by action recommendations.

5.1. Investment Game with Bounded Depths of Reasoning

Two players decide whether or not to invest, $\{I, N\}$, given uncertain state $\theta \in \{-1, 2\}$ and state distribution $\mu_0 = \text{Prob}(\theta = 2)$. The payoffs of the interaction are summarized in Table 2.

(u_1, u_2)	I	N
I	θ, θ	$\theta - 1, 0$
N	$0, \theta - 1$	$0, 0$

TABLE 2: The Investment Game.

To study bounded depths of reasoning, we use finitely many iterations of the ICR algorithm (Dekel, Fudenberg, and Morris (2007)) and resolve any remaining multiplicity via the selection criterion. The resulting solution concept ICR_k, formally defined below, is a natural choice in the context of information design when the designer is uncertain about the players' strategic sophistication. If the designer does not want to rely on the agents' abilities to reason with beliefs of orders higher than k, she may want to select an information structure that is *robust* to potential mistakes made using those higher-order beliefs. This solution concept, when paired with the min selection criterion, guarantees the designer a minimal payoff level, even if the agents turn out to be more sophisticated or coordinate on the worst possible outcome for the designer.

Given an information structure (S, π), let $R_{i,0}(s_i) := A_i$ for all $i \in N$ and $s_i \in T_i$. Then, for all $k \in \mathbb{N}$, define

$$R_{i,k}(s_i) := \left\{ a_i \in A_i : \begin{array}{l} \text{there exists } \nu \in \Delta(S_{-i} \times \Theta \times A_{-i}) \text{ s.t.:} \\ (1)\ \nu(s_{-i}, \theta, a_{-i}) > 0 \Rightarrow a_j \in R_{j,k-1}(s_j)\, \forall j \neq i \\ (2)\ a_i \in \text{argmax}_{a_i'} \sum_{s_{-i},\theta,a_{-i}} \nu(s_{-i}, \theta, a_{-i}) u_i(a_i', a_{-i}, \theta) \\ (3)\ \sum_{a_{-i}} \nu(s_{-i}, \theta, a_{-i}) = \mu_i(s_{-i}, \theta | s_i). \end{array} \right\}$$

ICR profiles at s are given by $R(s) = \prod_i R_i(s_i)$, where $R_i(s_i) := \cap_{k=0}^{\infty} R_{i,k}(s_i)$. Our choice of solution concept is $\text{ICR}_k := \prod_i R_{i,k}$ because $R_{i,k}$ involves only beliefs of up to order k. That is, when the above procedure is stopped at finite k, we are left with the set of predictions that correspond to actions that a type s_i, who can reason only through k-many levels, cannot rationally eliminate. In a recent paper, Germano, Weinstein, and Zuazo-Garin (2016) provide epistemic foundations for ICR_k. In particular, ICR_2 corresponds to common 1-belief of rationality.

Suppose that the designer would like to persuade both players to invest irrespective of the state: $v(a_1, a_2) = \mathbb{1}(a_1 = I) + \mathbb{1}(a_2 = I)$. Assume also that the designer adopts a robust approach to the problem in that she evaluates her payoffs at the *worst* k-rationalizable profile when there are many. Define $g : 2^{A_i} \to A_i$ as

$$g(A_i') = \begin{cases} \mathrm{I} & \text{if } A_i' = \{\mathrm{I}\}, \\ \mathrm{N} & \text{otherwise} \end{cases}$$

and then define the outcome correspondence for any $k \in \mathbb{N}$

$$O^k_{\mathrm{ICR}}(S,\pi) := \left\{ \gamma \in \Delta(A \times \Theta) : \gamma(a,\theta) := \sum_{s:a=\prod_i g(R_{i,k}(s_i))} \pi(s|\theta)\mu_0(\theta)\ \forall (a,\theta) \right\}. \quad (11)$$

5.2. Optimal Design

For simplicity, let us assume $k = 2$. For any i, let $\mu_i = \mathrm{Prob}_i(\theta = 2)$ be i's first-order belief, and notice that $a_i = I$ is dominant when $\mu_i > \frac{2}{3}$. Let $\lambda_i = \mathrm{Prob}_i(\mu_{-i} > \frac{2}{3})$ be i's second-order belief. Thus, we write

$$g(R_{i,2}(\cdot)) = \begin{cases} \mathrm{I} & \text{if } 3\mu_i - 2 + \lambda_i > 0 \\ \mathrm{N} & \text{otherwise.} \end{cases}$$

A player invests either if he is optimistic enough about the state (large μ_i) or optimistic enough that the other player is optimistic enough (large λ_i).

Maximization Within. For $\mu > \frac{2}{3}$, both players are optimistic enough to invest under no information, hence a public signal announcing μ—which is a minimal distribution of dimension 1×1—achieves the designer's maximal payoff. For $\mu \leq \frac{2}{3}$, the largest minimal distributions that we need to consider are of dimension 2×2: every player must have at least one first-order belief strictly above $\frac{2}{3}$ (so that the designer can leverage the second-order beliefs, i.e. $\lambda_i \geq 0$) and at least one below (so that Bayes plausibility is preserved). In the distribution below, let $\mu_1'' = \mu_2'' = 2/3 + \varepsilon$.[13]

e_μ	(μ_2', λ_2')	(μ_2'', λ_2'')
(μ_1', λ_1')	A	B
(μ_1'', λ_1'')	C	$1 - A - B - C$

From here, we try to achieve the maximal payoff—i.e., investment by both players with probability one—for as many priors $\mu \leq \frac{2}{3}$ as possible. We solve the following

[13]Here, $\varepsilon > 0$ is an arbitrarily small number needed because, in this example, the min-criterion will make $w(e^*_\mu)$ not upper-semicontinuous.

system in Appendix D:

$$\begin{cases} 3\mu_i' - 2 + \lambda_i' > 0 \ \ \forall i \\ \mu_i'' = \frac{2}{3} + \varepsilon \ \ \forall i \\ e_\mu \text{ is consistent,} \end{cases}$$

and obtain

e_μ^{**}	$\left(\frac{(2+3\varepsilon)\mu}{4-3\mu+6\varepsilon}, \frac{3\mu}{4-3\mu+6\varepsilon}\right)$	$\left(\frac{2}{3}+\varepsilon, 0\right)$
$\left(\frac{(2+3\varepsilon)\mu}{4-3\mu+6\varepsilon}, \frac{3\mu}{4-3\mu+6\varepsilon}\right)$	$1 - \frac{3\mu}{2+3\varepsilon}$	$\frac{3\mu}{4+6\varepsilon}$
$\left(\frac{2}{3}+\varepsilon, 0\right)$	$\frac{3\mu}{4+6\varepsilon}$	0

The solution $e_\mu^{**} \in \mathcal{C}$ is a belief-hierarchy distribution with the property that, for all $\mu > 8/15$, the strategy profile (I, I) is the only one that is ICR_2-rationalizable.[14] Since it yields a payoff of 2, it is therefore optimal for the designer to use e_μ^{**} in this range.

Remark. Under a direct approach based on the Revelation Principle, every player is given an action recommendation instead of the message inducing that action. In the above, for $\mu > 8/15$, the players would therefore always be told to invest, which is completely uninformative. While (I, I) still belongs to ICR_2 under no information, so does (NI, NI). This is why a direct approach fails to guarantee uniqueness, while (I, I) is *uniquely* rationalizable in e_μ^{**}.

For $\mu \leq 8/15$, the designer can no longer achieve investment by both players. She will therefore focus on maximizing the likelihood of one of the players choosing to invest. She can achieve that by talking privately to one of the players, which implies that she will use minimal distributions of dimension 1×2. For a given μ, any 1×2 minimal distribution can be parameterized as:[15]

e_μ	(μ_2', δ_{μ_1})	$(\mu_2'', \delta_{\mu_1})$
(μ_1, λ_1)	$1 - p$	p

In order to foster investment, the designer needs to maximize Player 1's second-order beliefs, $\lambda_1 = \text{Prob}(\mu_2 > \frac{2}{3})$. This in turn requires $\mu_2'' > \frac{2}{3}$. Moreover, the more likely μ_2'' is, while Bayes plausibility is satisfied, the larger λ_1 is. Hence, the designer will set $\mu_2' = 0$ and $\mu_2'' = \frac{2}{3} + \varepsilon$, with arbitrarily small $\varepsilon > 0$. Putting everything together, for $\mu \leq \frac{8}{15}$, the optimal private minimal distribution must be

[14] This is because e_μ^{**} induces either $\mu_i > \frac{2}{3}$, making $a_i = I$ dominant for player i, or a combination (μ_i, λ_i) of first and second-order belief such that $3\mu_i - 2 + \lambda_i > 0$.

[15] By symmetry, it is without loss to talk only to Player 2.

e^*_μ	$(0,0)$	$(\frac{2}{3}+\varepsilon, 0)$
$(\mu, \frac{3\mu}{2+3\varepsilon})$	$1-\frac{3\mu}{2+3\varepsilon}$	$\frac{3\mu}{2+3\varepsilon}$

Under e^*_μ, player 1 chooses I with certainty as long as

$$3\mu_1 - 2 + \lambda_1 = 3\mu - 2 + \frac{3\mu}{2+3\varepsilon} > 0 \quad \Leftrightarrow \quad \mu > \frac{4}{9} + \varepsilon',$$

where $\varepsilon' = \frac{2\varepsilon}{9(1+\varepsilon)}$, and chooses NI otherwise, while player 2 chooses I with probability $\frac{3\mu}{2+3\varepsilon}$. Hence,

$$w^*(\mu) = \sup_\varepsilon w(e^*_\mu) = \mathbb{1}\left(\mu \geq \frac{4}{9}\right) + \frac{3}{2}\mu.$$

Hence, maximization within gives the designer an expected payoff of

$$w^*(\mu) = \begin{cases} 2 & \text{if } \mu > 8/15 \\ \mathbb{1}\left(\mu \geq \frac{4}{9}\right) + \frac{3}{2}\mu & \text{otherwise,} \end{cases}$$

which is depicted as the dashed graph in Figure 1.

Maximization Between. Concavification completes the solution. For all $\mu_0 \leq \frac{8}{15}$, the designer puts probability $1 - \frac{15}{8+15\varepsilon'}\mu_0$ on e^*_0 and $\frac{15}{8+15\varepsilon'}\mu_0$ on $e^{**}_{\frac{8}{15}+\varepsilon'}$, giving her an expected payoff of $\frac{30}{8+15\varepsilon'}\mu_0$. The optimal information structure that corresponds to this is

$\pi(\cdot \mid \theta = -1)$	s'_2	s''_2
s'_1	$\frac{2-3\mu+3\varepsilon}{(2+3\varepsilon)(1-\mu)}$	$\frac{(1-3\varepsilon)\mu}{(4+6\varepsilon)(1-\mu)}$
s''_1	$\frac{(1-3\varepsilon)\mu}{(4+6\varepsilon)(1-\mu)}$	0

$\pi(\cdot \mid \theta = 2)$	s'_2	s''_2
s'_1	0	$\frac{1}{2}$
s''_1	$\frac{1}{2}$	0

By manipulating beliefs, the designer decreases the lowest prior μ_0 at which (I, I) can be uniquely rationalized from $\frac{2}{3}$ to $\frac{8}{15}$. She achieves her highest possible payoff for all $\mu_0 \in [\frac{8}{15}, \frac{2}{3}]$ by using 2×2 minimal distributions and informing both players. In the optimal information structure for these priors, each player receives one of two signals, good or bad. When a player receives the good signal, he believes that the state is high with probability $2/3 + \varepsilon$, and so he invests based on his first-order beliefs. When a player receives the bad signal, he also invests, but because he believes that the other player is likely enough to have received the good signal. In the latter, private information to player i fosters investment by player j via j's second-order beliefs, which is a form of *bandwagon effect*. Under ICR_2, this effect

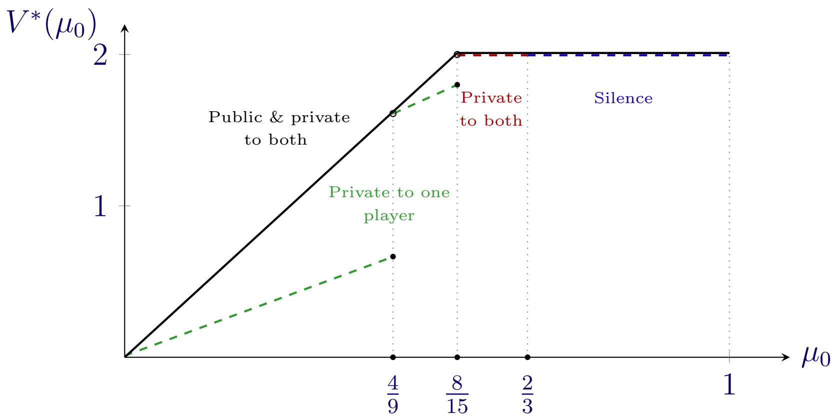

Figure 1: Value of maximization *within* (dashed) and *between* (solid) without constraint.

occurs via second-order beliefs only, but as players become more sophisticated, it extends to higher-order beliefs, giving more flexibility to incentivize investment.[16] Notice that none of the above would be possible under public information, since $\lambda_i > 0$ if and only if $\mu_i > 2/3$.

6. Structure of Optimal Solutions

In this section, we study a simple departure from the single-agent case that still captures a strategic interaction. Our two-step concavification approach separates the optimal solution into private and public signals in a way that is not obvious when using linear programming techniques. Furthermore, the simple framework presented below can serve as a blueprint for more general problems in which interactions between (groups of) players follow a similar structure. In these games, a first group of players cares only about each other's actions and the state; a second group cares only about each other's actions and those of the group above them, but not about the state or any other group's actions; and so on. These structures can be used to model many interesting interactions such as organizational design within firms, technology adoption, and financial markets.

[16]Indeed, we conjecture that the threshold μ above which the designer can obtain the maximal payoff with certainty is decreasing in k. This implies that the designer's expected payoff in this environment is increasing in the strategic sophistication of the players.

6.1. The Manager's Problem

Consider a situation in which a manager is in charge of two employees collaborating on a project: a supervisor (P) and a worker (W). Suppose that the project can be either easy ($\theta = 0$) or hard ($\theta = 1$), distributed according to $\mu_0 := \mu_0(\theta = 1)$, which is common knowledge. The supervisor and the worker simultaneously choose an effort level. The supervisor's choice can be interpreted as deciding whether to monitor the worker ($a_P = 1$) or not monitor him ($a_P = 0$). The worker can choose to either put forth high effort ($a_W = 1$) or low effort ($a_W = 0$).

The strategic interaction between the two employees is summarized in Table 3. On the one hand, the supervisor is only interested in monitoring the worker if the

$\theta = 0$	$a_W = 0$	$a_W = 1$
$a_P = 0$	$1, 1$	$1, 0$
$a_P = 1$	$0, 0$	$0, 1$

$\theta = 1$	$a_W = 0$	$a_W = 1$
$a_P = 0$	$0, 1$	$0, 0$
$a_P = 1$	$1, 0$	$1, 1$

Table 3: Game between supervisor (row) and worker (column).

project is hard, regardless of the worker's action choice. On the other hand, the worker only wants to exert high effort if he is being monitored, regardless of the state.

The manager decides on what information to disclose to her subordinates in a way that maximizes the expected value of her objective. We restrict attention to state-independent objectives $v : A \to \mathbb{R}$. The solution concept is $\Sigma = \text{BNE}$ and the selection rule is max so when there is multiplicity, players are assumed to coordinate on the manager's preferred equilibrium.

6.2. Solution

Our representation theorem prescribes maximizing first over minimal belief-hierarchy distributions, and then over distributions thereof. However, this problem can be daunting if we do not narrow down the set of minimal distributions to consider.[17] We must rely on the specific structure of the problem to identify the *relevant* minimal distributions over which to perform the maximization within. Using the structure of the Manager's Problem, we conclude that:

(i) without loss, we can restrict attention to distributions over first-order beliefs for the supervisor ($\mu_1 \in \Delta\Theta$) and second-order beliefs for the worker ($\lambda_2 \in$

[17] In Proposition 9 in Appendix F, we show that for every minimal distribution e, there exists an information design problem $\langle v, G\rangle$ such that e is uniquely optimal. This means that, a priori, no minimal distribution e can be discarded. This is no longer true once we fix an environment $\langle v, G\rangle$.

$\Delta\Delta\Theta$). This is proven in Propositions 5 and 6 in Appendix E. Therefore, the manager's maximization can take place over distributions η in $\mathcal{A} := \Delta^f(\Delta\Theta \times \Delta\Delta\Theta)$.[18] We denote $\eta_1 := \text{marg}_{\mu_1}\eta$ and $\eta_2 := \text{marg}_{\lambda_2}\eta$.

(ii) by an argument similar to the revelation principle, the optimal distribution will be of dimension 2×2 or smaller. That is, at most two different beliefs per player need to be induced in order to generate all possible equilibrium outcomes.

From (i) and (ii), it is not hard to show that every consistent η of dimension 2×2 can be generated as a convex combination of two smaller distributions that are in fact minimal. We thus identify two classes of *minimal consistent* distributions. The first class consists of distributions of dimension 1×1: one first-order belief μ_1 for the supervisor and one second-order belief $\lambda_2(\mu_1) = 1$ for the worker, where the latter satisfies consistency. The second class is generated via a private signal to the supervisor and the distributions in it are of dimension 2×1. This corresponds to two first-order beliefs, μ_1' and μ_1'', and one second-order belief, which by consistency needs to satisfy $\lambda_2(\mu_1') = \eta_1(\mu_1')$.

The manager's information design problem can be thus decomposed into *maximization within*

$$\begin{aligned} w^*(\mu) \quad &:= \quad \max_{\lambda_2 \in \Delta\Delta\Theta} \sum_{\text{supp}\,\lambda_2} v\big(a_1^*(\mu_1), a_2^*(\lambda_2)\big)\lambda_2(\mu_1) \\ &\qquad \text{subject to} \sum_{\text{supp}\,\lambda_2} \mu_1\lambda_2(\mu_1) = \mu \end{aligned} \tag{12}$$

and *maximization between*

$$\begin{aligned} V^*(\mu_0) \quad &:= \quad \max_{\eta_2 \in \Delta\Theta} \sum_{\text{supp}\,\eta_2} w^*(\mu)\eta_2(\mu) \\ &\qquad \text{subject to} \sum_{\text{supp}\,\eta_2} \mu\,\eta_2(\mu) = \mu_0. \end{aligned} \tag{13}$$

In the maximization within, the manager informs the supervisor optimally, thereby choosing the distribution of his first-order beliefs for any state distribution μ. By consistency, this distribution needs to give μ on average and also pins down the second-order beliefs of the worker.[19] In other words, this step is done by optimizing over the minimal consistent distributions of dimension 1×1 or 2×1 for each μ. In the maximization between, the manager then chooses the optimal randomization

[18] In what follows, we refer to η as a belief-hierarchy distribution, although technically we mean a belief-hierarchy distribution with marginal η over (μ_1, λ_2).

[19] Maximization within is not a concave envelope because λ_2 enters the objective function.

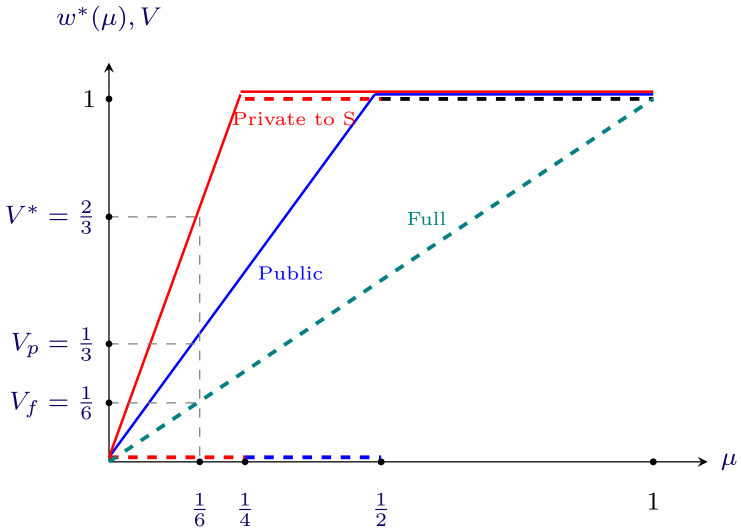

Figure 2: Value of information.

over the minimal distributions from the first step. This is equivalent to choosing optimally the distribution over the worker's second-order beliefs, η_2, which can be also interpreted as a distribution over the μ's of minimal distributions and therefore needs to satisfy Bayes plausibility with respect to the common prior μ_0.

To make things easier to illustrate, consider a manager who is interested only in the worker's action: $v(a, \theta) = a_W$.[20] The worker's equilibrium action is a function of his second-order beliefs. Denoting $\tilde{\lambda}_2 := \lambda_2(\{\mu_1 \geq 1/2\})$,

$$a_W^*(\tilde{\lambda}_2) = \begin{cases} 0 & \text{if } \tilde{\lambda}_2 < 1/2 \\ 1 & \text{if } \tilde{\lambda}_2 \geq 1/2. \end{cases}$$

Maximization within proceeds as follows:

- for $\mu \geq \frac{1}{2}$ it is optimal to not reveal any information and let both agents operate under the common prior. This translates into a minimal distribution of dimension 1×1 with $\mu_1 = \mu$ and $\lambda_2(\mu_1) = 1 = \tilde{\lambda}_2$, which gives $w^*(\mu) = 1$.

- for $\mu \in [\frac{1}{4}, \frac{1}{2})$, no minimal distribution of dimension 1×1 can achieve an expected payoff of 1. However, by informing the supervisor privately and sending him either to $\mu_1' = 0$ or to $\mu_1'' = \frac{1}{2}$, the manager can induce $\tilde{\lambda}_2 \geq \frac{1}{2}$ and hence $w^*(\mu) = 1$. Thus, the optimal minimal distribution is of dimension 2×1 with $\mu_1' = 0$, $\mu_1'' = \frac{1}{2}$, and $(1 - \lambda_2(\mu_1''))\mu_1' + \lambda_2(\mu_1'')\mu_1'' = \mu$. With these minimal distributions it is possible to obtain $\lambda_2(\mu_1'') = \tilde{\lambda}_2 \geq \frac{1}{2}$ as long as $\mu \geq \frac{1}{4}$, which ensures the above consistency equality can be satisfied.

[20]This type of objective function is reasonable when the worker is the main productive unit, as in the three-tier hierarchy model of Tirole (1986).

- for $\mu < \frac{1}{4}$ it must be that $\tilde{\lambda}_2 < \frac{1}{2}$ in any consistent minimal distribution and hence $w^*(\mu) = 0$. Therefore, the designer is indifferent between all consistent minimal distributions, and in particular all those of dimension 1×1 with $\mu_1 = \mu$ and $\lambda_2(\mu_1) = 1$.

Concavifying the thus constructed $w^*(\mu)$ (dashed red lines in Firgure 2), gives us the value that can be achieved by the designer when using both public and private signals optimally (solid red graph).

Consider a state distribution $\mu_0 = \frac{1}{6}$. Then the overall optimal distribution is constructed as follows: with probability $\frac{1}{3}$ send both players to a minimal distribution $e^*_{\mu=0}$ with $\mu = \mu_1 = \tilde{\lambda}_2 = 0$, and with probability $\frac{2}{3}$ send them to a minimal distribution with $\mu = \frac{1}{4}$, which involves a private signal to the supervisor that splits his first-order beliefs into $\mu_1' = 0$ and $\mu_1'' = \frac{1}{2}$, with $\tilde{\lambda}_2 = \frac{1}{2}$ by consistency. Notice that the Bayes plausibility requirement with respect to μ_0 is also satisfied: $0 \times \frac{1}{3} + \frac{1}{4} \times \frac{2}{3} = \frac{1}{6}$.

The overall optimal information design gives the manager an expected payoff of $V^*(\mu_0) = \frac{2}{3}$. As benchmark cases, we have also plotted the values for the case of public information ($V_p = \frac{1}{3}$) and full information ($V_f = \frac{1}{6}$). On the other hand, releasing no information at all will result in a payoff of 0 with certainty.

6.3. Structural Perspective

The information structure that induces the above optimal randomization over minimal distributions can be represented as an optimal public signal:

$\pi(\cdot \mid \theta = 0)$	t_0	t_1
t_0	$2/5$	0
t_1	0	$3/5$

$\pi(\cdot \mid \theta = 1)$	t_0	t_1
t_0	0	0
t_1	0	1

Table 4: Optimal Public Signal

and an optimal private signal which is sent only conditional upon a realization t_1 of the public signal:

In the above tables, the rows represent signal realizations observed by the supervisor, while the columns are those observed by the worker. Hence, t_0 and t_1 are publicly observed by both agents, while q_0 and q_1 are privately observed only by the supervisor. For example, when both agents observe t_0, they both commonly believe that the project is easy. On the other hand, when they both observe t_1, the worker

$\pi(\cdot\|t_1, \theta = 0)$	q_0	q_1
q_0	$2/3$	0
q_1	$1/3$	0

$\pi(\cdot\|t_1, \theta = 1)$	q_0	q_1
q_0	0	0
q_1	1	0

Table 5: Optimal Private Signal

knows that the supervisor will observe additional signals (q_0 or q_1) that will inform him better about the state.

Another optimal information structure can be computed directly, in "one block," using the linear program of the BCE approach: denoted by $\tilde{\pi}$, it is given in Table 6. While $\tilde{\pi}$ achieves the same equilibrium and optimal value for the designer, it amalgamates the public and private communication channels. However, we can apply our approach to this readily computed optimal information structure: first derive the belief-hierarchy distribution induced by $\tilde{\pi}$ (over first- and second-order beliefs); second, extremize it into minimal distributions, which also reveals the coefficients of the convex combination that merges them; finally, convert those into signals to recover Tables 4 and 5. By the Representation Theorem and its corollary, the private and the public signals so derived must be optimal. In this perspective, although our approach was not used to derive the optimal information structure in the first place, it is used to uncover its fundamental structure. Not only is this decomposition intuitive and practically instructive — regarding *how* the manager can communicate with his subordinates — but it also speaks to the general structure of solutions to information design problems by disentangling the public and the private communication channels.

$\tilde{\pi}(\cdot\|\theta = 0)$	s_0	s_1
s_0	$2/5$	$2/5$
s_1	0	$1/5$

$\tilde{\pi}(\cdot\|\theta = 1)$	s_0	s_1
s_0	0	0
s_1	0	1

Table 6: Optimal Information Structure Using BCE approach

7. Conclusion

This paper contributes to the foundations of information design. Our representation theorem formulates the belief-based approach to the problem and deconstructs it into maximization within and concavification. As a result, every optimal solution, can be presented as a combination of an optimal (purely) public and an optimal private signals, irrespective of the method used to compute it. This general result speaks to a fundamental distinction in information economics. In addition, our approach is flexible with respect to the solution concept and to the selection criterion resolving the multiplicity of outcomes, which is methodologically attractive in certain environments.

Our applications attempt to demonstrate both the structural and the methodological advantages of the results. In the first application, the revelation principle fails. Thus, working directly with beliefs is indispensable, and our method can be used to compute the optimal information structure. In the second application, our approach provides a simple and intuitive resolution of the problem. We show that the public-private distinction of the signals, achieved through our decomposition, can be useful, instructive, and not obvious when other available approaches are used.[21]

[21] In a three-agent Manager's Problem that would extend our application in the obvious way, the manager could talk to one agent while keeping the other two uninformed, talk to two agents while keeping the last one uninformed, or talk to all three simultaneously. Notice that as we add more agents following the same structure (each one caring only about the action of the previous one) the set of priors for which a payoff of 1 can be achieved with certainty also increases. For example, if we add a project manager who is in charge of supervising the supervisor, we can achieve a certain payoff of 1 for the manager through optimal information design for all priors greater than 1/8. We thank Margaret Meyer for pointing this out.

Appendix

A. Proof of Proposition 1

Proof. Let τ be induced by some (S,π), so that

$$\tau(t) = \sum_{\theta} \pi\left((h^{\mu})^{-1}(t)\big|\theta\right)\mu_0(\theta) \tag{14}$$

for all $t \in \operatorname{supp}\tau$. Define $p \in \Delta(\Theta \times \hat{T})$ as

$$p(\theta,t) = \pi\left((h^{\mu})^{-1}(t)\big|\theta\right)\mu_0(\theta) \tag{15}$$

for all θ and $t \in \operatorname{supp}\tau$. It is immediate from (14) and (15) that $\operatorname{marg}_T p = \tau$ and so $\operatorname{marg}_{T_i} p = \tau_i$ for all i. When any player i forms his beliefs $\mu_i : T_i \to \Delta(\Theta \times T_{-i})$ under information structure $(\operatorname{supp}\tau, \pi)$, he computes the conditional of p given t_i. That is, player i's belief hierarchies are derived from $p(\cdot|t_i)$ for all i and, thus,

$$p(\theta,t) = \beta_i^*(\theta, t_{-i}|t_i)\operatorname{marg}_{T_i} p(t_i)$$

for all i, θ, and $t \in \operatorname{supp}\tau$. We conclude that $\tau \in \mathcal{C}$. Finally,

$$\sum_{t_i \in \operatorname{supp}\tau_i} \beta_i^*(\theta|t_i)\tau_i(t_i) := \operatorname{marg}_\Theta p(\theta) = \sum_t \pi\left((h^{\mu})^{-1}(t)\big|\theta\right)\mu_0(\theta) = \mu_0(\theta)$$

for all θ, which proves Bayes plausibility.

Suppose now that $\tau \in \mathcal{C}$ and satisfies Bayes plausibility. Let us show that these conditions are sufficient for τ to be induced by some some (S,π). Define information structure $(\operatorname{supp}\tau, \pi_\tau)$ where

$$\pi_\tau(t|\cdot) : \theta \mapsto \frac{1}{\mu_0(\theta)}\beta_i^*(\theta, t_{-i}|t_i)\tau_i(t_i) \tag{16}$$

for all $t \in \operatorname{supp}\tau$, which is defined independently of the choice of i because $\tau \in \mathcal{C}$. First, let us verify that π_τ is a valid information structure. Bayes plausibility says

$$\sum_{t_i \in \operatorname{supp}\tau_i} \beta_i^*(\theta|t_i)\tau_i(t_i) = \mu_0(\theta),$$

which guarantees that

$$\pi_\tau(t|\theta) = \frac{1}{\mu_0(\theta)} \sum_{t_i \in \operatorname{supp}\tau_i} \beta_i^*(\theta|t_i)\tau_i(t_i) = 1,$$

and in turn that $\pi(\cdot|\theta)$ is a probability distribution for every θ. By construction, this information structure is such that, when any player j receives t_j, his beliefs are

$\mu_j(\cdot|t_j) = \beta_j^*(\cdot|t_j)$, also because $\tau \in \mathcal{C}$. To prove that π_τ generates τ, we need to check that

$$\tau(t) = \sum_\theta \pi(t|\theta)\mu_0(\theta) \tag{17}$$

for all $t \in \operatorname{supp}\tau$. By (16), the rhs of (17) is equal to $\beta_i^*(t_{-i}|t_i)\tau_i(t_i)$, which equals $\tau(\hat{T})$ because $\tau \in \mathcal{C}$ (in particular, because $\operatorname{marg}_\Theta p = \tau$). □

B. Solution Concepts

B.1. Assumptions

For any $\tau \in \mathcal{C}$, the pair $\mathcal{G} = \langle G, p_\tau \rangle$ describes a Bayesian game in which players behave according to solution concept $\Sigma^*(\mathcal{G}) \subseteq \{\sigma : \operatorname{supp}\tau \to \Delta A\}$, which results in an outcome correspondence

$$O_{\Sigma^*}(\mathcal{G}) := \left\{ \gamma \in \Delta(A \times \Theta) : \exists \sigma \in \Sigma^*(\mathcal{G}) \text{ s.t. } \gamma(a,\theta) = \sum_t \sigma(a|t) p_\tau(t,\theta)\, \forall (a,\theta) \right\}. \tag{18}$$

Again, for a fixed base game, we just write $\Sigma^*(\tau)$ and $O_{\Sigma^*}(\tau)$.

Assumption 1. *Given Σ, there is a solution concept Σ^* such that*

$$O_{\Sigma^*}(\tau) = \bigcup_{(S,\pi) \text{ induces } \tau} O_\Sigma(S,\pi), \quad \textit{and}$$

$$\forall \tau, \tau', \textit{ if } \sigma \in \Sigma^*(\tau) \textit{ then } \exists \sigma' \in \Sigma^*(\tau') \textit{ s.t. } \sigma(t) = \sigma'(t), \forall t \in \textit{supp}\,\tau \cap \textit{supp}\,\tau'.$$

Solution concept Σ^* is relevant only if it captures all outcomes from Σ, hence the first requirement. The second requirement is called **invariance**: it says that play at a profile of belief hierarchies t under Σ^* does not depend on the ambient distribution from which t is drawn. This property is important because:

Proposition 2. *If Σ^* is invariant, then O_{Σ^*} is linear.*[22]

It is not true for all Σ that $O_\Sigma(S,\pi) = O_\Sigma(\tau)$ whenever (S,π) induces τ. Indeed, the same solution concept may not generate the same outcome distributions whether it is applied to (S,π) or to its induced τ. In Bayes Nash equilibrium, for example, the existence of many messages s inducing the same profile t of hierarchies can create opportunities for correlation among those s and, hence, for play, that cannot

[22]For all $\tau', \tau'' \in \mathcal{C}$ and $\alpha \in [0,1]$, $\alpha O_{\Sigma^*}(\tau') + (1-\alpha) O_{\Sigma^*}(\tau'') = O_{\Sigma^*}(\alpha\tau' + (1-\alpha)\tau'')$, where $\alpha O_{\Sigma^*}(\tau) = \{\alpha\gamma : \gamma \in O_{\Sigma^*}(\tau)\}$.

be replicated by t alone. That said, for various solution concepts Σ, epistemic game theory has identified invariant Σ^*, sometimes $\Sigma^* := \Sigma$, such that $O_{\Sigma^*}(\tau) = O_\Sigma(S,\pi)$ whenever (S,π) induces τ. We borrow from that literature to define the appropriate Σ^* satisfying Assumption 1 for $\Sigma :=$ BNE. In Section 5, we use interim correlated rationalizability, which is invariant and satisfies (i) with $\Sigma^* = \Sigma$.

B.2. Illustration

For applications in Bayes Nash information design ($\Sigma :=$ BNE), it is useful to know which solution concept Σ^* satisfies $O_\Sigma(S,\pi) = O_{\Sigma^*}(\tau)$ whenever (S,π) induces τ, and whether Σ^* is invariant. This is done in Proposition 3.

Definition 2. Given $\tau \in \mathcal{C}$, the set of belief-preserving Bayes correlated equilibria in $\langle G, p_\tau \rangle$, denoted $\text{BCE}_B(\tau)$, consists of all $\sigma : \operatorname{supp}\tau \times \Theta \to \Delta A$ such that for all i,

$$\sum_{a_{-i},t_{-i},\theta} p_\tau(t,\theta)\sigma(a_i,a_{-i}|t_i,t_{-i},\theta)(u_i(a_i,a_{-i},\theta) - u_i(a_i',a_{-i},\theta)) \geq 0 \tag{19}$$

for all a_i, a_i' and $t_i \in \operatorname{supp}\tau_i$, and for all i,

$$\sigma_i(a_i|t_i,t_{-i},\theta) := \sum_{a_{-i}\in A_{-i}} \sigma(a_i,a_{-i}|t_i,t_{-i},\theta) \tag{20}$$

is independent of t_{-i} and θ.[23]

In a correlated equilibrium, player i of type t_i receives an action recommendation a_i that is incentive-compatible by (19). This condition alone defines a Bayes correlated equilibrium (see Bergemann and Morris (2016) and our conclusion). In addition, (20) requires that i's action recommendation reveal no more information to i about the other players' hierarchies and the state of the world than what is already contained in his hierarchy. Thus, after receiving the action recommendation, each player's belief hierarchy remains unaltered. Forges (2006) introduced a condition equivalent to (20), while Liu (2015) introduced the belief-preserving Bayes correlated equilibrium.

Proposition 3. *BCE_B is invariant and for all $\tau \in \mathcal{C}$,*

$$O_{BCE_B}(\tau) = \bigcup_{(S,\pi)\text{ induces }\tau} O_{BNE}(S,\pi).$$

[23]Note that here, we have extended the definition of the solution concept σ to allow for dependence on the state θ in addition to the type profile. This allows us to capture correlation between actions that could have only resulted from redundant types, which do not exist on the universal type space. We accordingly adapt the definition of invariance in the obvious way, i.e. the equivalence now must hold, in addition, for every $\theta \in \Theta$.

This proposition[24] demonstrates that BCE_B is invariant and that the set of BCE_B outcomes from τ is the union of all BNE outcomes for all information structures that induce τ. Then, as Bergemann and Morris (2016, p.507) put it, the belief-preserving BCE captures the implications of common knowledge of rationality and that players know exactly the information contained in τ (and no more) given the common prior assumption.

C. Proof of Theorem 1

Lemma 1. *$\mathcal{C}$ is convex.*

Proof. Take $\alpha \in [0,1]$ and $\tau', \tau'' \in \mathcal{C}$. By definition of $\mathcal{C}$, there are $p_{\tau'}$ and $p_{\tau''}$ such that $\mathrm{marg}_T p_{\tau'} = \tau'$ and $\mathrm{marg}_T p_{\tau'} = \tau'$ and

$$\begin{aligned} p_{\tau'}(\theta,t) &= \beta_i^*(\theta,t_{-i}|t_i)\tau_i'(t_i), \\ p_{\tau''}(\theta,t) &= \beta_i^*(\theta,t_{-i}|t_i)\tau_i''(t_i), \end{aligned} \tag{21}$$

for all θ, i and t. Define $\tau := \alpha\tau' + (1-\alpha)\tau''$ and note that $\tau_i = \alpha\tau_i' + (1-\alpha)\tau_i''$, by the linearity of Lebesgue integral. Define

$$p_\tau(\theta,t) := \beta_i^*(\theta,t_{-i}|t_i)\tau_{\alpha,i}(t_i)$$

for all i, θ, and $t \in \mathrm{supp}\,\tau$. Notice that p_τ is well-defined, because of (21). Thus,

$$\mathrm{marg}_T p_\tau = \alpha \mathrm{marg}_T p_{\tau'} + (1-\alpha)\mathrm{marg}_T p_{\tau''} = \alpha\tau' + (1-\alpha)\tau'' = \tau$$

and we conclude that $\tau \in \mathcal{C}$. □

Although $\mathcal{C}$ is convex, it is not closed because we can build sequences in $\mathcal{C}$ with growing supports, only converging to a belief-hierarchy distribution with an infinite support. Still, the next lemma proves that minimal (consistent) distributions are the extreme points of the set of consistent distributions.

Lemma 2. *$\mathcal{E} = \mathcal{C}^M$.*

Proof. An extreme point of $\mathcal{C}$ is a $\tau \in \mathcal{C}$ such that $\tau = \alpha\tau' + (1-\alpha)\tau''$ if and only if $\tau' = \tau'' = \tau$. We first show that if $\tau \in \mathcal{C}^{\mathrm{M}}$, then τ is an extreme point of $\mathcal{C}$. Suppose not. That is, fixing $\alpha \in (0,1)$, let $\tau = \alpha\tau' + (1-\alpha)\tau''$ for some $\tau' \neq \tau''$ and define $\tilde{\lambda} := \max\{\lambda \geq 0 : \tau + \lambda(\tau - \tau'') \in \mathcal{C}\}$. Then, it must be that $\tilde{\tau} := \tau + \tilde{\lambda}(\tau - \tau'') \neq \tau$. We want to show that $\mathrm{supp}\;\tilde{\tau} \subsetneq \mathrm{supp}\tau$. To see this, let $\tilde{t} \in \mathrm{supp}\;\tilde{\tau}$ and suppose $\tilde{t} \notin \mathrm{supp}\;\tau$. Then, by definition, $\tilde{\tau}(\tilde{t}) = -\lambda\tau''(\tilde{t}) \leq 0$, which

[24]The proof is available to the reader upon request.

is impossible. Moreover, there is a $t \in \operatorname{supp} \tau$, such that $t \notin \operatorname{supp} \tilde{\tau}$. If not, it would imply that for all $t \in \operatorname{supp} \tau$, $\tilde{\tau}(t) > 0$. When this is the case, however, we can find a $\lambda' > \tilde{\lambda}$, s.t. $\tau + \lambda'(\tau - \tau'') \in \mathcal{C}$, a contradiction on the fact that $\tilde{\lambda}$ is in fact the max. We conclude that $\operatorname{supp} \tilde{\tau} \subsetneq \operatorname{supp}\tau$ and thus $\tau \notin \mathcal{C}^{\mathrm{M}}$. Conversely, suppose τ is not minimal, i.e., there is a $\tau' \in \mathcal{C}$ such that $\operatorname{supp}\tau' \subsetneq \operatorname{supp}\tau$. Define $\tau'' \in \Delta T$ as $\tau''(\cdot) := \tau(\,\cdot \mid \operatorname{supp}\tau \setminus \operatorname{supp}\tau')$, the conditional distribution of τ given the subset $\operatorname{supp}\tau \setminus \operatorname{supp}\tau'$. Clearly

$$\tau = \alpha\tau' + (1-\alpha)\tau'' \tag{22}$$

where $\alpha = \tau(\operatorname{supp}\tau') \in (0,1)$. Since $\operatorname{supp}\tau'$ is belief-closed, so is $\operatorname{supp}\tau \setminus \operatorname{supp}\tau'$. Since τ'' is derived from a consistent τ and is supported on a belief-closed subspace, τ'' is consistent. Given that $\tau'' \neq \tau'$, (22) implies that τ is not an extreme point. $\square$

Proposition 4. *For any $\tau \in \mathcal{C}$, there exist unique $\{e_i\}_{i=1}^n \subseteq \mathcal{C}^M$ and weakly positive numbers $\{\alpha_i\}_{i=1}^n$ such that $\sum_{i=1}^n \alpha_i = 1$ and $\tau = \sum_{i=1}^n \alpha_i e_i$.*

Proof. Take any $\tau \in \mathcal{C}$. Either τ is minimal, in which case we are done, or it is not, in which case there is $\tau' \in \mathcal{C}$ such that $\operatorname{supp}\tau' \subsetneq \operatorname{supp}\tau$. Similarly, either τ' is minimal, in which case we conclude that there exists a minimal $e_1 := \tau'$ with support included in $\operatorname{supp}\tau$, or there is $\tau'' \in \mathcal{C}$ such that $\operatorname{supp}\tau'' \subsetneq \operatorname{supp}\tau'$. Given that τ has finite support, this procedure eventually delivers a minimal consistent belief-hierarchy distribution e_1. Since τ and e_1 are both consistent and hence, their supports belief-closed, $\operatorname{supp}(\tau \setminus e_1)$ must be belief-closed. To see why, note that for any $t \in \operatorname{supp}(\tau \setminus e_1)$, if there were $i, \hat{t} \in \operatorname{supp} e_1$ and $\theta \in \Theta$ such that $p_\tau(\theta, \hat{t}_{-i}|t_i) > 0$, then this would imply $p_\tau(\theta, \hat{t}_{-i}, t_i) > 0$ and, thus, $p_\tau(\theta, t_i, \hat{t}_{-(ij)}|\hat{t}_j) > 0$ (where $\hat{t}_{-(ij)}$ is the belief hierarchies of players other than i and j). As a result, player j would believe at $\hat{t}_j$ (a hierarchy that j can have in e_1) that i believes that players' types could be outside $\operatorname{supp} e_1$ (because $p_\tau(\theta, t) > 0$). Then, it would violate the fact that $\operatorname{supp} e_1$ is belief-closed, a contradiction. Given that $\operatorname{supp}(\tau \setminus e_1)$ is a belief-closed subset of $\operatorname{supp}\tau$ and τ is consistent, $\tau \setminus e_1$ is itself consistent under

$$p_{\tau\setminus e_1}(\theta, t) := \frac{p_\tau(\theta, t)}{\tau(\operatorname{supp}(\tau \setminus e_1))}$$

for all θ and $t \in \operatorname{supp}(\tau \setminus e_1)$. This follows immediately from the conditions that $p_\tau(\theta, t) = \beta_i^*(\theta, t_{-i}|t_i)\tau_i(t_i)$ for all θ, t and i, $\operatorname{marg}_T p_\tau = \tau$, and the definition of belief-closedness. Therefore, we can reiterate the procedure from the beginning and apply it to $\tau \backslash e_1$. After $\ell - 1$ steps, we obtain the consistent belief-hierarchy distributions $\tau\backslash\{e_1, \ldots, e_{\ell-1}\}$. Since τ has finite support, there must be ℓ large enough such that $\tau\backslash\{e_1, \ldots, e_{\ell-1}\}$ is minimal; when it happens, denote $e_\ell := \tau\backslash\{e_1, \ldots, e_{\ell-1}\}$. We conclude that

$$\tau = \sum_{i=1}^{\ell} \tau(\operatorname{supp} e_i) e_i$$

where $\tau(\operatorname{supp} e_i) \geq 0$ and $\sum_{i=1}^{\ell} \tau(\operatorname{supp} e_i) = \tau(\cup_{i=1}^{\ell} e_i) = \tau(\operatorname{supp} \tau) = 1$. □

Now, we prove linearity of w. The point is to show that the set of outcomes of a mixture of subspaces of the universal type space can be written as a similar mixture of the sets of outcomes of these respective subspaces.

Lemma 3. *The function w is linear over $\mathcal{C}^M$.*

Proof. Let $\tau', \tau'' \in \mathcal{C}^M$ and $\alpha \in [0,1]$. Define $\tau = \alpha\tau' + (1-\alpha)\tau''$. Proposition 2 shows linearity of O_{Σ^*}, so we have

$$\begin{aligned} w(\tau) &= \sum_{\theta,a} g(O_{\Sigma^*}(\tau))[\theta,a] v(a,\theta) \\ &= \sum_{\theta,a} f\Big(\alpha O_{\Sigma^*}(\tau') + (1-\alpha) O_{\Sigma^*}(\tau'')\Big)[\theta,a] v(a,\theta) \end{aligned}$$

Since g is linear, this becomes

$$\begin{aligned} & \alpha \sum_{\theta,a} f\big(O_{\Sigma^*}(\tau')\big)[\theta,a] v(a,\theta) + (1-\alpha) \sum_{\theta,a} f\big(O_{\Sigma^*}(\tau'')\big)[\theta,a] v(a,\theta) \\ = \; & \alpha w(\tau') + (1-\alpha) w(\tau''), \end{aligned}$$

which completes the proof. □

Proof of Theorem 1. Fix a prior $\mu_0 \in \Delta(\Theta)$ and take any information structure (S,π). From Proposition 1, it follows that (S,π) induces a consistent belief-hierarchy distribution $\tau \in \mathcal{C}$ such that $\operatorname{marg}_\Theta p_\tau = \mu_0$. By definition of Σ^* and w, we have $V(S,\pi) = w(\tau)$ and, thus, $\sup_{(S,\pi)} V(S,\pi) \leq \sup\{w(\tau) | \tau \in \mathcal{C} \text{ and } \operatorname{marg}_\Theta p_\tau = \mu_0\}$. Now, take $\tau \in \mathcal{C}$ such that $\operatorname{marg}_\Theta p_\tau = \mu_0$. By Proposition 1, we know that there exists an information structure (S,π) that induces τ and such that $V(S,\pi) = w(\tau)$. Therefore, $\sup_{(S,\pi)} V(S,\pi) \geq \sup\{w(\tau) | \tau \in \mathcal{C} \text{ and } \operatorname{marg}_\Theta p_\tau = \mu_0\}$. We conclude that

$$\sup_{(S,\pi)} V(S,\pi) = \sup_{\substack{\tau \in \mathcal{C} \\ \operatorname{marg}_\Theta p_\tau = \mu_0}} w(\tau). \tag{23}$$

By Proposition 4, there exists a unique $\lambda \in \Delta^f(\mathcal{C}^{\mathrm{M}})$ such that $\tau = \sum_{e \in \operatorname{supp} \lambda} \lambda(e) e$. Since p and marg are linear,

$$\operatorname{marg}_\Theta p_\tau = \operatorname{marg}_\Theta p_{\sum_e \lambda(e) e} = \sum_{e \in \operatorname{supp} \lambda} \lambda(e) \operatorname{marg}_\Theta p_e.$$

Then, by Lemma 3 and (23), we have

$$\sup_{(S,\pi)} V(S,\pi) = \sup_{\lambda\in\Delta^f(\mathcal{C}^{\mathrm{M}})} \sum_e w(e)\lambda(e) \qquad \text{subject to} \sum_e \mathrm{marg}_\Theta p_e\, \lambda(e) = \mu_0, \tag{24}$$

which concludes the proof. □

D. Investment Game

Consider a 2×2 minimal distribution of the form presented on p.13. To leverage second-order beliefs, we need to choose $\mu_1'' = \mu_2'' = \frac{2}{3}+\varepsilon$. Moreover, $1-A-B-C=0$ because each player i already invests at (μ_i'', λ_i'') based on his first-order beliefs alone. Additionally, symmetry is without loss due to the symmetry of the game and the objective. Hence, the relevant minimal distributions must be of the form:

e^*_μ	(μ_2', λ_2')	(μ_2'', λ_2'')
(μ_1', λ_1')	A	$\frac{1-A}{2}$
(μ_2'', λ_1'')	$\frac{1-A}{2}$	0

where $A\in[0,1]$. Next, we characterize consistency. The Bayes plausibility (BP) condition requires that

$$\frac{1+A}{2}\cdot\mu' + \frac{1-A}{2}\cdot\left(\frac{2}{3}+\varepsilon\right) = \mu. \tag{25}$$

Use first-order beliefs to designate the hierarchies (call this a player's type) and parameterize second-order beliefs as

$$\begin{array}{llll}
\lambda_i(\theta=2,\mu'|\mu') = & \mu'-x & \lambda_i(\theta=2,\mu'|\mu'') = & 2/3+\varepsilon \\
\lambda_i(\theta=-1,\mu'|\mu') = & 1-\mu'-y & \lambda_i(\theta=-1,\mu'|\mu'') = & 1/3-\varepsilon \\
\lambda_i(\theta=2,\mu''|\mu') = & x & \lambda_i(\theta=2,\mu''|\mu'') = & 0 \\
\lambda_i(\theta=-1,\mu''|\mu') = & y & \lambda_i(\theta=-1,\mu''|\mu'') = & 0
\end{array}$$

for a player of type μ' and $\mu''=2/3+\varepsilon$, respectively. Consistency requires that

$$\lambda_1(\theta,t_2|t_1)\tau(t_1) = \lambda_2(\theta,t_1|t_2)\tau(t_2)$$

for all (θ,t_1,t_2). This gives $x=\left(\frac{2}{3}+\varepsilon\right)\cdot\frac{1-A}{1+A}$ and $y=\left(\frac{1}{3}-\varepsilon\right)\cdot\frac{1-A}{1+A}$. It remains to make sure that all second-order beliefs are in $[0,1]$. This requires $\mu'\geq x$ and $y\leq 1-\mu'$, pinning down μ':

$$\mu' = \left(\frac{2}{3}+\varepsilon\right)\cdot\frac{1-A}{1+A}. \tag{26}$$

Substituting this into (25), we obtain $A = 1-\frac{3\mu}{2+3\varepsilon}$ and $\mu'=\frac{(2+3\varepsilon)\mu}{4-3\mu+6\varepsilon}$. Hence,

e^*_μ	$\mu' = \frac{(2+3\varepsilon)\mu}{4-3\mu+6\varepsilon}$	$\mu'' = 2/3+\varepsilon$
$\mu' = \frac{(2+3\varepsilon)\mu}{4-3\mu+6\varepsilon}$	$1 - \frac{3\mu}{2+3\varepsilon}$	$\frac{3\mu}{4+6\varepsilon}$
$\mu'' = 2/3+\varepsilon$	$\frac{3\mu}{4+6\varepsilon}$	0

is the minimal distribution that will ensure joint investment for the smallest prior. To compute that smallest prior, write down the condition that ensures investment for each type. When a player is of type $2/3+\varepsilon$, investment is guaranteed. When a player's type is type μ', he invests if:

$$3\mu' + (x+y) - 2 > 0 \tag{27}$$

which implies

$$\frac{1-A}{1+A} > 2 - 3\mu'. \tag{28}$$

By (26) and (28), we get $\mu > \frac{8+12\varepsilon}{15+9\varepsilon} \approx \frac{8}{15}$. Therefore, for all $\mu > 8/15$, e^*_μ ensures both players invest with probability one.

E. Manager's Problem

Proposition 5. *A belief-hierarchy distribution $\eta \in \mathcal{A}$ is consistent if and only if $\eta(\mu_1, \lambda_2) = \eta_2(\lambda_2)\lambda_2(\mu_1)$ for all $(\mu_1, \lambda_2) \in supp\,\eta$.*

Proof. ("only if") Let η be consistent. By (6), there is $p_\eta \in \Delta^f(\Theta \times \Delta\Theta \times \Delta\Delta\Theta)$ such that

$$\mathrm{marg}_{(\mu_1,\lambda_2)} p_\eta(\mu_1, \lambda_2, \theta) = \eta(\mu_1, \lambda_2) \tag{29}$$

for all $(\mu_1, \lambda_2) \in \mathrm{supp}\,\eta$. By consistency and definition of λ_2,

$$\lambda_2(\mu_1) = \frac{\sum_\theta p_\eta(\theta, \mu_1, \lambda_2)}{\eta_2(\lambda_2)}$$

and, therefore,

$$\eta(\mu_1, \lambda_2) = \sum_\theta p_\eta(\theta, \mu_1, \lambda_2) = \eta_2(\lambda_2)\lambda_2(\mu_1)$$

for all $\mu_1 \in \mathrm{supp}(\eta_1)$ and $\lambda_2 \in \mathrm{supp}(\eta_2)$, where the first equality follows by (29) and the second by the definition of λ_2.

("if") Take any $\eta \in \mathcal{A}$ and define p_η as

$$p_\eta(\mu_1, \lambda_2, \theta) := \mu_0(\theta)\frac{\eta_1(\mu_1)\mu_1(\theta)}{\mu_0(\theta)}\frac{\eta_2(\lambda_2)\lambda_2(\mu_1)}{\eta_1(\mu_1)} = \mu_1(\theta)\eta_2(\lambda_2)\lambda_2(\mu_1). \tag{30}$$

Therefore, $\mathrm{marg}_{(\mu_1,\lambda_2)} p_\eta(\mu_1, \lambda_2, \theta) = \eta(\mu_1, \lambda_2)$ by the condition of the proposition. Furthermore, it is easy to check that $p_\eta(\theta|\mu_1) = \mu_1(\theta)$ and $p_\eta(\mu_1|\lambda_2) = \lambda_2(\mu_1)$, that is, the first- and second-order beliefs (respectively, of player 1 and 2) supported by p_η are obtained by Bayesian updating of p_η, hence η is consistent. □

Definition 3. Let p_η be defined as in (30). A distribution $\tau \in \mathcal{C}$ induces $\eta \in \mathcal{A}$ if for all $(\mu_1, \lambda_2, \theta)$

$$p_\eta(\mu_1, \lambda_2, \theta) = \sum_{\substack{t_1 : \beta_1^*(\theta|t_1) = \mu_1(\theta) \\ t_2 : \beta_2^*(\{t_1 : \mathrm{marg}_\Theta \beta_1^*(\cdot|t_1) = \mu_1\}|t_2) = \lambda_2(\mu_1)}} p_\tau(t_1, t_2, \theta).$$

Definition 4. For any $\eta \in \mathcal{A}$, $\mathrm{BNE}(\eta)$ consists of all $\sigma = (\sigma_1(\cdot|\mu_1), \sigma_2(\cdot|\lambda_2))$ such that

$$\mathrm{supp}\,\sigma_1(\cdot|\mu_1) \subseteq \mathrm{argmax}_{a_1} \sum_\theta u_1(a_1, \theta)\mu_1(\theta)$$

for all $\mu_1 \in \mathrm{supp}\,\eta_1$, and

$$\mathrm{supp}\,\sigma_2(\cdot|\lambda_2) \subseteq \mathrm{argmax}_{a_2} \sum_{a_1, \mu_1} u_2(\sigma_1(a_1|\mu_1), a_2)\lambda_2(\mu_1)$$

for all $\mu_2 \in \mathrm{supp}\,\eta_2$.

We show next that, as far as distributions over equilibrium action profiles are concerned, we can work with distributions in $\mathcal{A}$ only. Define the set of action distributions under solution concept Σ to be

$$O_\Sigma^A(\tau) = \{\gamma_A \in \Delta(A) : \exists \gamma \in O_\Sigma(\tau) \text{ s.t. } \mathrm{marg}_A \gamma = \gamma_A\}.$$

Recall that, from Proposition 3, $O^A_{BCE_B}(\tau)$ captures all BNE distributions (over action profiles in $\langle G, (S, \pi)\rangle$ of all information structures.

Proposition 6. *If τ induces η, then $O^A_{BCE_B}(\tau) = O^A_{BNE}(\eta)$.*

Proof. In the game between P and W, $\mathrm{BCE}_B(\tau)$ requires for any τ that for all $t_1 \in T_1$, $a_1, a_1' \in A_1$,

$$\sum_{a_2, t_2, \theta} p_\tau(t, \theta)\sigma(a_1, a_2|t_1, t_2, \theta)(u_1(a_1, \theta) - u_1(a_1', \theta))$$
$$= \sigma_1(a_1|t_1) \sum_\theta p_\tau(t_1, \theta)(u_1(a_1, \theta) - u_1(a_1', \theta)) \geq 0$$

where we have used that $\sum_{a_2} \sigma(a_1, a_2|t_1, t_2, \theta) = \sigma_1(a_1|t_1)$ since σ is belief-preserving. Dividing both sides by $\sigma_1(a_1|t_1)\tau(t_1)$ and substituting in $\beta_1^*(\theta|t_1) = p_\tau(t_1, \theta)/\tau(t_1)$, we obtain

$$\sum_\theta \beta_1^*(\theta|t_1)(u_1(a_1, \theta) - u_1(a_1', \theta)) \geq 0$$

for all t_1. Since for all $t_1 \in T_1$, $\text{marg}_\Theta \beta_1^*(\cdot|t_1) = \mu_1$ for some $\mu_1 \in \Delta\Theta$, we can write

$$\text{supp}\,\sigma_1(\cdot|\mu_1) \subseteq \text{argmax}_{a_1} \sum_\theta u_1(a_1,\theta)\mu_1(\theta).$$

This conclusion also implies that $\sigma(a_1|a_2,t_1,t_2,\theta) = \sigma_1(a_1|t_1)$ for all $(a,t) \in A \times T$. Therefore, $\sigma(a_1,a_2|t_1,t_2,\theta) = \sigma_1(a_1|t_1)\sigma(a_2|t_1,t_2,\theta)$. Summing across all $a_1 \in A_1$ we get:

$$\sum_{a_1} \sigma(a_1,a_2|t_1,t_2,\theta) = \sigma(a_2|t_1,t_2,\theta) = \sigma_2(a_2|t_1,t_2,\theta) = \sigma_2(a_2|t_2)$$

where the last equality follows from the belief-preserving property of σ. Given $\sigma(a_1,a_2|t_1,t_2,\theta) = \sigma_1(a_1|t_1)\sigma_2(a_2|t_2)$, $\text{BCE}_B(\tau)$ requires that for all $t_2 \in T_2$, $a_2, a_2' \in A_2$,

$$\sigma_2(a_2|t_2) \sum_{a_1,t_1} p_\tau(t_1,t_2)\sigma_1(a_1|t_1)(u_2(a_1,a_2) - u_2(a_1,a_2')) \geq 0. \tag{31}$$

Since player 1's strategy is a function of μ_1, $\sigma_1(a_1|\mu_1)$, 2 formulates beliefs

$$\begin{aligned}\beta_2^*(\mu_1|t_2) &:= \beta_2^*(\{t_1 : \text{marg}_\Theta\beta_1^*(\cdot|t_1) = \mu_1\}|t_2) \\ &= \frac{p_\tau(\{t_1 : \text{marg}_\Theta\beta_1^*(\cdot|t_1) = \mu_1\}, t_2)}{\tau(t_2)}.\end{aligned}$$

Dividing (31) by $\sigma_2(a_2|t_2)\tau(t_2)$ and substituting in $\sigma_1(a_1|\mu_1)$ and $\beta_2^*(\mu_1|t_2)$ give

$$\sum_{a_1,\mu_1} \beta_2^*(\mu_1|t_2)\sigma_1(a_1|\mu_1)(u_2(a_1,a_2) - u_2(a_1,a_2')) \geq 0$$

for all $t_2 \in T_2$, $a_2, a_2' \in A_2$. For all $t_2 \in T_2$, there is $\lambda_2 \in \Delta\Delta\Theta$ such that $\beta_2^*(\cdot|t_2) = \lambda_2$, and so we can write

$$\text{supp}\,\sigma_2(\cdot|\lambda_2) \subseteq \text{argmax}_{a_2} \sum_{a_1,\mu_1} u_2(\sigma_1(a_1|\mu_1), a_2)\lambda_2(\mu_1)$$

for all λ_2. Hence, $(\sigma_1(\cdot|\mu_1), \sigma_2(\cdot|\lambda_2)) \in \text{BNE}(\eta)$. The equilibrium distribution over action profiles is given by

$$\begin{aligned}\sigma_\eta(a_1,a_2) &= \sum_{\mu_1,\lambda_2,\theta} p_\eta(\mu_1,\lambda_2,\theta)\sigma_1(a_1|\mu_1)\sigma_2(a_2|\lambda_2) \\ &= \sum_{\mu_1,\lambda_2,\theta} \sum_{\substack{t_1:\beta_1^*(\theta|t_1)=\mu_1(\theta) \\ t_2:\beta_2^*(\{t_1:\text{marg}_\Theta\beta_1^*(\cdot|t_1)=\mu_1\}|t_2)=\lambda_2(\mu_1)}} p_\tau(t_1,t_2,\theta)\sigma_1(a_1|t_1)\sigma_2(a_2|t_2) \\ &= \sum_{t_1,t_2,\theta} p_\tau(t_1,t_2,\theta)\sigma_1(a_1|t_1)\sigma_2(a_2|t_2) \\ &= \sigma_\tau(a_1,a_2),\end{aligned}$$

where we have used that τ induces η and the established equivalence between the σ_i's. Hence, $O^A_{BCE_B}(\tau) = O^A_{BNE}(\eta)$. □

F. On Minimal Consistent Distributions

Proposition 7.

(i) *Suppose that* $\tau \in \mathcal{C}$ *is conditionally independent. If* $\mu = marg_\Theta p_\tau$ *is not degenerate, then* τ *is minimal iff it is not perfectly informative. If* μ *is degenerate, then* τ *is minimal.*

(ii) *A public* $\tau \in \mathcal{C}$ *is minimal iff* $supp\,\tau$ *is a singleton.*

Proof. Part (i). Suppose that τ is conditionally independent and μ is non-degenerate. First, if τ is perfectly informative, then it can be written

$$\tau = \sum_\theta \mu(\theta)\tau_\theta,$$

where τ_θ is a distribution that gives probability 1 to belief hierarchies representing common knowledge that θ has realized. Given $\mu(\theta) \in (0, 1)$, τ is therefore a convex combination of belief-hierarchy distributions, hence it is not minimal.

Second, we show that if τ is non-minimal, then it must be perfectly informative. Let τ be non-minimal. By Lemma 2, there exist $\alpha \in (0, 1)$ and $\tau' \neq \tau''$ such that $\tau = \alpha\tau' + (1-\alpha)\tau''$. Without loss, we can assume $\text{supp}\ \tau' \cap \text{supp}\ \tau'' = \emptyset$. If this were not the case, we could find a consistent τ^* with $\text{supp}\,\tau^* = \text{supp}\ \tau' \cap \text{supp}\ \tau''$, in which case τ could be written as

$$\tau = \kappa\tau^* + (1-\kappa)\hat{\tau}$$

where $\kappa = \alpha q + (1-\alpha)r$, $q = \tau'(\text{supp}\,\tau^*)$, $r = \tau''(\text{supp}\,\tau^*)$ and[25]

$$\hat{\tau} = \frac{\alpha(1-q)}{1-\kappa}(\tau' \setminus \tau^*) + \frac{(1-\alpha)(1-q)}{1-\kappa}(\tau'' \setminus \tau^*).$$

Now, take $t' \in \text{supp}\ \tau'$, $t'' \in \text{supp}\ \tau''$ and note

$$p_\tau(t'_i, t''_{-i}|\theta) = \alpha p_{\tau'}(t'_i, t''_{-i}|\theta) + (1-\alpha)p_{\tau''}(t'_i, t''_{-i}|\theta) = 0 \tag{32}$$

for all θ and i. If τ were conditionally independent,

$$\begin{aligned}
& p_\tau(t'_i, t''_{-i}|\theta) \\
= \;& p_\tau(t'_i|\theta)\textstyle\prod_{j\neq i} p_\tau(t''_j|\theta) \\
= \;& \left(\alpha p_{\tau'}(t'_i|\theta) + (1-\alpha)p_{\tau''}(t'_i|\theta)\right)\textstyle\prod_{j\neq i}\left(\alpha p_{\tau'}(t''_j|\theta) + (1-\alpha)p_{\tau''}(t''_j|\theta)\right),
\end{aligned}$$

[25]For any $\tau, \tau' \in \mathcal{C}$ such that $\text{supp}\,\tau \cap \text{supp}\,\tau' \neq \emptyset$, we write $\tau \backslash \tau'$ to designate the unique consistent belief-hierarchy distribution with support $\text{supp}\,\tau \cap \text{supp}\,\tau'$.

which is strictly positive for some θ when τ is not perfectly informative, and thus contradicts (32). This implies that a non-minimal conditionally independent τ must be perfectly informative.

Part (ii). If τ is public, then every $\{t\}$ such that $t \in \operatorname{supp}\tau$ is a consistent distribution. Therefore, if $\operatorname{supp}\tau$ is a singleton, then it is clearly minimal. But if $\operatorname{supp}\tau$ is not a singleton, then τ is a convex combination of multiple consistent distribution, in which case τ is not minimal. □

Let $\mathcal{C}_\mu := \{\tau \in \mathcal{C} : \operatorname{marg}_\Theta p_\tau = \mu\}$ be the set of consistent distributions with posterior $\mu \in \Delta\Theta$ and let $E_\mu := \mathcal{C}_\mu \cap \mathcal{C}^{\mathrm{M}}$ denote the minimal ones among those. We next show that E_μ is dense in $\mathcal{C}_\mu$, although $\mathcal{C}^{\mathrm{M}}$ is small in a measure-theoretic sense relative to $\mathcal{C}$. Since there is no analog of the Lebesgue measure in infinite dimensional spaces, we use the notion of finite shyness proposed by Anderson and Zame (2001), which captures the idea of Lebesgue measure 0.

Definition 5. A measurable subset A of a convex subset C of a vector space $\mathcal{S}$ is finitely shy if there exists a finite dimensional vector space $V \subseteq \mathcal{S}$ for which $\lambda_V(C+s) > 0$ for some $s \in \mathcal{S}$, and $\lambda_V(A+s) = 0$ for all $s \in \mathcal{S}$, where λ_V is the Lebesgue measure defined on V.

Proposition 8. *$\mathcal{C}^{\mathrm{M}}$ is finitely shy in $\mathcal{C}$.*

Proof. By Lemma 1, $\mathcal{C}$ is a convex subset of the vector space $\mathcal{S}$ of all signed measures on T. Choose any distinct $e, e' \in \mathcal{C}^{\mathrm{M}}$ and let $V = \{\alpha(e-e') : \alpha \in \mathbb{R}\} \subseteq \mathcal{S}$. By construction, V is a one-dimensional subspace of $\mathcal{S}$. Let $\lambda_V \in \Delta V$ represent the Lebesgue measure on V. Notice that $\alpha(e-e') = \alpha e + (1-\alpha)e' - s$ for $s := e'$ and that $(\mathcal{C} - s) \cap V = \{\alpha(e-e') : \alpha \in [0,1]\}$ by convexity of $\mathcal{C}$. Hence, $\lambda_V(\mathcal{C}-s) > 0$. However, since $\mathcal{C}^M$ is the set of extreme points of $\mathcal{C}$, for every $s \in \mathcal{S}$, $(\mathcal{C}^M - s) \cap V$ contains at most two points. This gives $\lambda_V(\mathcal{C}^M - s) = 0$, since points have Lebesgue measure zero in V. □

Proposition 9. *Let Σ be BNE and suppose that the selection criterion is max. For any minimal belief-hierarchy distribution $e \in \mathcal{C}^{\mathrm{M}}$, there is an environment $\langle v, G\rangle$ for which $\lambda^* = \delta_e$ is the essentially unique optimal solution.*[26]

Proof. Fix $e \in \mathcal{C}^M$, $\varepsilon > 0$ and let $\mu_0 = p_e$. Denote by $G_\varepsilon(e) = (N, \{A_i, u_i\})$ the (base) game defined in Chen et al. (2010)'s Lemma 1 where $A_i \supseteq \operatorname{supp} e_i$. In this

[26]The proof establishes the stronger claim with $\Sigma :=$ ICR by using a result from Chen et al. (2010). A minimal distribution e is the essentially unique optimal solution if for all $\varepsilon > 0$, there is a game G_ε such that all $\tau \in \mathcal{C}$ with $d(e,\tau) > \varepsilon$ are strictly suboptimal (the metric is defined in (33)). By choosing a constant game or a constant designer's utility, it is easy to make all minimal distributions optimal, since the designer is indifferent among them. Uniqueness makes the result much stronger.

game, player i's actions include belief hierarchies from e_i. The (base) game $G_\varepsilon(e)$ is so conceived that, in the Bayesian game $(G_\varepsilon(e), e)$, every player has a strict incentive to truthfully report his true belief hierarchy, but for any τ that is suitably distant from e, some i in the Bayesian game $(G_\varepsilon(e), \tau)$ has a strict incentive *not* to report any $t_i \in \operatorname{supp} e_i$. For any i and $t_i, t_i' \in T_i$, let

$$d_i(t_i, t_i') := \sup_{k \geq 1} d^k(t_i, t_i')$$

where d^k is the standard metric (over k-order beliefs) that metrizes the topology of weak convergence. Let R_i and R be the ICR actions and profiles. Lemma 1 and Proposition 3 from Chen et al. (2010) imply that for every i and $t_i \in \operatorname{supp} e_i$,

$$t_i \in \operatorname*{argmax}_{a_i \in A_i} \sum_{\theta} \sum_{t_{-i} \in \operatorname{supp} e_{-i}} u_i(a_i, t_{-i}, \theta) \beta_i^*(\theta, t_{-i} | t_i)$$

and for every $\tau \in \mathcal{C}$ such that

$$d(e, \tau) := \max_i d_i^H(\operatorname{supp} e_i, \operatorname{supp} \tau_i) \geq \varepsilon, \tag{33}$$

where d_i^H is the standard Hausdorff metric, there exist i and $t_i' \in \operatorname{supp} \tau_i$ such that $\operatorname{supp} e_i \cap R_i(t_i') = \emptyset$. To see why, note that since e is minimal, there can be no sequence $(\tau_n) \subseteq \mathcal{C}$ such that $d(e, \tau_n) \geq \varepsilon$ for all n, while

$$\max_i \max_{t_i \in \operatorname{supp} \tau_{n,i}} \min_{t_i' \in \operatorname{supp} e_i} d_i(t_i, t_i') \to 0. \tag{34}$$

That is, for all τ such that $d(e, \tau) \geq \varepsilon$, there is $\delta > 0$ such that

$$\max_i \max_{t_i \in \operatorname{supp} \tau_{n,i}} \min_{t_i' \in \operatorname{supp} e_i} d_i(t_i, t_i') \geq \delta. \tag{35}$$

Put differently, there exist i and $t_i' \in \operatorname{supp} \tau_i$ such that $d_i(t_i, t_i') \geq \delta > 0$ for all $t_i \in \operatorname{supp} e_i$. From (the proof of) Proposition 3 in Chen et al. (2010), we conclude that $\operatorname{supp} e_i \cap R^k(t_i') = \emptyset$ for some k. Given that $R_i(t_i') = \cap_{k=1}^{\infty} R_i^k(t_i')$, we have $\operatorname{supp} e_i \cap R_i(t_i') = \emptyset$. Now, define the designer's utility as $v(a) := \mathbb{1}(a \in \operatorname{supp}\, e)$ for all $a \in A$. Then, the designer's expected payoff is

$$w(\tau) = \begin{cases} 1 & \text{if } \tau = e, \\ x & \text{if } d(e, \tau) < \varepsilon, \\ y & \text{if } d(e, \tau) \geq \varepsilon \end{cases}$$

where $x \leq 1$ and $y < 1$. When $d(e, \tau) < \varepsilon$, it is not excluded that $x = 1$, because all of $\operatorname{supp} \tau$, by virtue of being close to some hierarchy in $\operatorname{supp} e$, might report only in $\operatorname{supp} e$. However, whenever $d(e, \tau) \geq \varepsilon$, there is a hierarchy profile t occurring with positive probability that reports outside $\operatorname{supp} e$. Thus, the designer maximizes her expected payoff by setting $\lambda^* = \delta_e$, which is Bayes-plausible since $\mu_0 = p_e$. $\square$

References

Alonso, Ricardo and Odilon Câmara. 2015. "Persuading Voters." *American Economic Review* 106 (11):3590–3605.

Anderson, Robert M. and William Zame. 2001. "Genericity with Infinitely Many Parameters." *The B.E. Journal of Theoretical Economics* 1 (1):1–64.

Angeletos, George-Marios and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4):1103–1142.

Aumann, Robert J. 1976. "Agreeing to Disagree." *The Annals of Statistics* 4 (6):1236–1239.

Barelli, Paulo. 2009. "On the genericity of full surplus extraction in mechanism design." *Journal of Economic Theory* 144 (3):1320–1332.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2017. "First-Price Auctions with General Information Structures: Implications for Bidding and Revenue." *Econometrica* 85 (1):107–143.

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

Brandenburger, Adam and Eddie Dekel. 1993. "Hierarchies of Beliefs and Common Knowledge." *Journal of Economic Theory* 59 (1):189–198.

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

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang. 2016. "Pivotal Persuasion." Tech. rep. Working Paper.

Chen, Yi-Chun, Alfredo Di Tillio, Eduardo Faingold, and Siyang Xiong. 2010. "Uniform topologies on types." *Theoretical Economics* 5 (3).

Dekel, Edie, Drew Fudenberg, and Stephen Morris. 2007. "Interim Correlated Rationalizability." *Theoretical Economics* 2:15–40.

Ely, Jeffrey. 2017. "Beeps." *American Economic Review* 107 (1):31–53.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. 2015. "Suspense and Surprise." *Journal of Political Economy* 123 (1):215–260.

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

Germano, Fabrizio, Jonathan Weinstein, and Peio Zuazo-Garin. 2016. "Uncertain Rationality, Depth of Reasoning and Robustness in Games with Incomplete Information." Working Paper.

Heifetz, Aviad and Zvika Neeman. 2006. "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design." *Econometrica* 74 (1):213–233.

Inostroza, Nicolas and Alessandro Pavan. 2017. "Persuasion in Global Games with Application to Stress Testing." Working Paper, Northwestern University.

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

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. 2015. "Persuasion of a Privately Informed Receiver." Working paper.

Laclau, Marie and Ludovic Renou. 2016. "Public Persuasion." Working Paper.

Lipnowski, Elliot and Laurent Mathevet. 2015. "Disclosure to a Psychological audience." New York University.

Liu, Qingmin. 2015. "Correlation and common priors in games with incomplete information." *Journal of Economic Theory* 157 (C):49–75.

Mertens, Jean-Francois and Shmuel Zamir. 1985. "Formulation of Bayesian analysis for games with incomplete information." *International Journal of Game Theory* 14 (1):1–29.

Morris, Stephen and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review* 92 (5):1521–1534.

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

Rockafellar, R. Tyrrell. 1970. *Convex Analysis*. Princeton, New Jersey: Princeton University Press, 2nd ed.

Rubinstein, Ariel. 1989. "The Electronic Mail Game: Strategic Behavior under "Almost Common Knowledge."." *American Economic Review* 79 (3):385–91.

Taneva, Ina. 2014. "Information Design." Economics Department, Edinburgh University.

Vives, Xavier. 1988. “Aggregation of Information in Large Cournot Markets.” *Econometrica* 56 (4):851–876.

Yildiz, Muhamet. 2015. “Invariance to representation of information.” *Games and Economic Behavior* 94 (C):142–156.

Zhang, Jun and Junjie Zhou. 2016. “Information Disclosure in Contests: A Bayesian Persuasion Approach.” *Economic Journal* 126:2197–2217.