

Interim Correlated Rationalizability in Large Games*

Łukasz Balbus[†], Michael Greinecker[‡], Kevin Reffett[§], Łukasz Woźny[¶]

June 2025

Abstract

We provide general theoretical foundations for modeling strategic uncertainty in large distributional Bayesian games with general type spaces, using a version of interim correlated rationalizability. We then focus on the case in which payoff functions are supermodular in actions, as is common in the literature on global games. This structure allows us to identify extremal interim correlated rationalizable solutions with extremal interim Bayes–Nash equilibria. Notably, no order structure on types is assumed. We illustrate our framework and results using the large versions of the electronic mail game and a global game.

1 Introduction

Much of economic uncertainty is strategic uncertainty—uncertainty about what agents are going to do. However, much of applied economic modeling relies on equilibrium notions that are ill-suited to the study of strategic uncertainty. In equilibrium, every agent reacts optimally to what other agents are doing, which usually requires knowing what those other agents are doing. To circumvent this limitation imposed by equilibrium concepts, many researchers have introduced incomplete information into game-theoretic models. Under incomplete information, although every agent knows how every other agent would react to each possible piece of information, not knowing the actual information makes them uncertain about how others are going to act.

The program of studying strategic uncertainty by introducing incomplete information into the model and then studying the resulting equilibria was particularly successful in the literature on global games, starting with the pioneering

*We would like to thank Rabah Amir, Pierpaolo Battigalli, Martin Kaae Jensen, M. Ali Khan, Marcin Peński and Xavier Vives for helpful discussions during the writing of this paper.

[†]University of Zielona Góra, Poland.

[‡]ENS Paris-Saclay, France.

[§]Arizona State University, USA.

[¶]SGH Warsaw School of Economics, Poland.

work of [Carlsson and van Damme \(1993\)](#). Global games are a specific way to model incomplete information by introducing a small amount of uncertainty into the payoffs of a complete information game with strategic complementarities. In particular, nature draws a state from a common prior, and agents receive a private signal given by a commonly known signal technology. This perturbation, while typically small, can drastically change the behavior of players and lead to a unique equilibrium, unlike the often multiple equilibria seen in complete information games. Though the literature on global games started with two-player games, models with a continuum of players proved to be particularly useful for economic application; see for example the early survey of [Morris and Shin \(2003\)](#). Much of this literature has been quite open about the fact that the common prior and the prior signals serve mostly as a round-about way to introduce strategic uncertainty and break common certainty of economic behavior. Moreover, though the global games literature superficially studies mostly Bayes-Nash equilibria of games of incomplete information, many of the arguments in the background work by iteratively eliminating dominated strategies, as is done, for example, by [Atkeson \(2000\)](#). What drives the results in the global games literature seems to be a form of rationalizability. [Morris and Shin \(2003\)](#) go so far as to write that “[t]he natural way to understand the ‘trick’ to global games analysis is to go back and understand what is going on in terms of higher-order beliefs.” But higher-order beliefs derived from a common prior have a very special structure, and this structure clouds the role higher-order beliefs play. [Morris, Shin, and Yildiz \(2016\)](#) finally managed to analyze global games directly in terms of higher order beliefs, but their analysis is restricted to games with finitely many players.¹ The authors restrict themselves to studying games with finitely many players not because the logic of their argument breaks down otherwise but because they have no appropriate model of higher-order beliefs with a continuum of agents. This paper supplies the needed tools.

We first show that one can meaningfully formulate large Bayesian games in an interim way based on type spaces. We formulate large games in distributional form, following [Mas-Colell \(1984\)](#); they are not given in terms of an underlying set of players but are specified by a population distribution of characteristics. We then show that one can meaningfully study rationalizability in this setting. To do so, we adapt the notion of interim correlated rationalizability due to [Dekel, Fudenberg, and Morris \(2007\)](#). Interim correlated rationalizability is the solution concept corresponding to fixing a hierarchy of beliefs regarding the underlying state of nature, extending this hierarchy of beliefs to beliefs over action profiles, and imposing the requirement that everyone is rational and that everyone is commonly believed to be rational.² We closely follow the

¹[Mathevet \(2014\)](#) provides bounds on the set of rationalizable solutions in supermodular games with finitely many players but requires type spaces to have canonical orders.

²An explicit epistemic characterization along these lines is given by [Battigalli, Di Tillio,](#)

formulation of [Weinstein and Yildiz \(2017\)](#) who also allow for a continuum of actions and states of nature, but restrict themselves to working with finitely many players. Our first major result, [Theorem 1](#), shows the equivalence between two formulations of interim correlated rationalizability. Our official definition is based on iterative elimination of actions that are not best replies to any belief consistent with nobody playing actions that have been previously eliminated. The second formulation then characterizes interim correlated rationalizability as the most permissive self-enforcing (set-valued) theory of rational behavior.

Most games studied in the literature on global games exhibit strategic complementarities. Such games have very nice properties. There exists a largest and a smallest equilibrium, and all rationalizable solutions are sandwiched in between; see [Milgrom and Roberts \(1990\)](#). We show that a similar result holds for our large interim Bayesian games when we introduce strategic complementarities; this is our [Theorem 2](#). As a byproduct, we obtain the first existence result for Bayes-Nash equilibria in large games with a continuum of agents and states of nature. An important question in the global games literature is equilibrium uniqueness. Indeed, the fundamental result of [Carlsson and van Damme \(1993\)](#) shows that under certain conditions, there is a unique equilibrium in the limit as uncertainty vanishes. It follows from our result that there exists a unique interim Bayes-Nash equilibrium if and only if there exists a unique interim correlated rationalizable action for each player and each of their beliefs if and only if the largest and the smallest interim Bayes-Nash equilibrium coincide. We also provide sufficient epistemic conditions for equilibrium uniqueness.

Throughout, the type spaces that we use to model higher-order uncertainty about the state of nature are assumed to be compact metrizable. We show that this is essentially without loss of generality when the underlying states of nature form a compact metrizable space. In that case, we show (in [Appendix 3](#)) there exists a canonical type space inducing all (suitably coherent) hierarchies of beliefs. This result is of independent interest and opens the door to epistemic analyses of large games.

To illustrate the practicality of our approach, we provide two examples. First, we give a large game version of the coordinated attack problem and show how phenomena familiar from the electronic mail game of [Rubinstein \(1989\)](#) can occur. The application shows how one can embed the usual equilibrium framework with private information coming from a common prior in our framework. Second, we show that the epistemic condition given by [Morris, Shin, and Yildiz \(2016\)](#) for uniqueness in finite global games in terms of p -common certainty of uniform rank beliefs still works in large global games. In particular, our framework provides exactly the language to make the arguments work in large games.

Grillo, and Penta (2011). A different, more restrictive, notion of rationalizability for games of incomplete information has been proposed by [Ely and Pęski \(2006\)](#). Their solution concept requires a conditional independence condition that is less natural in our distributional framework, in which the correlation of beliefs between players cannot be formalized.

We want to point out how our work relates to some literature we have not mentioned so far. Rationalizability in large games of complete information has been studied by [Jara-Moroni \(2012\)](#), [Yu \(2014\)](#), and [Greinecker \(2017\)](#) in individualistic models with an explicit set of players. Much of the work in these papers is dedicated to establishing topological closure properties that are automatic in our distributional framework.

As mentioned above, rationalizable outcomes in supermodular games are sandwiched between extremal Bayes-Nash equilibria. We rely on methods that have been used to prove the existence of Bayes-Nash equilibria in supermodular games. In particular, we draw on the work of [Van Zandt \(2010\)](#), who studies the existence of interim Bayes-Nash equilibria in supermodular games of incomplete information with finitely many players, and [Balbus, Dziewulski, Reffett, and Woźny \(2015\)](#), who study existence in a distributional model that is similar to ours but much more restrictive in its modeling of uncertainty.

For applications, our toolkit makes it possible to study the rich applications of global games without the common prior assumption. [Izmalkov and Yildiz \(2010\)](#) and [Angeletos, Collard, and Dellas \(2018\)](#) construct specific parametric global games models without common priors; our tools make it possible to study higher-order uncertainty in full generality. Our mentioned adaptation of the uniqueness result of [Morris, Shin, and Yildiz \(2016\)](#) to a continuum of players shows that this can be done on a practical level.

2 Preliminaries

If (X, \mathcal{X}) is a measurable space, we let $\Delta(X)$ be the space of probability measures on (X, \mathcal{X}) endowed with the σ -algebra generated by functions of the form $\mu \mapsto \mu(E)$ with $E \in \mathcal{X}$. If X is Polish (separable and completely metrizable) and \mathcal{X} its Borel σ -algebra, our σ -algebra on $\Delta(X)$ is just the Borel σ -algebra induced by the topology of weak convergence of measures; see [Bertsekas and Shreve \(1978, Proposition 7.25\)](#). Generally, we endow the (at most) countable product of Polish spaces with the product topology, which is again Polish, and endow every Polish space with its Borel σ -algebra, and we endow products of measurable spaces with the product σ -algebra. These conventions are consistent. If X and Y are nonempty Polish spaces and ν is a Borel probability measure on X , we write $\Delta_\nu(X \times Y)$ for the space of Borel probability measures on $X \times Y$ with X -marginal ν . If Y is compact, then $\Delta_\nu(X \times Y)$ is compact too (in the topology of weak convergence). The proof is a simple application of Prohorov's "tightness"-characterization of relative compactness.

3 Large Bayesian games

Our game theoretical model will specify both the fundamentals, given by player characteristics, actions, states of nature, and payoffs, and the beliefs and higher order beliefs of players regarding the state of nature.

The model is a distributional model in which individual players are not explicitly specified; only their population distribution is part of the model. Implicitly, we assume that individual players are insignificant; their behavior does not influence any aggregates. The distributional approach to large games was first introduced by [Mas-Colell \(1984\)](#). Each player's fundamentals are specified by their characteristics. There is a Polish space C of *characteristics* whose population distribution is given by a Borel probability measure ν on C . All actions are included in a compact metrizable action space A . Not all actions need to be available to all players. There is an *action correspondence* $\mathcal{A} : C \rightarrow 2^A$ that specifies for a player with characteristic c the set $\mathcal{A}(c)$ of actions available to them. An *action profile* is an element of $\Delta_\nu(C \times A)$ supported on the graph of \mathcal{A} . There is also a compact metrizable space S of *states of nature*. To tie everything together, payoffs are encoded in a single *payoff function* $v : C \times A \times S \times \Delta_\nu(C \times A) \rightarrow \mathbb{R}$. In everything that follows, we assume that \mathcal{A} is upper hemicontinuous with nonempty and compact values and that v is continuous in every argument.

Remark 1. One could weaken the continuity assumption by merely assuming that \mathcal{A} is measurable and by requiring ν only to be measurable in C . This would not be more general; there would still exist a Polish topology on C under which \mathcal{A} is upper hemicontinuous (continuous, actually) and for which ν would be continuous. Indeed, one can take under the weaker assumption $\mathcal{A} : C \rightarrow 2^A$ to be a measurable function for the Hausdorff topology on the space of nonempty compact subsets of A by [Aliprantis and Border \(2006, Theorem 18.10\)](#) and identify ν with a measurable function $u : C \rightarrow C[A \times S \times \Delta_\nu(C \times A)]$ given by $u(c)(a, s, \mu) = v(c, a, s, \mu)$ by [Aliprantis and Border \(2006, Theorem 4.55\)](#). Now by [Aliprantis and Border \(2006, Theorem 4.59\)](#), there exists a finer Polish topology on C that introduces no new Borel sets such that the two measurable functions \mathcal{A} and u are continuous. Under this finer topology, \mathcal{A} is a continuous correspondence by [Aliprantis and Border \(2006, Theorem 17.15\)](#). It is also easy to see that v is continuous in all arguments if u is continuous.

To complete the specification of the Bayesian game, we have to specify beliefs and higher-order beliefs. We do this at the interim stage in terms of type spaces. For technical convenience, we follow [Böge and Eisele \(1979\)](#) and [Mertens and Zamir \(1985\)](#) and use spaces of states of the world that have no inherent product structure. A *type space* is a triple (T, σ, τ) with T a nonempty compact metrizable space of *states of the world* and two continuous functions $\sigma : T \rightarrow S$ and $\tau : T \rightarrow \Delta_\nu(C \times \Delta(T))$. The idea is that a state of the world specifies ev-

everything relevant in the model. In particular, it specifies the state of nature via the function σ . It also specifies the joint distribution of characteristics and beliefs. Since beliefs should be specified on everything relevant, they are specified for states of the world. Note that τ has values in $\Delta_V(C \times \Delta(T))$ and not in $\Delta_V(C \times T)$. From an element of $\Delta_V(C \times \Delta(T))$ we obtain a regular conditional probability as a function from C to $\Delta(\Delta(T))$. So for each characteristic, there is a distribution over beliefs. In particular, beliefs need not be a function of characteristics; we do not formally include beliefs as part of the characteristics. Note also that we interpret players as individually insignificant, so we need no introspection condition for type spaces; no player has explicit beliefs over their own beliefs.

The assumption that T is compact metrizable (and σ continuous) is essentially without loss of generality. Interim correlated rationalizability, as defined in the next section, is a monotone solution concept in the sense that having more hierarchies of beliefs allows for more interim correlated rationalizable solutions. One can show that a compact metrizable type space exists that includes all hierarchies of beliefs that satisfy minimal coherency conditions, provided that S is compact metrizable. Indeed, each state of the world t induces a unique state of nature $\sigma(s)$. So, a belief about the state of the world induces a first-order belief about S . A state of nature t also induces a population distribution $\tau(t)$ of beliefs about the state of the world and, thus, beliefs about the state of nature. Since we care about how beliefs and states of nature relate, a belief about states of the world induces a second-order joint belief about the state of nature and about the population distribution of first-order beliefs. Since second-order beliefs include first-order beliefs as their S -marginals, these derived first-order beliefs coincide with the actual first-order beliefs. A similar consistency condition applies to all higher-order beliefs, too; this is the condition that hierarchies of beliefs, the whole sequence of all higher-order beliefs, are *coherent*. Moreover, the hierarchies of beliefs induced by a type space can be identified with joint beliefs over the state of nature and coherent hierarchies of beliefs that, in turn, only consider coherent hierarchies of beliefs. And so on. We show in Appendix 3 that a canonical type space exists that includes all hierarchies of beliefs satisfying the restrictions imposed by such coherency conditions and, therefore, all hierarchies of beliefs occurring in any type space. If S is compact metrizable, so is the canonical type space.

4 Interim correlated rationalizability

Our first solution concept is interim correlated rationalizability. As we will see, there are two equivalent ways to define interim correlated rationalizability. Our official definition is based on iterative elimination of actions that are not best replies to beliefs that put no mass on any previously eliminated actions. The important property of interim correlated rationalizability is that otherwise

arbitrary correlation between action choices and states of nature are allowed. The restrictions on actions are given by a correspondence $\mathcal{S} : C \times \Delta(T) \rightarrow 2^A$. We can think of \mathcal{S} as a theory of behavior that specifies for each characteristic of a player and for each belief of that player what actions they could play. Provided \mathcal{S} has a measurable graph, we get a well-defined correspondence $\mathcal{D}_{\mathcal{S}} : T \rightarrow 2^{\Delta_v(C \times \Delta(T) \times A)}$ by specifying

$$\mathcal{D}_{\mathcal{S}}(t) = \left\{ \kappa \in \Delta_v(C \times \Delta(T) \times A) \mid \kappa \text{ is supported on the graph of } \mathcal{S} \right. \\ \left. \text{and has } C \times \Delta(T)\text{-marginal } \tau(t) \right\}.$$

We can interpret $\mathcal{D}_{\mathcal{S}}(t)$ as the space of all belief-action profiles feasible at the state of the world t if everyone behaves according to \mathcal{S} .

Lemma 1. If \mathcal{S} is an upper hemicontinuous correspondence with nonempty compact values, then so is $\mathcal{D}_{\mathcal{S}}$.

Let $\phi : \Delta_v(C \times \Delta(T) \times A) \rightarrow \Delta_v(C \times A)$ be the canonical marginal-function. If $\mathcal{D} : T \rightarrow 2^{\Delta_v(C \times \Delta(T) \times A)}$ has a measurable graph, we get a well-defined correspondence $\mathcal{S}_{\mathcal{D}} : C \times \Delta(T) \rightarrow 2^A$ by specifying

$$\mathcal{S}_{\mathcal{D}}(c, \beta) = \left\{ a \in A \mid \text{there is some } \mu \in \Delta(T \times \Delta_v(C \times \Delta(T) \times A)) \text{ with} \right. \\ \left. T\text{-marginal } \beta \text{ and supported on the graph of } \mathcal{D} \right. \\ \left. \text{such that } a \in \operatorname{argmax}_{A(c)} \int v(c, a, \sigma(t), \phi(\kappa)) \, d\mu(t, \kappa) \right\}.$$

Intuitively, $\mathcal{S}_{\mathcal{D}}(c, \beta)$ is the set of actions a rational player with characteristic c and belief β can play when being certain that everyone behaves according to \mathcal{D} .

Lemma 2. If \mathcal{D} is an upper hemicontinuous correspondence with nonempty compact values, then so is $\mathcal{S}_{\mathcal{D}}$.

We now define recursively two sequences of correspondences $\langle \mathcal{S}_m \rangle$ with $\mathcal{S}_m : C \times \Delta(T) \rightarrow 2^A$ and $\langle \mathcal{D}_m \rangle$ with $\mathcal{D}_m : T \rightarrow 2^{\Delta_v(C \times \Delta(T) \times A)}$. We let $\mathcal{S}_0(c, \beta) = A(c)$ for all (c, β) and let $\mathcal{D}_m = \mathcal{D}_{\mathcal{S}_m}$. Given that \mathcal{D}_m is defined, we define \mathcal{S}_{m+1} by $\mathcal{S}_{m+1} = \mathcal{S}_{\mathcal{D}_m}$. It follows from Lemma 1 and Lemma 2 that \mathcal{S}_m and \mathcal{D}_m are upper hemicontinuous correspondence with nonempty compact values. We let $\mathcal{S} : C \times \Delta(T) \rightarrow 2^A$ be defined by $\mathcal{S}(c, \beta) = \bigcap_m \mathcal{S}_m(c, \beta)$. It follows readily that \mathcal{S} is also an upper hemicontinuous correspondence with nonempty compact values from [Aliprantis and Border \(2006, Theorem 17.25\)](#). The correspondence \mathcal{S} represents our solution concept of *interim correlated rationalizability*. The action a is *interim correlated rationalizable* for a player with characteristics c and belief β if $a \in \mathcal{S}(c, \beta)$.

We now provide an alternative characterization of interim correlated rationalizability. An upper hemicontinuous nonempty and compact valued correspon-

dence $\mathcal{R} : C \times \Delta(T) \rightarrow 2^A$ *rationalizes itself* if $\mathcal{R} = \mathcal{S}_{\mathcal{D}_{\mathcal{R}}}$. Intuitively, a theory of behavior rationalizes itself if everyone who believes that everyone behaves according to the theory behaves rationally by following the theory.

Theorem 1. \mathcal{S} is the largest correspondence that rationalizes itself.

Note that interim correlated rationalizability is a monotone solution concept in that larger type spaces admit more interim correlated rationalizable solutions. Consequently, it is without loss of generality to work with the canonical type space that includes all combinations of a state of nature and a population distributions of hierarchies of beliefs from any type space. Further compactness and continuity requirements on types spaces we have imposed are, therefore, innocent.

5 Strategic complementarities and the interim Bayes-Nash equilibrium

We now impose more structure on the Bayesian games we study. We let \succeq be a lattice ordering on A under which the join and meet are continuous operations. It follows that (A, \succeq) is a complete lattice with a closed graph; see [Reny \(2011\)](#) for the details. We use \succeq also to denote the stochastic order on $\Delta_v(C \times A)$ and other closely related orders derived from it. This abuse of notation should not lead to confusion. (Background material on stochastic orderings is provided in Appendix 1).

We make the following additional assumptions:

- (i) $\mathcal{A}(c)$ is a sublattice of A for each $c \in C$ (closed under the original join and meet operators).
- (ii) The function v is supermodular in A , so for each $a, a' \in A$, $c \in C$, $s \in S$, and $\mu \in \Delta_v(C \times A)$

$$v(c, a \vee a', s, \mu) + v(c, a \wedge a', s, \mu) \geq v(c, a, s, \mu) + v(c, a', s, \mu).$$

- (iii) The function v has increasing differences in $A \times \Delta_v(C \times A)$, so for each $a \succeq a'$, the function

$$\mu \mapsto v(c, a, s, \mu) - v(c, a', s, \mu)$$

is \succeq -nondecreasing.

Note that $\Delta_v(C \times A)$ need not be a lattice even when A is; see [Kamae, Krengel, and O'Brien \(1977, page 901\)](#) for an example.

Under the new assumptions, we can show that there exists a largest and a smallest interim Bayes-Nash equilibrium and that every agent's interim correlated rationalizable actions are sandwiched in between.

For a correspondence ϕ with values in A , we let $\bigvee \phi$ and $\bigwedge \phi$ be the functions with values in A given by pointwise suprema and infima, respectively.

A strategy profile specifies for each characteristic of an agent and each of their beliefs the conditional distribution over actions. A symmetric strategy profile will be a strategy profile in which everyone with the same type and same belief will choose the same action. So a *symmetric strategy profile* is a measurable function $\zeta : C \times \Delta(T) \rightarrow A$ such that $\zeta(c, \beta) \in \mathcal{A}(c)$ for all c . In the following, we will always mean symmetric strategy profiles if we talk about a strategy profile.

If ζ is a strategy profile, we let $\zeta^* : \Delta_v(C \times \Delta(T)) \rightarrow \Delta_v(C \times \Delta(T) \times A)$ be the induced function given by

$$\zeta^*(\kappa)(B) = \int 1_B(c, \beta, \zeta(c, \beta)) \, d\kappa(c, \beta)$$

for each Borel set $B \subseteq C \times \Delta(T) \times A$.

The strategy profile ζ is an (*interim*) *Bayes-Nash equilibrium* if for all $c \in C$ and $\beta \in \Delta(T)$, $\zeta(c, \beta)$ is an element of

$$\operatorname{argmax}_{\mathcal{A}(c)} \int v(c, \cdot, s, \mu) \, d\beta \circ (\sigma, \phi \circ \zeta^* \circ \tau)^{-1}.$$

Theorem 2. The functions $\bigvee \mathcal{S}$ and $\bigwedge \mathcal{S}$ are both Bayes-Nash equilibria.

We defined Bayes-Nash equilibrium only for symmetric strategy profiles for the sake of convenience. Weakening the symmetry requirement would not change the result, but involve additional notational complications. It should be noted that the proof of Theorem 2 relies on methods from [Van Zandt \(2010\)](#) and [Balbus, Dziewulski, Reffett, and Woźny \(2015\)](#).

Remark 2. We work with large games in distributional form, but Theorem 2 also implies an existence theorem for large games with an explicit probability space of individual players. This is due to our equilibria being symmetric. Each individualistic large game induces a canonical distributional game. Taking the composition of the function from types to characteristics and a symmetric equilibrium provides us with a pure strategy equilibrium for the individualistic game.

6 The canonical type space

Throughout, we have assumed that T is compactly metrizable. Interim correlated rationalizability is a monotone solution concept in the sense that having more hierarchies of beliefs allows for more interim correlated rationalizable solutions. We now show that a compact metrizable type space exists that includes

all hierarchies of beliefs that satisfy minimal coherency conditions, provided that S is compact metrizable.

Let (T, σ, τ) be a type space. Each state of the world t induces a unique state of nature $\sigma(s)$. So, a belief about the state of the world induces a first-order belief about S . A state of the world t also induces a population distribution $\tau(t)$ of beliefs about the state of the world and, thus, beliefs about the state of nature. Since we care about how beliefs and states of nature relate, a belief about states of the world induces a second-order joint belief about the state of nature and about the population distribution of first-order beliefs. Since second-order beliefs include first-order beliefs as their S -marginals, these derived first-order beliefs coincide with the actual first-order beliefs. A similar consistency condition applies to all higher-order beliefs too; this is the condition that hierarchies of beliefs, the whole sequence of all higher-order beliefs, are *coherent*. Moreover, the hierarchies of beliefs induced by a type space can be identified with joint beliefs over the state of nature and coherent hierarchies of beliefs that, in turn, only consider coherent hierarchies of beliefs. And so on. We show in Appendix 3 that a canonical type space exists that includes all hierarchies of beliefs satisfying the restrictions imposed by such coherency conditions and, therefore, all hierarchies of beliefs occurring in any type space. If S is compact metrizable, so is the canonical type space.

7 Example: Coordinated attack

There is a continuum of players (normalized to 1) that need to coordinate to attack one of two spots (risky or safe). There are hence two actions, where $\alpha = 1$ denotes a decision to attack the risky and $\alpha = 0$ the safe spot. There are two states of nature determining the game that will be played. If $s = 0$ the payoffs are as follows:

$$\begin{cases} M(1 - \vartheta) & \text{if } \alpha = 0 \\ -L(1 - \vartheta) & \text{if } \alpha = 1, \end{cases}$$

where ϑ denotes the fraction of players choosing $\alpha = 1$ and M and L are parameters that satisfy $L > M > 0$. If $s = 1$ the payoffs are as follows:

$$\begin{cases} 0 & \text{if } \alpha = 0 \\ M\vartheta - L(1 - \vartheta) & \text{if } \alpha = 1. \end{cases}$$

For each state, players play a coordination game. Indeed, the marginal payoff from taking action 1 (versus 0) increases with ϑ for each state. Each game has two symmetric Nash equilibria: the greatest one, where all players attack the risky spot ($\vartheta^* = 1$) and the least one, in which none of players attacks the risky spot ($\vartheta^* = 0$). Players do not observe the state. There is a common prior with the probability that $s = 1$ given by π . For each $\pi \in [0, 1]$, again there are two Nash equilibria: $\vartheta^* = 1$ and $\vartheta^* = 0$ of the incomplete information game.

Every player has a position on the circle of unit circumference $[0, 1)$. Players at position 0 are perfectly informed of the true state and receive a positive signal if the state is 1. The signal travels along the circle, potentially passing the same position several times, until it dies at some random time. The mechanism is analogous to the one used by Rubinstein (1989). When we refer to the number of signals a player receives, we mean the number the signal passed them by.

Formally, we let $C = [0, 1)$ and ν be the uniform distribution. The only relevant characteristic of a player is their position. We take $S = A = \{0, 1\}$. The payoff function v is given by

$$v(c, a, s, \mu) = \begin{cases} M\mu_A(\{0\}) & \text{if } s = 0, a = 0 \\ -L\mu_A(\{0\}) & \text{if } s = 0, a = 1 \\ 0 & \text{if } s = 1, a = 0 \\ M\mu_A(\{1\}) - L\mu_A(\{0\}) & \text{if } s = 1, a = 1, \end{cases}$$

with μ_A denoting the A -marginal of $\mu \in \Delta_\nu(C \times A)$.

We let $T = \{0, 0\} \cup \{1\} \times [0, \infty]$. The second coordinate denotes the time when the signal dies. If this time is ∞ , the signal never dies, a situation corresponding to common certainty of the state of nature. The function σ is simply the projection onto S . The function τ will be defined below.

We assume the time the signal dies follows an exponential distribution (analogous to the geometric distribution used by Rubinstein) with intensity parameter $\alpha > 0$. The corresponding cumulative distribution function is given for nonnegative x by

$$F(x) = 1 - e^{-\alpha x}.$$

The expected time the signal dies is $1/\alpha$.

Next, we specify τ . A player's belief will be fully determined by the player's position and the time the signal dies. If a player positioned at i receives no signal, then either the state of nature is 0 or the state of nature is 1, but the signal died before reaching i . The conditional probability of latter happening is given by

$$\pi_i = \frac{\pi(1 - e^{-\alpha i})}{1 - \pi + \pi(1 - e^{-\alpha i})}.$$

The belief of a player who received no signal that the signal died before x conditional on the state of nature being 1 is given by

$$\frac{1 - e^{-\alpha x}}{1 - e^{-\alpha i}} \quad \text{for } x \in (0, i).$$

If a player receives at least one signal, then they are certain that the state of nature is 1. If a player positioned at i receives exactly $n > 0$ signals, the signal must have died in the interval $(ni, ni + 1)$. The conditional belief of such a player that the signal died after ni but before $ni + x$ is given by

$$\frac{1 - e^{-\alpha x}}{1 - e^{-\alpha}} \quad \text{for } x \in (0, 1).$$

These conditional probabilities are enough to specify the function τ .

The following proposition mirrors the central result of [Rubinstein \(1989\)](#).

Proposition 1. The equilibrium in which all players always play 0 is the only equilibrium in which everyone who did not observe a signal plays 0.

Proof. Suppose there were an equilibrium in which everyone who did not observe a signal plays 0 but in which action 1 is sometimes played. We say t is *active* if the player $t \bmod 1$ plays 1 after observing $\lfloor t \rfloor + 1$ signals. Let t^* be the essential infimum of all active times. By assumption, $t^* < \infty$. Let $i^* = t^* \bmod 1$. If i^* receives $\lfloor t^* \rfloor + 1$ signals, they know that the only players that might play 1 are those that have a larger position and received the same number or signals or a lower position and received one more signal. The expected fraction of such players is

$$\int_0^1 x \left(\frac{1 - e^{-\alpha x}}{1 - e^{-\alpha}} \right)' dx = \int_0^1 x \frac{\alpha e^{-\alpha x}}{1 - e^{-\alpha}} dx = \frac{1}{\alpha} - \frac{1}{e^{\alpha} - 1}.$$

Taking action 1 can only be optimal if $M\vartheta - (1 - L)\vartheta \geq 0$ or, equivalently

$$\vartheta \geq \frac{L}{M + L}.$$

We must have, therefore,

$$\frac{1}{\alpha} - \frac{1}{e^{\alpha} - 1} \geq \frac{L}{M + L} > \frac{1}{2}.$$

But the function

$$\alpha \mapsto \frac{1}{\alpha} - \frac{1}{e^{\alpha} - 1}$$

is strictly decreasing on the strictly positive real line has limit $1/2$ as α decreases to 0. So this inequality can never be satisfied. \square

8 Example: Equilibrium uniqueness in global games

We show here that the arguments of [Morris, Shin, and Yildiz \(2016\)](#) for uniqueness in finite global games in terms of higher order beliefs can be formulated in our setting; our framework is flexible enough.

In our example, characteristics play no role and are notationally suppressed. The action space is binary, $A = \{0, 1\}$ (“noninvest” and “invest”) and the states of nature is given by $S = [-1, 2]$. We write ϑ for the fraction of players playing 1. Payoff functions are given, by slight abuse of notation, by

$$v(a, s, \vartheta) = a \cdot (s + \vartheta - 1).$$

Under this specification, the game is supermodular. In particular, there is a largest and a smallest equilibrium, given by $\bigvee S : \Delta(T) \rightarrow A$ and $\bigwedge S : \Delta(T) \rightarrow$

A, respectively, and every rationalizable solution lies in between by Theorem 2. Since there are no proper characteristics, the extremal equilibria specify what an agent with a given belief might do. In particular, we are able to discuss when a belief gives rise to a unique rationalizable action, even in settings in which not everyone has a unique rationalizable action. This distinction is somewhat buried when there are only two players but is clear in our setting.

Our game is linear, and we make heavy use of the fact that usually only averages matter. For example, we let $x_\beta = x(\beta) = \int \sigma \, d\beta$ be the expected state under the belief β . Also, we let the expected fraction of players with a belief in $E \subseteq \Delta(T)$ under the belief β be given by $F_\beta(E) = \int \tau(E) \, d\beta$. For each player with belief β , the expected fraction of players that are less optimistic about the state of nature than them is their *rank*, formally defined as

$$R(\beta) = F_\beta(\{\beta' \in \Delta(T) \mid x_{\beta'} \leq x_\beta\}).$$

An important role in our uniqueness result will be played by the set beliefs of players who believe their rank to be close to the median rank. For $\epsilon > 0$, let

$$\text{URB}_\epsilon = \{\beta \in \Delta(T) \mid 1/2 - \epsilon \leq R(\beta) \leq 1/2 + \epsilon\}.$$

Another important set is the set of beliefs for which investing is “ ϵ -strictly risk dominant,” given by

$$\text{SRD}_\epsilon = \{\beta \in \Delta(T) \mid x_\beta > 1/2 + \epsilon\}.$$

Dually,

$$\text{nSRD}_\epsilon = \{\beta \in \Delta(T) \mid x_\beta < 1/2 - \epsilon\}.$$

is the set of beliefs for which noninvesting is strictly risk dominant. Our last ingredient is suitable (approximate) belief operators and certainty operators. For reasons we will explain below, we define these operators on $\Delta(T)$. For a measurable function $f : \Delta(T) \rightarrow \mathbb{R}$, define a corresponding belief operator by

$$B_f(E) = \{\beta \in E \mid F_\beta(E) \geq f(\beta)\}.$$

In particular, for $f = 1 - x$, we have

$$B_{1-x}(E) = \{\beta \in E \mid F_\beta(E) \geq 1 - x(\beta)\},$$

and for $f = x$, we have

$$B_x(E) = \{\beta \in E \mid F_\beta(E) \geq x(\beta)\}.$$

We identify a number p with the corresponding constant function $\Delta(T)$ when writing B_p . For each operator B_f , we recursively define its iterates by $B_f^1(E) = B_f(E)$ and $B_f^{n+1} = B_f(B_f^n(E))$, and the corresponding certainty operator by $C_f(E) = \bigcap_n B_f^n(E)$. We can use the operators C_{1-x} and C_x to characterize which actions could be played under a given belief.

Remark 3. These belief and certainty operators operate on Borel subsets of $\Delta(T)$ and not on Borel subsets of T , which might seem more natural. Belief operators usually satisfy an introspection condition. In our framework, there is no difference between believing that everyone believes something and believing that everyone else believes something. However, we need the former to hold, and that is why we define the belief and certainty operators the way we do. In general, one could define belief operators that operate on both beliefs and sets of states of the world simultaneously to achieve greater generality. In this application, there is no need for such generality.

Lemma 3. Investment can be optimal under β , $\bigvee \mathcal{S}(\beta) = 1$, if and only if $\beta \in C_{1-x}(\Delta(T))$. Similarly, noninvestment can be optimal under β , $\bigwedge \mathcal{S}(\beta) = 0$, if and only if $\beta \in C_x(\Delta(T))$.

Proof. We do the first case; the second case is analogous. Indeed, $\beta \in B_{1-x}(\Delta(T))$ is equivalent to $x_\beta \geq 0$, so investment is optimal if everyone else invests too. If everyone with beliefs in $B_{1-x}^n(\Delta(T))$ invests, then

$$B_{1-x}^{n+1}(\Delta(T)) = \left\{ F_\beta(B_{1-x}^n(\Delta(T))) \geq 1 - x(\beta) \right\}$$

contains the beliefs under which investment is optimal. \square

Importantly, since every player must either invest or not, investment is uniquely rationalizable if $\beta \notin C_x(\Delta(T))$, and noninvestment is uniquely rationalizable if $\beta \notin B_{1-x}(\Delta(T))$.

For a measurable set $E \subseteq \Delta(T)$ and $\epsilon \geq 0$, let

$$x_\epsilon^{**}(E) = \sup\{x_\beta \mid \beta \in E, x_\beta \leq R(\beta) + \epsilon\}$$

and

$$x_{**}^\epsilon(E) = \inf\{x_\beta \mid \beta \in E, x_\beta \geq R(\beta) - \epsilon\}.$$

Here, $x_\epsilon^{**}(E)$ is the highest value (of the state) for which any type within E has a value exceeding its rank by at most ϵ . Similarly, $x_{**}^\epsilon(E)$ is the lowest value (of the state) for which any type within E has a rank exceeding the corresponding value by at most ϵ .

Lemma 4. Let $E \subseteq \Delta(T)$ be closed and $E \subseteq B_p(E)$. Investment is uniquely rationalizable for any $\beta \in E$ such that

$$x_\beta > x_{1-p}^{**}(E).$$

Similarly, noninvestment is uniquely rationalizable for any $\beta \in E$ such that

$$x_\beta < x_{**}^{1-p}(E).$$

Proof. Again, we do the first case. By the last lemma, we have to show that the condition implies $\beta \notin C_x(\Delta(T))$. This is trivially the case if $C_x(\Delta(T)) \cap E = \emptyset$, so we can assume that $C_x(\Delta(T)) \cap E \neq \emptyset$. We show that $\{x_\beta \mid \beta \in C_x(\Delta(T)) \cap E\}$ has a maximum. Since the function $\beta \mapsto x_\beta$ is continuous and $\Delta(T)$ compact, it suffices to show that $C_x(H)$ is closed if H is, which reduces to showing that $B_x(H)$ is closed if H is. This, in turn, reduces to showing that $\beta \mapsto F_\beta(H)$ is upper-semicontinuous if H is closed, which follows from the Portmanteau theorem.

So let $\hat{\beta}$ be a maximizer of x on $C_x(\Delta(T)) \cap E$ and $\hat{x} = x_{\hat{\beta}}$ the corresponding value. By construction, if $x_\beta > \hat{x}$, then $\beta \notin C_x(\Delta(T)) \cap E$. If we know that $\beta \in E$, this means $\beta \notin C_x(\Delta(T))$ and investment is uniquely rationalizable at β . It suffices, therefore, to show that $x_{1-p}^{**}(E) \geq \hat{x}$.

Moreover,

$$\begin{aligned} F_{\hat{\beta}}(C_x(\Delta(T))) &= F_{\hat{\beta}}(C_x(\Delta(T)) \cap E) + F_{\hat{\beta}}(C_x(\Delta(T)) \setminus E) \\ &\leq F_{\hat{\beta}}(C_x(\Delta(T)) \cap E) + F_{\hat{\beta}}(\Delta(T) \setminus E) \\ &= F_{\hat{\beta}}(C_x(\Delta(T)) \cap E) + (1 - F_{\hat{\beta}}(E)) \\ &\leq F_{\hat{\beta}}(C_x(\Delta(T)) \cap E) + (1 - p), \end{aligned}$$

where the last inequality uses that $\hat{\beta} \in E$ and $E \subseteq B_p(E)$, which implies $F_{\hat{\beta}}(E) \geq p$. Now, using the fact that $\hat{\beta}$ is a maximizer, we get

$$\begin{aligned} F_{\hat{\beta}}(C_x(\Delta(T)) \cap E) + (1 - p) &\leq F_{\hat{\beta}}(\{\beta \in \Delta(T) \mid x_\beta \leq \hat{x}\}) + (1 - p) \\ &= R(\hat{\beta}) + (1 - p). \end{aligned}$$

Note that $B_x(C_x(\Delta(T))) = C_x(\Delta(T))$. Since $\hat{\beta} \in C_x(\Delta(T))$ and by the above inequalities:

$$\hat{x} = x_{\hat{\beta}} \leq R(\hat{\beta}) + (1 - p).$$

We hence have:

$$\hat{x} \in \{x_\beta \mid \beta \in \Delta(X), x_\beta \leq R(\beta) + (1 - p)\},$$

which implies

$$\hat{x} \leq \sup \{x_\beta \mid \beta \in E, x_\beta \leq R(\beta) + (1 - p)\} = x_{1-p}^{**}(E).$$

□

Proposition 2. If $C_{1-\epsilon}(\text{URB}_\epsilon)$ is closed, then investing is uniquely rationalizable for all

$$\beta \in \text{SRD}_{2\epsilon} \cap C_{1-\epsilon}(\text{URB}_\epsilon).$$

If $C_{1-\epsilon}(\text{URB}_\epsilon)$ is closed, then not investing is uniquely rationalizable for all

$$\beta \in \text{nSRD}_{2\epsilon} \cap C_{1-\epsilon}(\text{URB}_\epsilon).$$

Proof. Since, $C_{1-\epsilon}(\text{URB}_\epsilon) = B_{1-\epsilon}(C_{1-\epsilon}(\text{URB}_\epsilon))$, by Lemma 4, it suffices to show that $x_\beta > x_\epsilon^{**}(C_{1-\epsilon}(\text{URB}_\epsilon))$ for all β in the relevant intersection.

By definition, $C_{1-\epsilon}(\text{URB}_\epsilon) \subseteq \text{URB}_\epsilon$. Therefore,

$$x_\epsilon^{**}(C_{1-\epsilon}(\text{URB}_\epsilon)) \leq x_\epsilon^{**}(\text{URB}_\epsilon) = \sup\{x_\beta \mid \beta \in \text{URB}_\epsilon, x_\beta \leq R(\beta) + \epsilon\} \leq 1/2 + 2\epsilon.$$

Now,

$$\text{SRD}_{2\epsilon} = \{\beta \in \Delta(T) \mid x_\beta > 1/2 + 2\epsilon\},$$

so

$$\beta \in \text{SRD}_{2\epsilon} \cap C_{1-\epsilon}(\text{URB}_\epsilon)$$

implies $x_\beta > x_\epsilon^{**}(C_{1-\epsilon}(\text{URB}_\epsilon))$ and we are done. \square

Remark 4. As mentioned above, we have defined belief and certainty operators on $\Delta(T)$ instead of T . In our large game model, there is no difference between a player's belief about what all players believe and what all other players believe. In particular, the model doesn't directly express introspective beliefs; beliefs about a player's own beliefs. But it is exactly such introspective beliefs that occur in the relevant definition. For example, rank beliefs are about how a player ranks the optimism of their own beliefs relative to the rest of the population.

9 Discussion

In this paper, we have formulated large games of incomplete information and a suitable notion of interim correlated rationalizability, showed that the model is not overly restrictive by constructing the canonical type space, and showed that under strategic complementarities, all rationalizable outcomes are bracketed by extremal equilibria. As a byproduct, we have an existence result for pure-strategy equilibria in large games with nontrivial aggregate uncertainty and strategic complementarities. The model is flexible enough to subsume traditional approaches with a common prior, as we do in our coordinated attack problem, and to adapt interim arguments in terms of beliefs, as we do in our adaption of the uniqueness result of Morris, Shin, and Yildiz (2016). In this section, we want to discuss some technical aspects of our approach as well as some wider applications.

In our formulation, all players are implicitly negligible and small. There are applications in which it is natural to have large players such as governments and central banks, alongside the small players. Including such large players in the model would require a lot more notation, but would not give rise to any technical problems. We can embed a large player as an isolated point in the space of characteristics C . If $c \in C$ represents a large player, an element $\mu \in \Delta_v(C \times A)$ that represents a feasible joint distribution over types and actions has to satisfy $v(c) = \mu(c, a)$ for some $a \in A$ (abusing notation, we identify singletons and their elements here). To show that such conditions give rise to a closed subset, one can adapt the proof of [Levy \(2024, Proposition 7.2\)](#). For the construction of the canonical type space, one would need to impose a familiar introspection condition for large players.

We have assumed that A is compact to ensure that $\Delta_v(C \times A)$ is compact. However, we only need that the elements of $\Delta_v(C \times A)$ supported on the graph of \mathcal{A} are compact. If A is merely Polish but \mathcal{A} compact-valued, this holds by Prohorov’s characterization of relative compactness. If $K \subseteq C$ is a compact set such that $v(K) > 1 - \epsilon$, then $K \times \mathcal{A}(K)$ is compact, because \mathcal{A} is upper hemicontinuous and compact-valued, and $\mu(K \times \mathcal{A}(K)) > 1 - \epsilon$ for all $\mu \in \Delta_v(C \times A)$ supported on the graph of \mathcal{A} .

A more substantive restriction of our model is that it is entirely static. Global games have been extended to dynamic models such as the dynamic regime change model of [Angeletos, Hellwig, and Pavan \(2007\)](#). Extending our work to a dynamic setting would bring new technical and conceptual challenges. On the conceptual side, dynamic versions of rationalizability usually incorporate some form of forward or backward induction. On the technical side, dynamic games make it impossible to ignore probability-zero events completely. The existence of extremal equilibria in dynamic supermodular games with purely idiosyncratic uncertainty has been proven by [Balbus, Dziewulski, Reffett, and Woźny \(2022\)](#). Our approach allows for a much more flexible treatment of beliefs.

An important ingredient of our model is our model of beliefs, and we think it should be useful in wider settings. For example, [Morris \(1995\)](#) and [Ben-Porath and Heifetz \(2011\)](#) provide explicit foundations in terms of beliefs for competitive equilibria in economies with incomplete information. In particular, [Ben-Porath and Heifetz \(2011\)](#) study a competitive model with a continuum of agents. However, they use type spaces with a product structure, and this, together with the requirement of measurability, imposes economically overly restrictive type-symmetry conditions on beliefs. Our framework allows for a much more flexible treatment of beliefs.

There is an extensive literature on the robustness of equilibria to perturbations of higher-order beliefs; [Kajii and Morris \(1997\)](#) introduced the general program with ex-ante perturbations and [Weinstein and Yildiz \(2007\)](#) studied interim

perturbations. In our large player setting, perturbations can relate jointly to players and hierarchies of beliefs; topologies on individual hierarchies of beliefs are not sufficient to study perturbations of profiles of hierarchies of beliefs.

10 Appendix 1: Stochastic orderings

Let X be a Polish space and \succeq be a partial order on X with a closed graph. We can identify X with a closed subspace of $\Delta(X)$ by the embedding $x \mapsto \delta_x$. We can extend \succeq to all of $\Delta(X)$ by letting $\mu \succeq \mu'$ if $\int g d\mu \geq \int g d\mu'$ for every bounded measurable \succeq -nondecreasing function $g : X \rightarrow \mathbb{R}$. The following are equivalent:

1. $\mu \succeq \mu'$.
2. There exists $\lambda \in \Delta(X \times X)$ supported on the graph of \succeq such that the first marginal equals μ and the second marginal equals μ' .
3. $\mu(U) \geq \mu'(U)$ for every closed set $U \subseteq X$ that contains with every element also all \succeq -larger elements.
4. There exists a probability space and random variables m and m' defined on it with values in X such that m has distribution μ , m' has distribution μ' and $m \succeq m'$ holds almost surely.

For a proof, see [Kamae, Krengel, and O'Brien \(1977, Theorem 1\)](#). The \succeq relation is indeed a partial order on all of $\Delta(X)$; see [Kamae and Krengel \(1978, Theorem 2\)](#) and [Kamae, Krengel, and O'Brien \(1977, Proposition 3\)](#).

If \succeq is an order on X , there is a natural induced partial order on $C \times X$ such that, abusing notation again, $(c, x) \succeq (c', x')$ holds exactly when $c = c'$ and $x \succeq x'$. We will use this order to define stochastic orders on $\Delta_v(C \times X)$.

11 Appendix 2: Proofs

Proof of Lemma 1. We first show that the set

$$D = \left\{ \kappa \in \Delta_v(C \times \Delta(T) \times A) \mid \kappa \text{ is supported on the graph of } \mathcal{S} \right\}$$

is closed. Now $\Gamma(\mathcal{S})$, the graph of \mathcal{S} , is closed by the closed graph theorem. Let d be a bounded metric that metrizes $C \times \Delta(T) \times A$. Then

$$\Gamma(\mathcal{S}) = \left\{ (c, \beta, a) \mid d((c, \beta, a), \Gamma(\mathcal{S})) = 0 \right\}.$$

It follows that

$$D = \left\{ \kappa \in \Delta_v(C \times \Delta(T) \times A) \mid \int d(\cdot, \Gamma(\mathcal{S})) d\kappa = 0 \right\}$$

and this set is closed as the preimage of $\{0\}$ under a continuous function. Now the graph of \mathcal{D}_S is

$$\{(t, \kappa) \in T \times D \mid \kappa \text{ has } C \times \Delta(T)\text{-marginal } \tau(t)\}.$$

This graph is closed since both the “marginal-mapping” and τ are continuous functions. It follows that \mathcal{D}_S is upper hemicontinuous with compact values since $T \times \Delta_v(C \times \Delta(T) \times A)$ is compact. That all values are nonempty is obvious. \square

Proof of Lemma 2. First note that the set

$$\left\{ (c, a, \beta, \mu) \in C \times A \times \Delta(T) \times \Delta(T \times \Delta_v(C \times \Delta(T) \times A)) \mid \right. \\ \left. a \in \operatorname{argmax}_{A(c)} \int v(c, a, \sigma(t), \phi(\kappa)) \, d\mu(t, \kappa), \right. \\ \left. \mu \text{ has } T\text{-marginal } \beta. \right\}$$

is closed because expected utility is jointly continuous (see for example [Aliprantis, Glycopantis, and Puzzello \(2006\)](#)) and the marginal function is continuous. This set is the graph of a correspondence from $C \times \Delta(T)$ to the compact set $A \times \Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ and this correspondence is upper hemicontinuous by the closed graph theorem. Composing with the projection of $A \times \Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ onto A gives us the correspondence \mathcal{S}_D which is therefore upper hemicontinuous by [Aliprantis and Border \(2006, Theorem 17.23\)](#). Since forward images of compact sets under continuous functions are compact, \mathcal{S}_D is compact valued. \square

Proof of Theorem 1. Let $a \in \mathcal{S}(c, \beta)$. For each positive m , there exists $\mu_m \in \Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ with T -marginal β , supported on the graph of \mathcal{D}_{m-1} such that $a \in \operatorname{argmax}_{A(c)} \int v(c, a, \sigma(t), \phi(\kappa)) \, d\mu(t, \kappa)$. Since $\Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ is compact, there exists a subsequence converging to some μ . We claim that μ is supported on the graph of \mathcal{D}_S and that

$$a \in \operatorname{argmax}_{A(c)} \int v(c, a, \sigma(t), \phi(\kappa)) \, d\mu(t, \kappa).$$

Indeed, the latter part follows again from the joint continuity of expected utility.

To see that μ is supported on the graph of \mathcal{D}_S , we note that we only have to check that the $\Delta_v(C \times \Delta(T) \times A)$ -marginal is supported on the graph of \mathcal{S} . Let O_m be the complement of the graph of \mathcal{S}_m . Let μ_m^* and μ^* be the $\Delta_v(C \times \Delta(T) \times A)$ -marginal of μ_m and μ respectively. Clearly, $\mu_n^*(O_m) = 0$ for $n \geq m$. Therefore, $0 = \liminf_n \mu_n^*(O_m) \geq \mu^*(O_m) = 0$ by the Portmanteau theorem. Now $\bigcup_m O_m$ is exactly the complement of the graph of \mathcal{S} and by countable subadditivity, $\mu^*(\bigcup_m O_m) = 0$.

Now, if \mathcal{R} rationalizes itself, it follows from a simple induction argument that $\mathcal{R}(c, \beta) \subseteq \mathcal{S}_m(c, \beta)$ for all (c, β) , so \mathcal{S} is the largest correspondence that rationalizes itself. \square

For the proof of Theorem 2, we need some preliminary work. We define $V : C \times A \times \Delta(S \times \Delta_v(C \times A)) \rightarrow \mathbb{R}$ by

$$V(c, a, \kappa) = \int v(c, a, s, \mu) \, d\kappa(s, \mu).$$

Lemma 5. The function V is supermodular in A and has increasing differences in $A \times \Delta(S \times \Delta_v(C \times A))$.

Proof. Verifying that V is supermodular in A is straightforward. For the rest, assume that $a \succeq a'$ and $\kappa \succeq \lambda$. By the characterization of stochastic dominance in Appendix 1, there exists a probability space and random variables k and l defined on it with values in $S \times \Delta_v(C \times A)$ such that $k \succeq l$ almost surely and such that the distributions of k and l are κ and λ , respectively. By assumption, we have

$$v(c, a, k) - v(c, a', k) \geq v(c, a, l) - v(c, a', l)$$

almost surely. By the change of variables formula for push-forward measures, the integral of $v(c, a, k) - v(c, a', k)$ is $V(c, a, \kappa) - V(c, a', \kappa)$ and the integral of $v(c, a, l) - v(c, a', l)$ is $V(c, a, \lambda) - V(c, a', \lambda)$, so the result follows from integrating the inequality $v(c, a, k) - v(c, a', k) \geq v(c, a, l) - v(c, a', l)$. \square

Lemma 6. For all m , $\bigvee \mathcal{S}_m$ and $\bigwedge \mathcal{S}_m$ are measurable selections of \mathcal{S}_m .

Proof. We do the case of $\bigvee \mathcal{S}_m$ by induction. Note first that for every measurable correspondence W whose values are nonempty compact sublattices of A , the function $\bigvee W$ is a measurable selection of W , see Van Zandt (2010, Lemma 19).

So $\bigvee \mathcal{S}_0$ is a measurable selection of \mathcal{S}_0 since the values of \mathcal{A} are compact sublattices of A . Now assume that $\bigvee \mathcal{S}_m$ is a measurable selection of \mathcal{S}_m . Let $(c, \beta) \in C \times \Delta(T)$ and let $\mu \in \Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ have T -marginal β and be supported on the graph of \mathcal{D}_m . We show that $\mu \circ (\sigma, \phi)^{-1} \in \Delta(S \times \Delta_v(C \times A))$ satisfies

$$\beta \circ \left(\sigma, \phi \circ \bigvee \mathcal{S}_m^* \circ \tau \right)^{-1} \succeq \mu \circ (\sigma, \phi)^{-1}.$$

So let $g : \Delta(S \times \Delta_v(C \times A)) \rightarrow \mathbb{R}$ be a bounded measurable function increasing in $\Delta_v(C \times A)$. Let $r_\mu : T \rightarrow \Delta(\Delta_v(C \times \Delta(T) \times A))$ be a regular conditional probability for μ . We have

$$\int g(s, \kappa) \, d\beta \circ \left(\sigma, \phi \circ \bigvee \mathcal{S}_m^* \circ \tau \right)^{-1} = \int g\left(\sigma(t), \phi\left(\bigvee \mathcal{S}_m^*(\tau(t))\right)\right) \, d\beta(t)$$

and

$$\begin{aligned} \int g(s, \kappa) \, d\mu \circ (\sigma, \phi)^{-1} &= \int g(\sigma(t), \phi(\kappa)) \, d\mu(t, \kappa) \\ &= \iint g(\sigma(t), \phi(\kappa)) \, dr_\mu \, d\beta. \end{aligned}$$

Since μ is supported on the graph of \mathcal{D}_m , $r_\mu(t)$ must be supported on the t -section of this graph for β -almost all t . It follows that $r_\mu(t)$ assigns probability one to elements of $\Delta_v(C \times \Delta(T) \times A)$ with $C \times \Delta(T)$ -marginal $\tau(t)$ that are supported on $\mathcal{S}_m(\tau(t))$ for β -almost all t . But $\bigvee \mathcal{S}_m(\tau(t))$ is at least as large, so

$$\bigvee \mathcal{S}_m^* \circ \tau(t) \succeq \kappa$$

for $r_\mu(t)$ almost all κ . It follows then by integration and the appropriate monotonicity properties of ϕ and g that

$$\int g(s, \kappa) \, d\beta \left(\sigma, \phi \circ \bigvee \mathcal{S}_m^* \circ \tau \right)^{-1} \geq \int g(s, \kappa) \, d\mu \circ (\sigma, \phi)^{-1}.$$

By Lemma 5 and Topkis (1978, Corollary 4.1),

$$\operatorname{argmax}_{\mathcal{A}(c)} v \left(c, \cdot, \beta \circ \left(\sigma, \phi \circ \bigvee \mathcal{S}_m^* \circ \tau \right)^{-1} \right)$$

is a complete sublattice and clearly compact. By Van Zandt (2010, Lemma 19), its supremum, which is a measurable selection and therefore also a measurable selection of \mathcal{S}_{m+1} , is larger than every element of \mathcal{S}_{m+1} . But this means it must equal $\bigvee \mathcal{S}_{m+1}$ and $\bigvee \mathcal{S}_{m+1}$ is a measurable selection of \mathcal{S}_{m+1} . \square

Proof of Theorem 2. We do the first case. We first show that $\bigvee \mathcal{S}$ is a measurable selection of \mathcal{S} . Indeed, since $\langle \mathcal{S}_m \rangle$ is a decreasing sequence of correspondences with nonempty and compact values, and since $\bigvee \mathcal{S}_m$ is a selection of \mathcal{S}_m for each m by Lemma 6, $\lim_{m \rightarrow \infty} \bigvee \mathcal{S}_m(c, \beta)$ exists for all (c, β) and is an element of $\mathcal{S}(c, \beta)$. Since $\bigvee \mathcal{S}_m(c, \beta)$ is an upper bound of $\mathcal{S}(c, \beta)$ for all m and the order \succeq on A has a closed graph, $\bigvee \mathcal{S}(c, \beta) = \lim_{m \rightarrow \infty} \bigvee \mathcal{S}_m(c, \beta)$. It now follows that $\bigvee \mathcal{S}$ is measurable as the pointwise limit of a sequence of measurable functions. So $\bigvee \mathcal{S}$ is a measurable selection of \mathcal{S} .

It remains to show that the strategy profile $\bigvee \mathcal{S}$ is a best response to itself. Since \mathcal{S} rationalizes itself by Theorem 1, there exists for each (c, β) some $\mu \in \Delta(T \times \Delta_v(C \times \Delta(T) \times A))$ with T -marginal β and supported on the graph of $\mathcal{D}_\mathcal{S}$ such that

$$\bigvee \mathcal{S}(c, \beta) \in \operatorname{argmax}_{\mathcal{A}(c)} \int v(c, \cdot, \sigma(t), \phi(\kappa)) \, d\mu(t, \kappa).$$

We show that $\mu \circ (\sigma, \phi)^{-1} \in \Delta(\mathcal{S} \times \Delta_v(C \times A))$ satisfies

$$\beta \circ \left(\sigma, \phi \circ \bigvee \mathcal{S}^* \circ \tau \right)^{-1} \succeq \mu \circ (\sigma, \phi)^{-1}.$$

So let $g : \Delta(S \times \Delta_V(C \times A)) \rightarrow \mathbb{R}$ be a bounded measurable function increasing in $\Delta_V(C \times A)$. Let $r_\mu : T \rightarrow \Delta(\Delta_V(C \times \Delta(T) \times A))$ be a regular conditional probability for μ . We have

$$\int g(s, \kappa) d\beta \circ \left(\sigma, \phi \circ \bigvee S^* \circ \tau \right)^{-1} = \int g\left(\sigma(t), \phi\left(\bigvee S^*(\tau(t))\right)\right) d\beta(t)$$

and

$$\begin{aligned} \int g(s, \kappa) d\mu \circ (\sigma, \phi)^{-1} &= \int g(\sigma(t), \phi(\kappa)) d\mu(t, \kappa) \\ &= \iint g(\sigma(t), \phi(\kappa)) dr_\mu d\beta. \end{aligned}$$

Since μ is supported on the graph of \mathcal{D}_S , $r_\mu(t)$ must be supported on the t -section of this graph for β -almost all t . It follows that $r_\mu(t)$ assigns probability one to elements of $\Delta_V(C \times \Delta(T) \times A)$ with $C \times \Delta(T)$ -marginal $\tau(t)$ that are supported on $S(\tau(t))$ for β -almost all t . But $\bigvee S(\tau(t))$ is at least as large, so

$$\bigvee S^* \circ \tau(t) \succeq \kappa$$

for $r_\mu(t)$ almost all κ . It follows then by integration and the appropriate monotonicity properties of ϕ and g that

$$\int g(s, \kappa) d\beta \circ \left(\sigma, \phi \circ \bigvee S^* \circ \tau \right)^{-1} \geq \int g(s, \kappa) d\mu \circ (\sigma, \phi)^{-1}.$$

Now, for all $a \in \mathcal{A}(c)$,

$$\int v(c, a, \sigma(t), \phi(\kappa)) d\mu(t, \kappa) = \int v(c, a, \cdot) d\mu \circ (\sigma, \phi)^{-1} = V(c, a, \mu \circ (\sigma, \phi)^{-1})$$

by the change of variables formula for pushforward-measures. So

$$\bigvee S(c, \beta) \in \operatorname{argmax}_{\mathcal{A}(c)} V(c, a, \mu \circ (\sigma, \phi)^{-1}).$$

Let

$$a \in \operatorname{argmax}_{\mathcal{A}(c)} V\left(c, a, \beta \circ \left(\sigma, \phi \circ \bigvee S^* \circ \tau \right)^{-1}\right).$$

So by Lemma 5, Topkis (1978, Theorem 6.1), and what we have shown above,

$$a \vee \bigvee S(c, \beta) \in \operatorname{argmax}_{\mathcal{A}(c)} V\left(c, a, \beta \circ \left(\sigma, \phi \circ \bigvee S^* \circ \tau \right)^{-1}\right).$$

Since S rationalizes itself $a \vee \bigvee S(c, \beta) \in S(c, \beta)$. Therefore, $\bigvee S(c, \beta) \succeq a \vee \bigvee S(c, \beta)$. Since we, trivially, also have $a \vee \bigvee S(c, \beta) \succeq \bigvee S(c, \beta)$, we have $a \vee \bigvee S = \bigvee S$ and therefore

$$\bigvee S(c, \beta) \in \operatorname{argmax}_{\mathcal{A}(c)} V\left(c, a, \beta \circ \left(\sigma, \phi \circ \bigvee S^* \circ \tau \right)^{-1}\right),$$

which shows that $\bigvee S$ is a Bayes-Nash equilibrium. \square

12 Appendix 3: The canonical type space

We define coherent types first for abstract hierarchies while ignoring what these are about. We let Z_0, Z_1, \dots be a sequence of nonempty Polish spaces and write $X_n = \prod_{k=0}^n Z_k$ and $Z = \prod_{k=0}^{\infty} Z_k$. A nonempty product space is compact if and only if every factor space is, so Z is compact if and only if each Z_k is. An *abstract type* is an element of $T_0 = \prod_{k=1}^{\infty} \Delta(X_{k-1})$. The abstract type $\langle \delta_k \rangle_{k=1}^{\infty}$ is *coherent* if $\delta_k = \text{marg}_{X_{k-1}} \delta_{k+1}$ for all k . We write T_1 for the subspace of T_1 consisting of all coherent abstract types. The set of Borel probability measures on a Polish space is compact if and only if the underlying Polish space is. Consequently, T_0 is compact if and only if every Z_k is.

Lemma 7. The function $\chi : \Delta(Z) \rightarrow T_1$ given by

$$\chi(\delta) = \langle \text{marg}_{X_{k-1}} \delta \rangle_{k=1}^{\infty}$$

is a homeomorphism.

Proof. Clearly, every element of the range of χ is coherent by construction. The continuity of χ follows from the continuity of the marginal mappings. Since a probability measure on a product space is determined by the finite dimensional marginals (cylinder sets form a π -system), χ is injective. The Kolmogorov extension theorem, Aliprantis and Border (2006, Theorem 15.26), implies that χ is surjective. The continuity of χ^{-1} follows from Bogachev (2018, Corollary 2.4.8). \square

For generic $\langle \delta_k \rangle_{k=1}^{\infty} \in T_1'$, we write δ for the unique probability measure mapped to it by χ .

We let S and C be nonempty Polish spaces, ν a Borel probability measure on C and write $\Delta_{\nu}(C \times Y)$ for the set of all probability measures on $C \times Y$ with C -marginal ν . We now let $Z_0 = X_0 = S$ and define recursively

$$Z_{n+1} = \Delta_{\nu}(C \times \Delta(X_0) \times \dots \times \Delta(X_{n-1})) = \Delta_{\nu}\left(C \times \prod_{k=0}^{n-1} \Delta(X_k)\right).$$

We call abstract types in this setting simply *types*. A type represents an individual agent's hierarchy of beliefs on the population distribution of finite hierarchies of beliefs.

Next, we let $Z'_k = Z_{k+1}$ for all k and $Z' = \prod_{k=0}^{\infty} Z'_k$. T'_0 and T'_1 are defined correspondingly. By Lemma 7, there is a canonical homeomorphism between $\Delta(Z')$ and T'_1 . We can exploit the special structure of T'_1 to prove the following:

Lemma 8. The function

$$\chi' : \Delta_v(C \times T_0) \rightarrow T'_1$$

given by

$$\chi'(\lambda) = \left\langle \text{marg}_{C \times \prod_{j=0}^{k-1} \Delta(X_j)} \lambda \right\rangle_{k=1}^{\infty}$$

is a homeomorphism.

Proof. The proof follows the proof of Lemma 7 almost verbatim, the only difference is in how one applies the Kolmogorov extension theorem. \square

Lemma 8 allows us to identify T'_1 with a population distribution over types. These types need not be coherent. We want to incorporate the restrictions that the types are coherent, that they are certain that the population distribution of types consists only of coherent types, and so on. Dealing with this issue is the next step. For generic $\langle \lambda_k \rangle_{k=1}^{\infty} \in T'_1$, we write λ for the unique probability measure mapped to it by χ' . For a measurable set $E \subseteq T_0$, let

$$B(E) = \{ \langle \lambda_k \rangle_{k=1}^{\infty} \in T'_1 \mid \lambda(C \times E) = 1 \}.$$

If E is closed, so is $B(E)$. We can now recursively define T_n for positive n by

$$T_n = \left\{ \langle \delta_k \rangle_{k=1}^{\infty} \in T_1 \mid \delta(S \times B(T_{n-1})) = 1 \right\}$$

and let $T_{\infty} = \bigcap_n T_n$.

Lemma 9. The set T_{∞} is nonempty, closed, and

$$T_{\infty} = \left\{ \langle \delta_k \rangle_{k=1}^{\infty} \in T_1 \mid \delta(S \times B(T_{\infty})) = 1 \right\}.$$

Proof. We first show that $\langle T_n \rangle$ is a decreasing sequence. Clearly, $T_0 \supseteq T_1$ are closed sets. Since B is an increasing operator that maps closed sets to closed sets, $T_{n-1} \supseteq T_n$ implies

$$\begin{aligned} T_{n+1} &= \left\{ \langle \delta_k \rangle_{k=1}^{\infty} \in T_1 \mid \delta(S \times B(T_n)) = 1 \right\} \subseteq \\ &\left\{ \langle \delta_k \rangle_{k=1}^{\infty} \in T_1 \mid \delta(S \times B(T_{n-1})) = 1 \right\} = T_n. \end{aligned}$$

Consequently, we have for each $\langle \delta_k \rangle_{k=1}^{\infty} \in T_{\infty}$ that $\delta(S \times B(T_n)) = 1$ holds for all n and, therefore, also $\delta(S \times B(T_{\infty})) = 1$. Conversely, $\delta(S \times B(T_{\infty})) = 1$ implies trivially that $\delta(S \times B(T_n)) = 1$ for all n . This gives us the equality

$$T_{\infty} = \left\{ \langle \delta_k \rangle_{k=1}^{\infty} \in T_1 \mid \delta(S \times B(T_{\infty})) = 1 \right\}.$$

Also, T_{∞} is closed as an intersection of closed sets. To see that T_{∞} is nonempty, take some $s \in S$ and replace S by the singleton $\{s\}$. For the resulting model, there is a single hierarchy in T_0 , which is trivially in every T_n , even for the original S . \square

Proposition 3. T_∞ is the homeomorphic image of $\Delta(S \times \Delta_v(C \times T_\infty))$ under the mapping given by

$$\alpha \mapsto \chi(\alpha \circ (\text{id}_S, \chi')^{-1}).$$

Proof. By Lemma 9, T_∞ is the homeomorphic image of $\Delta(S \times B(T_\infty))$ under χ and, by definition,

$$B(T_\infty) = \{\langle \lambda_k \rangle_{k=1}^\infty \in T'_1 \mid \lambda(C \times T_\infty) = 1\},$$

so $B(T_\infty)$ is the homeomorphic image of $\Delta_v(C \times T_\infty)$ under χ' . \square

We use these results to construct a particular type space. Let $T^* = S \times \Delta_v(C \times T_\infty)$ and let $\sigma^* : T^* \rightarrow S$ be the projection onto S . Finally, let $\tau^* : T^* \rightarrow \Delta_v(C \times \Delta(T^*))$ be the composition of the inverse of the homeomorphism defined in the statement of Proposition 3 with the projection onto the second coordinate. We call (T^*, σ^*, τ^*) the *canonical type space*. Note that σ^* and τ^* are continuous, and T^* is compact whenever S is.

References

- Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis. A hitchhiker's guide*. Springer, Berlin, third edition, 2006.
- Charalambos D. Aliprantis, Dionysius Glycopantis, and Daniela Puzzello. The joint continuity of the expected payoff functions. *Journal of Mathematical Economics*, 42(2):121–130, 2006.
- George-Marios Angeletos, Christian Hellwig, and Alessandro Pavan. Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, 75(3):711–756, 2007.
- George-Marios Angeletos, Fabrice Collard, and Harris Dellas. Quantifying confidence. *Econometrica*, 86(5):1689–1726, 2018.
- Andrew Atkeson. [rethinking multiple equilibria in macroeconomic modeling]: Comment. *NBER macroeconomics annual*, 15:162–171, 2000.
- Łukasz Balbus, Paweł Dziewulski, Kevin Reffett, and Łukasz Woźny. Differential information in large games with strategic complementarities. *Economic Theory*, 59(1):201–243, 2015.
- Łukasz Balbus, Paweł Dziewulski, Kevin Reffett, and Łukasz Woźny. Markov distributional equilibrium dynamics in games with complementarities and no aggregate risk. *Theoretical Economics*, 17(2):725–762, 2022.
- Pierpaolo Battigalli, Alfredo Di Tillio, Edoardo Grillo, and Antonio Penta. Interactive epistemology and solution concepts for games with asymmetric information. *The BE Journal of Theoretical Economics*, 11(1), 2011.
- Elchanan Ben-Porath and Aviad Heifetz. Common knowledge of rationality and market clearing in economies with asymmetric information. *Journal of Economic Theory*, 146(6):2608–2626, 2011.

- Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic optimal control. The discrete time case*, volume 139 of *Mathematics in Science and Engineering*. 1978.
- Vladimir I. Bogachev. *Weak convergence of measures*, volume 234 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2018.
- Werner Böge and Theodor Eisele. On solutions of Bayesian games. *International Journal of Game Theory*, 8(4):193–215, 1979.
- Hans Carlsson and Eric van Damme. Global games and equilibrium selection. *Econometrica*, 61(5):989–1018, 1993.
- Eddie Dekel, Drew Fudenberg, and Stephen Morris. Interim correlated rationalizability. *Theoretical Economics*, 2:15–40, 2007.
- Jeffrey C. Ely and Marcin Peřski. Hierarchies of belief and interim rationalizability. *Theoretical Economics*, 1:19–65, 2006.
- Michael Greinecker. Typology of beliefs and rationalizability with many players. 2017.
- Sergei Izmalkov and Muhamet Yildiz. Investor sentiments. *American Economic Journal: Microeconomics*, 2(1):21–38, 2010.
- Pedro Jara-Moroni. Rationalizability in games with a continuum of players. *Games and Economic Behavior*, 74(2):556–571, 2012.
- Atsushi Kajii and Stephen Morris. The robustness of equilibria to incomplete information. *Econometrica*, 65(6):1283–1309, 1997.
- Teturo Kamae and Ulrich Krengel. Stochastic partial ordering. *The Annals of Probability*, 6(6):1044–1049, 1978.
- Teturo Kamae, Ulrich Krengel, and George L. O’Brien. Stochastic inequalities on partially ordered spaces. *The Annals of Probability*, 5(6):899–912, 1977.
- Yehuda John Levy. Bayesian equilibrium: From local to global. *Journal of Mathematical Economics*, 113:103012, 2024.
- Andreu Mas-Colell. On a theorem of Schmeidler. *Journal of Mathematical Economics*, 13(3):201–206, 1984.
- Laurent Mathevet. Beliefs and rationalizability in games with complementarities. *Games and Economic Behavior*, 85:252–271, 2014.
- Jean-François Mertens and Shmuel Zamir. Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, 14(1):1–29, 1985.
- Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, 1990.
- Stephen Morris. Justifying rational expectations. 1995.
- Stephen Morris and Hyun Song Shin. *Global Games: Theory and Applications*, volume 1 of *Econometric Society Monographs*, page 56–114. Cambridge University Press, 2003.

- Stephen Morris, Hyun Song Shin, and Muhamet Yildiz. Common belief foundations of global games. *Journal of Economic Theory*, 163:826–848, 2016.
- Philip J. Reny. On the existence of monotone pure-strategy equilibria in Bayesian games. *Econometrica*, 79(2):499–553, 2011.
- Ariel Rubinstein. The electronic mail game: Strategic behavior under “almost common knowledge”. *The American Economic Review*, 79(3):385–391, 1989.
- Donald M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26(2):305–321, 1978.
- Timothy Van Zandt. Interim Bayesian Nash equilibrium on universal type spaces for supermodular games. *Journal of Economic Theory*, 145(1):249–263, 2010.
- Jonathan Weinstein and Muhamet Yildiz. A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica*, 75(2):365–400, 2007.
- Jonathan Weinstein and Muhamet Yildiz. Interim correlated rationalizability in infinite games. *Journal of Mathematical Economics*, 72(C):82–87, 2017.
- Haomiao Yu. Rationalizability in large games. *Economic Theory*, 55(2):457–479, 2014.