Hellman, Ziv; Levy, Yehuda John

Article

Bayesian games with a continuum of states

Theoretical Economics

Provided in Cooperation with:
The Econometric Society

Theoretical Economics 12 (2017), 1089–1120 1555-7561/20171089

Bayesian games with a continuum of states

Ziv Hellman
Department of Economics, Bar Ilan University

Yehuda John Levy
Department of Economics and Nuffield College, University of Oxford

We show that every Bayesian game with purely atomic types has a measurable Bayesian equilibrium when the common knowledge relation is smooth. Conversely, for any common knowledge relation that is not smooth, there exists a type space that yields this common knowledge relation and payoffs such that the resulting Bayesian game does not have any Bayesian equilibrium. We show that our smoothness condition also rules out two paradoxes involving Bayesian games with a continuum of types: the impossibility of having a common prior on components when a common prior over the entire state space exists, and the possibility of interim betting/trade even when no such trade can be supported ex ante.

Keywords. Bayesian games, Bayesian equilibrium, common priors, continuum of states.

JEL classification. C72.

1. Introduction

When are Bayesian games guaranteed to have Bayesian equilibria? One answer to that question was given in Harsanyi (1967), the same work that introduced the common prior assumption, by reducing the question of the existence of Bayesian equilibria to the question of the existence of Nash equilibria in an associated game of complete information. As the latter always exist, so do Bayesian equilibria.

Harsányi's theorem on the existence of Bayesian equilibria, however, was proved only for Bayesian games in which all variables are finite. That is, it holds for games with a finite number of players, finite action spaces, finite payoff parameters, and a finite number of possible types.

When state spaces have continuum many states, Harsányi's theorem no longer holds. Simon (2003) presented an example of a three-player Bayesian game over

Ziv Hellman: ziv.hellman@mail.huji.ac.il
Yehuda John Levy: john.calculus@gmail.com
The research of the first author was supported in part by the European Research Council under the European Commission's Seventh Framework Programme (FP7/2007–2013)/ERC Grant Agreement 249159, and in part by Israel Science Foundation Grants 538/11 and 212/09. The research of the second author was supported in part by Israel Science Foundation Grant 1596/10.

DOI: 10.3982/TE1544

a continuum of states with no measurable Bayesian equilibrium.[1] This was extended in Hellman (2014a), which contains an example of a two-player Bayesian game over a continuum of states with no Bayesian ε-equilibrium for sufficiently small ε.

The main results of this paper deal with conditions for the existence of Bayesian equilibria and common priors when the state and type spaces are infinite. Theorem 1 provides sufficient conditions for the existence of measurable Bayesian equilibria within the class of Bayesian games over a continuum of states. These conditions are that (i) in every state of the world, each individual belief is purely atomic, and (ii) the common knowledge relation is *smooth*, that is, the common knowledge components are precisely the level sets of some measurable function. The theorem also establishes a certain necessity of the smoothness, showing that when it does not hold, there are payoffs and types with precisely this common knowledge structure for which measurable Bayesian equilibria do not exist.

Continuum state spaces also present challenges relating to the concept of the common prior. Simon (2000) presented an example in which the very existence of a common prior depends on whether one is looking at the ex ante stage or the interim stage. That is, common priors exist globally in the full state space in the ex ante stage but do not exist over any common knowledge component (i.e., in the interim stage). This is so counterintuitive that Heifetz (2006) conjectured (using the concept of common improper priors) that despite the lack of consistency in the existence of common priors in such examples, there would still be *behavioral* consistency in terms of agreement, i.e., agents would consistently agree not to trade in both the ex ante stage and the interim stage.

This leads us to the other results of this paper: in Theorems 2 and 3 we show that exactly the same smoothness conditions that characterize which games necessarily possess Bayesian equilibria also provide necessary and sufficient conditions for ex ante/interim stage consistency of the existence of common priors and no trade/no betting theorems in continuum of states. In particular, the conjecture of Heifetz (2006) is wrong: there *are* examples of behavioral inconsistency, with agents unable to agree to trade ex ante but agreeing to trade in the interim stage whenever the common knowledge relation is nonsmooth.

Smoothness is a critical condition in the main results here, both those relating to measurable Bayesian equilibria and consistency of common prior existence. Intuitively, nonsmooth models appear to be rare and unusual cases, requiring a conscious effort to conjure up. It would appear that most applications of Bayesian games with a continuum of states in which one would be interested, such as models of profits, elapsed time, accumulated resources, and so on, are more likely than not to satisfy the conditions for the existence of measurable equilibria.

With respect to the existence of measurable Bayesian equilibria in games with continuum-many types, a seminal paper by Milgrom and Weber (1985) proved that such

[1]Restricting attention to the question of the existence of measurable equilibria is not truly restrictive: given a game without measurable Bayesian equilibria, one can always construct another game, with an additional player whose payoffs depend on the strategies of the players in the original game, that has no well defined equilibria at all.

equilibria exist when players have absolutely continuous information. However, the class of smooth games that are established to have equilibria in this paper are a subset of a different class $\mathcal{CIC}$, the class of games with a subset of *continuously distributed informational commonality* introduced by Stinchcombe (2011a), which in turn is a subclass of $\mathcal{DIS}$, games with *discontinuous information structures*. All the results in Milgrom and Weber (1985) assume the condition of absolutely continuous information, which means that they are disjoint from $\mathcal{DIS}$ and hence not applicable at all to the games here.

In his paper, Stinchcombe (2011a) shows that for games in $\mathcal{CIC}$ with generic payoff functions, the expected payoffs of the players are *not* continuous as functions of the strategy profiles. When the expected payoffs are continuous as a function of the players' *behavioral strategies*, which are functions from players' knowledge to actions, Bayesian equilibria exist. The smooth games in this paper are elements of $\mathcal{CIC}$, leading to the conclusion that although generically the expected payoffs of those games are not continuous, they are nevertheless guaranteed by Theorem 1 to have Bayesian equilibria. More details may be found in Section 6.3.

We conclude by noting that during the composition of this paper, some perhaps surprising parallels between concepts used in game theory and descriptive set theory concepts were uncovered. For example, a regular conditional distribution t of a probability measure μ parallels the posterior t of a prior μ with respect to a knowledge structure; the saturation of a point with respect to a countable Borel equivalence relation corresponds to the knowledge component of a state in an epistemic game theoretic model. Hopefully, these sorts of parallels can be deepened in future research, leading to more new results.

2. Preliminaries and the model

2.1 *Smoothness*

As is standard, a Polish space is a separable, completely metrizable space. Measurability without further qualification in this paper, in the context of a Polish space X, is understood to mean measurability with respect to the Borel σ-algebra of X.

A relation $\mathcal{E}$ on a Polish space Ω is said to be *Borel* if it is Borel as a subset of $\Omega \times \Omega$. In other words, the relation is Borel if the set $\{(x, y) \in \Omega \times \Omega \mid x \,\mathcal{E}\, y\}$ is a Borel subset of $\Omega \times \Omega$. It is said to be *countable* if each equivalence class, referred to as *classes* or *atoms*, is countable. We abbreviate *countable Borel equivalence relation* as CBER.

A very central definition from descriptive set theory that is used extensively in this paper follows.

Definition 1. A Borel equivalence relation $\mathcal{E}$ on a Polish space Ω is *smooth* if there is a Polish space Y and a Borel function $\psi : \Omega \to Y$ such that for all $x, y \in \Omega$,

$$x \,\mathcal{E}\, y \quad \Longleftrightarrow \quad \psi(x) = \psi(y)$$

(i.e., the classes of $\mathcal{E}$ are precisely the level sets of ψ).

If $\mathcal{E}$ is the common knowledge relation, a function ψ witnessing the smoothness of the relation can be thought of as an auxiliary tool that enables us to ascertain when x and y are in the same common knowledge component: that occurs if and only if $\psi(x) = \psi(y)$.

A transversal of $\mathcal{E}$ is a set $T \subseteq X$ that intersects each $\mathcal{E}$ equivalence class at exactly one point. It is easy to see that if a Borel $\mathcal{E}$ has a Borel transversal, then it is smooth: intuitively, the map "$\psi(x) =$ the only element of T that is $\mathcal{E}$-equivalent to x" witnesses the smoothness of $\mathcal{E}$. For CBERs, the converse is true as well.

From this, one can show that if every equivalence class of $\mathcal{E}$ is finite, then $\mathcal{E}$ is smooth. We use this fact repeatedly. Consider the set

$$T = \{x \in X \mid \forall y \in X, x \,\mathcal{E}\, y \implies x \leq y\},$$

i.e., the set of the $\leq$ elements of the $\mathcal{E}$-equivalence classes, for any Borel linear order on the domain of $\mathcal{E}$; such a set exists by a theorem of Kuratowski. This T is seen to be Borel and a transversal of $\mathcal{E}$, hence finiteness of the $\mathcal{E}$-classes is sufficient for smoothness; for details, see, e.g., Example 6.1 of Kechris and Miller (2004). However, for CBERs, which are the focus of much of the material of this paper, matters are not so simple.

When $\mathcal{E}$ is an equivalence relation, for each $x \in \Omega$, one may consider the class containing x, which we denote by $[x]_{\mathcal{E}}$. A set $B \subseteq \Omega$ is said to be *saturated* with respect to $\mathcal{E}$ if it is the union of $\mathcal{E}$-equivalence classes, i.e., if there is a set $A \subseteq \Omega$ such that $B = [A]_{\mathcal{E}} := \bigcup_{x \in A} [x]_{\mathcal{E}}$. The collection of all the Borel $\mathcal{E}$-saturated sets of a Borel equivalence relation $\mathcal{E}$ forms a σ-algebra, denoted $\sigma(\mathcal{E})$.

2.2 *Proper regular conditional distributions*

Game theorists are used to working with priors and posteriors. The appropriate generalization to the context of the structures in this paper makes use of the concept of proper regular conditional distributions.

For a Polish space X, let $\Delta(X)$ denote the space of regular Borel probability distributions on X, with the topology of weak convergence of probability measures, and let $\Delta_f(X) \subseteq \Delta(X)$ (resp. $\Delta_a(X) \subseteq \Delta(X)$) denote the subspace of finitely supported (resp. purely atomic) measures. The space $\Delta(X)$ is itself a Polish space.

If $(\Omega, \mathcal{B})$ is a measurable space, $\mu \in \Delta(\Omega)$, and $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{B}$, then (see Blackwell and Ryll-Nardzewski 1963) a *proper regular conditional distribution* (henceforth, proper RCD) of μ, given $\mathcal{F}$, is a mapping $t : \Omega \times \mathcal{B} \to [0, 1]$ such that for each $B \in \mathcal{B}$, $\omega \to t_\omega(B)$ is Borel, and such that

$$\mu(B) = \int_\Omega t_\omega(B)\, d\mu(\omega) \quad \text{for all } B \in \mathcal{B} \tag{1}$$

and

$$t_\omega(A) = 1 \quad \text{for } \mu\text{-a.e. } \omega \in A \in \mathcal{F}.$$

It can be shown that (1) implies that for every $T \in \mathcal{B}$,

$$t_\omega(T) = E_\mu[1_T \mid \mathcal{F}](\omega), \quad \mu\text{-a.e. } \omega \in \Omega.$$

In terms that may be more familiar for game theorists, a proper RCD t of a probability measure μ may be thought of as the posterior t of a prior μ with respect to a knowledge structure $\mathcal{F} = \sigma(\mathcal{E})$.

2.3 *Knowledge spaces*

Most game theory models[2] work with partitionally generated type spaces. In such models, where Ω is finite or countable, each player i has a partition Π^i of Ω. This approach suffers from a difficulty in the case of a continuum of states, since the partition has to "agree" with the measurable structure. In addition, in the continuum case, one cannot work with arbitrary unions of partitions elements; only Borel unions are admissible. Our approach differs from the more classical approach given in Nielson (1984) and Brandenburger and Dekel (1987)—see Section 6.2 for more details on this—in favor of defining knowledge via relations (instead of σ-algebras), which is better suited for the class of purely atomic types that concern us. Our approach also differs from the "types" approach of Milgrom and Weber (1985); see Section 6.1.

We work in general with a nonempty, finite set of players I and a Polish space of *states* Ω. With each player i we associate a Borel equivalence relation over Ω, denoted $\mathcal{E}^i$, called i's *knowledge relation*. Intuitively, the unions of classes of $\mathcal{E}^i$ represent the events that player i can identify; hence, $\sigma(\mathcal{E}^i)$ is the set of Borel events that player i can identify. (In the discrete setting in which knowledge spaces are generated from knowledge partitions Π^i of Ω, $\sigma(\mathcal{E}^i)$ would be given by the unions of elements of Π^i.)

Adopting the convention that $\mathscr{E}$ stands for the profile of knowledge relations $(\mathcal{E}^i)_{i\in I}$, a *knowledge space* is then a triple $(\Omega, I, \mathscr{E})$. Given a knowledge space $(\Omega, I, \mathscr{E})$, the equivalence relation *induced* by $\mathscr{E}$, which we denote by $\mathcal{E}$, is the transitive closure of the union $\bigcup_{i\in I} \mathcal{E}^i$; i.e., the smallest equivalence relation containing each element in $\mathscr{E}$. Observe that $\sigma(\mathcal{E}) = \bigcap_{i\in I} \sigma(\mathcal{E}^i)$. In terms that may be more familiar, $\mathcal{E}$ is the *common knowledge equivalence relation*. The class of the common knowledge relation $\mathcal{E}$ containing ω is called the *common knowledge component* containing ω, and is denoted $\mathcal{C}(\omega)$.

2.4 *Type spaces and priors*

Fix a knowledge space $(\Omega, I, \mathscr{E})$. For each $i \in I$, a *type function* t^i is a mapping $t^i : \Omega \to \Delta(\Omega)$ that is $\sigma(\mathcal{E}^i)$-measurable and satisfies $t^i_\omega(A) = 1$ whenever $\omega \in A \in \sigma(\mathcal{E}^i)$.

Adopting the convention that t stands for the tuple $(t^i)_{i\in I}$, a triple (Ω, I, t) is called a *type space*. A type space implicitly defines the knowledge relations $\mathcal{E}^i$ underlying the type functions: $\omega \; \mathcal{E}^i \; \omega'$ (i.e., $(\omega, \omega') \in \mathcal{E}^i$) if and only if $t^i_\omega = t^i_{\omega'}$. Intuitively, $t^i_\omega(B)$ is the probability player i associates to the set B in state ω.

A measure $\mu^i \in \Delta(\Omega)$ such that t^i is a proper RCD for μ^i given $\sigma(\mathcal{E}^i)$ is a *prior* for t^i. A *common prior* is a measure μ that is a prior for the type functions of all the players $i \in I$.

[2] The definition can be broadened to "nearly all models in the economics, game theory, and the decision theory literature."

2.5 *Purely atomic positive spaces*

We adopt two main restrictive assumptions on type spaces. The assumption of positivity, meaning that every state in a player's knowledge component is ascribed nonzero probability by his type in said component, is convenient but not really necessary for our results. In contrast, the other supposition, of countable support for every type, is a substantively needed assumption.

Definition 2. A type space satisfying the conditions that $t^i_\omega \in \Delta_f(\Omega)$,[3] for all $i \in I$ and all $\omega \in \Omega$, is called a *finitely supported* type space.

Definition 3. A type space such that for all $i \in I$ and all $\omega \in \Omega$, the type t^i_ω is purely atomic (i.e., has countable support), is called a *purely atomic* type space.

We always assume that types are purely atomic. (Occasionally, we also require them to be finitely supported.) In addition, we henceforth assume all type spaces satisfy *positivity*.

Definition 4. A type space such that for all $i \in I$ and all $\omega \in \Omega$, $t^i_\omega[\omega] > 0$ is called *positive*. (More generally, if t is a proper RCD of μ with respect to $\mathcal{F}$, t is called positive if $t_\omega[\omega] > 0$ for all $\omega \in \Omega$.)[4]

When positivity does not hold, the set of states that violated it are of measure zero under any prior (see Proposition 7 in Appendix A); P hence, positivity is a fairly benign and merely technical assumption. In combination, pure atomicity and positivity imply that each class of each player's knowledge relation is countable, and hence so are the classes of the common knowledge equivalence relation.[5] In this case, the knowledge relation of each player is always smooth and the knowledge sets of each player are the level sets of his type function.

Definition 5. A CBER is *belief induced* if there are finitely many smooth CBERs that generate it.

[3]Recall that $\Delta_f(\Omega)$ is the set of finitely supported measures over Ω; hence, finite support is equivalent to each player ascribing positive probability only to a finite number of elements in all knowledge components. Type spaces with finite fan-out, as defined in Simon (2003), in which each partition element of the underlying partitionally based knowledge space contains only a finite number of elements, are a special case of this, although these classes coincidence if positivity (see Definition 4) is assumed.

[4]The restriction to positive type spaces is not a serious one and is implemented largely for convenience and simplicity. Theorem 2 only relies on this assumption to guarantee that a weaker assumption holds, namely that the knowledge classes of each player are countable. Theorem 3(i), which follows from Theorem 2, could similarly be proved under this weaker assumption. Theorems 1(ii) and 3(ii) deliver constructions in which positivity is guaranteed anyway. Only the proof of Theorem 1(i) makes direct use of positivity instead of the above weaker condition, in particular Proposition 11; this, too, can be relaxed at the expense of a more complicated proof.

[5]This can be shown easily if one builds the common knowledge relation inductively from the players' knowledge relations, as done in Section A.1.

Lemma 19 implies that a CBER is belief induced if and only if it is the common knowledge relation of some finitely supported positive type space.[6] Not all CBERs are belief induced.[7] We elaborate on this in Appendix B.

2.6 *Bayesian games and Bayesian equilibria*

A *Bayesian game* $\Gamma = (\Omega, I, t, A, r)$ consists of the following components:

- The set (Ω, I, t) forms a type space (with knowledge relations $\mathcal{E}$ understood implicitly as generated by t).
- The equality $A = (A^i)_{i \in I}$ is a tuple consisting of a finite action set for each player $i \in I$.
- The relationship $r : \Omega \times \prod_{i \in I} A^i \to \mathbb{R}^I$ is a bounded measurable payoff function, with r^i then being the resulting payoff to player i. The payoff function r extends multilinearly to mixed actions in the usual manner.

A *strategy* of a player $i \in I$ is a mapping $s^i : \Omega \to \Delta(A^i)$ that is constant on each player's knowledge component. In other words, if $\omega, \omega' \in \Omega$ are in the same atom of $\mathcal{E}^i$, i.e., $t^i_\omega[\omega'] > 0$, then $s^i(\omega) = s^i(\omega')$.

A *Bayesian ε-equilibrium* with $\varepsilon \geq 0$ is a profile of strategies $s = (s^i)_{i \in I}$ such that for each $i \in I$, all $\omega \in \Omega$,[8] and each alternative strategy $x \in \Delta(A^i)$ of player i,[9]

$$\sum_{\{\omega' | t^i_\omega[\omega'] > 0\}} r^i(\omega', s(\omega')) t^i_\omega[\omega'] + \varepsilon \geq \sum_{\{\omega' | t^i_\omega[\omega'] > 0\}} r^i(\omega', x, s^{-i}(\omega')) t^i_\omega[\omega'].$$

When a Bayesian ε-equilibrium s satisfies the condition that each s^i is Borel measurable,[10] s is said to be a *measurable*[11] Bayesian ε-equilibrium (ε-MBE). When $\varepsilon = 0$, we refer simply to an MBE instead of a 0-MBE.

[6]On a knowledge relation with finite classes, one can define a type function that is uniform over each class.

[7]The authors are grateful to Benjamin Weiss for pointing this out.

[8]If there is a common prior, the definition can be modified to require ε-optimality of the strategy in *almost every* state.

[9]Recall that we have assumed types are purely atomic and payoffs are bounded.

[10]The combination of being Borel measurable and being constant in each knowledge component of player i is equivalent to requiring that s^i is $\sigma(\mathcal{E}^i)$-measurable.

[11]It is possible for a game to have Bayesian ε-equilibria that are not measurable as in, for example, Simon (2003). However, for our purposes it suffices to concentrate on characterizing the existence of measurable ε-equilibria, because given a game Γ that admits only nonmeasurable equilibria, it is always possible to create another game Γ' that has no equilibria at all. This is accomplished by adding an additional player k to Γ' who is not in the player set of Γ. The payoffs of players $i \neq k$ in Γ' are defined to be exactly identical to their payoffs in Γ, while k's payoff is given by an integral over the actions of the players $i \neq k$. But if the equilibrium strategies of the players $i \neq k$ are nonmeasurable, at equilibrium player k cannot even define a payoff, much less an optimal strategy. See Hellman (2014a) for an explicit example of such a construction.

3. Results

3.1 *Measurable Bayesian equilibria*

Our main claim is that smoothness of type spaces is crucial for many results, including the existence of measurable equilibria in Bayesian games, the persistence of common priors over components, and consistency of no betting in the ex ante and interim stages. These are all detailed in this section.

Definition 6. A purely atomic positive type space whose common knowledge relation is smooth is called a *smooth* type space. A Bayesian game whose underlying type space is smooth is called a *smooth Bayesian game*.

Sometimes we want to specify exactly how a type space fails to be smooth.

Definition 7. Let $\mathcal{E}$ be a nonsmooth belief-induced CBER. Then a type space τ whose underlying common knowledge relation is $\mathcal{E}$ is called an $\mathcal{E}$*-nonsmooth* type space. A Bayesian game whose underlying type space is $\mathcal{E}$-nonsmooth is called an $\mathcal{E}$-*nonsmooth* Bayesian game.

Theorem 1 extends Harsányi's theorem, essentially stating that (within the class of purely atomic type spaces) a Bayesian game is guaranteed to have a measurable Bayesian equilibrium if and only if it is smooth. Theorem 1 also resolves the paradox appearing in Section 4.2.1.

Theorem 1. *(i) Every smooth Bayesian game has an MBE.*

(ii) Conversely, for every nonsmooth belief-induced CBER $\mathcal{E}$, there is an $\mathcal{E}$-nonsmooth Bayesian game Γ that has no MBE.

In part (ii), the types can be constructed to be positive, have finitely supported types, and a common prior, and such that the game Γ in fact does not possess an ε-MBE for $\varepsilon > 0$ small enough.

To prove Theorem 1(i), we proceed in three steps. First, we develop a notion of the space of all (positive) Bayesian games with countably many states S, player set I, and set of actions A (Proposition 11), which we denote by $\mathfrak{B}(S, I, A)$ (or just $\mathfrak{B}$ for short). Second, we then prove the existence of a Bayesian equilibrium *selection* for this class of games (Corollary 13). Finally, we show that one can measurably map the games induced on each common knowledge component of a general game into the space of games on countably many states S (Proposition 14). The composition of this mapping and the Bayesian equilibrium selection from the second step give us the required global Bayesian equilibrium.

We can construct such a mapping because the smoothness, it turns out, allows us measurably to enumerate the elements of each atom, and once we have this enumeration, we can map the game on each atom to its appropriate game in the space $\mathfrak{B}$; when we lack such an enumeration, this cannot be done because we have no canonical way to select the mapping. Details are given in Section A.3.

For the proof of Theorem 1(ii), we embed the game given in Hellman (2014a), which does not have an ε-MBE, into the given structure, using known theorems on embedding countable Borel equivalence relations into each other.

3.2 *Common priors over components*

Let $\tau = (\Omega, I, t)$ be a type space with common knowledge relation $\mathcal{E}$. Let $K \in \sigma(\mathcal{E})$, that is, K is a common knowledge event. Writing t_K for the restriction of the profile of types to K, $(t^i|_K)_{i\in I}$, one has that $\tau_K = (K, I, t_K)$ is a well defined type space over K. This is true in particular if K is an atom of $\mathcal{E}$. Similarly, we write $\mathcal{E}|_X$ for the restriction of the equivalence relation $\mathcal{E}$ to X; formally, $\mathcal{E}|_X = (X \times X) \cap \mathcal{E}$.

Definition 8. If μ is a common prior, then we say that a property holds for *almost every common knowledge component* if the set of components for which it does not hold is contained in a μ-null set.

Theorem 2 essentially states that given a type space τ with a common prior, the type space τ_K for any common knowledge component K is guaranteed also to have a common prior if and only if the underlying common knowledge relation is smooth almost everywhere. Theorem 2 also resolves the paradox that appears in Section 4.2.2.

Theorem 2. *Let τ be a type space with a common prior μ and common knowledge relation $\mathcal{E}$. The following conditions are equivalent:*

(i) There exists $X \in \sigma(\mathcal{E})$ with $\mu(X) = 1$ such that $\mathcal{E}|_X$ is smooth.

(ii) For almost every common knowledge component K, the type space τ_K has a common prior.

(iii) There is a proper RCD t of μ given $\sigma(\mathcal{E})$ such that for almost every common knowledge component K and each $x \in K$, t_x is a common prior for τ_K.

The proof of Theorem 2 is given in Section A.4. The main step is to show that (i) implies (ii). The key here is to show that for each player i, if one first takes the regular conditional distribution of μ with respect to $\sigma(\mathcal{E})$ and then from *that* one takes the conditional distribution with respect to player i's knowledge structure, one recovers i's original type.

3.3 *No betting*

For a type space τ, a *bet* is a list of $(f^i)_{i\in I}$ of bounded[12] random variables $f^i : \Omega \to \mathbb{R}$ such that $\sum_{i\in I} f^i(\omega) = 0$ for all $\omega \in \Omega$. An *acceptable bet* is a bet that satisfies the condition that

$$E^i\big[f^i \mid \omega\big] = \int_\Omega f^i(s)\, dt^i_\omega[s] > 0 \quad \text{for all } i \in I,\ \omega \in \Omega. \tag{2}$$

[12]We assume boundedness to avoid anomalies; see Feinberg (2000) and Hellman (2014b).

In words, an acceptable bet is a bet that each player believes, based on his type function, that he is guaranteed to win no matter what the true state of the world is, despite the fact that the bet is zero sum at each state. If a type space admits no acceptable bets, then we say that there is *no betting* over the type space.

Theorem 3 essentially states that we are only guaranteed that almost all common knowledge components possess no acceptable bets (i.e., there are no acceptable bets at the interim stage) when the common knowledge relation is smooth. Theorem 3 also resolves the paradox that appears in Section 4.2.3.

The condition that guarantees consistency in agreeing to trade between the ex ante and interim stages thus turns out to be virtually identical to the condition that guarantees the existence of measurable equilibria in Bayesian games (as in Theorem 1). On the one hand, smoothness guarantees common priors on components, which excludes admissible bets. On the other hand, when smoothness fails, games in which betting is admissible on components, like that in Section 4.2.3, can be embedded in the structure.

Theorem 3. *(i) A smooth type space with a common prior admits no betting over almost every common knowledge component.*

(ii) Conversely, for every nonsmooth belief-induced CBER $\mathcal{E}$, there is an $\mathcal{E}$-nonsmooth (positive, finitely supported) type space with a common prior such that on almost every common knowledge component there exists an acceptable bet.

The proof of Theorem 3(i) is given in Section A.4 as an immediate corollary of Theorem 2; the proof of Theorem 3(ii) is given in Section A.5. Like the proof of Theorem 1(ii), the proof Theorem 3(ii) involves embedding a game in which acceptable bets exist on the common knowledge components (the example of the two-player game from Lehrer and Samet (2011) that appears in Section 4) into the given structure, and then showing that the acceptable bet for those two players on the image of the embedding can be extended (on each component individually) to an acceptable bet for all players (however many needed to give the desired common knowledge structure) on the entire component.

4. Motivating paradoxes

4.1 *The ex ante vs. interim stage*

Theorems 1, 2, and 3 may be motivated by several paradoxes related to Bayesian equilibria and common priors, some of which previously appeared in the literature. They all revolve around the distinction between the ex ante stage and the interim stage. Indeed, the main theorems here resolve these paradoxes by providing full characterizations of when these pathologies may occur and when they are guaranteed not to appear.

The full state space, over which priors are defined, is usually taken to be the ex ante stage, while the common knowledge component represents the interim stage after each player receives a signal. According to a widely accepted view, in reality there is no chance move that selects a player's type. The true situation the players face is the interim stage after the vector of types has been selected. However, incomplete information requires us to consider the ex ante stage so as to understand how the players make their choices in the interim stage, even though it is a fiction and there is no actual distinction between the different stages.

The paradoxes here challenge that view, because in these examples player behavior is different depending on whether we are in the ex ante stage or the interim stage. This is particularly striking in the third example. The concept of "no acceptable bets" can be extended to "no trading" (Milgrom and Stokey (1982)); the third example then shows that one can construct knowledge structures in which players cannot measurably agree to trade in the ex ante stage but may agree to trade in the interim stage.

4.2 *Paradoxes*

The first two paradoxes—on Bayesian games and common priors over continuum type spaces—are well known in the literature for about a decade. The third paradox—on no betting—is new.

4.2.1 *The "now you see it, now you don't" Bayesian equilibrium* Simon (2003) and Hellman (2014a) present examples of Bayesian games over a continuum of states that have no Bayesian equilibria. In those games, there is no profile of measurable strategies of the players (with those strategies having as their domain the entire state space Ω) that forms a Bayesian equilibrium.

However, in these games each common knowledge component $\mathcal{C}(\omega_0)$, for any state ω_0, is countable. Hence, the Bayesian game restricted to each common knowledge component *does* have a Bayesian equilibrium, by standard arguments (e.g., Simon 2003, Proposition 1). In summary, there is no ex ante Bayesian equilibrium, but there do exist interim Bayesian equilibria.

4.2.2 *The "now you see it, now you don't" common prior* This paradox was first noted in Simon (2000). We present here a slight variation of a version appearing in Lehrer and Samet (2011).

Consider the following type space over a state space Ω, as depicted in Figure 1. The state space Ω is constructed out of four disjoint subsets of $\mathbb{R}^2$, labelled A_j for $j \in \{1, 2, 3, 4\}$:

$$A_1 = \{(x, x+1) \mid -1 \le x < 0\},$$
$$A_2 = \{(x, x) \mid -1 \le x < 0\},$$
$$A_3 = \{(x, x-1) \mid 0 \le x \le 1\},$$
$$A_4 = \{(x, \psi(x)) \mid 0 \le x \le 1\},$$

where $\psi(x) = x - c (\mathrm{mod}\ 1)$ for a fixed irrational c in $(0, 1)$.

Player 1 is informed of the first coordinate of the state and player 2 is informed of the second coordinate. Thus, the class of $\mathcal{E}^1$ containing ω—denote it $\mathcal{E}^1(\omega)$—consists of the

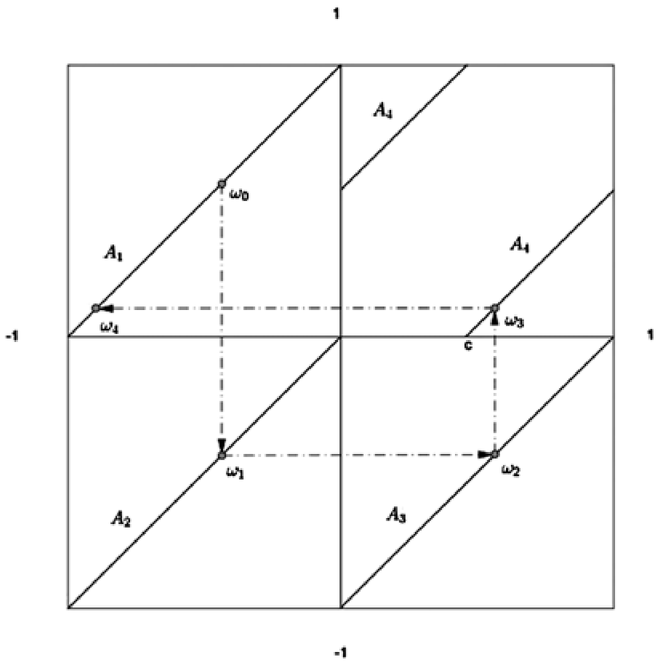

FIGURE 1. The state space consists of the three diagonals A_1, A_2, and A_3, and of A_4. The latter is obtained by a rightward shift of the top-right diagonal by an irrational number c.

two points on the vertical line that contains the state ω. Similarly, $\mathcal{E}^2(\omega)$ contains the two points on the horizontal line that include the state ω.[13]

The posterior t^i_ω for each of the two points in $\mathcal{E}^i(\omega)$ is $\frac{1}{2}$. Furthermore, let μ be the probability measure $\frac{1}{4}\sum_{j=1}^4 \psi_j$, where ψ_j is the Lebesgue measure over A_j. Lehrer and Samet (2011) show that measurability conditions are satisfied by the posteriors and that μ is a common prior for (t^1, t^2).

However, although the entire space Ω has a well defined common prior, if we again concentrate on the common knowledge component $\mathcal{C}(\omega_0)$ of any arbitrary state ω_0 (fixing the posteriors), then there is *no* common prior[14] over $\mathcal{C}(\omega_0)$. The reason for this is that $\mathcal{C}(\omega_0)$ is a doubly infinite countable sequence

$$\ldots, \omega_{-(k+1)}, \omega_{-k}, \ldots, \omega_{-1}, \omega_0, \omega_1, \ldots, \omega_k, \omega_{k+1}, \ldots \tag{3}$$

such that for all $k \in \mathbb{Z}$, $(\omega_k, \omega_{k+1}) \in \mathcal{E}_1$ and $(\omega_k, \omega_{k-1}) \in \mathcal{E}_2$ or vice versa. Any common prior ν over $\mathcal{C}(\omega_0)$ must satisfy the condition that $\nu(\omega_k) = \nu(\omega_{k+1})$ for all k. Thus all the countably many states in $\mathcal{C}(\omega_0)$ must have the same probability, which is impossible.

[13] Formally, the knowledge relations are defined by

$$(x, y)\, \mathcal{E}^1\, (x', y') \leftrightarrow x = x', \qquad (x, y)\, \mathcal{E}^2\, (x', y') \leftrightarrow y = y'.$$

[14] There may, however, be a *common improper prior* over $\mathcal{C}(\omega_0)$. An improper prior allows for the possibility that the total measure it defines over a space diverges.

In summary, there is an ex ante common prior but there does not exist an interim common prior. In particular, in light of Theorem 2, it follows that the common knowledge relation $\mathcal{E}$ generated in Figure 1 is not smooth. This could be seen by more elementary means: the restriction of $\mathcal{E}$ to any one of the sets A_1, A_2, A_3, A_4 is equivalent to the equivalence relation induced by an irrational rotation of the circle (i.e., $x \to x - c \bmod 1$, c being irrational) and this relation is well known to be nonsmooth.

4.2.3 *The "now you see it, now you don't" acceptable bet* Equation (2) states that when a bet is acceptable, there is common knowledge everywhere that every player has an expectation of positive gain, even though a bet is everywhere zero sum by definition.

If there is a common prior over the entire space, then summing the integrals and integrating over the entire space shows that there is no acceptable bet (cf. a similar argument in Hellman 2014b). By Theorem 7 in Feinberg (2000) (see also Heifetz 2006), if Ω is compact and we allow only continuous bets, then the converse also holds, that is, if there is no acceptable bet whose domain is the entire state space Ω, then there must be a common prior over Ω.

As there is a common prior over the entire space in the example depicted in Figure 1, there can be no acceptable bet over the entire space. However, one *can* construct acceptable bets on each common knowledge component in this example. We concentrate on a particular state $\omega_0 \in A_1$ (as in the figure) and the common knowledge component $\mathcal{C}(\omega_0)$ containing it; hence $(\omega_{2k}, \omega_{2k+1}) \in \mathcal{E}_1$ and $(\omega_{2k}, \omega_{2k-1}) \in \mathcal{E}_2$ for all $k \in \mathbb{Z}$ (where the enumeration follows the arrows in Figure 1). A variation of a construction from Hellman (2014b), using $\mathcal{C}(\omega_0)$ as in (3), defines the following function $f : \mathcal{C}(\omega_0) \to \mathbb{R}$:

$$f(\omega_n) = \begin{cases} 0 & \text{if } n = 0, \\ (-1)^{n+1} \cdot \displaystyle\sum_{i=1}^{n} \frac{1}{2^i} & \text{if } n > 0, \\ (-1)^{n} \cdot \displaystyle\sum_{i=1}^{-n} \frac{1}{2^i} & \text{if } n < 0. \end{cases}$$

The function f is presented graphically in Figure 2. It is easy to check that $(f, -f)$ is an acceptable bet over $\mathcal{C}(\omega_0)$, even though there is no globally acceptable bet over the entire space Ω. In summary, there is no ex ante betting, but there is interim betting.

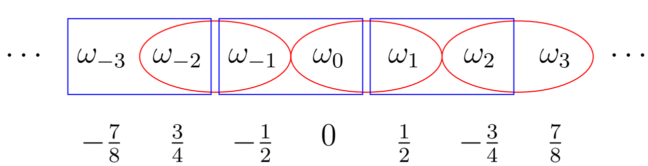

FIGURE 2. Graphical depiction of an acceptable bet over a common knowledge component of the type space depicted in Figure 1. Player 1's knowledge is represented by ellipses; player 2's knowledge is represented by rectangles.

5. Examples of smooth and nonsmooth structures

We present here some examples to illustrate the concept of smooth equivalence relations. These examples are conceptually simpler than the examples underlying the paradoxes of Bayesian games given in Section 4.2. We repeatedly use the criterion mentioned earlier and given by Proposition 1: smoothness of a CBER $\mathcal{E}$ is equivalent to the existence of a Borel set $B \subseteq \Omega$, known as a *Borel transversal*, which intersects each atom of $\mathcal{E}$ at exactly one point.

We begin with some examples of smooth relations.

Example 1. The state space is $\Omega = \mathbb{R}$, with relation $x \sim y$ if and only if $x - y$ is an integer. In other words, a player is informed only of the values after the decimal point for any $x \in \mathbb{R}$, with the integer value hidden. This relation is smooth.[15] The equality $B = [0, 1)$ is a Borel transversal. $\diamondsuit$

Example 2. We have $\Omega = \mathbb{R}$, with the atoms of the common knowledge class of the form $\{\pm x + n \mid n \in \mathbb{Z}\}$; this is induced if one player's knowledge relation is $x \sim_1 -x$ and the other player's knowledge relation is as in the first example, i.e., $x \sim_2 y$ if and only if $x - y$ is an integer. The induced common knowledge relation is again easily seen to be smooth by taking the Borel transversal $B = [0, \frac{1}{2}]$. $\diamondsuit$

Example 3. If $\Omega = \Omega_1 \times \Omega_2$, consider the relations $(x, y) \sim_1 (x', y')$ if and only if $x = x'$ and $(x, y) \sim_2 (x', y')$ if and only if $y = y'$ are smooth; take the transversals $\Omega_1 \times \{y_0\}$ and $\{x_0\} \times \Omega_2$ for some $x_0 \in \Omega_1$, $y_0 \in \Omega_2$. This common knowledge relation refers to the case in which there is common knowledge about one aspect of the state of nature but no knowledge about the other. $\diamondsuit$

Theorems 1, 2, and 3 then guarantee that for any type space with these common knowledge relations, common priors exist on all components, Bayesian equilibria exist regardless of the payoffs, and there are no agreeable bets on any components. Finding the common priors and Bayesian equilibria, however, can be quite cumbersome, hence the advantage of possessing general existence theorems such as those we present here.

We now turn to some nonsmooth examples.

Example 4. We have $\Omega = \mathbb{R}$, with the common knowledge relation $x \sim y$ if and only if $x - y$ is rational. This common knowledge relation can be induced if a third player is added to the structure in Example 2, with the knowledge relation of this newly added third player given by $x \sim_3 y$ if and only if $x = my$, where m is an integer.[16] It is well known that in this structure there does not exist a Borel transversal (see, for example, Rudin 1986, Chapter 2). $\diamondsuit$

[15] In fact, one can prove a general theorem: if Ω is a Polish space with metric d, and $\mathcal{E}$ is a CBER such that for each class A of $\mathcal{E}$, $\inf_{x \neq y, x, y \in A} d(x, y) > 0$, then $\mathcal{E}$ is smooth. In other words, as long as in each atom the elements "keep their distance" and do not get "bunched up," the relation is smooth.

[16] For example, the third player may not be sure if the value made known to him is the total amount or the amount per person in a group of unknown size after the amount has been divided.

Example 5. We have $\Omega = A^{\mathbb{Z}}$ for some finite A and the relation is given by $(x_j)_{j\in\mathbb{Z}} \sim (y_k)_{k\in\mathbb{Z}}$ if and only if $\exists m \in \mathbb{Z}$, $\forall n$, $x_n = y_{n+m}$, i.e., the relation induced by the shift. This common knowledge relation is induced when the state of nature is represented by a doubly infinite sequence of data elements in A but there is uncertainty as to where the "middle" point is. This relation is well known to be nonsmooth. ◊

Since the common knowledge structures above are not smooth, common knowledge components may not possess common priors, Bayesian equilibria may not exist for certain priors and payoffs, and for certain type structures—even those induced by a common prior—we may find acceptable bets on components even though there are no globally acceptable bets. Again, constructively coming to these conclusions in each separate case could be extremely cumbersome, while Theorem 1 guarantees the existence of such cases in a general manner.

We end with a generalization of an Example appearing in Section 4. We include Example 6 here to show that even when it is the case that when a player observes his own type he knows that the types of the other players are limited to a finite number of possible points, the resulting knowledge structure may be nonsmooth.

Example 6. Let there be N players and let $\Omega \subseteq \mathbb{R}^N$ be a finite union $\bigcup_{j=1}^{n} \Omega_j$ such that each Ω_j is a subset of a plane P_j of dimension $n-1$ not parallel to any axis; i.e., each P_j is of the form

$$\left\{x \in \mathbb{R}^n \mid \sum a_i x_i = c\right\} \quad \text{for some } a_1, \dots, a_n, c \in \mathbb{R}, \text{ such that } a_i \neq 0 \text{ for all } i.$$

Assume there is a common prior μ on each Ω_j that is absolutely continuous with respect to the $n-1$-dimensional Lebesgue measure on P_j. The information structure is such that for each i, player i is informed of the ith coordinate, i.e.,

$$\mathcal{E}^i := \left\{\left((x_1, \dots, x_N), (y_1, \dots, y_N)\right) \in \Omega \times \Omega \mid x_i = y_i\right\}.$$

For each player, the knowledge classes are finite. The resulting knowledge structure may be smooth, but, as the examples of Section 4 show, may also be nonsmooth. ◊

6. Relationship to the literature

6.1 *Type spaces*

Many papers on games of incomplete information, such as Milgrom and Weber (1985), model players' information by *types*. In such modelling, each player i has a type space Ω_i with measurable structure $\mathcal{F}_i$, and the set of states of the world in their framework is $\Omega := \prod_i \Omega_i$ with a nonatomic common prior μ defined on a σ-algebra $\mathcal{F}$ containing $\bigotimes_i \mathcal{F}_i$. Players in this framework are told their own signals and then use that to deduce a distribution on the states of the world via Bayesian updating with respect to a common prior μ; we omit the technical details.

The model of Bayesian games studied in this paper can also be formulated using type spaces. In our framework, each knowledge relation $\mathcal{E}_i$ is smooth and the classes are level sets of the type function. Hence, by Proposition 1 in Appendix A, the quotient space $\Omega_i := \Omega/\mathcal{E}_i$, which is the collection of classes of $\mathcal{E}_i$ with the quotient σ-algebra, is standard Borel, i.e., a Polish space in some topology consistent with its measurable structure. Equivalently, we can take Ω_i to be the range of player i's type function. A common prior μ on Ω then induces a common prior on $\prod_i \Omega_i$, itself a quotient space of Ω.

6.2 *Knowledge*

An alternative approach to modelling knowledge and type spaces on a continuum of states, going back to Nielson (1984) (see also Brandenburger and Dekel 1987) is to model information using σ-algebras. That is, player i's knowledge is represented by a σ-algebra $\mathcal{F}^i$, where it is understood that two elements x, y are in the same information component for player i whenever $x \in B \in \mathcal{F}^i$ implies that $y \in B$. (The connection to a knowledge equivalence relation $(\mathcal{E}^i)_{i\in I}$ from our framework is given by $\mathcal{F}^i = \sigma(\mathcal{E}^i)$.) We have avoided this approach for several reasons.

First of all, while that approach may be more appropriate for general knowledge structures, when the types are purely atomic (and hence knowledge classes are countable), it is more intuitive, in our opinion, to express the knowledge of the players, as well as the common knowledge structures, using equivalence relations. While it is true that our framework is set upon the background of measurable structures, it still echoes the standard partitional approach to knowledge. The only reason we cannot directly extend the standard approach is because we restrict ourselves to measurable sets.

Second, if one wishes to model continuum knowledge spaces based directly on knowledge of σ-algebras, various serious technical problems arise. For example, there is no guarantee that each player's knowledge components are measurable or, more generally, that the saturation of a measurable set with respect to a player's knowledge (or the induced common knowledge relation) is measurable. These problems are partially overcome by identifying sets that differ by a set of measure 0, but, other than the fact that this requires a common prior at the onset, this fix is useless for our purposes, as we need to look at the individual common knowledge components, which are generically of null measure.

6.3 *Equilibrium existence*

Given the technical difficulties discussed in the previous subsection that are encountered in dealing with general knowledge structures, results on equilibrium existence in general knowledge structures are almost nonexistent. One seminal positive result that does establish the existence of equilibrium is Milgrom and Weber (1985).

As discussed in Section 6.1, Milgrom and Weber (1985) model incomplete information using types, and our framework can directly be translated into that framework. The notion of strategies in Milgrom and Weber (1985) also differs from the definition used here; they use *distributional strategies* in contrast to the definition of strategies given above in Section 2.6. This is also not a serious difference; Balder (1988) similarly proves equilibrium existence for the same class of games in the class of strategies of the "more classical" sense used here.

However—and here is where the real substantive difference lies—Milgrom and Weber (1985) then go on to assume that the common prior μ is absolutely continuous with respect to $\bigotimes_i \mu_i$, with μ_i being the marginal on Ω_i. Unless the common prior is purely atomic, this assumption clearly implies that types cannot be purely atomic and, hence, this assumption, which guarantees the existence of Bayesian equilibrium, is *not* satisfied in our framework.

Following Stinchcombe (2011a), denoting the class of games in which μ is *not* absolutely continuous with respect to a product measure by $\mathcal{DIS}$ (for discontinuous information structure), the results of this paper relate solely to a subclass of $\mathcal{DIS}$. In contrast, the class of games in Milgrom and Weber (1985) is precisely those that are disjoint from $\mathcal{DIS}$.

Stinchcombe (2011a) also studies structures within the $\mathcal{DIS}$ class. That paper defines a structure of beliefs to have a subset satisfying *continuously distributed informational commonality* (CIC) if on a nonnull subset at least two players can agree on a nonatomically distributed variable; formally, if there is a set $B \subseteq \Omega$ that is common knowledge (i.e., $B \in \bigcap_j \sigma(\mathcal{E}_j)$, recalling notation from Section 6.1) with $\mu(B) > 0$, players $i, j \in I$, and a Borel mapping $\phi : B \to [0, 1]$ that is $\sigma(\mathcal{E}_i) \cap \sigma(\mathcal{E}_j)$-measurable (in B), such that $\phi_* \circ \mu := \mu \circ \phi^{-1}$ is nonatomic.[17] Denoting the class of information structures that contain CIC's by $\mathcal{CIC}$, Theorem A of Stinchcombe (2011a) shows that $\mathcal{CIC} \subseteq \mathcal{DIS}$.[18] By definition, smooth games are in $\mathcal{CIC}$.[19]

Nonmembership in the class $\mathcal{DIS}$ guarantees equilibrium by Milgrom and Weber (1985). However, within $\mathcal{DIS}$, membership or nonmembership in $\mathcal{CIC}$ has no direct implications for the existence of MBE. On the one hand, a two-player game that is ergodic (that is, the only common knowledge sets are of μ-measure 0 or 1) is *not* in $\mathcal{CIC}$; an example of this is the game in Hellman (2014b), which is ergodic and has no equilibria. On the other hand, it is easy to construct games with no equilibria in which two players have identical knowledge, which places them within $\mathcal{CIC}$; e.g., to the game from Hellman (2014b), add dummy players with perfect information.

Theorem B of Stinchcombe (2011a) shows that for games in $\mathcal{CIC}$ with generic payoff functions, the expected payoffs of the players are not continuous as functions of the strategy profiles (where the strategies, viewed as maps from type spaces to mixed actions that are measurable with respect to players' information, are endowed with the weak-$*$ topology induced by μ). As previously mentioned, every smooth Bayesian game is in $\mathcal{CIC}$, while Theorem 1 of this paper shows that every such game has an MBE. Thus, games with purely atomic types and smooth common knowledge relations form an interesting class: although generically the expected payoffs of such games are not continuous, measurable Bayesian equilibria are still guaranteed to exist. The relationships between $\mathcal{DIS}$ and $\mathcal{CIC}$, and their implications for Bayesian equilibrium existence are summarized graphically in Figure 3.

[17]That is, for each $x \in [0, 1]$, $\mu(\{\omega \in B \mid \phi(\omega) = x\}) = 0$.

[18]The theorem requires that the knowledge σ-algebras of at least two players support nonatomic measures; this holds by construction in our setup as all player's knowledge relations are smooth; see Proposition 1.

[19]All Polish spaces are Borel isomorphic.

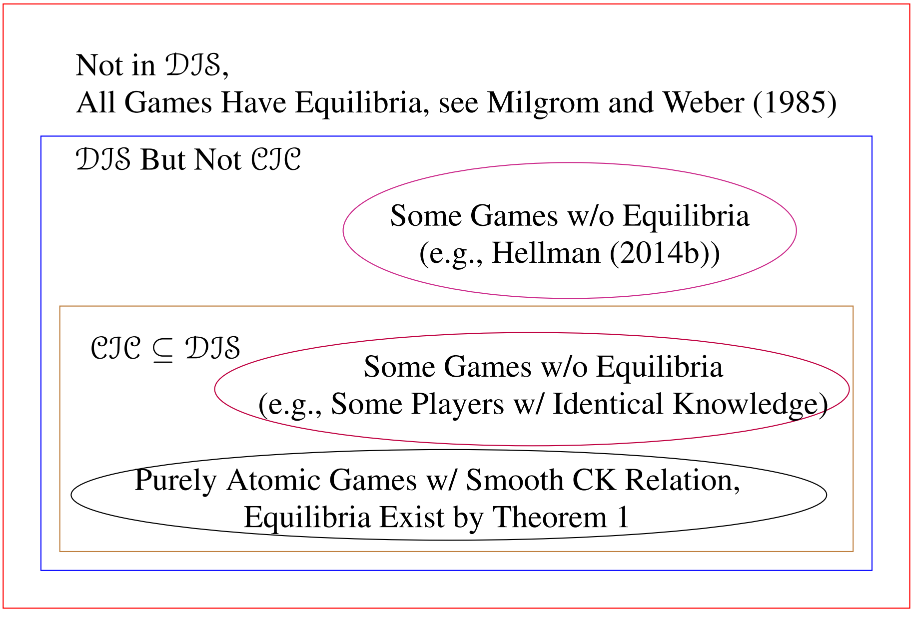

Figure 3. A representation of the interrelationships between $\mathcal{DIS}$ and $\mathcal{CIC}$ and the implications of these properties for the existence of equilibria.

We also mention in passing Stinchcombe (2011b), which shows the existence of correlated equilibrium. In that framework, actions are a function not only of types, but also of a public signal chosen uniformly in $[0, 1]$. Interestingly, the proof there shows that MBE exists in all Bayesian games if one allows for *saturated* measurable structures, i.e., structures constructed using nonstandard analysis. Such extensions, however, are highly nonconstructive.

Appendix A: Tools and proofs

A.1 *Mathematical tools*

Recall that $[T]_{\mathcal{E}}$ denotes the *saturation* of T with respect to a CBER $\mathcal{E}$, i.e., the smallest union of classes of $\mathcal{E}$ containing T. In terms that may be more familiar to game theorists used to working with finite partitions, $[\omega]_{\mathcal{E}}$ is the knowledge component containing ω. Recall that a *transversal* of an equivalence relation is a set that intersects each equivalence class at exactly one point.

Given a Polish Ω and a CBER $\mathcal{E}$, we let $\Omega/\mathcal{E}$ denote the quotient space whose elements are the equivalence classes by $\mathcal{E}$, and the induced σ-algebra consists of precisely the images of the $\mathcal{E}$-saturated sets in Ω under the quotient map.

We make repeated use of the following proposition, which follows from Propositions 6.3 and 6.4 of Kechris and Miller (2004) and the discussion preceding them.

Proposition 1. *The following conditions are equivalent for a CBER $\mathcal{E}$ on a Polish space Ω:*

(a) *The CBER $\mathcal{E}$ is smooth.*

(b) *There is a Borel transversal for $\mathcal{E}$.*

(c) *The quotient space $\Omega/\mathcal{E}$ is standard Borel.*[20]

Proposition 2 follows from[21] Theorem 1 of Blackwell and Ryll-Nardzewski (1963).

Proposition 2. *If $\mathcal{E}$ is a smooth CBER on a Polish space Ω and $\mu \in \Delta(\Omega)$, then there exists a proper RCD t of μ given $\sigma(\mathcal{E})$.*

Proposition 3 is a slight variation of the Lusin–Novikov theorem (e.g., Kechris 1995, Theorem 18.10).

Proposition 3. *Let $\mathcal{E}$ be a CBER on a Polish space Ω. Then there are Borel functions $(f_n)_{n=1}^{\infty} : \Omega \to \Omega$ such that for all $\omega \in \Omega$, $\{f_n(\omega)\}_{n\in\mathbb{N}} = [\omega]_{\mathcal{E}}$.*

From this (or related results) one can deduce by standard techniques; see, e.g., Dougherty et al. (1994, Theorem 5.1).

Proposition 4. *Let $\mathcal{E}$ be a smooth CBER on a Polish space Ω and let S be a countably infinite set. Then there is a Borel mapping $\Phi : \Omega \to S$ such that for each $\mathcal{E}$-class C of Ω, the restriction $\Phi|_C : C \to S$ is injective and $\Phi(C) = S$ if C is infinite.*

We also recall the following well known result, of which we make repeated implicit use.

Proposition 5. *Let X, Y be Polish spaces and let $f : X \to Y$ be Borel such that for each $y \in Y$, $f^{-1}(y)$ is at most countable (i.e., the map is countable-to-one). Then for each Borel $B \subseteq X$, $f(B)$ is Borel.*

Let τ be a type space with knowledge relations $(\mathcal{E}^i)_{i\in I}$. For each $i \in I$ and each set $N \subseteq \Omega$, let $C^i(N)$ denote the saturation of N with respect to $\mathcal{E}^i$, i.e., $C^i(N) = [N]_{\mathcal{E}^i}$. For each finite sequence $\hat{i} = (i_1, \dots, i_k) \in I^* = \bigcup_{n\geq 0} I^n$ and $N \subseteq \Omega$, let

$$C^{\hat{i}}(N) = C^{i_k}\left(C^{i_{k-1}}\left(\cdots\left(C^{i_1}(N)\right)\cdots\right)\right)$$

and

$$\mathcal{C}(N) = \bigcup_{\hat{i}\in I^*} C^{\hat{i}}(N),$$

which is the smallest common knowledge set containing N. Since, by Propositions 4 and 5, the saturation of Borel sets under a CBER is also Borel, we have the following lemma.

[20] That is, there is a measurable bijection between it and a Polish space.

[21] The condition given there for the existence of proper RCDs is easily seen to follow from the existence of a Borel transversal, which, by Proposition 1, follows from smoothness.

Lemma 6. *If N is Borel, then so is $C^i(N)$ for each $i \in I$ and so is $\mathcal{C}(N)$.*

The following proposition justifies our restriction to positive type spaces.

Proposition 7. *If t is a proper RCD of μ with respect to $\sigma(\mathcal{E})$ for a CBER $\mathcal{E}$, then*

$$\mu(\{\omega \mid t_\omega[\omega] = 0\}) = 0.$$

In particular, if τ is a type space (not necessarily positive) with a common prior μ, then for each $i \in I$,

$$\mu(\{\omega \in \Omega \mid t^i_\omega[\omega] = 0\}) = 0.$$

Proof. Denote $N = \{\omega \mid t_\omega[\omega] = 0\}$. Since $t_\omega[\omega'] = 0$ if $(\omega, \omega') \notin \mathcal{E}$ and since $t_\omega[\omega'] = 0$ for all the countably many ω' with $(\omega, \omega') \in \mathcal{E}$ since $\omega \to t_\omega$ is $\sigma(\mathcal{E})$-measurable, we see that $t_\omega[N] = 0$ for all $\omega \in \Omega$. The proposition follows from the definition of an RCD. □

The following definition and its properties can be found in Dougherty et al. (1994, Section 3).

Definition 9. Let $\mathcal{E}$ be a CBER on a Polish space Ω. A measure $\mu \in \Delta(\Omega)$ is called *$\mathcal{E}$-quasi-invariant* if for any Borel set $A \subseteq \Omega$, $\mu(A) = 0$ if and only if $\mu([A]_\mathcal{E}) = 0$.

Lemma 8. *Let $\mathcal{E}$ be a CBER on a Polish space Ω, and let t be a proper RCD of $\mu \in \Delta(\Omega)$ given $\sigma(\mathcal{E})$. Then t is positive on an $\mathcal{E}$-saturated set of full measure if and only if μ is $\mathcal{E}$-quasi-invariant.*

Proof. Assume μ is $\mathcal{E}$-quasi-invariant. Denoting $N = \{\omega \mid t_\omega[\omega] = 0\}$, Proposition 7 implies that $\mu(N) = 0$, so $\mu([N]_\mathcal{E}) = 0$. Therefore, t is positive on the $\mathcal{E}$-saturated set of full measure $\Omega \setminus [N]_\mathcal{E}$.

Conversely, if t is positive on a $\mathcal{E}$-saturated set of full measure X, then for Borel $A \subseteq \Omega$ with $\mu(A) = 0$, denoting $B = A \cap X$,

$$0 = \mu(A) \geq \mu(B) = \int_\Omega t_\omega(B)\, d\mu(\omega) = \int_{[B]_\mathcal{E}} t_\omega(B)\, d\mu(\omega) \tag{A.1}$$

since $t_\omega(B) = 0$ for $\omega \notin [B]_\mathcal{E}$. Since t is positive in X, $t_\omega(B) > 0$ in B and, hence, in $[B]_\mathcal{E}$ as $\omega \to t_\omega$ is $\sigma(\mathcal{E})$-measurable. Hence, by (A.1), $\mu([B]_\mathcal{E}) = 0$ and, hence, finally $\mu([A]_\mathcal{E}) = 0$, as $[A]_\mathcal{E} \subseteq [B]_\mathcal{E} \cup (\Omega \setminus X)$. □

If $(\Omega, \mathcal{E})$ and $(\Lambda, \mathcal{D})$ are Polish spaces with induced Borel equivalence relations $\mathcal{E}$ and $\mathcal{D}$, $(\Omega, \mathcal{E})$ is said to be *embeddable* into $(\Lambda, \mathcal{D})$ if there is an injective Borel mapping $\psi : \Omega \to \Lambda$ such that for all $\omega, \eta \in \Omega$, $\omega\, \mathcal{E}\, \eta \iff \psi(\omega) \mathcal{D} \psi(\eta)$; in this case, we denote $(\Omega, \mathcal{E}) \sqsubseteq (\Lambda, \mathcal{D})$.

A CBER is said to be *hyperfinite* (Dougherty et al. (1994)) if it is induced by the action of a Borel $\mathbb{Z}$ action on Ω; i.e., if there is a bijective[22] Borel mapping $T : \Omega \to \Omega$ such that $x\, \mathcal{E}\, y \iff \exists n \in \mathbb{Z}, T^n(x) = y$.

[22]If a Borel mapping between Polish spaces is injective, then by Proposition 5, its inverse is also Borel.

PROPOSITION 9. *Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be nonsmooth CBERs on Polish spaces Ω_1 and Ω_2, with $\mathcal{E}_1$ being hyperfinite. Then $(\Omega_1, \mathcal{E}_1) \sqsubset (\Omega_2, \mathcal{E}_2)$.*

PROOF. Since $\mathcal{E}_2$ is nonsmooth, the Glimm–Effros dichotomy for CBERs (see Harrington et al. 1990 or Dougherty et al. 1994, Theorem 3.4) implies that there is a universal hyperfinite CBER[23] $(\Omega_0, \mathcal{E}_0)$ such that $(\Omega_0, \mathcal{E}_0) \sqsubset (\Omega_2, \mathcal{E}_2)$. By Theorem 7.1 of Dougherty et al. (1994), any two nonsmooth hyperfinite equivalence relations can be embedded into each other; hence, $(\Omega_1, \mathcal{E}_1) \sqsubset (\Omega_0, \mathcal{E}_0)$. Hence, $(\Omega_1, \mathcal{E}_1) \sqsubset (\Omega_2, \mathcal{E}_2)$. $\square$

A.2 *Embedding of games*

Proposition 10 is the primary tool needed for the proofs of Theorems 1(ii) and 3(ii).

PROPOSITION 10. *Let Ω and X be Polish spaces, and let $\mathcal{E}$ be a CBER on Ω that is nonsmooth and belief induced. Let $\tau_X = (X, J, (t^j_X)_{j\in J})$ be an everywhere finitely supported and positive type space with a common prior μ_X, and assume that its common knowledge equivalence relation $\mathcal{E}_X$ is hyperfinite. Then one can construct a Borel embedding $\psi : X \to \Omega$ and a set of players I, such that $J \subset I$, along with an everywhere finitely supported and positive type space $\tau = (\Omega, I, (t^i)_{i\in I})$ possessing a common prior μ for which the following statements hold:*

- *The CBER $\mathcal{E}$ is the common knowledge relation induced by the type space τ.*
- *We have $t^j_{\psi(\cdot)} = \psi_*((t^j_X)_{(\cdot)})$ for each $j \in J$ in X; explicitly, for $\omega, \omega' \in Y$, $t^j_{\psi(\omega)}[\psi(\omega')] = (t^j_X)_\omega[\omega']$.*
- *For $j \in J$ and $\omega \notin \psi(X)$, $t^j_\omega = \delta_\omega$, i.e., the Dirac measure at ω.*
- *We have $\psi_*(\mu_X) := \mu_X \circ \psi^{-1} \ll \mu$ and $\psi_*(\mu_X)(A) = \mu(A)$ for $A \in \sigma(\mathcal{E})$.*

The middle two points say that for each player $j \in J$, his type function on X becomes his type function on $\psi(X)$, and he has perfect knowledge on $\Omega \setminus \psi(X)$. Note in particular that if $\omega_1 \; \mathcal{E}_X \; \omega_2$, then $\psi(\omega_1) \; \mathcal{E} \; \psi(\omega_2)$. The last point says that the prior induced on Ω by μ_X is absolutely continuous with respect to the final common prior μ, and they agree on $\mathcal{E}$-saturated sets.

PROOF OF PROPOSITION 10. Since $\mathcal{E}_X$ is hyperfinite, by Proposition 9 there is an embedding $\psi : (X, \mathcal{E}_X) \sqsubseteq (\Omega, \mathcal{E})$. Denote $\Phi = \psi(X)$ and $\Omega_0 = [\Phi]_\mathcal{E} = [\psi(X)]_\mathcal{E}$; both Φ and Ω_0 are Borel by Propositions 3 and 5.

Since $\mathcal{E}$ is belief induced, by Proposition 18 there exist a set of players K and smooth CBERs $(\mathcal{E}^k)_{k\in K}$ with finite classes such that $\mathcal{E}$ is induced by $(\mathcal{E}^k)_{k\in K}$. For $j \in J$, define the knowledge relations

$$\mathcal{E}^j = \{(x, y) \mid (x = y) \vee \big(x, y \in \psi(X) \wedge \big(\psi^{-1}(x), \psi^{-1}(y)\big) \in \mathcal{E}^j_X\big)\},$$

[23] Given $\Omega_0 = 2^\mathbb{N}$, $\mathcal{E}_0$ is the tail equivalence relation; Dougherty et al. (1994, Corollary 8.2) shows this to be hyperfinite.

i.e., $\mathcal{E}^j$ is induced by $\mathcal{E}^j_X$ on $\psi(X)$ and player j has perfect knowledge outside of Φ. Define $I = K \cup J$. By construction, $\mathcal{E}$ is induced by $(\mathcal{E}^i)_{i\in I}$, since it is induced by $(\mathcal{E}^i)_{i\in I\setminus J}$ and $\mathcal{E}^j$ refines $\mathcal{E}$ for $j \in J$.

Denote by $\mathcal{E}_0 = \mathcal{E}_{\Omega_0}$ and $\mathcal{E}_\Phi = \mathcal{E}|_\Phi$ the restrictions of $\mathcal{E}$ to $\Omega_0 = [\Phi]_\mathcal{E}$ and Φ, respectively. For brevity, let $\hat{\mu} = \psi_*(\mu_X) := \mu_X \circ \phi^{-1}$; $\hat{\mu}$ is then $\mathcal{E}_\Phi$-quasi-invariant by the assumption that μ_X is positive and by Proposition 8. By Dougherty et al. (1994, Proposition 3.1), there exists an $\mathcal{E}_0$-quasi-invariant measure ν on Ω_0 satisfying $\hat{\mu} \ll \nu$ and satisfying $\hat{\mu}(A) = \nu(A)$ for $A \in \sigma(\mathcal{E})$.

Observe the following string of implications for a set $A \subseteq \Omega$,

$$\hat{\mu}(A) = 0 \to \hat{\mu}\big([A\cap\Phi]_{\mathcal{E}_\Phi}\big) = 0 \to \hat{\mu}\big([A\cap\Phi]_{\mathcal{E}}\big) = 0 \to \nu\big([A\cap\Phi]_{\mathcal{E}}\big) = 0, \tag{A.2}$$

where the first implication is because $\hat{\mu}$ is $\mathcal{E}_\Phi$-quasi-invariant, the second is because $[A\cap\Phi]_\mathcal{E} \subseteq [A\cap\Phi]_{\mathcal{E}_\Phi} \cup (\Omega\setminus\Phi)$, and the last is since ν, $\hat{\mu}$ agree on $\mathcal{E}$-saturated sets. Also,

$$\nu(A\setminus\Phi) = 0 \to \nu\big([A\setminus\Phi]_\mathcal{E}\big) = 0 \to \hat{\mu}\big([A\setminus\Phi]_\mathcal{E}\big) = 0, \tag{A.3}$$

where the first implication is because ν is $\mathcal{E}$-quasi-invariant, and the last is since $\hat{\mu} \ll \nu$.

We deal now with two cases to construct μ to be quasi-invariant:

If $\nu(\Phi) = 1$, let $\mu = \hat{\mu}$. Then if $A \subseteq \Omega$ with $\mu(A) = 0$, then $\nu(A\setminus\Phi) \le \nu(\Omega\setminus\Phi) = 0$ and $\hat{\mu}(A) = 0$. It follows that since $[A]_\mathcal{E} = [A\cap\Phi]_\mathcal{E} \cup [A\setminus\Phi]_\mathcal{E}$, we have $\mu([A]_\mathcal{E}) = 0$ by (A.2) and (A.3).

Otherwise, if $\nu(\Phi) < 1$, set

$$\mu = \frac{1}{2}\hat{\mu} + \frac{1}{2}\nu(\cdot \mid \Omega_0\setminus\Phi).$$

If $\mu(A) = 0$, then clearly we also have $\hat{\mu}(A) = 0$ and $\nu(A\setminus\Phi) = 0$, and then, as above, $\mu([A]_\mathcal{E}) = 0$.

Either way, μ is $\mathcal{E}_0$-quasi-invariant. Now, for $i \in I\setminus J$, let t^i be a proper RCD of μ with respect to $\mathcal{E}^i$, which exists by Proposition 2. By Lemma 8, it is positive on an $\mathcal{E}^i$-saturated set of full μ-measure and can be modified on a μ-null set to be positive everywhere.[24] Clearly, for $j \in J$, t^j as defined in the statement of the proposition is a positive proper RCD of μ with respect to $\mathcal{E}^j$, since $\mu(\cdot|\Phi) = \hat{\mu}(\cdot \mid \Phi)$. □

A.3 *Proof of Theorem 1*

Fix a countably infinite set S. Let $\mathfrak{B}$ denote the collection of all I-tuples $(s^i, g^i)_{i\in I}$ for which $(S, I, A, (s^i, g^i)_{i\in I})$ constitutes a positive Bayesian game, with (s^i) denoting the types and (g^i) denoting the payoff functions. Collection $\mathfrak{B}$ is endowed with the topology of pointwise convergence:[25] for each α in a directed set, denote by Υ_α a pair $(s^i_\alpha, g^i_\alpha)_{i\in I}$ of I-tuples of types and payoff functions associated to α by a net. Then $\Upsilon_\alpha \to \Upsilon = (s^i, g^i)_{i\in I}$ in $\mathfrak{B}$ if for every player $i \in I$, every $\omega \in S$, and every pure action profile $a \in \prod_{i\in I} A^i$ one has $g^i_\alpha(\omega, a) \to g^i(\omega, a)$ and $s^i_\alpha(\omega) \to s^i(\omega)$.

[24] Since $\mathcal{E}^i$ has finite classes, over a μ-null set of $\mathcal{E}^i$ classes let t^i be uniform in each class.

[25] We define the topology in terms of nets.

Proposition 11. *Collection $\mathfrak{B}$ is homeomorphic to a Borel subset of $\Xi := (\Delta(S) \times \mathbb{R}^{\prod_{i\in I} A^i})^{S\times I}$ and, hence, is Polish (in some topology that induces the same Borel structure).*

The simple intuition is that for each player and state pair $(s,i) \in S \times I$, we need to specify an element in $\Delta(S)$ as well as an element of $\mathbb{R}^{\prod_{i\in I} A^i}$, which specifies what payoff that player receives as a result of each possible action profile.

Henceforth, we identify $\mathfrak{B}$ with some such fixed subset of Ξ.

Proof of Proposition 11. Write $\mathfrak{B} = \prod_{i\in I}(\mathfrak{B}^i_s \times \mathfrak{B}^i_g)$, where $\mathfrak{B}^i_s$ (resp. $\mathfrak{B}^i_g$) denotes the projection of $\mathfrak{B}$ to the space of types (resp. payoffs) for player i, with the induced topologies. It suffices to show that $\mathfrak{B}^i_s$ is homeomorphic to a Borel subset of $(\Delta(S))^S$ and that $\mathfrak{B}^i_g$ is homeomorphic to a Borel subset of $\mathbb{R}^{S\times\prod_{i\in I} A^i}$.

The latter claim is trivial once one notices that for any countable set C, the set of bounded functions in $\mathbb{R}^C$ is Borel, as it can be written

$$\bigcup_{n\in\mathbb{N}}\bigcap_{c\in C}\{a \in \mathbb{R}^C \mid |a_c| \leq n\},$$

and that the Tychonoff topology is indeed the required topology of pointwise convergence.

We next turn to the former claim. As mentioned above, the intuition describing the map from $\mathfrak{B}^i_s$ to $(\Delta(S))^S$ is to specify the beliefs of player i in each state. Hence, the image of $\mathfrak{B}^i_s$ under such a map is given by the subset of Ξ defined by two conditions: they satisfy positivity and they are constant over their support. Mathematically, these conditions are, respectively,

$$\begin{aligned}&\bigcap_{\omega\in S}\{s^i \in (\Delta(S))^S \mid s^i_\omega[\omega] > 0\}\\ &\quad\cap \bigcap_{\omega,\eta,\zeta\in S}\{s^i \in (\Delta(S))^S \mid s^i_\omega[\eta] > 0 \to s^i_\omega(\zeta) = s^i_\eta[\zeta]\},\end{aligned}$$

and again the topology is the topology of pointwise convergence. □

Denote by $\Sigma = \prod_{i\in I}\Sigma^i$ the space of strategy profiles over (S, I, A). The space Σ^i of strategies for player i on the countable space S is clearly a compact subspace of $(\Delta(A^i))^S$; hence, Σ is a compact space in the induced topology.

Proposition 12. *The Bayesian equilibrium correspondence $BE : \mathfrak{B} \to \Sigma$ has a Borel graph and takes on compact nonempty values.*

Proof. The fact that every Bayesian game with a countable state space has at least one Bayesian equilibrium follows from standard fixed point arguments; see, e.g., Simon (2003). The fact that the set of Bayesian equilibria is compact also follows by standard arguments. To show that the graph G of the BE correspondence is Borel, note that[26]

$$G = \Big\{ \big((s^i, g^i)_{i \in I}, \sigma \big) \in \mathfrak{B} \times \Sigma \mid \forall \omega \in S, \forall i \in I, \forall b \in A^i, \\ \sum_{v \in S} g^i\big(v, \sigma(\omega)\big) s^i_\omega[v] \geq \sum_{v \in S} g^i\big(v, b, \sigma^{-p}(\omega)\big) s^i_\omega[v] \Big\}. \qquad \square$$

Corollary 13. *There exists a Borel mapping $\psi : \mathfrak{B} \to \Sigma$ such that for all $\Lambda \in \mathfrak{B}$, $\psi(\Lambda)$ is a Bayesian equilibrium of Λ.*

The proof of Corollary 13 follows immediately from Proposition 12 and the selection theorem of Kuratowski and Ryll-Nardzewski (see, e.g., Aliprantis and Border 2006, Theorem 18.13, and Himmelberg 1975).

Recalling that $\mathcal{C}(\omega)$ is the common knowledge component containing a state ω, denote by $s|_{\mathcal{C}(\omega)}$ the restriction of the collection of types to the domain $\mathcal{C}(\omega)$ and denote by $g|_{\mathcal{C}(\omega)}$ the same restriction with respect to the payoff functions. Given two Bayesian games

$$(S, I, A, s_S, g_S) \quad \text{and} \quad (T, I, A, s_T, g_T)$$

with finite or countable state spaces and the same player and action sets, an *embedding* from S to T is an injective mapping $\phi : S \to T$ such that the following statements hold:

- For all $\omega \in S$, $i \in I$, and pure action profiles x, $g^i_S(\omega, x) = g^i_T(\phi(\omega), x)$.
- For all $\omega, \eta \in S$ and $i \in I$, $(s^i_S)_\omega[\eta] = (s^i_T)_{\phi(\omega)}[\phi(\eta)]$.

Note that $\psi(S)$ is then common knowledge in the type space s_T. Intuitively, the Bayesian game (S, I, A, s_S, g_S) is copied isomorphically to the Bayesian game $(\phi(S), I, A, s|_{\phi(S)}, g|_{\phi(S)})$.

Proposition 14. *Let $\Gamma = (\Omega, I, A, t, r)$ be a Bayesian game such that the common knowledge relation $\mathcal{E}$ is smooth. Let $\mathfrak{B} = \mathfrak{B}(S, I, A)$ be the set of Bayesian games that all share some same countable state space S and the same player and action space as Γ. Then there is a Borel map $\Phi : \Omega \to S$ and a Borel map*[27] *$\Lambda : \Omega/\mathcal{E} \to \mathfrak{B}$ such that for each $\omega \in \Omega$, if we denote*

$$\Gamma_\omega = \big(\mathcal{C}(\omega), I, A, t|_{\mathcal{C}(\omega)}, r|_{\mathcal{C}(\omega)}\big),$$

then $\Phi|_{\mathcal{C}(\omega)}$ is an embedding of Γ_ω in $\Lambda(\mathcal{C}(\omega))$.

See Figure 4. Note that for some ω, $\mathcal{C}(\omega)$ may be finite, in which case $\Phi|_{\mathcal{C}(\omega)}$ is not surjective.

[26]The quantifiers are all countable here.
[27]By Proposition 11, $\Omega/\mathcal{E}$ is standard Borel.

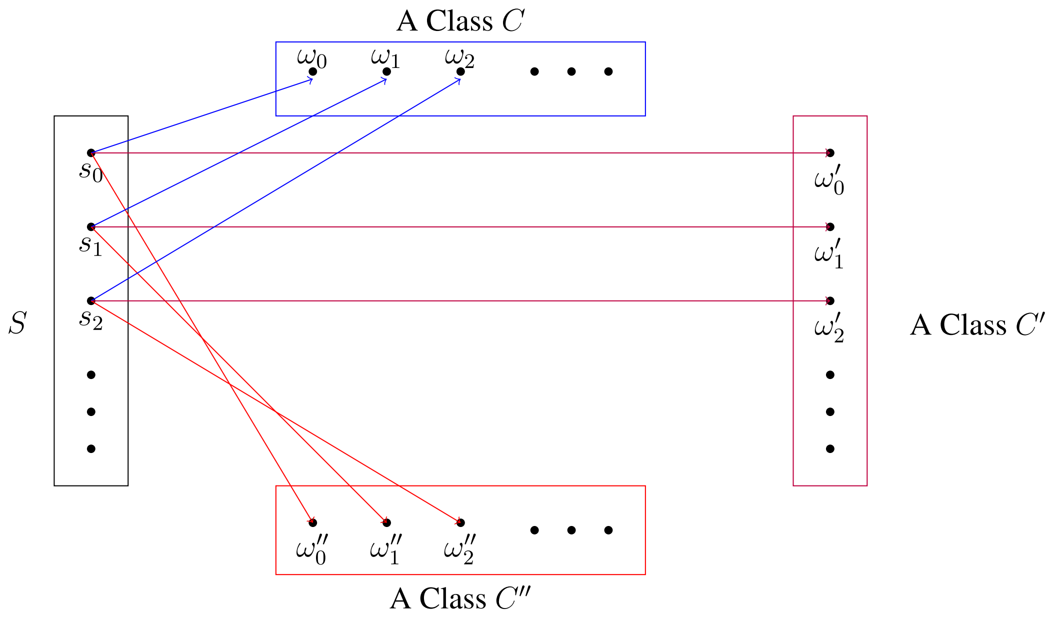

FIGURE 4. The mapping Φ.

PROOF OF PROPOSITION 14. Let $\Phi : \Omega \to S$ be as in Proposition 4; i.e., Borel and injective in each equivalence class.[28] We can then define $\Lambda(q) = (g_q^i, s_q^i)_{i\in I}$ for each common knowledge component $q \in \Omega/\mathcal{E}$ by

$$g_q^i\big(\Phi(\omega), x\big) = r^i(\omega, x),$$
$$g_q^i(s, x) \equiv 0 \quad \forall s \in S \setminus \Phi(q)$$

and

$$\big(s_q^i\big)_{\Phi(\omega)}\big[\Phi(\eta)\big] = t_\omega^i[\eta],$$
$$\big(s_q^i\big)_s = \delta_s \quad \forall s \in S \setminus \Phi(q),$$

where δ_s denotes the Dirac measure at s, i.e., players get payoff 0 and have perfect knowledge outside of the image of Φ on q. It is straightforward to check that Φ and Λ so defined satisfy the requirements. $\square$

PROOF OF THEOREM 1(I). Let $\psi : \mathfrak{B} \to \Sigma = (\prod_{i\in I} \Delta(A^i))^S$ be a Bayesian equilibrium selection as in Corollary 13. Let Φ, Λ be as in Proposition 14. For each $\omega \in \Omega$, define

$$\sigma(\omega) = \psi\big(\Lambda\big(\mathcal{C}(\omega)\big)\big)\big(\Phi(\omega)\big),$$

i.e., at stage ω, σ plays as the equilibrium that ψ chooses for the game $\Lambda(\mathcal{C}(\omega))$ at state $\Phi(\omega)$. Such σ is then an MBE. $\square$

[28] The surjectivity in each infinite equivalence class is not needed here.

Proof of Theorem 1(ii). Let $\Gamma_X = (X, \{1,2\}, \{L,R\} \times \{L,R\}, t^1_X, t^2_X, r^1_X, r^2_X)$, with $X = \{-1,1\} \times 2^{\mathbb{N}}$, be the two-player game presented in Hellman (2014a), with state space $X = \{-1,1\} \times 2^{\mathbb{N}}$, which does not possess an ε-MBE for sufficiently small ε; fix some such $\varepsilon > 0$. We recall the common knowledge relation $\mathcal{E}_X$ in this game: Define $S_X : X \to X$ by $S_X(x_0, x_1, x_2, x_3, \ldots) = (-x_0, x_2, x_3, x_4, \ldots)$. Then $\mathcal{E}_X$ is the equivalence relation on X given by

$$\mathcal{E}_X = \{(x,y) \mid \exists k, m \geq 0, S^k_X(x) = S^m_X(y)\}.$$

This relation is hyperfinite by Corollary 8.2 of Dougherty et al. (1994).

Let $\psi : X \to \Omega$ denote an embedding as in Proposition 10 with the induced positive type space $\tau = (\Omega, I, (t^i)_{i\in I})$ on Ω and a common prior μ satisfying $\psi_*(\mu_X) \ll \mu$. Define the payoffs

$$r^j(\omega, x) = \begin{cases} r^j_X(\psi^{-1}(\omega, x)) & \text{if } j = 1,2 \text{ and } \omega \in \psi(X), \\ 0 & \text{otherwise.} \end{cases}$$

By the properties of Γ_X and ε listed above, and the properties of ψ listed in Proposition 10, the game

$$\Gamma = (\Omega, \{L,R\}^I, I, t^1, t^2, t^3, \ldots, t^I, r^1, r^2, r^3, \ldots, r^I)$$

does not possess an ε-MBE. □

A.4 *Proving Theorem 2*

Recall the notion of a restriction of a type space from Section 3.2. Also recall that a subset A of a Polish space X is *analytic* if it the projection of a Borel set in $X \times Y$ for some Polish space Y.

Lemma 15. *The correspondence $\Omega \to \Delta(\Omega)$ given by*

$$\Psi(\omega) = \{\nu \in \Delta_a(\Omega) \mid \nu(\mathcal{C}(\omega)) = 1 \textit{ and } \nu|_{\mathcal{C}(\omega)} \textit{ is a common prior for } \tau_{\mathcal{C}(\omega)}\}$$

has an analytic graph in $\Omega \otimes \Delta(\Omega)$ and $|\Psi(\omega)| \leq 1$ for all $\omega \in \Omega$.

Proof. The fact that $|\Psi(\omega)| \leq 1$ (i.e., that on a countable space in which no proper nonempty subset is common knowledge there exists at most one common prior) is precisely Proposition 3 of Hellman and Samet (2012).

Let (f_n) be as in Proposition 3. Define $\Phi^i : \Omega \to \Delta(\Omega)$ by

$$\Psi^i(\omega) = \operatorname{conv}\{t^i(f_j(\omega)) \mid j \in \mathbb{N}\}$$

to be the convex hull of the priors for player i over those knowledge components contained in the common knowledge component of ω. Then $Gr(\Psi^i)$ is the projection of the

$$\begin{tikzcd}
\mu \arrow[r, "\text{RCD w.r.t. } \sigma(\mathcal{E})"] \arrow[dr, "\text{RCD w.r.t. } \sigma(\mathcal{E}^i)"'] & t \arrow[d, "\text{RCD w.r.t. } \sigma(\mathcal{E}^i|_C) \text{ for each } \mathcal{E}\text{-class } C"] \\
 & t^i
\end{tikzcd}$$

FIGURE 5. The commutative diagram.

Borel set $A^i \subseteq \Omega \times \Delta(\mathbb{N}) \times \Delta(\Omega)$ given by

$$A^i = \left\{(\omega, \eta, \nu) \in \Omega \times \Delta(\mathbb{N}) \times \Delta(\Omega) \mid \nu = \sum_{j \in \mathbb{N}} \eta[j] \cdot t^i\big(f_j(\omega)\big)\right\}.$$

Finally, it is known (e.g., Samet 1998) that $\Psi(\omega) = \bigcap_i \Psi^i(\omega)$; hence, $Gr(\Psi) = \bigcap_i Gr(\Phi^i)$; and the finite intersection of analytic sets is analytic. □

PROOF OF THEOREM 2. Clearly, property (iii) implies property (ii). Suppose (ii) holds; then, for Ψ as in Lemma 15, $|\Psi(\omega)| = 1$ for μ-a.e. $\omega \in \Omega$ and Ψ is clearly constant on each class of $\mathcal{E}$. Since $Gr(\Psi)$ is analytic, Ψ is a μ-measurable function on $\{\omega \mid \Psi(\omega) \neq \varnothing\}$, e.g., Kechris (1995, Theorem 21.10). Hence, after restricting Ψ to some $X \in \sigma(\mathcal{E})$ of full μ-measure, the graph of Ψ defines a Borel function $\psi : X \to \Delta(\Omega)$, constant on each $\mathcal{E}$-class but different on different $\mathcal{E}$-classes. Hence, $\mathcal{E}|_X$ is smooth.

Finally, assume property (i) holds and assume without loss of generality (w.l.o.g.) that $\Omega = X$. By Proposition 2, there is a μ-a.e. proper RCD t for μ given $\sigma(\mathcal{E})$. The claim that t is a common prior on a μ-a.e. component follows now from Proposition 16 below, which states formally the commutativity of the diagram in Figure 5. □

PROPOSITION 16. *Let $\mathcal{E}$, $\mathcal{D}$ be smooth CBERs on a Polish space Ω, with $\mathcal{D}$ refining $\mathcal{E}$ (that is, $\mathcal{D} \subseteq \mathcal{E}$; the classes of $\mathcal{D}$ are contained in classes of $\mathcal{E}$), let μ be a regular Borel probability measure on Ω, and let $t^{\mathcal{E}}$ and $t^{\mathcal{D}}$ be proper RCDs of μ with respect to $\sigma(\mathcal{E})$ and $\sigma(\mathcal{D})$ respectively. Then, for μ-a.e. $\omega \in \Omega$, for all Borel $A \subseteq \Omega$,*

$$t^{\mathcal{E}}_{\omega}\big(A \mid [\omega]_{\mathcal{D}}\big) = t^{\mathcal{D}}_{\omega}(A).$$

PROOF. Example 4 of Chang and Pollard (1997, p. 297) shows that for μ-a.e. $\mathcal{E}$-class C, $t^{\mathcal{D}}|_C$ (the type function $t^{\mathcal{D}}$ restricted to C) is a proper RCD of $t^{\mathcal{E}}|_C$ given $\sigma(\mathcal{D}|_C)$ (where $\mathcal{D}|_C$ is the restriction of the relation $\mathcal{D}$ to C); this clearly implies the proposition. □

A.5 *Proving Theorem 3*

PROOF OF THEOREM 3(I). By Theorem 2, almost every common knowledge component K has a common prior of the restricted type space τ_K. This is sufficient, by Theorem 1.a in Hellman (2014b), to conclude that there can be no acceptable bet over τ_K for such generic K. □

PROOF OF THEOREM 3(II). We first note that the conclusion of Theorem 3(ii) applies to the example Γ_X given in Section 4.2.3, as explained there, and in fact using only two players. Denote the type space there $(X, \{1,2\}, (t^1_X, t^2_X))$ with knowledge relations $(\mathcal{E}^1_X, \mathcal{E}^2_X)$. The common knowledge equivalence relation $\mathcal{E}_X$ is hyperfinite, as it is clearly induced by a $\mathbb{Z}$ action.

Suppose a countable space K is given with a type space. For our construction, we want to work a bit more generally with acceptable bets on arbitrary subsets $L \subseteq K$ of states and subset $J \subseteq I$ of players. To this end, for a bounded function $f : L \to \mathbb{R}$, denote by $\overline{f}$ the extension of f to K by $\overline{f}(\omega) = 0$ for $\omega \notin L$. Then we say that $(f^j)_{j\in J}$ is an acceptable bet for J on L if (compare with (2))

$$\sum_{i\in J} f^i(\omega) = 0, \qquad E^j\big[\overline{f}^j \mid \omega\big] > 0 \quad \forall j \in J, \omega \in L.$$

Now returning to the example Γ_X and given a nonsmooth belief-induced CBER $\mathcal{E}$ on Ω, let $\psi : X \to \Omega$ denote an embedding as in Proposition 10 with induced finitely supported positive type space $\tau = (\Omega, I, t)$ on Ω and common prior μ satisfying $\psi_*(\mu_X) \ll \mu$. Let K be a common knowledge component in τ such that $\psi^{-1}(K)$ is one of those components in τ_X on which there is an acceptable bet (f_1, f_2) (with $f_2 = -f_1$). By assumption, this is true for μ-almost every component K, since it is true for μ_X-almost every component in X (in fact, for every component), and μ and $\psi_*(\mu_X)$ have the same $\mathcal{E}$-saturated null sets. Hence, we also have an acceptable bet for these players, $L := \{1, 2\}$ on $\psi(\psi^{-1}(K)) \subseteq K$, and we need to show that there is an acceptable bet on the *entire* component K for *all* players:

PROPOSITION 17. *Let $\Gamma_K = (K, I, t)$ be a countable type space,*[29] *with finitely supported types,*[30] *such that K does not strictly contain any nonempty common knowledge component. Let $L \subseteq K$ and $J \subseteq I$, and suppose there is an acceptable bet $(g^j)_{j\in J}$ for J on L. Then there is an acceptable bet for all the players in I on all of K.*

PROOF. First, we show how to define an acceptable bet on the subset L for *all* players (in case $I \neq J$). Fix a player $i_0 \in J$. Choose a strictly positive function $h : L \to \mathbb{R}$ such that

$$E^{i_0}[h \mid \omega] < E^{i_0}\big[f^{i_0} \mid \omega\big] \quad \forall \omega \in L.$$

It is easy to see that this is possible, as types are finitely supported and $0 < E^{i_0}[f^{i_0} \mid \omega]$ for all $\omega \in L$. Then define $(g'^i)_{i\in I}$ on L by

$$g'^i = \begin{cases} g^i & \text{if } i \in J, i \neq i_0, \\ g^i - h & \text{if } i = i_0, \\ h \cdot \dfrac{1}{|I| - |J|} & \text{if } i \in I, i \notin J. \end{cases}$$

It is easy to check that $(g'^i)_{i\in I}$ is an acceptable bet on L.

[29] Positivity is not needed for this proposition.

[30] Indeed, in our case, the types are finitely supported and positive, as they are derived via Proposition 10 and the example of Section 4.2.3.

Now we proceed inductively and keep enlarging the domain L of $(g'^i)_{i\in I}$: Let $\omega_0 \in K \setminus L$ and $i_0 \in I$ be such that $C^{i_0}(\omega_0) \cap L \neq \varnothing$, where we recall that $C^{i_0}(\omega_0)$ is the knowledge component of i_0 containing ω_0. If there are no such i_0 and ω_0, then we are done by the assumption that there are no proper subsets that are common knowledge. Let $\gamma > 0$ be such that

$$t^{i_0}(\omega_0)[\omega_0] \cdot \gamma < E^{i_0}\big[\overline{g'}^j | \omega_0\big],$$

which exists as $E^{i_0}[\overline{g'}^j | \omega_0] > 0$, and define $(g'^i)_{i\in I}$ at ω_0 by

$$g'^i(\omega_0) = \begin{cases} -\gamma & \text{if } i = i_0, \\ \dfrac{\gamma}{|I|-1} & \text{if } i \neq i_0. \end{cases}$$

The term $(g'^i)_{i\in I}$ is an acceptable bet on $L \cup \{\omega_0\}$. Now repeat the procedure with $L \cup \{\omega_0\}$ replacing L and so forth until no states are left. The resulting profile is clearly zero sum but with positive expectation for each player at each stage. $\triangleleft$

$\square$

Appendix B: Belief-induced relationships

We have in several places relied on the fact that a belief-induced relation can always be assumed to be generated by types that are finitely supported. We formally state and prove this proposition here.

Proposition 18. *A CBER $\mathcal{E}$ is belief induced if and only if there are CBERs $\mathcal{E}^1, \dots, \mathcal{E}^m$ with finite classes that generate $\mathcal{E}$.*

It suffices to show this for each player separately (in a profile of players whose beliefs induce $\mathcal{E}$).

Lemma 19. *If $\mathcal{E}$ is a smooth CBER, then there are CBERs $\mathcal{E}_1$, $\mathcal{E}_2$ with finite classes that generate $\mathcal{E}$.*

Proof. It follows from Proposition 4 that there is Borel $\Phi : \Omega \to \mathbb{N}$ such that for each class C of $\mathcal{E}$, $\Phi|_C : C \to \mathbb{N}$ is an injection. Let $\mathcal{D}^1$ and $\mathcal{D}^2$ be two finite equivalence relations on $\mathbb{N}$ such that the relation generated by $\mathcal{D}^1$ and $\mathcal{D}^2$ has only one class. For example,

$$\mathcal{D}^1 = \big\{(0,1), (2,3), (4,5), \dots\big\}, \qquad \mathcal{D}^2 = \big\{(0), (1,2), (3,4), \dots\big\}$$

and define

$$\mathcal{E}^j = \big\{(x,y) \in \mathcal{E} \mid \big(\Phi(x), \Phi(y)\big) \in \mathcal{D}^j\big\}.$$ $\square$

As previously mentioned, there are CBERs that cannot be induced by finitely many smooth CBERs. This can be shown using the concept of the *cost* of a CBER $\mathcal{E}$ with an invariant[31] measure μ. We briefly recall this concept; for a more comprehensive treatment, see Kechris and Miller (2004).

A Borel graph G on a Polish space Ω is a Borel relation on Ω (i.e., a Borel subset of $\Omega \times \Omega$) that is irreflexive and symmetric. A Borel graph G induces a Borel equivalence relation $\mathcal{E}$ on Ω: $\mathcal{E}$ is the reflexive and transitive closure of G. We say that G *spans* $\mathcal{E}$. Given such a graph, for each $v \in \Omega$, let $d_G(v) \in \{0, 1, 2, \dots, \infty\}$ denote the *degree* of v, i.e., the cardinality of the set $\{w \in \Omega \mid (v, w) \in G\}$. Clearly, if $d_G(v)$ is countable for all $v \in \Omega$, then so is the induced CBER $\mathcal{E}$, and every CBER is spanned by some Borel graph with vertices of countable degree (the relation itself).

The *cost* of a CBER $\mathcal{E}$ (with respect to an invariant measure μ) is defined as

$$C_\mu(\mathcal{E}) := \inf\left\{\frac{1}{2}\int_\Omega d_G(\omega)\, d\mu(\omega) \mid G \text{ spans } \mathcal{E}\right\}.$$

A result of Levitt (e.g., Kechris and Miller 2004, Chapter 20) is that if T is a Borel transversal for a CBER $\mathcal{E}$ with an $\mathcal{E}$-invariant measure μ, then $C_\mu(\mathcal{E}) = \mu(\Omega \setminus T)$; in particular, if μ is finite and $\mathcal{E}$ is smooth (and hence possesses a Borel transversal by Proposition 1), then $C_\mu(\mathcal{E}) < \infty$.

Suppose that $\mathcal{E}$ is a CBER and, furthermore, that $\mathcal{E}$ is generated by $\mathcal{E}_1, \dots, \mathcal{E}_n$ (that is, the coarsest equivalence relation that each $\mathcal{E}_k$ refines). Let μ be $\mathcal{E}$-invariant and finite. Then μ is clearly also $\mathcal{E}_k$-invariant for each $k = 1, \dots, n$, and it is easy to see that[32]

$$C_\mu(\mathcal{E}) \le \sum_{k=1}^{n} C_\mu(\mathcal{E}_k).$$

Combining this observation with the result of Levitt, we see that if $\mathcal{E}_1, \dots, \mathcal{E}_n$ are smooth, $C_\mu(\mathcal{E})$ is finite.

Hence, to show a non-belief-induced CBER, it suffices to find a CBER with infinite cost with respect to some invariant measure $\mathcal{E}$ on it. A result of Gaboriau (see Kechris and Miller 2004, Corollary 27.10) states that if $\mathcal{E}$ is a CBER with finite invariant measure μ and $\mathcal{T}$ is a Borel tree[33] that spans $\mathcal{E}$, then $C_\mu(\mathcal{E}) = \frac{1}{2}\int_\Omega d_{\mathcal{T}}(\omega)\, d\mu(\omega)$.

Now let F_∞ denote the free (nonabelian) group with countably many generators. This group acts on 2^{F_∞} via $(f(x))(g) = x(f \cdot g)$ for $x \in 2^{F_\infty}$, $f, g \in F_\infty$, and induces a CBER by $x \sim y$ if and only if $\exists g \in F_\infty$ with $g \cdot x = y$. From this, one deduces easily that if $\mu = \prod_{f \in F_\infty}(\frac{1}{2}, \frac{1}{2})$ (which is clearly $\mathcal{E}$-invariant), then $C_\mu(\mathcal{E}) = \infty$ by Gaboriau's result.

References

Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third edition. Springer-Verlag, Berlin. [1112]

[31] A (not necessarily finite) measure is $\mathcal{E}$-invariant if for every Borel bijection $f : \Omega \to \Omega$ satisfying $f(\omega)\,\mathcal{E}\,\omega$ for all $\omega \in \mathcal{E}$, it holds that for all Borel $A \subseteq \Omega$, $\mu(f^{-1}(A)) = \mu(A)$.

[32] If $G_1, \dots, G_n$ span $\mathcal{E}_1, \dots, \mathcal{E}_n$, respectively, then $G = \bigcup_{k=1}^n G_n$ spans $\mathcal{E}$.

[33] A Borel tree is a Borel graph with no cycles.

Balder, Erik J. (1988), "Generalized equilibrium results for games with incomplete information." *Mathematics of Operations Research*, 13, 265–276. [1104]

Blackwell, David and Czesław Ryll-Nardzewski (1963), "Non-existence of everywhere proper conditional distributions." *The Annals of Mathematical Statistics*, 34, 223–225. [1092, 1107]

Brandenburger, Adam and Eddie Dekel (1987), "Common knowledge with probability 1." *Journal of Mathematical Economics*, 16, 237–245. [1093, 1104]

Chang, Joseph T. and David Pollard (1997), "Conditioning as disintegration." *Statistica Neerlandica*, 51, 287–317. [1115]

Dougherty, Randall, Stephen C. Jackson, and Alexander S. Kechris (1994), "The structure of hyperfinite Borel equivalence relations." *Transactions of the American Mathematical Society*, 341, 193–225. [1107, 1108, 1109, 1110, 1114]

Feinberg, Yossi (2000), "Characterizing common priors in the form of posteriors." *Journal of Economic Theory*, 91, 127–179. [1097, 1101]

Harrington, Leo A., Alexander S. Kechris, and Alain Louveau (1990), "A Glimm–Effros dichotomy for Borel equivalence relations." *Journal of the American Mathematical Society*, 3, 903–928. [1109]

Harsanyi, John C. (1967), "Games of incomplete information played by Bayesian players, part I: The basic model." *Management Science*, 14, 159–182. [1089]

Heifetz, Aviad (2006), "The positive foundation of the common prior assumption." *Games and Economic Behavior*, 56, 105–120. [1090, 1101]

Hellman, Ziv (2014a), "A game with no Bayesian approximate equilibrium." *Journal of Economic Theory*, 153, 138–151. [1090, 1095, 1097, 1099, 1114]

Hellman, Ziv (2014b), "Countable spaces and common priors." *International Journal of Game Theory*, 43, 193–213. [1097, 1101, 1105, 1115]

Hellman, Ziv and Dov Samet (2012), "How common are common priors?" *Games and Economic Behavior*, 74, 517–525. [1114]

Himmelberg, Charles J. (1975), "Measurable relations." *Fundamenta Mathematicae*, 87, 53–72. [1112]

Kechris, Alexander S. (1995), *Classical Descriptive Set Theory*, volume 156 of Graduate Texts in Mathematics. Springer-Verlag, New York. [1107, 1115]

Kechris, Alexander S. and Benjamin D. Miller (2004), *Topics in Orbit Equivalence*, volume 1852 of Lecture Notes in Mathematics. Springer-Verlag, Berlin. [1092, 1106, 1118]

Lehrer, Ehud and Dov Samet (2011), "Agreeing to agree." *Theoretical Economics*, 6, 269–287. [1098, 1099, 1100]

Milgrom, Paul R. and Nancy Stokey (1982), "Information, trade and common knowledge." *Journal of Economic Theory*, 26, 17–27. [1099]

Milgrom, Paul R. and Robert J. Weber (1985), "Distributional strategies for games with incomplete information." *Mathematics of Operations Research*, 10, 619–632. [1090, 1091, 1093, 1103, 1104, 1105]

Nielson, Lars T. (1984), "Common knowledge, communication, and convergence of beliefs." *Mathematical Social Sciences*, 8, 1–14. [1093, 1104]

Rudin, Walter (1986), *Real and Complex Analysis*, third edition. Higher Mathematics Series. McGraw-Hill, New York. [1102]

Samet, Dov (1998), "Common priors and separation of convex sets." *Games and Economic Behavior*, 24, 172–174. [1115]

Simon, Robert (2000), "The common prior assumption in belief spaces: An example." Discussion Paper #228, Center for the Study of Rationality, Hebrew University, Jerusalem. [1090, 1099]

Simon, Robert Samuel (2003), "Games of incomplete information, ergodic theory, and the measurability of equilibria." *Israel Journal of Mathematics*, 138, 73–92. [1089, 1094, 1095, 1099, 1111]

Stinchcombe, Maxwell B. (2011a), "Balance and discontinuities in infinite games with type-dependent strategies." *Journal of Economic Theory*, 146, 656–671. [1091, 1105]

Stinchcombe, Maxwell B. (2011b), "Correlated equilibrium existence for infinite games with type-dependent strategies." *Journal of Economic Theory*, 146, 638–655. [1106]

Co-editor Faruk Gul handled this manuscript.

Manuscript received 27 May, 2013; final version accepted 30 August, 2016; available online 12 September, 2016.