CONDITIONAL BELIEFS AND HIGHER-ORDER PREFERENCES*

Byung Soo Lee†

This version: 2013-09-09

Abstract

In this paper, we provide the Bayesian foundations of type structures—such as those used for epistemic analysis of iterated admissibility by Brandenburger *et al.* (2008)—in which beliefs are LPS's (lexicographic probability systems) rather than standard probability measures as in Mertens and Zamir (1985). This turns out to be a setting in which the distinction between preference hierarchies (Epstein and Wang, 1996) and belief hierarchies is meaningful and the former has conceptual advantages. In particular, using preference hierarchies allows us to identify conditions under which the distinction between LPS beliefs about types and LCPS (lexicographic conditional probability system) beliefs about types is a meaningful one. Furthermore, we construct "universal" LPS/LCPS type structures and find that they describe the same finite-order preferences even though the universal LPS type structure describes more hierarchies. Finally, we give an epistemic condition for iterated admissibility using coherent hierarchies that cannot be types.

KEYWORDS: Preference hierarchy, universal type structure, lexicographic probability system, iterated admissibility, epistemic game theory

JEL CLASSIFICATION: C72, D80.

*I thank Amanda Friedenberg for her helpful correspondence relating to this project. I am also grateful to Emiliano Catonini, Geir Asheim, and Aviad Heifetz for their informative comments.

†E-mail: byungsoolee@rotman.utoronto.ca / Address: Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6.

1 INTRODUCTION

A lexicographic probability system (LPS), as described by Blume *et al.* (1991a), is a finite list of probability measures called "theories". The decision maker starts by evaluating her choices under the primary theory. If the evaluation results in more than one optimal choice, she moves onto the next theory to break those ties. If these theories are mutually exclusive—e.g., have disjoint supports—then the LPS can be interpreted as a conditional probability system (CPS). An LPS comprised of mutually exclusive theories is aptly called a *lexicographic* conditional probability system (LCPS).

Blume *et al.* (1991b) showed that Nash equilibrium refinements such as admissible equilibrium, perfect equilibrium (Selten, 1975) and proper equilibrium (Myerson, 1978) could be simply characterized using LPS's instead of convergent sequences of probability measures. Kohlberg and Mertens (1986) showed that "given a game tree, a proper equilibrium of its normal form will give a sequential equilibrium in any variant of that tree obtained by applying any of the... inessential transformations." As one might expect, LPS's have since been added to the game theorist's toolkit for normal-form analysis of extensive-form games (cf. Asheim, 2001; Brandenburger and Friedenberg, 2007). In fact CPS's, as used by Myerson (1986) and Battigalli and Siniscalchi (1999) to describe beliefs in extensive-form games, can also be viewed as LPS's satisfying some restrictions on the set of events on which beliefs can be conditioned (cf. Hammond, 1994; Halpern, 2010).

More recently, Brandenburger *et al.* (2008)—henceforth BFK—solved a long-standing puzzle by giving an epistemic characterization of iterated admissibility.[1] Their solution involves using LPS's to resolve the sharp tension between caution and common belief of rationality that was identified by Samuelson (1992). However, some nontrivial conceptual questions have been raised by BFK's use of LCPS type structures—i.e., type structures in which types are mapped to LCPS beliefs about other types.

In order to answer these questions, the foundations of LPS type structures and their connection to belief hierarchies must be established in the same way that papers such as Mertens and Zamir (1985) and Brandenburger and Dekel (1993) did for standard probability type structures. The related issues are briefly enumerated and summarized below.

Issue 1 BFK's epistemic conditions are stated using LCPS type structures. However, the finite-order beliefs implied by such type structures are LPS's but not necessarily LCPS's. In fact, a "rational" player in such environments will necessarily have finite-order beliefs that are not LCPS's unless the underlying game is trivial. An immediate implication of this is that we must first define spaces of finite-order LPS beliefs in order to construct canonical LCPS type structures.

[1]i.e., iterated elimination of weakly dominated strategies.

Issue 2 Infinitely many LPS's can represent the same lexicographic expected utility (LEU) preference relation.[2] In the terminology of our paper, the space of LPS's contains redundant preference representations. Straightforwardly extending Brandenburger and Dekel (1993)'s construction of hierarchies—as Catonini (2012) does—results in a type space with many redundant types. To put it more precisely, for each LPS belief hierarchy constructed in this manner, there will be infinitely many distinct LPS belief hierarchies that represent the same *preference hierarchy*.

Issue 3 Such redundant types are problematic for three reasons. Firstly, the distinction between LCPS's and LPS's becomes purely cosmetic when the underlying space of uncertainty contains duplicate elements that represent the same descriptions of reality. Given that BFK's epistemic characterization of strategies that survive finitely many rounds of elimination of inadmissible strategies hold in any belief-complete LCPS type structure—even those that contain preference-redundant types of the sort mentioned above—it is unclear to what extent those results depend on the use of LCPS type structures versus LPS type structures, if at all.

Issue 4 Secondly, it is known that redundant types can be used as implicit coordination devices by the players. As demonstrated by Liu (2009), a redundant type structure can also be viewed as a nonredundant type structure with an expanded space of fundamental uncertainty—i.e., one with additional random variables that may serve as explicit coordination devices. The presence of redundant types can therefore change what is meant by players acting independently.

Issue 5 Thirdly, because spaces of finite-order LPS beliefs also contain redundant representations of each finite-order preference relation, finite-order LPS beliefs are more informative about higher-order preferences than they should be. Ann's first-order belief represents her first-order preference relation over some space of acts. It seems unreasonable that we can deduce more about Ann's higher-order preferences from Ann's first-order belief than we can from Ann's first-order preference relation. However, due to the aforementioned redundancies, marginal beliefs are more informative about joint beliefs than they should be.

Issue 6 If the belief hierarchy approach leads to type spaces with redundant preference hierarchies, then it makes sense to construct preference hierarchies directly.[3] Epstein and Wang (1996) provide a template for constructing such "beliefs about beliefs without probabilities". However, Epstein and Wang (1996) do not work with

[2]Redundant representations of LEU preferences in the space of LPS beliefs are given more detailed treatment in Lee (2013).

[3]Di Tillio (2008) explicitly constructs hierarchies of preference relations. However, his techniques resist straightforward generalization to our setting because they rely on the set of m^{th}-order preferences being finite for all m. I thank Pierpaolo Battigalli for this reference.

preference relations directly, but rather use utility functions over acts as proxies for preference relations. Ganguli and Heifetz (2013) take an analogous approach by, roughly speaking, using LEU functions—which are families of linear functionals—as proxies for LEU preferences.[4] Unfortunately, even after normalizations, such an approach cannot prevent the sort of preference-redundancies responsible for Issues 3–5.

To address these issues, we begin by constructing the set of all coherent hierarchies of LEU *preferences* so that no two elements of the set represent the same preference hierarchy. We then obtain a somewhat surprising result: There are LEU preference hierarchies that cannot be LPS belief hierarchies. In particular, there are some LEU preferences hierarchies that cannot be represented by types! Like Epstein and Wang (1996), we do not model the preferences relations directly but instead use an abstract space of objects that represent them.

We also construct two canonical type structures that may be called "universal". The first induces all preference hierarchies that can be described by LPS type structures. The second induces all preference hierarchies that can be described by all nonredundant—i.e., meaningful—LCPS type structures.

The canonical LPS type structure is more powerful in the sense that it induces more preference hierarchies than the canonical LCPS type structure. However, we also find that both induce all finite-order preferences. It follows that the distinction between LPS types and LCPS types is an infinite-order notion rather than a finite-order one. This suggests that finite-order epistemic conditions that can be stated in LCPS type structures have equivalent "translations" in LPS type structures. For example, Dekel *et al.* (2013) show that BFK's epistemic characterization of finitely many rounds of weak dominance elimination in LCPS type structures carries over to LPS type structures.

Furthermore, if redundant types are permitted, a hierarchy can be described by an LPS type structure if and only if it can be described by an LCPS type structure. In other words, the class of LPS type structures and the class of LCPS type structures are equally expressive in some sense when redundant types are present.

One of the main findings of Keisler and Lee (2012) was that there exist belief-complete LCPS type structures in which *rationality and common assumption of rationality (RCAR)*—BFK's analog of *rationality and common belief of rationality (RCBR)*—characterizes iterated admissibility. This was surprising in light of BFK's Theorem 10.1, which says that, in a belief-complete LCPS type structure with continuous type-belief maps, no state of the world can satisfy RCAR. The two contrasting results raise the question of which "large" type structure is the most appropriate model for the analysis of RCAR.

In this paper, we attempt to bypass the issue altogether by giving an epistemic characterization of iterated admissibility within an explicit model of preference hier-

[4]To be more precise, the setup in Ganguli and Heifetz (2013) permits more general preferences than just LEU preferences.

archies as opposed to in an implicit model such as a type structure. Our condition, which we also call RCAR, is a close analog of BFK's RCAR in explicit hierarchy state spaces. Our proof relies on the use of preference hierarchies that cannot be types, which seems to be consistent with BFK's earlier intuition about the nonexistence of RCAR in complete and continuous type structures.

The remainder of this paper is organized as follows: Section 2 contains mathematical preliminaries. Section 3 formally defines LPS's, LCPS's, and our nonredundant space of LEU preference representations. Some of the key issues mentioned above are also given slightly more detailed treatment. Finite-order LEU preferences and coherent LEU preference hierarchies are constructed in Section 4. Section 5 follows up by constructing our canonical LPS/LCPS type structures. Results about LCPS type structures and their relation to LPS type structures can be found in Section 6. Section 7 gives our epistemic characterization of iterated admissibility. Finally, Section 8 concludes. Proofs that are either too long or add little to the essential content of the paper have been postponed to the appendices.

2 MATHEMATICAL PRELIMINARIES

Definition 2.1. A topological space is an ordered pair $\langle X, \mathscr{T} \rangle$, where X is a set and $\mathscr{T}$ is a topology on X.

Whenever there is no risk of confusion, we will refer to the topological space $\langle X, \mathscr{T} \rangle$ as X for the sake of brevity. Furthermore, we may also refer to a topological space X without specifying the associated topology. In such cases, we will let $\mathscr{T}(X)$ denote this unspecified topology on X.

Definition 2.2. Let $\langle X, \mathscr{T} \rangle$ be a topological space and let $Y \subseteq X$. The subspace topology on Y (relative to X) is the collection $\mathscr{T}|Y \equiv \{U \cap Y : U \in \mathscr{T}\}$. The topological space $\langle Y, \mathscr{T}|Y \rangle$ is called a topological subspace of $\langle X, \mathscr{T} \rangle$.

Whenever there is no risk of confusion, we will refer to the topological space $\langle Y, \mathscr{T}|Y \rangle$ as the subspace Y of X for the sake of brevity.

Definition 2.3. A topological space $\langle X, \mathscr{T} \rangle$ is a Polish space if $\mathscr{T}$ is separable and completely metrizable.

Definition 2.4. A standard Borel space is an ordered pair $\langle X, \mathscr{B} \rangle$, where X is a set and $\mathscr{B}$ is a σ-algebra on X such that there exists some Polish topology $\mathscr{T}$ that generates $\mathscr{B}$. Elements of $\mathscr{B}$ are called Borel subsets of X.

Whenever there is no risk of confusion, we will refer to the standard Borel space $\langle X, \mathscr{B} \rangle$ as X for the sake of brevity. Furthermore, we may also refer to a standard Borel space X without specifying the associated Borel σ-algebra. In such cases, we will let $\mathscr{B}(X)$ denote this unspecified Borel σ-algebra.

Definition 2.5. Let $\langle X, \mathscr{B} \rangle$ be a standard Borel space and let $Y \in \mathscr{B}$. The subspace σ-algebra on Y (relative to X) is the collection $\mathscr{B}|Y \equiv \{U \cap Y : U \in \mathscr{B}\}$. The Borel subspace $\langle Y, \mathscr{B}|Y \rangle$ is itself a standard Borel space.

3 BELIEFS AND PREFERENCES

3.1 BELIEFS

Definition 3.1. Let X be a standard Borel space. The set of all Borel probability measures defined on X is denoted by $\mathsf{P}(X)$. $\mathsf{P}(X)$ is itself a standard Borel space when it is endowed with the Borel σ-algebra generated sets of the form

(1) $\{\mu \in \mathsf{P}(X) \colon \mu(E) > p\}$, where $E \in \mathscr{B}(X)$ and $p \in [0, 1]$

Definition 3.2. Let X be a standard Borel space. The set of all lexicographic probability systems (LPS's) on X is denoted by $\mathsf{LPS}(X)$ and defined as follows.

(2) $$\mathsf{LPS}_m(X) \equiv \prod_{j=1}^{m} \mathsf{P}(X) \qquad\qquad \mathsf{LPS}(X) \equiv \bigcup_{m \geq 1} \mathsf{LPS}_m(X)$$

When they are endowed with the usual product/union σ-algebras, $\mathsf{LPS}_m(X)$ and $\mathsf{LPS}(X)$ are standard Borel spaces.

Definition 3.3. Let X be a standard Borel space. Probability measures $\mu, \nu \in \mathsf{P}(X)$ are mutually singular if there exist disjoint Borel sets U, V in X such that $\mu(U) = \nu(V) = 1$. We write $\mu \perp \nu$ to indicate that μ and ν are mutually singular.[5]

Definition 3.4. Let X be a standard Borel space and let $\sigma = (\mu_1, \ldots, \mu_m) \in \mathsf{LPS}_m(X)$, where $m \geq 1$. The LPS σ is a lexicographic conditional probability system (LCPS) if it is comprised of pairwise mutually singular probability measures. Equivalently, σ is an LCPS if there exists pairwise-disjoint Borel sets $U_1, \ldots, U_m$ such that

(3) $\mu_1(U_1) = \cdots = \mu_m(U_m) = 1$

The set of all LCPS's on X is denoted by $\mathsf{LCPS}(X)$. Because $\mathsf{LCPS}(X)$ is a subset of $\mathsf{LPS}(X)$, it is itself a standard Borel space when viewed as a Borel subspace of $\mathsf{LPS}(X)$. For all $m \geq 1$, $\mathsf{LCPS}_m(X) \equiv \mathsf{LPS}_m(X) \cap \mathsf{LCPS}(X)$.

3.2 PREFERENCES AND REDUNDANCY

For Definitions 3.5–3.8, let X be a standard Borel space.

[5] A probability measure can be viewed as an element of some linear space. A pair of mutually singular probability measures can be viewed as a pair of orthogonal ($\perp$) elements in this space.

Definition 3.5. An act defined on X is a Borel map $f\colon X \to [0,1]$. The set of all acts defined on X is denoted by $\mathsf{ACT}(X)$.

Definition 3.6. Let $f \in \mathsf{ACT}(X)$ and let $\sigma = (\mu_1, \ldots, \mu_m) \in \mathsf{LPS}(X)$, where $m \geq 1$. The lexicographic expected utility (LEU) of f under σ is the following m-tuple of expected utilities (EUs).

$$(4) \quad \mathbf{E}_\sigma[f] \equiv \left(\int_X f \, d\mu_1, \ldots, \int_X f \, d\mu_m \right) = (\mathbf{E}_{\mu_1}[f], \ldots, \mathbf{E}_{\mu_m}[f])$$

Definition 3.7. Let $\sigma = (\mu_1, \ldots, \mu_n) \in \mathsf{LPS}(X)$ be an LPS. The preference relation $\succsim^\sigma$ on $\mathsf{ACT}(X)$ is defined as follows, where $\geq^{\mathsf{L}}$ is the lexicographic order.[6]

$$(5) \quad \forall f, g \in \mathsf{ACT}(X) \quad f \succsim^\sigma g \iff \mathrm{E}_\sigma[f] \geq^{\mathsf{L}} \mathrm{E}_\sigma[g]$$

The space $\mathsf{P}(X) = \mathsf{LPS}_1(X)$ of probability measures has the nice property that, for all $\mu, \nu \in \mathsf{P}(X)$, $\succsim^\mu = \succsim^\nu$ if and only if $\mu = \nu$. In other words, each EU preference relation is represented by exactly one element of $\mathsf{P}(X)$. Unfortunately, the analogous property does not hold for LPS's. In order to carefully handle the hazards arising from this fact, the following equivalence notion is needed.

Definition 3.8. Let $\rho, \sigma \in \mathsf{LPS}(X)$. We say that ρ and σ are preference-equivalent if $\succsim^\rho = \succsim^\sigma$ and write $\rho \cong \sigma$.

Each LPS $\sigma \in \mathsf{LPS}(X)$ is preference-equivalent to an uncountable number of LPS's in $\mathsf{LPS}(X)$. This poses nontrivial conceptual challenges to meaningfully defining beliefs about beliefs—i.e., LPS's about LPS's—that can be interpreted as conditional probability systems—i.e., those that can be written as LCPS's. This is because the mutual singularity of probability measures loses its incisiveness if the space of uncertainty contains redundant elements. The following simple example illustrates this issue.

Consider a game environment with two players, who are called "Ann" and "Bob". Ann and Bob each have finite strategy sets, which are respectively denoted $S_a = \{U, D\}$ and $S_b = \{L, R\}$. Each player has a first-order LPS belief about what strategy his/her opponent will use. Suppose we are interested in Ann's second-order beliefs—i.e., her beliefs about Bob's strategy and first-order belief—that can be written as LCPS's. Such second-order beliefs belong to the set $\mathsf{LCPS}(S_b \times \mathsf{LPS}(S_a))$.

Let $\rho, \rho' \in \mathsf{LPS}(S_a)$ be preference-equivalent first-order beliefs of Bob such that $\rho \neq \rho'$. We can construct a second-order belief $\sigma = (\mu_1, \mu_2)$ for Ann, where

$$\mu_1(\{(L, \rho)\}) = 1 = \mu_2(\{(L, \rho')\})$$

[6]The lexicographic order $\geq^{\mathsf{L}}$ is the union of the relations $>^{\mathsf{L}}$ and $=^{\mathsf{L}}$, which are defined as follows. $(a_1, \ldots, a_m) =^{\mathsf{L}} (b_1, \ldots, b_m)$ if and only if $(a_1, \ldots, a_m) = (b_1, \ldots, b_m)$; $(a_1, \ldots, a_m) >^{\mathsf{L}} (b_1, \ldots, b_m)$ if and only if there exists some k such that $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$ and $a_{k+1} > b_{k+1}$.

The belief σ is clearly an LCPS because μ_1 and μ_2 are mutually singular. However, are (L, ρ) and (L, ρ') different descriptions of reality?

If the subjectivist interpretation of probability—and beliefs in general—is taken seriously, it might be argued that the substantive content of an LPS is entirely captured by the associated preference relation. As such, one might then argue that beliefs about beliefs have content only to the extent that they describe beliefs about preferences. The LPS $\sigma' = (\mu_1, \mu_1)$ has the same content as the LCPS σ in that regard, but σ' is not an LCPS.

It follows that the beliefs in $\mathsf{LCPS}(S_b \times \mathsf{LPS}(S_a))$ and $\mathsf{LPS}(S_b \times \mathsf{LPS}(S_a))$ are descriptively equivalent under the Savage paradigm. As such, having a well-behaved space of LPS beliefs that contains exactly one representation of each LEU preference relation is a necessary prerequisite to defining meaningful LCPS second-order beliefs. The following result from Lee (2013) delivers precisely such a space.

Proposition 1. *There exists a Borel subspace* $\mathsf{LEU}(X) \subseteq \mathsf{LPS}(X)$ *and a surjective Borel map* $\varsigma_X : \mathsf{LPS}(X) \to \mathsf{LEU}(X)$ *such that*

(6) $\quad \forall \sigma, \sigma' \in \mathsf{LPS}(X) \quad \varsigma_X(\sigma) \cong \varsigma_X(\sigma') \iff \sigma \cong \sigma'$

(7) $\quad \forall \sigma \in \mathsf{LPS}(X) \quad \sigma \cong \varsigma_X(\sigma)$

Furthermore, $\mathsf{LEU}(X)$ *only contains minimal-length*[7] *representations of LEU preferences.*

Definition 3.9. For each standard Borel space X, fix the ordered pair $\langle \mathsf{LEU}(X), \varsigma_X \rangle$ that exists by Proposition 1. For all $m \geq 1$, let $\mathsf{LEU}_m(X) \equiv \mathsf{LEU}(X) \cap \mathsf{LPS}_m(X)$.

3.3 MARGINAL BELIEFS VERSUS MARGINAL PREFERENCES

For Definitions 3.10–3.11, let X, Y be nonempty standard Borel spaces. The following is a naive but prima facie natural way to define a marginal operator on LPS's.

Definition 3.10. Let $\sigma = (\mu_1, \ldots, \mu_m) \in \mathsf{LPS}(X \times Y)$, where $m \geq 1$. The belief-marginal of σ on X is written as $\operatorname{marg}_X \sigma$ and defined as follows.

(8) $\quad \operatorname{marg}_X \sigma \equiv (\operatorname{marg}_X \mu_1, \ldots, \operatorname{marg}_X \mu_m)$

Unfortunately, $\operatorname{marg}_X \sigma$ reveals more information about σ than simply how $\succsim^\sigma$ ranks acts that are measurable with respect to X. More precisely speaking, the belief-marginal operator *violates* the following property.

(9) $\quad \forall \rho, \sigma \in \mathsf{LPS}(X \times Y) \quad \operatorname{marg}_X \rho \cong \operatorname{marg}_X \sigma \iff \operatorname{marg}_X \rho = \operatorname{marg}_X \sigma$

Note that the analogous property holds for probability measures.

(10) $\quad \forall \mu, \nu \in \mathsf{P}(X \times Y) \quad \operatorname{marg}_X \mu \cong \operatorname{marg}_X \nu \iff \operatorname{marg}_X \mu = \operatorname{marg}_X \nu$

[7] An LPS σ has minimal length if there is no shorter LPS ρ such that $\sigma \cong \rho$.

The belief hierarchies constructed using this naive marginal operator exhibit perverse irregularities. For example, knowing Ann's first-order *belief* will permit us rule out many second-order *preferences* that are consistent with it. To see this more clearly, let us once again consider a game played by Ann and Bob. Ann and Bob each have finite strategy sets, which are respectively denoted $S_a = \{U, D\}$ and $S_b = \{L, R\}$.

$$\begin{array}{c|c|c|} & L & R \\ \hline \rho & 1 & 0 \\ \hline \rho' & 0 & 0 \\ \hline \end{array} \qquad\qquad \begin{array}{c|c|c|} & L & R \\ \hline \rho & 0 & 0 \\ \hline \rho' & 1 & 0 \\ \hline \end{array}$$

$$\mu_1 \qquad\qquad\qquad\qquad \mu_2$$

Figure 1: Second-order beliefs $\sigma = (\mu_1, \mu_2), \sigma' = (\mu_1)$

Each player has a first-order LPS belief about what strategy his/her opponent will use. Let $\rho, \rho' \in \mathsf{LPS}(S_a)$ be two possible first-order beliefs of Bob such that $\rho \ncong \rho'$. Ann's second-order beliefs belong to $\mathsf{LPS}(S_b \times \mathsf{LPS}(S_a))$. Consider the second-order beliefs σ and σ', which are defined in Figure 1.

The belief-marginals represent the same preference relation over $\mathsf{ACT}(S_b)$—i.e., $\operatorname{marg}_{S_b} \sigma \cong \operatorname{marg}_{S_b} \sigma'$—which is evident from the following chain of easily verified equalities.

$$\begin{aligned} \operatorname{marg}_{S_b} \sigma &= (\operatorname{marg}_{S_b} \mu_1, \operatorname{marg}_{S_b} \mu_2) \\ &= (\operatorname{marg}_{S_b} \mu_1, \operatorname{marg}_{S_b} \mu_1) \cong (\operatorname{marg}_{S_b} \mu_1) = \operatorname{marg}_{S_b} \sigma' \end{aligned}$$

However, if the analyst knows that Ann's first-order belief is equal to $\operatorname{marg}_{S_b} \sigma'$, then she can immediately rule out the possibility that Ann's second-order belief is σ despite the fact that σ and σ' induce the same first-order preferences—i.e., preferences over $\mathsf{ACT}(S_b)$. It seems eminently unreasonable to argue that LPS's (ν, ν) and (ν)—where $\nu = \operatorname{marg}_{S_b} \mu_1$—could be viewed as representing different beliefs about Bob's strategies.

Proposition 1 allows us to define a marginal operator that avoids these irregularities. We will call this the preference-marginal operator to distinguish it from the belief-marginal operator.

Definition 3.11. Let $\sigma = (\mu_1, \ldots, \mu_m) \in \mathsf{LPS}(X \times Y)$, where $m \geq 1$. The preference-marginal of σ on X is written as $\operatorname{margp}_X \sigma$ and defined as follows.

(11) $\operatorname{margp}_X \sigma \equiv \varsigma_X(\operatorname{marg}_X \sigma)$

The mapping $\sigma \mapsto \operatorname{margp}_X \sigma$ is Borel since it is a composition of Borel maps. Furthermore, it is immediately seen that the following property holds.

$$\forall \rho, \sigma \in \mathsf{LPS}(X \times Y) \quad \operatorname{margp}_X \rho \cong \operatorname{margp}_X \sigma \iff \operatorname{margp}_X \rho = \operatorname{margp}_X \sigma$$

4.1 BASIC ENVIRONMENT

Consider a game setting where $I = \{a, b\}$ denotes the set of all human players. Players a and b are respectively called Ann and Bob. As usual, we let $-a \equiv b$ and $-b \equiv a$. With apologies to Bob, we also adopt the convention of using female pronouns when we refer to generic players—e.g., Player i knows *her* own preferences. Player i's strategy set is denoted by S_i and it is a nonempty standard Borel space. We assume that each player has two or more strategies. The basic uncertainty in this environment is about the strategies that are played.[8]

4.2 FINITE-ORDER PREFERENCES

We construct the spaces of finite-order preferences by induction. We begin by defining the base cases.[9]

$$X_i^0 \equiv S_i \qquad\qquad X_i^1 \equiv X_i^0 \times \mathsf{LEU}(X_{-i}^0)$$

Let $\pi_i^0 \colon X_i^1 \to X_i^0$ and $\varrho_i^1 \colon X_i^1 \to \mathsf{LEU}(X_i^0)$ be the natural projections.

$$\pi_i^0 \equiv x_i^1 \mapsto \operatorname{proj}_{X_i^0} x_i^1 \qquad\qquad \varrho_i^1 \equiv x_i^1 \mapsto \operatorname{proj}_{\mathsf{LEU}(X_{-i}^0)} x_i^1$$

Ann's space of n^{th}-order uncertainty will be X_b^{n-1} and her n^{th}-order preferences will belong to $\mathsf{LEU}(X_b^{n-1})$. For $n \geq 1$, the tuple $(X_i^{n+1}, \pi_i^n, \varrho_i^{n+1}, \hat{\pi}_i^{n-1})$ of objects is defined inductively. First, let $\hat{\pi}_i^{n-1} \colon \mathsf{LEU}(X_i^n) \to \mathsf{LEU}(X_i^{n-1})$ be the preference-marginal operation.

$$\hat{\pi}_i^{n-1} \equiv \sigma \mapsto \operatorname{margp}_{X_i^{n-1}} \sigma$$

We define X_i^{n+1} as the set of ordered pairs in $X_i^n \times \mathsf{LEU}(X_{-i}^n)$ that are coherent in a specific sense.

$$X_i^{n+1} \equiv \{(x_i^n, h_i^{n+1}) \in X_i^n \times \mathsf{LEU}(X_{-i}^n) \colon \varrho_i^n(x_i^n) = \hat{\pi}_i^{n-1}(h_i^{n+1})\}$$

Once we define $\pi_i^n \colon X_i^{n+1} \to X_i^n$ and $\varrho_i^{n+1} \colon X_i^{n+1} \to \mathsf{LEU}(X_{-i}^n)$ to be the natural projections, the diagram in Figure 2 commutes.

Lemma 4.1. *For all $n \geq 0$, X_i^n and $\mathsf{LEU}(X_i^n)$ are nonempty standard Borel spaces.*

Proof of Lemma 4.1. See Appendix A. □

[8] Our definitions and results easily extend to environments with incomplete information and any finite number of players. The simplification herein reduces notational clutter without sacrificing essential content.

[9] The language and notation in some parts borrow liberally—sometimes verbatim—from the elegant setup of Di Tillio (2008).

$$
\begin{tikzcd}
X_i^{n+1} \arrow[r, "\varrho_i^{n+1}"] \arrow[d, "\pi_i^n"] & \mathsf{LEU}(X_{-i}^n) \arrow[d, "\hat{\pi}_i^{n-1}"] \\
X_i^n \arrow[r, "\varrho_i^n"] & \mathsf{LEU}(X_{-i}^{n-1})
\end{tikzcd}
$$

Figure 2: Coherency

4.3 COHERENT HIERARCHIES

Finally, we can define coherent preference hierarchies (or simply hierarchies), which are the objects of principal interest to us.

Definition 4.1. A coherent preference hierarchy of Player i is a sequence that belongs to the set H_i, which is defined as follows.

$$H_i \equiv \left\{ (h_i^1, h_i^2, \dots) \in \prod_{n \geq 0} \mathsf{LEU}(X_{-i}^n) \colon \forall n \geq 1 \quad \hat{\pi}_i^{n-1}(h_i^{n+1}) = h_i^n \right\}$$

Definition 4.2. A state of Player i is a sequence that belongs to the set X_i, which is defined as follows.

$$X_i \equiv \varprojlim_n X_i^n = \left\{ (x_i^0, x_i^1, \dots) \in \prod_{n \geq 0} X_i^n \colon \forall n \geq 0 \quad \pi_i^n(x_i^{n+1}) = x_i^n \right\}$$

Take any $x_i = (x_i^0, x_i^1, \dots) \in X_i$. The first coordinate of

$$x_i^{n+1} = (\overbrace{\pi^n(x_i^{n+1})}^{=x_i^n}, \varrho_i^{n+1}(x_i^{n+1}))$$

exactly duplicates the information described by x_i^n. It follows that the descriptive content of x_i and $\left(\varpi_i^0(x_i), (\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i))_{n\geq 0}\right)$ are equal. It follows that the map

$$x_i \mapsto \left(\varpi_i^0(x_i), (\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i))_{n\geq 0}\right)$$

is a natural Borel isomorphism from X_i into $X_i^0 \times H_i$ that exactly preserves all information about Player i. As such, with apologies for the abuse of notation, we use the expressions X_i and $X_i^0 \times H_i$ interchangeably in this paper depending on which form is more convenient for the task at hand.

Lemma 4.2. *The following map is a Borel isomorphism.*

$$X_i \to X_i^0 \times H_i \qquad\qquad x_i \mapsto \left(\varpi_i^0(x_i), (\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i))_{n\geq 0}\right)$$

Proof of Lemma 4.2. See Appendix A. $\square$

Definition 4.3. For every $n \geq 0$, $\varpi_i^n \colon X_i \to X_i^n$ denotes the natural projection $x_i \mapsto \operatorname{proj}_{X_i^n} x_i$.

Lemma 4.3. *X_i and H_i are nonempty standard Borel spaces.*

Proof of Lemma 4.3. See Appendix A. $\square$

Definition 4.4. For all $n \geq 1$, the maps $\mathbb{bh}_i^n$ and $\mathbb{bh}_i$ are defined as follows.

$$\mathbb{bh}_i^n \colon \mathsf{LPS}(X_{-i}) \to \mathsf{LEU}(X_i^{n-1}) \qquad \mathbb{bh}_i^n(\sigma) \equiv \operatorname{margp}_{X_{-i}^{n-1}} \sigma \tag{12}$$

$$\mathbb{bh}_i \colon \mathsf{LPS}(X_{-i}) \to H_i \qquad \mathbb{bh}_i(\sigma) \equiv (\mathbb{bh}_i^n(\sigma))_{n\geq 1} \tag{13}$$

The maps $\mathbb{bh}_i^n$ and $\mathbb{bh}_i$ respectively impute Ann's finite-order preference and preference hierarchy to her belief about Bob's states. The notation is intended to remind the reader that $\mathbb{bh}_i$ is a *belief-to-hierarchy* map.

Lemma 4.4. *Let $n \geq 1$. The maps $\mathbb{bh}_i^n$ and $\mathbb{bh}_i$ are Borel. Furthermore, each LEU preference relation over $\mathsf{ACT}(X_{-i})$ is uniquely identified by a hierarchy in H_i—i.e., the following statement holds.*

$$\forall \rho, \sigma \in \mathsf{LPS}(X_{-i}) \quad \rho \cong \sigma \iff \mathbb{bh}_i(\rho) = \mathbb{bh}_i(\sigma) \tag{14}$$

Proof of Lemma 4.4. See Appendix A. $\square$

Definition 4.5. Let X be a standard Borel space and $E \subseteq \mathsf{LPS}(X)$. The set E is said to be preference-repetitive (or simply repetitive) if there exist distinct LPS's $\rho, \sigma \in E$ that represent the same preference relation. Such ρ and σ are said to be preference-redundant (or simply redundant) in E.

An immediate consequence of Lemma 4.4 is that the restriction of $\mathbb{bh}_i$ to any non-repetitive subspace of $\mathsf{LPS}(X_{-i})$ is necessarily one-to-one. For example, $\mathbb{bh}_i|_{\mathsf{LEU}(X_{-i})}$ and $\mathbb{bh}_i|_{\mathsf{LCPS}(X_{-i})}$ are both one-to-one maps.

Corollary 4.1. *Image of any non-repetitive Borel set $E \subseteq \mathsf{LPS}(X_{-i})$ under $\mathbb{bh}_i$ is Borel.*

Proof of Corollary 4.1. The map $\mathbb{bh}_i$ is one-to-one when its domain is restricted to E. It follows that $\mathbb{bh}_i|E$ is a Borel isomorphism from E into $\mathbb{bh}_i(E)$. $\square$

Corollary 4.2. *A hierarchy can be described by an LPS in $\mathsf{LPS}(X_{-i})$ if and only if it can be described by an LPS in $\mathsf{LEU}(X_{-i})$. In other words, the following equality holds.*

$$\mathbb{bh}_i(\mathsf{LPS}(X_{-i})) = \mathbb{bh}_i(\mathsf{LEU}(X_{-i}))$$

Furthermore, the set of all such hierarchies—i.e., $\mathbb{bh}_i(\mathsf{LPS}(X_{-i}))$—is Borel in H_i.

Proof of Corollary 4.2. By Lemma 4.4 and the definition of $\mathsf{LEU}(X_{-i})$, the following is true.

$$\forall \rho \in \mathsf{LPS}(X_{-i})\ \exists \sigma \in \mathsf{LEU}(X_{-i}) \quad \rho \cong \sigma \wedge \mathbb{h}_i(\rho) = \mathbb{h}_i(\sigma)$$

From this, we can deduce the inclusion $\mathbb{h}_i(\mathsf{LPS}(X_{-i})) \subseteq \mathbb{h}_i(\mathsf{LEU}(X_{-i}))$.

$$\begin{aligned} \mathsf{LPS}(X_{-i}) \supseteq \mathsf{LEU}(X_{-i}) \qquad & \therefore \quad \mathbb{h}_i(\mathsf{LPS}(X_{-i})) \supseteq \mathbb{h}_i(\mathsf{LEU}(X_{-i})) \\ & \therefore \quad \mathbb{h}_i(\mathsf{LPS}(X_{-i})) = \mathbb{h}_i(\mathsf{LEU}(X_{-i})) \end{aligned}$$

By Corollary 4.1, the set $\mathbb{h}_i(\mathsf{LPS}(X_{-i})) = \mathbb{h}_i(\mathsf{LEU}(X_{-i}))$ is Borel. □

From the fact that $\mathbb{h}_a$ is well-defined, we can see that Ann's belief about Bob's states can always be written as a hierarchy. However, the converse is not always true. It is not necessarily true that Ann's hierarchy can be written as a belief about Bob's states. This is the content of Theorem 4.1.

Theorem 4.1. $\mathbb{h}_i(\mathsf{LPS}(X_{-i})) \neq H_i$.

Proof of Theorem 4.1. The strategy of the proof is to construct a hierarchy

$$h_i = (h_i^1, h_i^2, \dots)$$

such that, for all $n \geq 1$, h_i^n has length n.

We start by fixing some $y_0 \in X_{-i}$. For each $n \geq 0$, we choose[10] and fix some $y_{n+1} \in X_{-i}$ such that $\varpi_{-i}^n(y_{n+1}) = \varpi_{-i}^n(y_n)$ and $\varpi_{-i}^{n+1}(y_{n+1}) \neq \varpi_{-i}^{n+1}(y_n)$. Because y_n and y_{n+1} are coherent states of Player $-i$, the following can be concluded.

$$\begin{aligned} &\forall m \leq n \quad \varpi_{-i}^m(y_{n+1}) = \varpi_{-i}^m(y_n) \\ &\forall m > n \quad \varpi_{-i}^m(y_{n+1}) \neq \varpi_{-i}^m(y_n) \end{aligned}$$

The following statements therefore hold jointly for all $m, n \geq 0$.

$$\begin{aligned} &m \leq n \implies \forall \hat{m} \leq m \quad \varpi_{-i}^{\hat{m}}(y_m) = \varpi_{-i}^{\hat{m}}(y_n) \\ &m > n \implies \forall \hat{n} > n \quad \varpi_{-i}^{\hat{n}}(y_m) \neq \varpi_{-i}^{\hat{n}}(y_n) \end{aligned}$$

We now define a sequence $(\mu_n)_{n \geq 0}$ of probability measures in $\mathsf{P}(X_{-i}) \subseteq \mathsf{LPS}(X_{-i})$ such that $\mu_n(y_n) = 1$ for all $n \geq 0$. We use this sequence to define the hierarchy $(h_i^1, h_i^2, \dots) \in H_i$ as follows.

$$\forall n \geq 1 \quad h_i^n \equiv (\mathbb{h}_i^n(\mu_{n-1}), \mathbb{h}_i^n(\mu_{n-2}), \dots, \mathbb{h}_i^n(\mu_0))$$

[10] We take advantage of the fact that, for each $x_j \in X_j$ and $n \geq 0$, we can alway find some $x_j' \in X_j$ such that $\varpi_j^n(x_j) = \varpi_j^n(x_j')$ and $\varpi_j^{n+1}(x_j) \neq \varpi_j^{n+1}(x_j')$.

The measures that comprise h_i^n are pairwise mutually singular. This is easily verified from the following set of equalities, which arise from the definition of $(\mu_n)_{n\geq 0}$.

$$\forall m < n \quad \operatorname{supp} \mathbb{bh}_i^n(\mu_m) = \varpi_{-i}^{n-1}(y_m)$$

It follows that h_i^n is an LCPS and has minimal length. Furthermore, h_i is a coherent hierarchy because the following holds for all $n \geq 1$.

$$\begin{aligned}
\hat{\pi}^n(h_i^{n+1}) &\cong \overbrace{(\operatorname{marg}_{X_{-i}^n} \mathbb{bh}_i^{n+1}(\mu_n), \operatorname{marg}_{X_{-i}^n} \mathbb{bh}_i^{n+1}(\mu_{n-1}), \ldots, \operatorname{marg}_{X_{-i}^n} \mathbb{bh}_i^{n+1}(\mu_0))}^{n+1 \text{ measures}} \\
&= (\overbrace{\mathbb{bh}_i^n(\mu_n), \mathbb{bh}_i^n(\mu_{n-1})}^{\text{equal}}, \ldots, \mathbb{bh}_i^n(\mu_0)) \\
&\cong \overbrace{(\mathbb{bh}_i^n(\mu_{n-1}), \ldots, \mathbb{bh}_i^n(\mu_0))}^{n \text{ measures}} = h_i^n
\end{aligned}$$

We know that $\mathbb{bh}_i^n(\mu_n) = \mathbb{bh}_i^n(\mu_{n-1})$ because

$$\operatorname{supp} \mathbb{bh}_i^n(\mu_n) = \varpi_{-i}^{n-1}(y_n) = \varpi_{-i}^{n-1}(y_{n-1}) = \operatorname{supp} \mathbb{bh}_i^n(\mu_{n-1}).$$

It follows that $h_i \in H_i$ since $\hat{\pi}^n(h_i^{n+1}) = h_i^n$ for all $n \geq 1$.

Suppose by way of contradiction that there exists an LPS $\sigma \in \mathsf{LPS}(X_{-i})$ such that $\mathbb{bh}_i(\sigma) = h_i$. Its length must be some $N \geq 1$. It follows that $\mathbb{bh}_i^{N+1}(\sigma)$ is preference-equivalent to an LPS of length N—namely, $\operatorname{marg}_{X_i^{N+1}} \sigma$. However, h_i^{N+1}, which has length $N+1$, cannot be represented by a shorter LPS. Since $h_i^{N+1} = \mathbb{bh}_i^{N+1}(\sigma) \cong \operatorname{marg}_{X_i^{N+1}} \sigma$, this yields a contradiction. $\square$

5 TYPE STRUCTURES

5.1 FROM TYPES TO HIERARCHIES

Definition 5.1. An LPS-type structure is a tuple $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ such that the following holds for all $i \in I$:

1. T_i is a nonempty standard Borel space; and
2. $\beta_i \colon T_i \to \mathsf{LPS}(S_{-i} \times T_{-i})$ is a Borel map.

T_i is called Player i's type space and its elements are called her types. The map β_i is called Player i's type-belief map.

Definition 5.2. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. $\mathbf{T}$ is

1. called an LCPS-type structure if $\beta_i(T_i) \subseteq \mathsf{LCPS}(S_{-i} \times T_{-i})$ for all $i \in I$; and
2. called a P-type structure if $\beta_i(T_i) \subseteq \mathsf{P}(S_{-i} \times T_{-i})$ for all $i \in I$.

The familiar type structures of Mertens and Zamir (1985); Brandenburger and Dekel (1993); Tan and Werlang (1988) are P-type structures. The lexicographic type structures in BFK are LCPS-type structures.

It obvious that the higher-order preferences, and therefore the coherent hierarchy, implied by a given type can be recovered using the type-belief maps. Before we do so, it is useful to first extend the notion of pushforward measures to LPS's.

Definition 5.3. Let X, Y be standard Borel spaces and $f\colon X \to Y$ a Borel map.

1. Given a $\mu \in \mathsf{P}(X)$, the pushforward[11] belief $f\mu \in \mathsf{P}(Y)$ is defined as follows.

   (15) $\forall E \in \mathscr{B}(Y) \quad f\mu(E) \equiv \mu(f^{-1}(E))$

2. Given a $\sigma = (\mu_1, \ldots, \mu_n) \in \mathsf{LPS}(X)$, the pushforward belief $f\sigma \in \mathsf{LPS}(Y)$ is defined as follows.

   (16) $f\sigma \equiv (f\mu_1, \ldots, f\mu_n)$

3. Given a $\sigma = (\mu_1, \ldots, \mu_n) \in \mathsf{LPS}(X)$, the pushforward preference $f * \sigma \in \mathsf{LEU}(Y)$ is defined as follows.

   (17) $f * \sigma \equiv \varsigma_X(f\sigma)$

Definition 5.4. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. Define the maps $\mathtt{tx}_{i[\mathbf{T}]}, \mathtt{tx}^0_{i[\mathbf{T}]}, \mathtt{tx}^1_{i[\mathbf{T}]}, \mathtt{tx}^2_{i[\mathbf{T}]}, \ldots$ as follows.

(18) $\mathtt{tx}^0_{i[\mathbf{T}]}\colon X^0_i \times T_i \to X^0_i \quad \equiv \quad (x^0_i, t_i) \mapsto (x^0_i)$

(19) $\mathtt{tx}^{n+1}_{i[\mathbf{T}]}\colon X^0_i \times T_i \to X^{n+1}_i \quad \equiv \quad (x^0_i, t_i) \mapsto (\mathtt{tx}^n_{i[\mathbf{T}]}(x^0_i, t_i), \mathtt{tx}^n_{-i[\mathbf{T}]} * \beta_i(t_i))$

(20) $\mathtt{tx}_{i[\mathbf{T}]}\colon X^0_i \times T_i \to X_i \quad \equiv \quad (x^0_i, t_i) \mapsto (x^0_i, (\mathtt{tx}^n_{-i[\mathbf{T}]} * \beta_i(t_i))_{n \geq 0})$

Definition 5.5. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. Define the map $\mathtt{th}_{i[\mathbf{T}]}$—called Player i's type-hierarchy map associated with $\mathbf{T}$—as follows.

(21) $\mathtt{th}_{i[\mathbf{T}]}\colon T_i \to H_i \quad \equiv \quad t_i \mapsto (\mathtt{tx}^n_{-i[\mathbf{T}]} * \beta_i(t_i))_{n \geq 0}$

The $(n+1)^{\text{th}}$-order preference induced by type $t_i \in T_i$ is given by the pushforward preference $\mathtt{tx}^n_{-i[\mathbf{T}]} * \beta_i(t_i)$. The preference hierarchy induced by t_i is given by $\mathtt{th}_{i[\mathbf{T}]}(t_i)$.

Lemma 5.1. *Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. Every map in the following set is Borel.*

$$\{\mathtt{th}_{i[\mathbf{T}]}\colon i \in I\} \cup \{\mathtt{tx}_{i[\mathbf{T}]}\colon i \in I\} \cup \{\mathtt{tx}^n_{i[\mathbf{T}]}\colon i \in I \wedge n \geq 0\}$$

[11] Etymology: A probability measure μ on X being "pushed forward" to a belief on Y by the map $f\colon X \to Y$.

Harsanyi (1967)'s insight that hierarchies, which are cumbersome objects, can be represented by types, which are comparatively simple objects, inspired numerous papers establishing the foundations of various type structures. A recurring theme in these investigations is whether there is a type structure that can describe "all higher-order beliefs"—a notion that varies according to the context in which the question is asked. Following several recent papers, we use *terminality* in this paper as an umbrella term to describe such properties of type structures.[12] The following common variants of the terminality question are considered.

Definition 5.6. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. $\mathbf{T}$ is said to be

1. **strongly terminal** if it describes every preference hierarchy, i.e.,

$$\forall i \in I \quad \mathtt{th}_{i[\mathbf{T}]}(T_i) = H_i$$

2. **weakly terminal** in a given family $\mathcal{F}$ of LPS-type structures if it describes every preference hierarchy that can be described by type structures in $\mathcal{F}$, i.e.,

$$\forall i \in I \quad \forall \hat{\mathbf{T}} \in \mathcal{F} \quad \mathtt{th}_{i[\hat{\mathbf{T}}]}(T_i) \subseteq \mathtt{th}_{i[\mathbf{T}]}(T_i)$$

3. **finitely terminal** if it describes all finite-order preferences, i.e.,

$$\forall i \in I \quad \forall n \geq 0 \quad \{\mathtt{tx}^n_{-i[\mathbf{T}]} * \beta_i(t_i) \colon t_i \in T_i\} = \mathsf{LEU}(X^n_i)$$

Theorem 5.1. *No strongly terminal* LPS*-type structure exists.*

Proof of Theorem 5.1. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. For each $t_i \in T_i$,

$$\mathtt{bh}_i(\mathtt{tx}_{-i[\mathbf{T}]}\beta_i(t_i)) = \mathtt{th}_i(t_i),$$

where $\mathtt{tx}_{-i[\mathbf{T}]}\beta_i(t_i)$ and $\mathtt{th}_i(t_i)$ respectively correspond to the belief about X_{-i} and preference hierarchy induced by t_i. The hierarchies that can be described by $\mathbf{T}$ can also be described as beliefs about the other players' states—i.e., beliefs about X_{-i}. Because Theorem 4.1 states that some hierarchies cannot be described in that form, it follows that some hierarchies cannot be described by $\mathbf{T}$. □

5.2 CANONICAL TYPE STRUCTURES

We construct two canonical type structures by first defining the following sequences of sets of coherent hierarchies.[13]

$$T^{\mathsf{LEU}}_{i[0]} \equiv H_i \qquad\qquad T^{\mathsf{LCPS}}_{i[0]} \equiv H_i$$
$$T^{\mathsf{LEU}}_{i[n+1]} \equiv \mathtt{bh}_i(\mathsf{LEU}(X^0_{-i} \times T^{\mathsf{LEU}}_{-i[n]})) \qquad\qquad T^{\mathsf{LCPS}}_{i[n+1]} \equiv \mathtt{bh}_i(\mathsf{LCPS}(X^0_{-i} \times T^{\mathsf{LEU}}_{-i[n]}))$$

[12]The word *universality* has also been used frequently.

[13]For each Borel subset $E \neq \varnothing$ of X_{-i}, we slightly abuse notation by letting $\mathsf{LPS}(E)$ denote the set $\{\sigma \in \mathsf{LPS}(X) \colon \sigma(E) = \vec{1}\}$. The sets $\mathsf{P}(E)$, $\mathsf{LEU}(E)$, $\mathsf{LCPS}(E)$ are defined mutatis mutandis.

Definition 5.7. The canonical LPS type structure is the tuple $\mathbf{T}^{\mathsf{LEU}} \equiv \langle T_i^{\mathsf{LEU}}, \beta_i^{\mathsf{LEU}} \rangle_{i \in I}$ such that $T_i^{\mathsf{LEU}} \equiv \bigcap_{n \geq 0} T_{i[n]}^{\mathsf{LEU}}$ and $\beta_i^{\mathsf{LEU}} \colon T_i^{\mathsf{LEU}} \to \mathsf{LEU}(S_{-i} \times T_{-i}^{\mathsf{LEU}})$ is defined as follows.

$$\forall t_i \in T_i^{\mathsf{LEU}} \quad \beta_i^{\mathsf{LEU}}(t_i) \equiv \mathtt{bh}_i^{-1}(t_i)$$

Definition 5.8. The canonical CPS type structure is the tuple $\mathbf{T}^{\mathsf{LCPS}} \equiv \langle T_i^{\mathsf{LCPS}}, \beta_i^{\mathsf{LCPS}} \rangle_{i \in I}$ such that $T_i^{\mathsf{LCPS}} \equiv \bigcap_{n \geq 0} T_{i[n]}^{\mathsf{LCPS}}$ and $\beta_i^{\mathsf{LCPS}} \colon T_i^{\mathsf{LCPS}} \to \mathsf{LCPS}(S_{-i} \times T_{-i}^{\mathsf{LCPS}})$ is defined as follows.

$$\forall t_i \in T_i^{\mathsf{LCPS}} \quad \beta_i^{\mathsf{LCPS}}(t_i) \equiv \mathtt{bh}_i^{-1}(t_i)$$

Lemma 5.2. $\mathbf{T}^{\mathsf{LEU}}$ *and* $\mathbf{T}^{\mathsf{LCPS}}$ *are* LPS*-type structures and their type-belief maps are Borel isomorphisms.*

Proof of Lemma 5.2. By Corollary 4.1, images of non-repetitive Borel sets under $\mathtt{bh}_i$ are Borel sets. For any nonempty standard Borel space X, $\mathsf{LCPS}(X)$ and $\mathsf{LEU}(X)$ are non-repetitive subsets of $\mathsf{LPS}(X)$. For all $n \geq 0$, $T_{i[n+1]}^{\mathsf{LEU}}$ and $T_{i[n+1]}^{\mathsf{LCPS}}$ are Borel because they are images of non-repetitive Borel sets under $\mathtt{bh}_i$. It follows that T_i^{LEU} and T_i^{LCPS} are Borel sets in H_i.

For any nonempty standard Borel space X, $\mathsf{P}(X) \subseteq \mathsf{LCPS}(X) \subseteq \mathsf{LPS}(X)$. It follows that both $\mathbf{T}^{\mathsf{LEU}}$ and $\mathbf{T}^{\mathsf{LCPS}}$ contain the P-type structure of Brandenburger and Dekel (1993), which is nonempty.

LEU preferences satisfy the limit closure property—i.e., if a sequence of Borel sets are 1-believed in the sense that their complements are Savage-null, then the intersection of those sets is also 1-believed in the same sense.[14] Therefore, the equalities below hold.

$$T_i^{\mathsf{LEU}} = \mathtt{bh}_i(\mathsf{LEU}(X_{-i}^0 \times T_{-i}^{\mathsf{LEU}})) \qquad T_i^{\mathsf{LCPS}} = \mathtt{bh}_i(\mathsf{LCPS}(X_{-i}^0 \times T_{-i}^{\mathsf{LCPS}}))$$

The sets $\mathsf{LEU}(X_{-i}^0 \times T_{-i}^{\mathsf{LEU}})$ and $\mathsf{LCPS}(X_{-i}^1 \times T_{-i}^{\mathsf{LCPS}})$ are non-repetitive and Borel. It follows from Lemma 4.4 that the restriction of $\mathtt{bh}_i$ to those sets must be one-to-one and Borel. The type-belief maps β_i^{LEU} and β_i^{LCPS}, being the functional inverses of those restrictions, are Borel isomorphisms since they are one-to-one and surjective Borel mappings. $\square$

Theorem 5.2. $\mathbf{T}^{\mathsf{LEU}}$ *is weakly terminal in the class of* LPS*-type structures.*

Proof of Theorem 5.2. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_i$ be an LPS-type structure. We want to show that

$$\forall t_i \in T_i \quad \mathtt{th}_{i[\mathbf{T}]}(t_i) \in T_i^{\mathsf{LEU}}.$$

[14] In the literature, 1-belief, or simply belief, of an event E corresponds to belief with probability 1. Extending this notion to LPS's, an event E is 1-believed under LPS $(\mu_1, \ldots, \mu_n)$ if $\mu_1(E) = \cdots = \mu_n(E) = 1$.

For each $t_i \in T_i$, the hierarchy $\mathtt{th}_{i[\mathbf{T}]}(t_i)$ can also be represented by the belief

$$\varsigma_{X_{-i}}(\mathtt{tx}_{-i[\mathbf{T}]}\beta_i(t_i)) \in \mathsf{LEU}(X_{-i}) = \mathsf{LEU}(X^0_{-i} \times H_{-i}).$$
$$\therefore \mathtt{th}_{i[\mathbf{T}]}(t_i) \in T^{\mathsf{LEU}}_{i[1]} = \mathtt{bh}_i(\mathsf{LEU}(X^0_{-i} \times H_{-i}))$$

Furthermore, $\mathtt{th}_{i[\mathbf{T}]}(t_i)$ belongs to $T^{\mathsf{LEU}}_{i[2]}$—i.e., the set of hierarchies that can be represented as beliefs that 1-believe $X^0_{-i} \times T^{\mathsf{LEU}}_{-i[1]}$—because $\mathtt{th}_{-i[\mathbf{T}]}(t_{-i}) \in T^{\mathsf{LEU}}_{-i[1]}$ for all $t_{-i} \in T_{-i}$. By applying this line of argument inductively, it is shown that

$$\mathtt{th}_{i[\mathbf{T}]}(t_i) \in \bigcap_{n \geq 0} T^{\mathsf{LEU}}_{i[n]} = T^{\mathsf{LEU}}_i$$

In the literature, the map $\mathtt{th}_{i[\mathbf{T}]}$ is often called Player i's unique type morphism from $\mathbf{T}$ to $\mathbf{T}^{\mathsf{LEU}}$. □

6 HIERARCHIES THAT ARE CONDITIONAL BELIEFS

6.1 REDUNDANT TYPES

Consider an LPS-type structure $\langle T_i, \beta_i \rangle_{i \in I}$. Each $t_i \in T_i$ represents a belief $\beta_i(t_i)$ about $X^0_{-i} \times T_{-i}$. As briefly discussed in Section 3.2, the statement that $\beta_i(t_i)$ is an LCPS loses its meaning when the space of uncertainty—i.e., $X^0_{-i} \times T_{-i}$—contains elements that are redundant with respect to the uncertainty that we wish to model. In the case of type structures, what we wish to model is uncertainty about preference hierarchies. As such, the set T_{-i} is descriptively useful only to the extent that it describes hierarchies in H_{-i}. We can define when a type structure is redundant in that regard.

Definition 6.1. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be an LPS-type structure. A type $t_i \in T_i$ is redundant if there is some $t_i \neq \hat{t}_i$ such that $\mathtt{th}_{i[\mathbf{T}]}(t_i) = \mathtt{th}_{i[\mathbf{T}]}(\hat{t}_i)$. We say that $\mathbf{T}$ is redundant if there is some $i \in I$ such that redundant types exist in T_i.

Any preference hierarchy that can be described by an LPS-type structure can also be described by a possibly redundant LCPS-type structure. This is an implication of the following result.

Theorem 6.1. *There exists an* LCPS*-type structure that is weakly terminal in the class of* LPS*-type structures.*

Proof of Theorem 6.1. We wish to construct an LCPS-type structure $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ that generates the same set of hierarchies as $\mathbf{T}^{\mathsf{LEU}}$ does. Let $T_i = \mathbb{N} \times T^{\mathsf{LEU}}_i$. We will define the map $\beta_i \colon T_i \to \mathsf{LCPS}(X^0_{-i} \times T_{-i})$ in a piecewise fashion.

First, for each $m \in \mathbb{N}$, the space $\{m\} \times T_i$ admits the following countable partition into Borel sets.

$$\Pi(m) = \{P_{(m,n)} \colon n \geq 1\},$$
$$\text{where} \quad P_{(m,n)} = \{(m, t_i) \in T_i \colon \beta_i^{\mathsf{LEU}}(t_i) \in \mathsf{LPS}_n(X_{-i}^0 \times T_{-i})\}.$$

For any (m, n), the function β_i can be defined on the sub-domain $P_{(m,n)} \in \Pi(m)$ as follows.

$$\beta_i(m, t_i) \equiv (f_{-i[1]}\mu_1, \ldots, f_{-i[n]}\mu_n) \in \mathsf{LCPS}(X_{-i}^0 \times T_{-i}),$$
$$\text{where} \quad \beta_i^{\mathsf{LEU}}(t_i) = (\mu_1, \ldots, \mu_n) \quad \text{and}$$
$$f_{-i[k]} \colon X_{-i}^0 \times T_{-i}^{\mathsf{LEU}} \to X_{-i}^0 \times \{k\} \times T_{-i}^{\mathsf{LEU}} \equiv (x_{-i}^0, t_{-i}) \mapsto (x_{-i}^0, k, t_{-i})$$

The map β_i is clearly Borel on $P_{(m,n)}$ for all (m, n). It follows that β_i is Borel on $\bigcup \Pi(m) = \bigcup_{n \geq 1} P_{(m,n)}$ because $\Pi(m)$ is countable. Therefore, β_i is Borel on each member of the countable partition $\{\bigcup \Pi(m) \colon m \in \mathbb{N}\}$ of T_i. It follows that β_i is a Borel map.

Finally, $\mathbf{T}$ generates the same hierarchies as $\mathbf{T}^{\mathsf{LEU}}$ because the following equality holds for all $(m, t_i) \in T_i$ by construction.

$$\operatorname{marg}_{X_{-i}^0 \times T_{-i}^{\mathsf{LEU}}} \beta_i(m, t_i) = \beta_i^{\mathsf{LEU}}(t_i)$$

Since $\mathbf{T}^{\mathsf{LEU}}$ is weakly terminal, so is $\mathbf{T}$. $\square$

In contrast, not every preference hierarchy that can be described by an LPS-type structure can also be described by a *nonredundant* LCPS-type structure. Furthermore, the canonical LCPS-type structure describes precisely the set of hierarchies that are described by nonredundant LCPS-type structures.

Theorem 6.2. *The canonical* LCPS*-type structure* $\mathbf{T}^{\mathsf{LCPS}}$ *is weakly terminal in the class of nonredundant* LCPS*-type structures.*

Proof of Theorem 6.2. Let $\mathbf{T} = \langle T_i, \beta_i \rangle_{i \in I}$ be a nonredundant LCPS-type structure. We want to show that $\mathtt{th}_{i[\mathbf{T}]}(t_i) \in T_i^{\mathsf{LCPS}}$ for all $t_i \in T_i$.

For each $t_i \in T_i$, the hierarchy $\mathtt{th}_{i[\mathbf{T}]}(t_i)$ can also be represented by the belief

$$\mathtt{tx}_{-i[\mathbf{T}]}\beta_i(t_i) \in \mathsf{LCPS}(X_{-i}) = \mathsf{LCPS}(X_{-i}^0 \times H_{-i}).$$

The pushforward $\mathtt{tx}_{-i[\mathbf{T}]}\beta_i(t_i)$ is an LCPS because $\beta_i(t_i)$ is an LCPS and $\mathtt{tx}_{-i[\mathbf{T}]}$ is one-to-one when $\mathbf{T}$ is nonredundant. It follows that $\mathtt{th}_{i[\mathbf{T}]}(t_i) \in T_{i[1]}^{\mathsf{LCPS}} = \mathtt{h}_i(\mathsf{LCPS}(X_{-i}^0 \times H_{-i}))$. Furthermore, $\mathtt{th}_{i[\mathbf{T}]}(t_i)$ belongs to $T_{i[2]}^{\mathsf{LCPS}}$—i.e., the set of hierarchies that can be represented as beliefs that 1-believe $X_{-i}^0 \times T_{-i[1]}^{\mathsf{LCPS}}$—because $\mathtt{th}_{-i[\mathbf{T}]}(t_{-i}) \in T_{-i[1]}^{\mathsf{LCPS}}$ for all $t_{-i} \in T_{-i}$. By applying this line of argument inductively, it is shown that

$$\mathtt{th}_{i[\mathbf{T}]}(t_i) \in \bigcap_{n \geq 0} T_{i[n]}^{\mathsf{LCPS}} = T_i^{\mathsf{LCPS}}.$$

$\square$

Theorem 6.3. *The canonical* LCPS*-type structure* $\mathbf{T}^{\mathsf{LCPS}}$ *is not weakly terminal in the class of* LCPS*-type structures.*[15]

Proof of Theorem 6.3. Due Theorem 5.2, we know that $T_i^{\mathsf{LCPS}} \subseteq T_i^{\mathsf{LEU}}$. Note that there exists some $\sigma \in \mathsf{LEU}(X_{-i}^0 \times T_{-i}^{\mathsf{LCPS}})$ such that σ is not preference-equivalent to any belief in $\mathsf{LCPS}(X_{-i}^0 \times T_{-i}^{\mathsf{LCPS}})$. By Lemma 4.4, we can then conclude that σ does not induce a hierarchy in T_i^{LCPS}. Such σ—being a belief about $X_{-i}^0 \times T_{-i}^{\mathsf{LEU}} \supseteq X_{-i}^0 \times T_{-i}^{\mathsf{LCPS}}$— induces the hierarchy $\mathtt{bh}_i(\sigma) \in T_i^{\mathsf{LEU}}$, which can be represented in some LCPS-type structure due to Theorem 6.1. $\square$

In light of the preceding results, it makes sense to restrict our attention to nonredundant LCPS-type structures when we are interested in LCPS beliefs about hierarchies.

6.2 NONREDUNDANT LCPS TYPES ARE ALMOST LPS TYPES

Although nonredundant LCPS-type structures describe a strict subset of the hierarchies that are described by LPS-type structures, the two nevertheless have equal descriptive power in the following important way.

Theorem 6.4. *The canonical* LCPS*-type structure* $\mathbf{T}^{\mathsf{LCPS}}$ *is finitely terminal.*

Proof of Theorem 6.4. We want to show that the following holds for all $n \geq 0$

$$\{\mathtt{tx}^n_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(t_i) \colon t_i \in T_i\} = \mathsf{LEU}(X^n_{-i})$$

Due the definition of $\mathbf{T}^{\mathsf{LCPS}}$, $\mathtt{tx}_{-i[\mathbf{T}^{\mathsf{LCPS}}]}$ is the identity mapping. Therefore, $\mathtt{tx}^n_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(\cdot) = \mathrm{margp}_{X^n_{-i}} \beta_i^{\mathsf{LCPS}}(\cdot)$ for all $n \geq 0$. Furthermore, $\mathbf{T}^{\mathsf{LCPS}}$ is a belief-complete LCPS-type structure in the sense that $\beta_i^{\mathsf{LCPS}}(T_i^{\mathsf{LCPS}}) = \mathsf{LCPS}(X^0_{-i} \times T_{-i}^{\mathsf{LCPS}})$. From these facts, we get the following for all $i \in I$.

$$\begin{aligned}
\varpi^0_{-i}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i}) &= \mathrm{proj}_{X^0_{-i}} X^0_{-i} \times T^{\mathsf{LCPS}}_{-i} = X^0_{-i} \\
\mathtt{tx}^0_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(t_i) &= \{\mathrm{margp}_{X^0_{-i}} \sigma \colon \sigma \in \mathsf{LCPS}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i})\} \\
&= \varsigma_{X^0_{-i}} \left(\{\mathrm{marg}_{X^0_{-i}} \sigma \colon \sigma \in \mathsf{LCPS}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i})\} \right)
\end{aligned}$$

For all nonempty standard Borel spaces X, Y, U such that $U \subseteq X \times Y$ and $\mathrm{proj}_X U = X$, $\{\mathrm{marg}_X \sigma \colon \sigma \in \mathsf{LCPS}(U)\} = \mathsf{LPS}(X)$. Therefore, we can simplify the expression above.

$$= \varsigma_{X^0_{-i}} \left(\mathsf{LPS}(X^0_{-i}) \right) = \mathsf{LEU}(X^0_{-i})$$

For the induction step, let the following hold for all $m < n$ and $i \in I$, where $n > 1$.

$$\begin{aligned}
\varpi^m_{-i}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i}) &= \mathrm{proj}_{X^m_{-i}} X^0_{-i} \times T^{\mathsf{LCPS}}_{-i} = X^m_{-i} \\
\{\mathtt{tx}^m_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(t_i) \colon t_i \in T_i\} &= \mathsf{LEU}(X^m_{-i})
\end{aligned}$$

[15] It is also therefore not weakly terminal in the class of LPS-type structures.

Recall Definition 5.5 and note that $\mathtt{tx}^{n-1}_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(t_i)$ is equal to the $\mathsf{LEU}(X^{n-1}_{-i})$-coordinate of the hierarchy t_i. The following equalities then follow from the induction hypothesis.

$$\begin{aligned} \operatorname{proj}_{X_i^{n-1}} X_i^0 \times T_i^{\mathsf{LCPS}} &= X_i^{n-1} \\ \operatorname{proj}_{\mathsf{LEU}(X^{n-1}_{-i})} X^0_{-i} \times T^{\mathsf{LCPS}}_{-i} &= \mathsf{LEU}(X^{n-1}_{-i}) \end{aligned}$$

Because $X^0_{-i} \times T^{\mathsf{LCPS}}_{-i} \subseteq X_{-i}$, the consistency requirement built into the definition of X_{-i} allows us to deduce the first equality below from the two above.[16]

$$\varpi^n_{-i}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i}) = \operatorname{proj}_{X^n_{-i}} X^0_{-i} \times T^{\mathsf{LCPS}}_{-i} = X^n_{-i} \subseteq X_i^{n-1} \times \mathsf{LEU}(X^{n-1}_{-i}) \tag{22}$$

$$\mathtt{tx}^{n}_{-i[\mathbf{T}^{\mathsf{LCPS}}]} * \beta_i^{\mathsf{LCPS}}(t_i) = \{\operatorname{margp}_{X^n_{-i}} \sigma \colon \sigma \in \mathsf{LCPS}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i})\} \tag{23}$$

$$= \varsigma_{X^n_{-i}} \Big(\{\operatorname{marg}_{X^n_{-i}} \sigma \colon \sigma \in \mathsf{LCPS}(X^0_{-i} \times T^{\mathsf{LCPS}}_{-i})\}\Big) \tag{24}$$

Mutatis mutandis, the arguments used in the analogous step of the base case imply the following simplification because (22) holds.

$$= \varsigma_{X^n_{-i}} \big(\mathsf{LPS}(X^n_{-i})\big) = \mathsf{LEU}(X^n_{-i}) \tag{25}$$

□

An important implication of Theorem 6.4 is that, for the purposes of analyzing epistemic conditions involving finite-order beliefs, there is effectively no difference between LPS-type structures and nonredundant LCPS-type structures. To put it another way, every coherent preference hierarchy can be approximated by a sequence of types in nonredundant LCPS-type structures.[17]

7 COMMON ASSUMPTION OF RATIONALITY

7.1 ADMISSIBILITY

Let $G = \langle S_a, S_b, u_a, u_b \rangle$ be a finite game of complete information. The symbols S_i and u_i respectively denote Player i's strategy set and utility function.

Definition 7.1. A strategy $s_i \in S_i$ is admissible against $S'_{-i} \subseteq S_{-i}$ if it is optimal with respect to some belief σ such that $\operatorname{supp}\sigma = S'_{-i}$—i.e.,

$$\exists \sigma \in \mathsf{LPS}(S_{-i}) \quad \big[\operatorname{supp}\sigma = S'_{-i} \wedge \forall s'_i \in S_i \quad \mathbf{E}_\sigma u_i(s_i, \cdot) \geq^{\mathsf{L}} \mathbf{E}_\sigma u_i(s'_i, \cdot)\big] \tag{26}$$

[16] In other words, $\operatorname{proj}_{X_i^n} X_i^0 \times T_i^{\mathsf{LCPS}} = X_i^n$ whenever $\operatorname{proj}_{\mathsf{LEU}(X^{n-1}_{-i})} T_i^{\mathsf{LCPS}} = \mathsf{LEU}(X^{n-1}_{-i})$ because T_i^{LEU} is a set of *coherent* hierarchies.

[17] This is true provided that we endow X_i with a product topology as is usual in the literature.

Definition 7.2. Let the sets $S_i^0, S_i^1, S_i^2, \ldots, S_i^\infty$ be defined as follows.

$$S_i^0 \equiv S_i$$
$$S_i^{m+1} \equiv \{s_i \in S_i^m \colon s_i \text{ is admissible against } S_{-i}^m\} \qquad S_i^\infty \equiv \bigcap_{m \geq 0} S_i^0$$

For all $m \geq 1$, S_i^m is called Player i's m-admissible strategy set and S_i^∞ is called her iteratively admissible (IA) strategy set.

7.2 EPISTEMIC CONDITION FOR ITERATED ADMISSIBILITY

Let $\langle S_a, S_b, \pi_a, \pi_b \rangle$ be a finite game as before. The fundamental uncertainty of each player concerns the strategy played by her opponent—i.e., $X_i^0 = S_i$. The preceding sections have adopted a *topology-free* approach to higher-order beliefs to the extent that X_i^n is discussed as a Borel space without reference to the topology that generates its Borel sets. The choice of topology on X_i^n does not alter the results and definitions of the previous sections as long as it generates a standard Borel σ-algebra.

In this section, we assume that the finite set $X_i^0 = S_i$ is endowed with the discrete topology. Furthermore, for all $n \geq 1$, X_i^n is endowed with a separable and metrizable topology $\mathscr{T}(X_i^n)$ that generates the standard Borel algebra $\mathscr{B}(X_i^n)$ and is consistent with $\mathscr{T}(X_i^{n-1})$.[18] The notions of assumption and full-support beliefs, which are central to this section, cannot be defined without reference to topologies. Nevertheless, the topology-free approach of the preceding sections is partially maintained—at least in spirit—to the extent that the results of this section do not depend on *other* specific details of these topologies.

Definition 7.3. Let Ω be a topological space. An LPS $\sigma = (\mu_1, \ldots, \mu_n) \in \mathsf{LPS}(\Omega)$ is a full-support LPS if, for every open set $U \in \mathscr{T}(\Omega)$, there exists some j such that $\mu_j(U) > 0$.

Definition 7.4. Let Ω be a separable and metrizable topological space. Let $\mathsf{LPS}^+(\Omega)$ denote the set of all full-support LPS's in $\mathsf{LPS}(\Omega)$. We also define the following sets of full-support LPS's for all $m \geq 1$. These sets are Borel in $\mathsf{LPS}(\Omega)$.

$$\mathsf{LPS}_m^+(\Omega) \equiv \mathsf{LPS}_m(\Omega) \cap \mathsf{LPS}^+(\Omega)$$
$$\mathsf{LCPS}^+(\Omega) \equiv \mathsf{LCPS}(\Omega) \cap \mathsf{LPS}^+(\Omega) \qquad \mathsf{LEU}^+(\Omega) \equiv \mathsf{LEU}(\Omega) \cap \mathsf{LPS}^+(\Omega)$$
$$\mathsf{LCPS}_m^+(\Omega) \equiv \mathsf{LCPS}(\Omega) \cap \mathsf{LPS}_m^+(\Omega) \qquad \mathsf{LEU}_m^+(\Omega) \equiv \mathsf{LEU}(\Omega) \cap \mathsf{LPS}_m^+(\Omega)$$

Definition 7.5. The sequence $h_i = (h_i^1, h_i^2, \ldots) \in H_i$ is a full-support hierarchy if each finite-order belief in the sequence has full-support—i.e., $h_i^{n+1} \in \mathsf{LPS}^+(X_{-i}^n)$ for all $n \geq 0$. We let H_i^+ denote the set of Player i's full-support hierarchies.

[18] i.e., for any open $U \in \mathscr{T}(X_i^{n-1})$, the set $\{(x_i^{n-1}, h_i^n) \in X_i^n \colon x_i^{n-1} \in U\}$ is open in X_i^n. This guarantees that full-support beliefs on X_i^n will have full-support marginals on X_i^{n-1}.

The following definition of assumption is from Lee (2013). It extends BFK's definition of assumption for LCPS's to LPS's.

Definition 7.6. Let Ω be a separable and metrizable topological space, $\sigma \in \mathcal{L}(\Omega)$ an LPS, and $E \subseteq \Omega$ a nonempty Borel set. We say that E is assumed under σ if there exists some LPS $\rho = (\nu_1, \ldots, \nu_n) \cong \sigma$ and $j \geq 1$ such that

1. for all $i \leq j$, $\nu_i(E) = 1$; and
2. for all $i > j$, $\nu_i(E) = 0$; and
3. for all $U \in \mathscr{T}(\Omega)$ such that $U \cap E \neq \varnothing$, $\nu_k(U_E) > 0$ for some k.

Definition 7.7. The player state $(s_i, h_i) \in S_i \times H_i$ is said to be rational if $h_i \in H_i^+$ and s_i is optimal—i.e., maximizes LEU—w.r.t. to the first-order belief $\varrho_i^1 \circ \varpi_i^1(s_i, h_i)$. Let R_i^1 denote the set of all rational player states in $S_i \times H_i$.

Definition 7.8. Let $m \geq 1$. A player state satisfies rationality and m^{th}-order assumption of rationality (RmAR) if it belongs to the following set.

$$R_i^{m+1} \equiv R_i^m \cap \{(s_i, (h_i^n)_n) \in S_i \times H_i \colon \varpi_{-i}^m(R_{-i}^m) \text{ is assumed under } h_i^{m+1}\}$$

Definition 7.9. A player state satisfies rationality and common assumption of rationality (RCAR) if it belongs to the following set.

$$R_i^\infty \equiv \bigcap_{m \geq 0} R_i^m$$

RmAR and RCAR are analogs of the identically named epistemic conditions in BFK. The key difference is that our definition is stated in the space of player states rather than in type structures.

Theorem 7.1. *For all $m \geq 1$, $\operatorname{proj}_{S_i} R_i^m = S_i^m$.*

Proof of Theorem 7.1. See Appendix B. □

Theorem 7.2. *$R_i^\infty \neq \varnothing$ and $\operatorname{proj}_{S_i} R_i^\infty = S_i^\infty$.*

Proof of Theorem 7.2. See Appendix B. Here, we offer a sketch of the main idea behind the proof. □

Because S_i is a finite set, there exists some M such that $S_a^\infty = S_a^m = S_a^M$ and $S_b^\infty = S_b^m = S_b^M$ for all $m \geq M$. A strategy belongs to S_i^∞ if and only if it maximizes LEU with respect to some LPS in the following set.

$$\{(\mu_1, \ldots, \mu_{M+1}) \in \mathsf{LPS}(S_{-i}) \colon \operatorname{supp} \mu_k = S_{-i}^{M+1-k}\}$$

The proof shows that each LPS belonging to this set represents the first-order preferences of some player state that satisfies RCAR. The player state we construct has a hierarchy such that, for all $m \geq M$, the m^{th}-order preferences are represented by a minimal-length LPS that is at least $(m+1)$-long. Note that such a hierarchy cannot be represented in an LPS-type structure because it cannot be summarized by a single LPS, which must necessarily have finite-length.

BFK show that their version of RCAR cannot be satisfied in any complete *and* continuous LCPS-type structure. The basic idea is that a type in such type structures that m^{th}-order assumes of rationality must be mapped to an LCPS of length greater than or equal to $(m+1)$. Because every LCPS has finite length, no single type in such epistemic models can m^{th}-order assume rationality for *all* m. Our idea of using hierarchies that cannot be types is inspired by this intuition from BFK.

Let $m \geq M$ and let $h_i^{m+1} = (\mu_1^{M+1}, \ldots, \mu_{M+1}^{M+1}) \in \mathsf{LEU}^+(X_{-i}^m)$ be an LPS that satisfies the following properties.

$$\begin{aligned} &\mu_1^{m+1} \in \mathsf{P}^+(\varpi_{-i}^m(R_{-i}^m)) \\ \forall k > 1 \quad &\mu_k^{m+1} \in \mathsf{P}^+(\varpi_{-i}^M(R_{-i}^{m+1-k} \setminus R_{-i}^{m+2-k})) \end{aligned}$$

Furthermore, suppose that, for any such h_i^{m+1}, we can find some $h_i^{m+2} \in \mathsf{LEU}^+(X_{-i}^{m+1})$ such that $h_i^{m+2} = (\mu_1^{m+2}, \ldots, \mu_{m+2}^{m+2})$ and

$$\begin{aligned} \operatorname{margp}_{X_{-i}^m} h_i^{m+2} &= h_i^{m+1} \\ &\mu_1^{m+2} \in \mathsf{P}^+(\varpi_{-i}^{m+1}(R_{-i}^{m+1})) \\ \forall k > 1 \quad &\mu_k^{m+2} \in \mathsf{P}^+(\varpi_{-i}^{m+1}(R_{-i}^{m+2-k} \setminus R_{-i}^{m+3-k})). \end{aligned}$$

Let $(h_i^{M+k})_{k\geq 1}$ be a sequence that is inductively constructed in this way. If we let $h_i^k = \operatorname{margp}_{X_{-i}^k} h_i^{M+1}$ for all $k \leq M$, then the sequence $(h_i^n)_{n\geq 1}$ is a full-support hierarchy. It follows that there is some s_i such that $x_i \equiv (s_i, (h_i^n)_{n\geq 1}) \in R_i^1$. It can also be verified from definitions that $x_i \in R_i^{m+1}$ for all $m \geq M$, which means that $x_i \in R_i^\infty$. The main difficulty lies in showing the inductive step.

8 DISCUSSION

BFK formulated the notion of LCPS-type structures and applied it toward epistemic analysis that had evaded iterated admissibility until that point. Like P-type structures, LCPS-type structures were intended to capture beliefs hierarchies. The existence of belief-complete LCPS-type structures is demonstrated in BFK, but determining the precise content of this class of epistemic models was left as an open question.

Furthermore, BFK's RCAR[19] can be satisfied in some belief-complete LCPS-type structures but not in others. Keisler and Lee (2012) showed that BFK's epistemic

[19] Note that the RCAR defined in Section 7 is analogous to, but not equivalent to, BFK's RCAR.

analysis of iterated admissibility depends on type-structure-specific attributes unrelated to belief hierarchies, which demonstrated the desirability of a canonical model in which such attributes are determined by explicit construction of belief hierarchies.

These questions turned out to be nontrivial ones from both conceptual and technical viewpoints. Just as probability measures are EU preference representations that take the form of beliefs, LPS's are LEU preference representations that take the form of beliefs. In the case of EU preferences, the spaces of preference hierarchies and belief hierarchies are isomorphic. One might expect the same to be the case for LEU preferences, but this is not the case.

Epstein and Wang (1996) provide a method for generating preference hierarchies from a broad class of well-behaved preferences that includes EU preferences but not LEU preferences. A key motivation for preference-based epistemic analysis comes from the difficulty of defining higher-order beliefs when beliefs do not take the convenient form of probability measures. However, in the case of LEU preferences, defining higher-order LPS beliefs is straightforward but leads to the kind of conceptually perverse behavior discussed in Section 3—e.g., preference-redundancy and 4^{th}-order beliefs that are overly informative about 99^{th}-order beliefs. Therefore, the preference hierarchy approach in our setting is driven by conceptual issues.

Our results show that, as models of preference hierarchies, LCPS-type structures and LPS-type structures have equal descriptive power unless we rule out preference-redundant types. Given that BFK's setting does not rule out preference-redundant types, this suggests that there is no compelling reason to insist that the canonical model for our epistemic analysis be a LCPS-type structure instead of an LPS-type structure.

As a candidate for the canonical model that was sought earlier, we construct a universal LPS-type structure. However, our answer is a partial one for two reasons. First, we show that there are some preference hierarchies that cannot be described by any LPS-type structure. Note that preference hierarchies are the primitive objects of interest and type structures are inventions of game theorists designed to make the analysis of the former more tractable. As we see in the proof of Theorem 7.2, these missing hierarchies can be of interest to epistemic analysis.

Second, the canonical type spaces we construct are Borel subsets of the preference hierarchy spaces. It is not obvious how the hierarchy spaces should be topologized because we rely on a result from Lee (2013), which shows the existence of a well-behaved and nonredundant LEU preference space but says little about the precise shape of this space. However, we believe that this is a somewhat superficial view and that the primary source of difficulty arises from a more fundamental problem. BFK topologizes the space $\mathsf{LPS}(\Omega)$ as follows when Ω is a Polish space.

1. First, let $\mathsf{P}(\Omega)$ be endowed with the topology of weak convergence.

2. Second, let $\mathsf{LPS}_m(\Omega) = \prod_{k=1}^{m} \mathsf{P}(\Omega)$ be endowed with the product topology.

3. Finally, let $\mathsf{LPS}(\Omega) = \bigcup_{m \geq 1} \mathsf{LPS}_m(\Omega)$ be a topological union.

At first glance, this topology is a natural and straightforward extension of the usual topology on $\mathsf{P}(\Omega)$. However, the convergence of LPS beliefs under this topology exhibits strange behavior.

For example, let μ and ν be uniform measures on the intervals $[0,1]$ and $[1,2]$, respectively. The sequence $(\sigma_n)_{n\geq 1} = \big((1-\frac{1}{n})\mu + \frac{1}{n}\nu\big)_{n\geq 1}$ of LPS's converges to $\mu \in \mathsf{LPS}_1(\mathbb{R})$ in BFK's topology. An intuitively appealing argument can be made that this the sequence should converge to an LPS that places an infinitesimal, but nonzero, weight on the theory ν—namely, the LPS $\sigma = (\mu,\nu) \in \mathsf{LPS}_2(\mathbb{R})$. In fact, Blume *et al.* (1991b, See Propositions 1 and 2) show that $\succsim^\sigma$ can be described in terms of the sequence $(\sigma_n)_{n\geq 1}$. That said, defining a well-behaved topology that captures such notions of convergence is a nontrivial problem that may require techniques from nonstandard analysis.

In the absence of a canonical topology on the space of preference hierarchies, we can still engage in epistemic analysis by showing results that hold under all topologies that satisfy some weak regularity conditions. For example, Theorem 7.2 gives an epistemic characterization of IA that is valid for a large class of topologies on the hierarchy space.[20]

APPENDIX A HIGHER-ORDER PREFERENCES

Proof of Lemma 4.1. The proof is by induction.

Base case ($n = 0, 1$) We begin with the fact that $X_i^0 = S_i$ is a standard Borel space for all $i \in I$. If Ω is a standard Borel space, then $\mathsf{LEU}(\Omega)$ is a standard Borel space as well. It follows that $\mathsf{LEU}(X_{-i}^0)$ is a standard Borel space. $X_i^1 = X_i^0 \times \mathsf{LEU}(X_{-i}^0)$ is a product of standard Borel spaces and is therefore itself a standard Borel space. $\mathsf{LEU}(X_{-i}^1)$ is a standard Borel space because X_{-i}^1 is a standard Borel space.

Inductive hypothesis Let $n \geq 1$. For all $m \leq n$, let X_i^m and $\mathsf{LEU}(X_{-i}^{m-1})$ be standard Borel spaces.

Inductive step $\mathsf{LEU}(X_{-i}^n)$ is a standard Borel space because X_{-i}^n is a standard Borel space. We need to show that X_i^{n+1} is a Borel subset of $X_i^n \times \mathsf{LEU}(X_{-i}^n)$. Let the map κ_i be defined as follows.

$$\begin{aligned} \kappa_i \colon X_i^n \times \mathsf{LEU}(X_{-i}^n) &\to \mathsf{LEU}(X_{-i}^{n-1}) \times \mathsf{LEU}(X_{-i}^{n-1}) \\ (x_i^n, h_i^{n+1}) &\overset{\kappa_i}{\mapsto} (\varrho_i^n(x_i^n), \hat{\pi}_i^{n-1}(h_i^{n+1})) \end{aligned}$$

The map is Borel because each coordinate of its output is given by a Borel map. The diagonal set $D_i^n \equiv \{(h_i^n, h_i^n) \colon h_i^n \in \mathsf{LEU}(X_{-i}^{n-1})\}$ is Borel. It follows that $\kappa_i^{-1}(D_i^n) = X_i^{n+1}$ is Borel. □

[20] For example, any countably generated metrizable product topology will do.

Lemma A.1. *Let* $x_i = (x_i^0, x_i^1, \dots) \in X_i$. *Then*

$$(x_i^0, (\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i))_{n\geq 0}) = (x_i^0, (\varrho_i^{n+1}(x_i^{n+1}))_{n\geq 0}) \in X_i^0 \times H_i.$$

Proof of Lemma A.1. Because $x_i \in X_i$, we have $x_i^0 \in X_i^0$. By definition, $\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i)$ belongs to $\mathsf{LEU}(X_{-i}^n)$ for all $n \geq 0$. Let $h_i^{n+1} \equiv \varrho_i^{n+1}(x_i^{n+1})$ for all $n \geq 0$.

We only need to show that $\hat{\pi}_i^{n-1}(h_i^{n+1}) = h_i^n$ for all $n \geq 1$. Because $x_i \in X_i$, we have $\pi_i^n(x_i^{n+1}) = x_i^n$ for all $n \geq 0$. Due to the consistency requirements built into the definition of X_i^{n+1}, we therefore have $\hat{\pi}_i^{n-1} \circ \varrho_i^{n+1}(x_i^{n+1}) = \varrho_i^n(x_i^n)$. Substituting with h_i^n where possible, we have $\hat{\pi}_i^{n-1}(h_i^{n+1}) = h_i^n$ for all $n \geq 1$ □

Lemma A.2. *Let* $(x_i^0, h_i) = (x_i^0, h_i^1, h_i^2, \dots) \in X_i^0 \times H_i$. *Then*

$$(x_i^0, x_i^1, \dots) \in X_i, \quad \textit{where } x_i^{n+1} = (x_i^n, h^{n+1}) \textit{ for all } n \geq 0.$$

Proof of Lemma A.2. Because $h_i \in H_i$, we have $\hat{\pi}_i^{n-1}(h_i^{n+1}) = h_i^n$ for all $n \geq 1$. We want to show that $(x_i^n, h_i^{n+1}) \in X_i^{n+1}$.

It is trivial that $(x_i^0, h_i^1) \in X_i^0 \times \mathsf{LEU}(X_{-i}^0) = X_i^1$. Suppose that $x_i^m \in X_i^m$ for all $m \leq n$. We can show that $x_i^{n+1}(x_i^n, h_i^{n+1})$ satisfies the consistency requirement via the sequence of substitutions below.

$$\varrho_i^n \circ \pi^n(x_i^{n+1}) = \varrho_i^n(x_i^n) = h_i^n = \hat{\pi}_i^{n-1}(h_i^{n+1}) = \hat{\pi}_i^{n-1} \circ \varrho^{n+1}(x_i^{n+1})$$

□

Lemma A.3. *The map*

$$X_i \to X_i^0 \times \prod_{n\geq 0} \mathsf{LEU}(X_{-i}^n) \qquad (x_i^0, x_i^1, \dots) \mapsto (x_i^0, (\varrho_i^{n+1} \circ \varpi_i^{n+1}(x_i))_{n\geq 0})$$

is injective and Borel. Furthermore, image of X_i *under this map is* $X_i^0 \times H_i$.

Proof of Lemma A.3. The map is Borel because each coordinate of the output is given by a Borel map. The rest follows immediately from Lemmas A.1 and A.2. □

Proof of Lemma 4.3. By definition, X_i is the inverse limit of the inverse system $(X_i^n, \pi_i^n)_{n\geq 0}$, where X_i^n is Borel and the Borel map $\pi_i^n \colon X_i^{n+1} \to X_i^n$ is surjective.[21] It follows that the inverse limit X_i is a nonempty standard Borel space when endowed with the subspace Borel σ-algebra of the product $\prod_{n\geq 0} X_i^n$. See 17.16 in Kechris (1995).

Fix some $x_i^0 \in X_i^0$ and let f_i denote the map in Lemma A.3.

$$f_i^{-1}(\{x_i^0\} \times H_i) = \{x_i \in X_i \colon \varpi_i^0(x_i) = x_i^0\}$$

Since f_i is an injective Borel map, images of Borel sets are Borel. The set

$$\{x_i \in X_i \colon \varpi_i^0(x_i) = x_i^0\}$$

is Borel in X_i. It follows that $\{x_i^0\} \times H_i$ is a Borel subset of $X_i^0 \times \prod_{n\geq 0} \mathsf{LEU}(X_{-i}^n)$. H_i is a standard Borel space because it is isomorphic to $\{x_i^0\} \times H_i$. □

[21] π_i^n is defined in Section 4.2.

Proof of Lemma 4.2. Lemmas A.3 and 4.3 collectively imply the desired result. $\square$

Proof of Lemma 4.4. The map $\mathbb{bh}_i^n$ is Borel for all $n \geq 1$ because the preference-marginal operation is a Borel map on the relevant domain. The map $\mathbb{bh}_i$ is then Borel because each coordinate of the output is given by a Borel map.

Let $\rho, \sigma \in \mathsf{LPS}(X_{-i})$. That $\rho \cong \sigma \implies \mathbb{bh}_i(\rho) = \mathbb{bh}_i(\sigma)$ follows naturally from the definitions of $\mathbb{bh}_i$ and the margp operator—if two preferences are identical, then so should the marginals of those preferences. We need to show the converse, i.e., $\mathbb{bh}_i(\rho) = \mathbb{bh}_i(\sigma) \implies \rho \cong \sigma$.

Without loss of generality, let ρ and σ be minimal-length LPS's, where $m \geq 1$.

$$\rho = (\mu_1, \ldots, \mu_m) \qquad\qquad \sigma = (\nu_1, \ldots, \nu_m)$$

Because ρ and σ are both minimal-length LPS's, they are made up of a linearly independent components—i.e.,

$$\forall j \leq m \quad [\mu_j \notin \operatorname{span}(\{\mu_k \colon j \neq k\}) \quad \wedge \quad \nu_j \notin \operatorname{span}(\{\nu_k \colon j \neq k\})]\,.$$

Denote the belief-marginals of ρ and σ as follows for all $n \geq 1$.

$$\rho^n \equiv \operatorname{marg}_{X_{-i}^{n-1}} \rho \equiv (\mu_1^n, \ldots, \mu_m^n) \qquad\qquad \sigma^n \equiv \operatorname{marg}_{X_{-i}^{n-1}} \sigma \equiv (\nu_1^n, \ldots, \nu_m^n)$$

By definition, the following holds for all $n \geq 1$ due to the transitivity of the preference-equivalence relation.

$$\rho^n \cong \mathbb{bh}_i^n(\rho) = \mathbb{bh}_i^n(\sigma) \cong \sigma^n \qquad \therefore \rho^n \cong \sigma^n$$

Because ρ and σ are minimal-length LPS's, there must exist $M \geq 1$ such that ρ^N and σ^N are minimal-length LPS's for all $N \geq M$. Fix some $N \geq M$. We then have

$$\forall j \leq m \quad \exists \langle \alpha_1^j, \ldots, \alpha_j^j \rangle \in \mathbb{R}^j \quad \left[\alpha_j^j > 0 \ \wedge \ \nu_j^N = \sum_{k=1}^{j} \alpha_k^j \mu_k^N \right]$$

The coherency of X_{-i} implies the following for all $n \leq N$.

$$\forall j \leq m \qquad \mu_j^n = \operatorname{marg}_{X_{-i}^{n-1}} \mu_j = \operatorname{marg}_{X_{-i}^{n-1}} \mu_j^N$$

$$\begin{aligned}
\forall j \leq m \qquad \nu_j^n &= \operatorname{marg}_{X_{-i}^{n-1}} \nu_j = \operatorname{marg}_{X_{-i}^{n-1}} \nu_j^N = \operatorname{marg}_{X_{-i}^{n-1}} \sum_{k=1}^{j} \alpha_k^j \mu_k^N \\
&= \sum_{k=1}^{j} \alpha_k^j \operatorname{marg}_{X_{-i}^{n-1}} \mu_k^N = \sum_{k=1}^{j} \alpha_k^j \mu_k^n
\end{aligned}$$

Induction on N shows that

$$\forall j \leq m \quad \forall n \geq 1 \quad \nu_j^n = \sum_{k=1}^{j} \alpha_k^j \mu_k^n.$$

Finally, we want to show that $\nu_j = \sum_{k=1}^{j} \alpha_k^j \mu_k$ for all $j \le m$. By definition, ν_j and μ_j are probability measures on the inverse limit space $X_{-i} = \varprojlim_n X_{-i}^n$ that respectively extend the sequences $(\nu_j^1, \nu_j^2, \dots)$ and $(\mu_j^1, \mu_j^2, \dots)$—i.e.,

$$\nu_j = \varprojlim_n \nu_j^n \qquad\qquad \mu_j = \varprojlim_n \mu_j^n.$$

The measure $\sum_{k=1}^{j} \alpha_k^j \mu_k = \varprojlim_n (\sum_{k=1}^{j} \alpha_k^j \mu_k^n)$ because

$$\operatorname{marg}_{X_{-i}^{n-1}} \sum_{k=1}^{j} \alpha_k^j \mu_k = \sum_{k=1}^{j} \alpha_k^j \operatorname{marg}_{X_{-i}^{n-1}} \mu_k = \sum_{k=1}^{j} \alpha_k^j \mu_k^n$$

By the Kolmogorov Consistency Theorem (cf. 17.16. Kechris, 1995), such extensions must be unique. Because $(\sum_{k=1}^{j} \alpha_k^j \mu_k^n)_{n\ge1} = (\nu_j^1, \nu_j^2, \dots)$, the extensions of the two sequences to $\mathsf{P}(X_{-i})$ must be equal. □

APPENDIX B ADMISSIBILITY

Proof of Theorem 7.1. The theorem is an immediate corollary of Lemma B.3. □

Proof of Theorem 7.2. By Lemma B.5 there exists some $(m_0, m_1, \dots)$ such that

$$\text{(27)} \quad \forall k \ge 0 \quad \forall M \ge m_k \quad \forall i \in I \quad \varpi_i^k(R_i^M) = \varpi_i^k(R_i^M \setminus R_i^{M+1}) = \varpi_i^k(R_i^{M+1})$$

We can therefore glue together the sequence $(\varpi_i^k(R_i^{m_k}))_{k\ge0}$ by taking the projective/inverse limit of those spaces. The inverse limit is a nonempty standard Borel space because the projection maps—$\pi_i^k \colon X_i^{k+1} \to X_i^k$—are measurable and

$$\text{(28)} \quad \forall k \ge 0 \quad \pi_i^k(\varpi_i^{k+1}(R_i^{m_{k+1}})) = \varpi_i^k(R_i^{m_k}).$$

due to Lemma B.5. The inverse limit $\varprojlim_k \varpi_i^k(R_i^{m_k}) \subseteq R_i^\infty$. Furthermore, because of (27) and (28), we have

$$S_i^\infty = \varpi_i^0(R_i^{m_0}) = \operatorname{proj}_{X_i^0}\left[\varprojlim_k \varpi_i^k(R_i^{m_k})\right] \subseteq \varpi_i^0(R_i^\infty) \subseteq \bigcap_{m\ge0} \varpi_i^0(R_i^m) = S_i^\infty.$$

It follows immediately that $\varpi_i^0(R_i^\infty) = S_i^\infty$. □

Lemma B.1. *Let $G = \langle S_a, S_b, u_a, u_b\rangle$ be a finite game, $m \ge 1$, $s_a \in S_a$. The following set is Borel in $\mathsf{LPS}(S_b)$.*

$$\{\sigma \in \mathsf{LPS}(S_{-i}) \colon s_i \text{ is optimal w.r.t. } \sigma\} \\ = \{\sigma \in \mathsf{LPS}_m(S_b) \colon \forall s_a' \in S_a \quad \mathbf{E}_\sigma u_a(s_a, \cdot) \ge^{\mathsf{L}} \mathbf{E}_\sigma u_b(s_a', \cdot)\}$$

Proof of Lemma B.1. This subset of $\mathbb{R}^{m\cdot|S_b|}$ is a finite union of sets that satisfy a finite number of linear inequalities. It is therefore Borel. This is essentially the same argument found in the proof of Lemma C.4 in BFK. $\square$

Lemma B.2. *For all $m \geq 1$, R_i^m is Borel in H_i.*

Proof of Lemma B.2. Define the following objects for all $m \geq 1$ and $s_i \in S_i$.

$$\begin{aligned} O_i^1(s_i) &\equiv \{s_i\} \times \{\sigma \in \mathsf{LEU}^+(S_{-i})\colon s_i \text{ is optimal w.r.t. } \sigma\} \subseteq X_i^1 \\ \hat{R}_i^m &\equiv \bigcup_{s_i\in S_i} O_i^m(s_i) \\ O_i^{m+1}(s_i) &\equiv (\pi_i^m)^{-1}(O_i^m(s_i)) \cap \\ &\qquad\qquad (\varrho_i^{m+1})^{-1}(\{h_i^{m+1} \in \mathsf{LEU}^+(S_{-i})\colon \hat{R}_{-i}^m \text{ is assumed under } h_i^{m+1}\}) \end{aligned}$$

Base case By Lemma B.1, $O_i^1(s_i)$ is a Borel set for all $s_i \in S_i$. It follows that $\hat{R}_i^1$, which is a finite union of Borel sets, is a Borel set as well.

Inductive hypothesis Let $M \geq 1$. For all $s_i \in S_i$ and $m \leq M$, $O_i^m(s_i)$ and R_i^m are Borel.

Inductive step The maps π_i^M and ϱ_i^{M+1} are Borel. The set of LPS's that assume a given Borel set—such as the set inside the parentheses of the expression $(\varrho_i^{m+1})^{-1}(\cdot)$ from the definition of $O_i^{m+1}(s_i)$ above—is Borel (See Lee, 2013). By the inductive hypothesis, $O_i^{M+1}(s_i)$ is Borel because it is an intersection of two preimages of Borel sets under Borel maps. It follows that $\hat{R}_i^{M+1}$, which is a finite union of Borel sets, is a Borel set as well.

Finally, we want to show that R_i^m is Borel for all $m \geq 1$. We can rearrange the definition of R_i^{m+1} as follows by simply substituting equivalent expressions.

$$\begin{aligned} R_i^{m+1} &\equiv R_i^m \cap \{(s_i,(h_i^n)_n) \in S_i \times H_i\colon \varpi_{-i}^m(R_{-i}^m) \text{ is assumed under } h_i^{m+1}\} \\ &= R_i^1 \cap \bigcap_{k=1}^m \{(s_i,(h_i^n)_n) \in S_i \times H_i\colon \varpi_{-i}^k(R_{-i}^k) \text{ is assumed under } h_i^{k+1}\} \\ &= R_i^1 \cap \bigcap_{k=1}^m \{(s_i,(h_i^n)_n) \in S_i \times H_i\colon \hat{R}_{-i}^k \text{ is assumed under } h_i^{k+1}\} \end{aligned}$$

The set R_i^1 can be rewritten as the following finite intersection of Borel sets by simply substituting equivalent expressions. H_i^+ is Borel because it is a countable intersection of Borel sets.

$$R_i^1 = (S_i \times H_i^+) \cap (\varpi_i^1)^{-1}\left(\bigcup_{s_i\in S_i} O_i^1(s_i)\right)$$

It follows that R_i^m can be rewritten as the following finite intersection of Borel sets by simply substituting equivalent expressions.

$$
\begin{aligned}
R_i^m &= (S_i \times H_i^+) \cap \bigcap_{k=1}^{m} (\varpi_i^k)^{-1} \left(\bigcup_{s_i \in S_i} O_i^k(s_i) \right) \\
&= (S_i \times H_i^+) \cap (\varpi_i^m)^{-1} \left(\bigcup_{s_i \in S_i} O_i^m(s_i) \right)
\end{aligned}
$$

□

Lemma B.3. *Let $S_i^0 = S_i$ and $R_i^0 \equiv S_i^0 \times H_i$. Then, the following holds for all $m \geq 0$.*

$$\operatorname{proj}_{S_i}(R_i^m \setminus R_i^{m+1}) = S_i^m = \operatorname{proj}_{S_i} R_i^m$$

Proof of Lemma B.3. Note that $\operatorname{proj}_{S_i}(R_i^m \setminus R_i^{m+1}) = \varpi_i^0(R_i^m \setminus R_i^{m+1})$. The proof is by induction.

Base case: That $\varpi_i^0(R_i^0 \setminus R_i^1) = S_i^0$ is trivial and immediate. Because $\varrho_i^1 \circ \varpi_i^1(R_i^1) = \mathsf{LEU}^+(S_{-i}) \neq \mathsf{LEU}(S_{-i})$, there exists some $h_i \in H_i \setminus H_i^+$ and $S_i^0 \times \{h_i\} \subseteq R_i^0 \setminus R_i^1$.

Inductive hypothesis Let $M > 0$. Let the following hold for all $m < M$.

$$\varpi_i^0(R_i^m \setminus R_i^{m+1}) = S_i^m = \varpi_i^0(R_i^m)$$

Inductive step We want to show that $\varpi_i^0(R_i^M \setminus R_i^{M+1}) = S_i^M = \varpi_i^0(R_i^M)$. Define the set E_i^k as follows for all k.

$$E_i^k \equiv \varpi_i^M(R_i^k)$$

By the inductive hypothesis, $\operatorname{proj}_{S_i} E_i^k \setminus E_i^{k+1} = S_i^k$ for all $k < M$ because the definition of RmAR does not depend on beliefs of order higher than $m+1$.

There exist beliefs $\mu_2, \ldots, \mu_M \in \mathsf{P}(X_{-i}^M)$ such that $\mu_{M-k} \in \mathsf{P}^+(E_{-i}^k \setminus E_{-i}^{k+1})$ for all $k < M$. Let μ_1 be a strictly convex combination of probability measures in $\mathsf{P}^+(E_{-i}^{M-1})$ and $\mathsf{P}^+(E_{-i}^M)$. It must be the case that $\mu_1 \in \mathsf{P}^+(E_{-i}^{M-1})$ because $E_{-i}^M \subseteq E_{-i}^{M-1}$. The LPS $\sigma = (\mu_1, \ldots, \mu_M)$ cannot assume $E_{-i}^M = \varpi_i^M(R_i^M)$. However, it does assume E_{-i}^k for all $k < M$.

Because σ is an LCPS as it is defined, we can take advantage of existing results.[22] For any $(\nu_1', \ldots, \nu_M') \in \mathsf{LPS}(S_{-i})$ such that $\operatorname{supp} \nu_k' = \operatorname{marg}_{S_{-i}} \mu_k$ for all k, there exists an LCPS $\sigma' = (\mu_1', \ldots, \mu_M')$ such that, for all k,

1. $\operatorname{marg}_{S_{-i}} \mu_k' = \nu_k'$; and

[22]Technical intermediate results about the existence of LCPS's that assume the same events can be found in the appendices of BFK and Keisler and Lee (2012).

2. μ'_k and μ_k have the same null sets.

LCPS's of equal length such that μ'_k and μ_k have the same null sets for all k assume the same events.

It follows that $\varpi_i^0(R_i^M \setminus R_i^{M+1})$ includes all $s_i \in S_i$ that are optimal with respect to LPS's belonging to the following set.

$$\{(\nu_1, \ldots, \nu_M) \in \mathsf{LPS}^+(S_{-i})\colon \forall k \quad \operatorname{supp} \nu_k = S_i^{M-k}\}$$

Therefore, $S_i^M \subseteq \varpi_i^0(R_i^M \setminus R_i^{M+1})$. By the induction hypothesis, $\varpi_i^0(R_i^M \setminus R_i^{M+1}) \subseteq \varpi_i^0(R_i^{M-1}) = S_i^{M-1}$. We know that any h_i^M that assumes $\varpi_{-i}^{M-1}(R_{-i}^{M-1})$ must have an initial segment that belongs to $\mathsf{LPS}^+(\varpi_{-i}^{M-1}(R_{-i}^{M-1}))$. The belief-marginal on S_{-i} of this initial segment must have support equal to S_{-i}^{M-1}. It follows that $S_i^M \supseteq \varpi_i^0(R_i^M \setminus R_i^{M+1})$. Finally, we have $S_i^M = \varpi_i^0(R_i^M \setminus R_i^{M+1})$. $\square$

Lemma B.4. *Let $U' \subseteq U \subseteq \Omega \subseteq \Omega_1 \times \Omega_2 \times \Omega_3$, where $\Omega_j = \operatorname{proj}_{\Omega_j} \Omega$ is endowed with a separable and metrizable topology that generates the standard Borel σ-algebra. If U and U' are nonempty Borel sets such that $\operatorname{proj}_{\Omega_1} U = \operatorname{proj}_{\Omega_1} U' = \operatorname{proj}_{\Omega_1} U \setminus U'$ and $\operatorname{proj}_{\Omega_1} \Omega_{12} \setminus U_{12} = \Omega_1$ then*

$$\{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A' \cap A\} = \{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A \setminus A'\} = \{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A\},$$

where

$$\begin{aligned} A' &= \{\sigma \in \mathsf{LPS}(\Omega)\colon \sigma \text{ assumes } U'\} \quad \text{and} \\ A &= \{\sigma \in \mathsf{LPS}(\Omega)\colon (\operatorname{marg}_{\Omega_1 \times \Omega_2} \sigma) \text{ assumes } (\operatorname{proj}_{\Omega_1 \times \Omega_2} U)\} \\ \Omega_{12} &= \operatorname{proj}_{\Omega_1 \times \Omega_2} \Omega \quad \text{and} \quad U_{12} = \operatorname{proj}_{\Omega_1 \times \Omega_2} U. \end{aligned}$$

Proof of Lemma B.4. Let $U_1 = \operatorname{proj}_{\Omega_1} U$. First, $\sigma \in A$ if and only if σ is preference-equivalent to some LPS in the following set.

$$\{\sigma'.\sigma''\colon \sigma' \in \mathsf{LPS}^+(U_{12}) \wedge \sigma'' \in \mathsf{LPS}(\Omega_{12} \setminus U_{12})\}$$

It follows that $\nu \in \{\operatorname{marg}_{\Omega_1} \sigma\colon \sigma \in A\}$ if and only if it is preference-equivalent to some LPS in the following set.

$$\begin{aligned} &\{\operatorname{marg}_{\Omega_1} \sigma'.\sigma''\colon \sigma' \in \mathsf{LPS}^+(U_{12}) \wedge \sigma'' \in \mathsf{LPS}(\Omega_{12} \setminus U_{12})\} \\ &\qquad = \{(\operatorname{marg}_{\Omega_1} \sigma').(\operatorname{marg}_{\Omega_1} \sigma'')\colon \sigma' \in \mathsf{LPS}^+(U_{12}) \wedge \sigma'' \in \mathsf{LPS}(\Omega_{12} \setminus U_{12})\} \\ &\qquad = \{\nu'.\nu''\colon \nu' \in \mathsf{LPS}^+(U_1) \wedge \nu'' \in \mathsf{LPS}(\Omega_1)\} \end{aligned}$$

Given that

$$\{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A' \cap A\} \subseteq \{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A\} \supseteq \{\operatorname{margp}_{\Omega_1} \sigma\colon \sigma \in A \setminus A'\},$$

we will only need to show that

$$\{\operatorname{margp}_{\Omega_1}\sigma\colon \sigma\in A'\cap A\}\supseteq\{\operatorname{margp}_{\Omega_1}\sigma\colon \sigma\in A\}\subseteq\{\operatorname{margp}_{\Omega_1}\sigma\colon \sigma\in A\setminus A'\}.$$

Take any $\nu'.\nu''\in\{\nu'.\nu''\colon \nu'\in\mathsf{LPS}^+(U_1)\wedge\nu''\in\mathsf{LPS}(\Omega_1)\}$. We want to show that there exists some LPS in $\sigma\in A'\cap A$ such that $\operatorname{marg}_{\Omega_1}\sigma\cong\nu'.\nu''$. Because $U_1=\operatorname{proj}_{\Omega_1}U=\operatorname{proj}_{\Omega_1}U'=\operatorname{proj}_{\Omega_1}U\setminus U'$, there exist LPS's $\zeta\in\mathsf{LPS}^+(U')$ and $\zeta'\in\mathsf{LPS}^+(U\setminus U')$ such that $\nu=\operatorname{marg}_{\Omega_1}\zeta=\operatorname{marg}_{\Omega_1}\zeta'$. Furthermore, because $\operatorname{proj}_{\Omega_1}\Omega\setminus U=\Omega_1$ (this follows from $\operatorname{proj}_{\Omega_1}\Omega_{12}\setminus U_{12}=\Omega_1$), there exists some $\zeta''\in\mathsf{LPS}(\Omega\setminus U)$ such that $\nu'=\operatorname{marg}_{\Omega_1}\zeta''$. Let $\sigma=\zeta.\zeta'.\zeta''$. Then $\sigma\in A\cap A'$ and $\operatorname{marg}_{\Omega_1}\sigma=\nu.\nu.\nu'\cong\nu.\nu'$. Therefore, $\{\operatorname{margp}_{\Omega_1}\sigma\colon\sigma\in A'\cap A\}\supseteq\{\operatorname{margp}_{\Omega_1}\sigma\colon\sigma\in A\}$.

Finally, we want to show that there exists some LPS in $\sigma\in A\setminus A'$ such that $\operatorname{marg}_{\Omega_1}\sigma\cong\nu'.\nu''$. Fix ζ,ζ',ζ'' as in the previous paragraph. Let $\sigma=\zeta'.\zeta.\zeta''$. Then $\sigma\in A\setminus A'$ and $\operatorname{marg}_{\Omega_1}\sigma=\nu.\nu.\nu'\cong\nu.\nu'$. Therefore, $\{\operatorname{margp}_{\Omega_1}\sigma\colon\sigma\in A\}\subseteq\{\operatorname{margp}_{\Omega_1}\sigma\colon\sigma\in A\setminus A'\}$. □

Lemma B.5. *For all $k\geq 0$, there exists some $m\geq 0$ such that*

$$\forall M\geq m\quad\forall i\in I\quad \varpi_i^k(R_i^M)=\varpi_i^k(R_i^M\setminus R_i^{M+1})=\varpi_i^k(R_i^{M+1})$$

Proof of Lemma B.5. The proof is by induction.

Base case We want to show that exists some m such that, the following holds.

$$\forall M\geq m\quad\forall i\in I\quad \varpi_i^0(R_i^M)=\varpi_i^0(R_i^M\setminus R_i^{M+1})=\varpi_i^0(R_i^{M+1})$$

Because S_i is a finite set, there exists some m such that $S_i^m=S_i^M$ for all $M\geq m$. By Lemma B.3, the following holds for all $M\geq m$.

$$S_i^m=S_i^M=\varpi_i^0(R_i^M\setminus R_i^{M+1})=\varpi_i^0(R_i^M)$$

Inductive hypothesis There exists some m such that

$$\forall M\geq m\quad\forall i\in I\quad \varpi_i^k(R_i^M)=\varpi_i^k(R_i^M\setminus R_i^{M+1})=\varpi_i^k(R_i^{M+1})$$

Induction step Let

$$A'=\{\sigma\in\mathsf{LPS}(X_{-i}^{m+1})\colon \sigma\text{ assumes }\varpi_{-i}^{m+1}(R_{-i}^{m+1})\}\quad\text{and}$$
$$A=\{\sigma\in\mathsf{LPS}(X_{-i}^{m+1})\colon (\operatorname{marg}_{X_{-i}^m}\sigma)\text{ assumes }\varpi_{-i}^{m}(R_{-i}^{m})\}.$$

By applying Lemma B.4 under the induction hypothesis, we can see that the following is true.

$$\{\operatorname{margp}_{X_{-i}^k}\sigma\colon\sigma\in A'\cap A\}=\{\operatorname{margp}_{X_{-i}^k}\sigma\colon\sigma\in A\setminus A'\}=\{\operatorname{margp}_{X_{-i}^k}\sigma\colon\sigma\in A\}$$

From this, the following equalities can be deduced from the definition of RmAR in this paper.

$$\begin{aligned}\{\operatorname{margp}_{X^k_{-i}} \sigma \colon \sigma \in \varrho_i^{m+1} \circ \varpi_i^{m+1}(R_i^{m+2})\} \\ &= \{\operatorname{margp}_{X^k_{-i}} \sigma \colon \sigma \in \varrho_i^{m+1} \circ \varpi_i^{m+1}(R_i^{m+1} \setminus R_i^{m+2})\} \\ &= \{\operatorname{margp}_{X^k_{-i}} \sigma \colon \sigma \in \varrho_i^{m+1} \circ \varpi_i^{m+1}(R_i^{m+1})\}\end{aligned}$$

It immediately follows that $\varpi_i^{k+1}(R_i^{m+2}) = \varpi_i^{k+1}(R_i^{m+1} \setminus R_i^{m+2}) = \varpi_i^{k+1}(R^{m+1})$. The induction hypothesis implies the existence of a number m that satisfies some property. Note that, if this property is satisfied for $m = n$, then it is also satisfied for all $m \geq n$. By changing the arguments above where necessary, it is immediately shown that

$$\forall M \geq m+1 \quad \varpi_i^{k+1}(R_i^{M+1}) = \varpi_i^{k+1}(R_i^M \setminus R_i^{M+1}) = \varpi_i^{k+1}(R^M).$$

$\square$

REFERENCES

Asheim, G. B. (2001). Proper rationalizability in lexicographic beliefs. *International Journal of Game Theory*, **30**, 453–478.

Battigalli, P. and Siniscalchi, M. (1999). Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory*, **88** (1), 188–230.

Blume, L., Brandenburger, A. and Dekel, E. (1991a). Lexicographic probabilities and choice under uncertainty. *Econometrica*, **59** (1), 61–79.

—, — and — (1991b). Lexicographic probabilities and equilibrium refinements. *Econometrica*, **59** (1), 81–98.

Brandenburger, A. and Dekel, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, **59** (1), 189–198.

— and Friedenberg, A. (2007). The relationship between rationality on the matrix and the tree, http://pages.stern.nyu.edu/~abranden/papers.html.

—, — and Keisler, H. J. (2008). Admissibility in games. *Econometrica*, **76** (2), 307–352.

Catonini, E. (2012). Common assumption of cautious rationality and iterated admissibility, mimeo.

Dekel, E., Friedenberg, A. and Siniscalchi, M. (2013). Assumption, mimeo.

Di Tillio, A. (2008). Subjective expected utility in games. *Theoretical Economics*, **3** (3), 287–323.

Epstein, L. G. and Wang, T. (1996). Beliefs about beliefs without probabilities. *Econometrica*, **64** (6), 1343–1373.

Ganguli, J. and Heifetz, A. (2013). Universal interactive preferences, mimeo.

Halpern, J. Y. (2010). Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, **68** (1), 155–179.

Hammond, P. J. (1994). Elementary non-Archimedean representations of probability for decision theory and games. In P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher*, vol. 1, Dordrecht, Netherlands: Kluwer, pp. 25–59.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I-III. part I. the basic model. *Management Science*, **14** (3), 159–182.

Kechris, A. S. (1995). *Classical Descriptive Set Theory*. New York: Springer-Verlag.

Keisler, H. J. and Lee, B. S. (2012). Common assumption of rationality, mimeo.

Kohlberg, E. and Mertens, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, **54** (5), 1003–1037.

Lee, B. S. (2013). A space of lexicographic preferences, mimeo.

Liu, Q. (2009). On redundant types and bayesian formulation of incomplete information. *Journal of Economic Theory*, **144** (5), 2115–2145.

Mertens, J.-F. and Zamir, S. (1985). Formulation of bayesian analysis for games with incomplete information. *International Journal of Game Theory*, **14** (1), 1–29.

Myerson, R. B. (1978). Refinements of the nash equilibrium concept. *International Journal of Game Theory*, **7** (2), 73–80.

— (1986). Multistage games with communication. *Econometrica*, **54** (2), 323–358.

Samuelson, L. (1992). Dominated strategies and common knowledge. *Games and Economic Behavior*, **4** (2), 284–313.

Selten, R. (1975). Refinements of the nash equilibrium concept. *International Journal of Game Theory*, **4** (1), 25–55.

Tan, T. C.-C. and Werlang, S. (1988). The Bayesian foundations of solution concepts of games. *Journal of Economic Theory*, **45** (2), 370–391.