

Lexicographic Beliefs and Assumption*

Eddie Dekel[†] Amanda Friedenberg[‡] Marciano Siniscalchi[§]

May 5, 2016

Abstract

Foundations for iterated admissibility (i.e., the iterated removal of weakly dominated strategies) need to confront a fundamental challenge. On the one hand, admissibility requires that a player consider every strategy of their opponents possible. On the other hand, reasoning that the opponents are rational requires ruling out certain strategies. Brandenburger, Friedenberg and Keisler’s (BFK, *Econometrica*, 2008) foundations for iterated admissibility address this challenge with two ingredients: lexicographic beliefs and the concept of “assumption.” However, BFK restrict attention to lexicographic beliefs whose supports are essentially disjoint. This restriction does not have a compelling behavioral rationale, or a clear intuitive interpretation. At the same time, it plays a crucial role in BFK’s foundations for iterated admissibility—specifically, in their analysis of assumption. We provide an alternate characterization of assumption, which applies to all lexicographic beliefs. We also characterize two variants of assumption, based on two extensions of ‘weak dominance’ to infinite state spaces. These notions of assumption coincide with BFK’s notion when the state space is finite and lexicographic beliefs have disjoint support; but they are different in more general settings. Leveraging these characterization results, we show that disjoint supports do not play a role in the foundations for iterated admissibility.

JEL classification: C72; D81; D83

Keywords: Weak dominance, Weak dominance in infinite games, Iterated admissibility, Lexicographic probability systems, Assumption, Epistemic game theory

*Eddie Dekel gratefully acknowledges NSF grant SES-1227434 and Amanda Friedenberg gratefully acknowledges NSF grant SES-1358008. We thank John Farragut, Egor Starkov, and Jaber Zarezadeh for excellent research assistance.

[†]Northwestern and Tel Aviv Universities, email: dekel@northwestern.edu

[‡]Arizona State University, email: amanda.friedenberg@asu.edu

[§]Northwestern University, email: marciano@northwestern.edu

1 Introduction

Lexicographic beliefs (henceforth ℓ -beliefs) have become a relatively standard tool, both for studying refinements and for providing epistemic characterizations of solution concepts.¹ The appeal of ℓ -beliefs is that they can be used to address a tension between being certain that an opponent is rational and having full-support beliefs about opponents' actions. To clarify, suppose that, if Bob is rational, he will not play specific actions. Can Ann be certain that Bob is rational, and at the same time be cautious and assign non-zero probability to all of Bob's actions? The answer is no if Ann has standard probabilistic beliefs. Suppose instead that Ann has ℓ -beliefs. That is, she has a vector $(\mu_0, \dots, \mu_{n-1})$ of probabilities over the relevant space of uncertainty, S_b (Bob's strategy space) and uses them lexicographically to determine her preferences over her own strategies: Ann first ranks her strategies using μ_0 ; if that leads to more than one best reply for Ann, she uses μ_1 to rank them, and so on. If the union of the supports of the probabilities μ_i is all of S_b , then Ann's beliefs have, in a sense, full support. At the same time, Ann can still be confident in Bob's rationality, for example in the sense that the primary hypothesis μ_0 assigns positive probability only to strategies of Bob that are rational.

There are two notions of ℓ -beliefs that have been studied and used in the literature: lexicographic conditional probability systems (henceforth LCPSs) in which, loosely speaking, the supports of the different beliefs (i.e., the μ_i 's) are disjoint, and the more general class of lexicographic probability systems (LPSs) in which this disjointedness condition is not imposed. In particular, LCPSs are used by [Brandenburger, Friedenberg and Keisler \(2008, henceforth, BFK\)](#) to provide an epistemic characterization of iterated admissibility—thereby answering a long-standing open question.²

However, there are reasons not to find the restriction to LCPSs appealing. First, while [Blume, Brandenburger and Dekel \(1991a\)](#) provide an axiom that characterizes LCPSs within the class of LPSs, their axiom has a flavor of reverse-engineering: it says no more than the probabilities in the LPS have disjoint support; it offers no further normative or other appeal. Indeed, the interpretation of LPSs is quite natural and intuitive. The probability μ_0 is the player's primary hypothesis, in the sense that she is (almost fully) confident in it. The probability μ_1 is her secondary hypothesis: she is willing to entertain it as an alternative,

¹See, for example, [Blume, Brandenburger and Dekel \(1991b\)](#), [Brandenburger \(1992\)](#), [Stahl \(1995\)](#), [Mailath, Samuelson and Swinkels \(1997\)](#), [Rajan \(1998\)](#), [Asheim \(2002\)](#), [Govindan and Klumpp \(2003\)](#), [Brandenburger, Friedenberg and Keisler \(2008\)](#), [Keisler and Lee \(2010\)](#), [Lee \(2015a\)](#), [Yang \(2013\)](#), and [Catonini and De Vito \(2014\)](#) amongst many others.

²To be more precise: BFK provide an epistemic characterization of m rounds of deleting inadmissible strategies, for any finite m . Their epistemic conditions involve finite-order reasoning. However, they show an "impossibility result" for common reasoning—that is, common reasoning is impossible in their model.

but considers it “infinitely” less plausible than μ_0 ; and so on. There is no reason that primary and secondary hypotheses must have disjoint supports. For instance, one may be confident that a coin is fair, but entertain the secondary hypothesis that it is biased towards falling on heads.³ Second, the marginal of an LCPS need not be an LCPS. For example, suppose that two players are playing the game in Figure 1.1, where the pairs of actions A, B for each constitute a zero-sum matching pennies game, (A, C) and (C, A) give $(-2, 3)$ and $(3, -2)$ respectively and anything else gives $(-4, -4)$. Consider the ℓ -belief over this game where μ_0 is that the players are playing the equilibrium of the matching pennies game while μ_1 is that they are playing the Pareto superior outcome that requires correlation of (A, C) and (C, A) with probability one half each. The beliefs on the joint action space are clearly an LCPS, but the marginal on one player’s actions has the first belief being that A and B are equally likely while the second belief is that A and C are equally likely, which is clearly not an LCPS. Thus, if one takes a small-worlds approach in which the beliefs we use to study a particular game are the marginals of some belief on a larger space, then the beliefs in the game need not be an LCPS (even if one requires that the overall belief be an LCPS). For these reasons we find LPSs more suitable for the study of refinements than LCPSs.

	A	B	C
A	-1,1	1,-1	-2,3
B	1,-1	-1,1	-4,-4
C	3,-2	-4,-4	-4,-4

Figure 1.1: The marginal of an LCPS may not be an LCPS

This raises the question of whether BFK’s characterization of iterated admissibility (IA) requires the use of LCPSs. We show that it does not: There is an analogue of BFK’s characterization of IA for the more general notion of LPSs.⁴ As we will discuss below, this result is important for evaluating the epistemic foundations of IA. Showing the result requires two steps.

First, BFK define what it means for one player to “be certain” that another is rational. A key feature of their approach is that they do so in terms of the player’s preferences. This decision-theoretic approach is analogous to that taken in [Morris \(1997\)](#) and, more recently, [Asheim and S¸ovik \(2005\)](#). One advantage is that such preference-based definitions can be evaluated on their own merits, independently of the (arbitrary) choice of a particular representation.

³Of course one may instead have the secondary hypothesis that the coin will fall on an edge, which would have disjoint support, but that does not seem like the only story one could tell.

⁴In fact, we show that each of BFK’s three main results all hold for the more general notion of LPSs.

BFK introduce the notion of “assumption.” An event E is *assumed* if it is “infinitely more likely” than its complement. They formalize assumption as a requirement on preferences. When the state space is finite and beliefs have full support, the requirement can be stated as:

(*) whenever a player prefers an act x to an act y conditional on E (loosely speaking, if she were to be informed of E), she also prefers x to y unconditionally (i.e., without this information).⁵

In the usual case where the player has one level of beliefs, this corresponds exactly to probability-1 belief. (See Section 3.1 for precise statements.) BFK show that, with LCPS beliefs, condition (*) is (essentially) equivalent to the following:

there is a belief level j such that:

(BFK-i) for all $i \leq j$ we have $\mu_i(E) = 1$ and

(BFK-ii) for all $i > j$ we have $\mu_i(E) = 0$.

However, the equivalence between the preference-based condition (*) and its ℓ -belief counterpart, conditions (BFK-i) and (BFK-ii), *only* holds if the player’s beliefs are represented by an LCPS. We illustrate this in Examples 3.1-3.2, which also show that the problem lies with condition (BFK-ii) above.

Therefore, to state an analogue of BFK’s epistemic conditions with unrestricted beliefs, it is necessary to first characterize condition (*) for general LPSs. Our main result, Theorem 3.2, does just that. It provides the precise weakening of condition (BFK-ii) required for the equivalence. The key idea is that condition (BFK-ii) implies:

(†) the payoffs at states in E “do not matter” as far as the probabilities of level greater than j are concerned; that is, the ranking of acts by probabilities above j are unaffected by the payoffs at states in E .

However, there is another way in which (†) can hold: if, for $k > j$, the restriction of μ_k to E is a linear combination of the lower-level probabilities μ_0, \dots, μ_j . With LCPSs, we need not worry about this possibility, since the supports of $\mu_0, \dots, \mu_j, \mu_k$ must be disjoint. But if we drop the disjointness requirement, we need to allow for this possibility.

Second, BFK provide an epistemic characterization of IA (and self-admissible sets, or SASs; see Definition 6.4), using the LCPS formulation of assumption, i.e., conditions (BFK-i) and (BFK-ii). We show that, when players’ beliefs are represented by unrestricted LPSs,

⁵We emphasize that, as in Savage, there is no real “information” in our static setting; this is just suggestive language.

the very same epistemic conditions continue to characterize IA (and SASs), provided we use our LPS formulation of assumption.⁶

Note that we allow players to hold a larger set of beliefs than do BFK. This immediately implies that more strategies are consistent with rationality, as there are more beliefs to justify a given strategy. However, as BFK point out, allowing more beliefs may lead to *fewer* strategies being consistent with rationality and mutual assumption thereof. (This is because assumption is non-monotonic, i.e., a set can be assumed even though a larger set is not assumed.) Despite these two opposing forces, our analysis still gives an epistemic characterization of SASs (Theorem 6.1) and IA (Theorem 6.2). It also retains BFK’s so-called impossibility theorem (Theorem 6.3).

Our result addresses a tension in BFK’s analysis. In particular, a key ingredient in their epistemic characterization of IA is a “completeness” requirement: the analysis takes place in a type structure that represents every possible belief about the opponent’s strategies and beliefs. But, because they restrict attention to LCPSs, BFK’s completeness requirement is that the type structure represents every possible *LCPS* belief. However, since the marginal of an LCPS need not be an LCPS, players can have first-order beliefs (about the strategies played) which are not LCPSs. In fact, BFK show that, for the purposes of providing an epistemic characterization of IA, the type structure *must* include first-order beliefs that are non-LCPSs. This leads to a tension: On the one hand, BFK’s analysis insists on LCPS beliefs on the full space of uncertainty; on the other hand, their analysis requires inclusion of non-LCPS beliefs on the first-order space of uncertainty (i.e., on the strategies of the opponent). Our Theorem 6.2 resolves this tension by considering type structures that represent all possible LPS beliefs—not just LCPS beliefs.

The preceding informal discussion imposed two implicit restrictions—that the state space was finite and that beliefs have full support. However, BFK’s epistemic conditions for iterated admissibility requires a “complete type structure,” which induces an uncountable state space. (See Section 6.) Thus, when we turn to the formal analysis, it is important to consider uncountable state spaces. This requires care: Condition (*) defines assumption for a finite state space and full-support beliefs. As BFK observe, for uncountable state spaces, it is no longer a suitable definition of assumption, even if there are full-support beliefs. (Correspondingly, conditions (BFK-i) and (BFK-ii) do not characterize assumption on uncountable state spaces, even for full-support LCPSs. See BFK’s Supplemental Appendix S.1, where they discuss this point.) Because our analysis accommodates arbitrary (uncountable) state spaces and we do not require full-support beliefs, we will need to follow BFK in modifying

⁶The proofs of these epistemic results follow BFK closely; the only significant modification is in establishing measurability. See the Appendix.

Condition (*).⁷

Return to the introductory case, in which the state space is finite and beliefs are represented by full-support LCPSs. There, condition (*) has an alternative, equivalent, formulation:

(**) whenever an act x weakly dominates y on E (i.e., x is at least as good as y in every state in E , and strictly better at some state in E), the player prefers x to y unconditionally.

Condition (**) was proposed by [Asheim and Dufwenberg \(2003\)](#).⁸ It can be seen as a variant of assumption. [Asheim and Sørvik \(2005\)](#) established the aforementioned equivalence between conditions (*) and (**). However, these conditions are *not* equivalent for LPSs, even if the state space is finite (Example 4.2).

This raises the question of whether (**) can be used to provide an epistemic characterization of IA. A central challenge in addressing the question is extending (**) to arbitrary state spaces—specifically, extending the notion of weak dominance to arbitrary state spaces. We consider two such extensions; see Section 4.1 for details. We provide an LPS-based characterization of these versions of assumption (Theorem 4.1). We show that either variant can be used to provide an epistemic characterization of IA (Theorem 6.2). Moreover, BFK’s impossibility theorem is retained under both variants (Theorem 6.3).

Section 2 introduces the framework. Section 3 reviews the definition of assumption and BFK’s characterization for LCPSs (Section 3.1), motivates the modifications needed for LPSs (Section 3.2), and provides the behavioral characterization thereof (Section 3.3). Section 4 provides two definitions of weak dominance for arbitrary state spaces (Section 4.1), uses them to extend (**) (Section 4.2), and provides the characterizations thereof (Section 4.3). Section 5 provides the proofs of the characterization theorems. Section 6 applies these results to the epistemic characterizations of iterated admissibility and SASs. Section 7 discusses the closely related work of [Lee \(2013, 2015a\)](#). The Appendix provides proofs not included in the body.

2 Preliminaries

Let (Ω, \mathcal{S}) be a Polish space, where \mathcal{S} is the Borel σ -algebra on Ω . We call the elements of \mathcal{S} “Borel sets” or “events.” Write $\mathcal{P}(\Omega)$ for the set of probability measures on Ω and endow $\mathcal{P}(\Omega)$ with the topology of weak convergence, so that it is also a Polish space.

⁷We thank an anonymous referee for suggesting that we drop the full-support requirement.

⁸We thank an anonymous referee for suggesting studying (**) in the context of BFK’s epistemic analysis.

Denote by \geq^L the lexicographic order on \mathbb{R}^n . That is, given vectors $u = (u_0, \dots, u_{n-1})$ and $v = (v_0, \dots, v_{n-1})$ in \mathbb{R}^n , $u \geq^L v$ if and only if $u_j < v_j$ implies $u_k > v_k$ for some $k < j$.

A **lexicographic probability system (LPS)** on Ω will be some $\sigma = (\mu_0, \dots, \mu_{n-1})$ where each $\mu_i \in \mathcal{P}(\Omega)$. Call an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ a **lexicographic conditional probability system (LCPS)** if there are Borel sets U_0, \dots, U_{n-1} such that, for all i , $\mu_i(U_i) = 1$ and $\mu_j(U_i) = 0$ for $j \neq i$.⁹ Notice that, if Ω is finite, this simply requires that the measures μ_0, \dots, μ_{n-1} have disjoint supports. An LPS $\sigma = (\mu_0, \dots, \mu_n)$ has **full support** if $\bigcup_{i=0}^{n-1} \text{supp } \mu_i = \Omega$.

Let \mathcal{A} be the set of all measurable functions from Ω to $[0, 1]$. A particular function $x \in \mathcal{A}$ is an act. For $c \in [0, 1]$, write \vec{c} for the constant act associated with c , i.e. $\vec{c}(\Omega) = \{c\}$. Given acts $x, z \in \mathcal{A}$ and a Borel subset E in Ω , write $(x_E, z_{\Omega \setminus E})$ for the act $y \in \mathcal{A}$ with

$$y(\omega) = \begin{cases} x(\omega) & \text{if } \omega \in E \\ z(\omega) & \text{if } \omega \in \Omega \setminus E. \end{cases}$$

When $\Omega = \{\omega_0, \omega_1, \dots, \omega_K\}$, write (x_0, x_1, \dots, x_K) for an act x with $x(\omega_k) = x_k$. In this case, we also write $\mu = (\mu(\omega_0), \mu(\omega_1), \dots, \mu(\omega_K))$ for some $\mu \in \mathcal{P}(\Omega)$.

Given an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ on Ω , define a preference relation \succsim^σ on \mathcal{A} where $x \succsim^\sigma y$ if and only if

$$\left(\int_{\Omega} x(\omega) d\mu_i(\omega) \right)_{i=0}^{n-1} \geq^L \left(\int_{\Omega} y(\omega) d\mu_i(\omega) \right)_{i=0}^{n-1}.$$

Write \succ^σ for the associated strict preference relation. Given a Borel set E , define the conditional preference given E in the usual way, i.e., $x \succsim_E^\sigma y$ if for some act $z \in \mathcal{A}$, $(x_E, z_{\Omega \setminus E}) \succsim^\sigma (y_E, z_{\Omega \setminus E})$. (Since \succsim^σ satisfies independence, the choice of the act z does not affect the conditional preference relation.) Write \succ_E^σ for the associated strict preference relation and \sim_E^σ for the associated indifference relation.

An event E is \succsim^σ -**null** if $x \sim^\sigma y$, for all $x, y \in \mathcal{A}$ that coincide on $\Omega \setminus E$ (i.e., with $x(\omega) = y(\omega)$ for all $\omega \in \Omega \setminus E$). Equivalently, E is \succsim^σ -null if $x \sim_E^\sigma y$ for all $x, y \in \mathcal{A}$.

Remark 2.1 Fix some LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. A Borel set E is \succsim^σ -null if and only if $\mu_i(E) = 0$ for all i .

Proof. If $\mu_i(E) = 0$ for each i , then $x \sim_E^\sigma y$. Conversely, if $\mu_i(E) > 0$ for some i , then $\vec{1} \succ_E^\sigma \vec{0}$. ■

⁹ This terminology is due to [Blume, Brandenburger and Dekel, 1991a](#), who define LCPSs for finite state spaces. The present definition is Definition 4.1 in BFK.

3 BFK-Assumption

3.1 Definition and LCPS-based characterization

BFK define assumption in terms of the preference relation \succsim^σ associated with an LCPS σ ; their definition equally applies to the case where σ is an LPS. The informal idea is that an event E is assumed if states in E “determine” strict preference.

Definition 3.1 (BFK, Definition A.3) *Say a set E is **BFK-assumed** under \succsim^σ if E is Borel and the following hold:*

Non-Triviality: *for each open set U , if $E \cap U \neq \emptyset$, then $E \cap U$ is not \succsim^σ -null.*

Strict Determination: *for all $x, y \in \mathcal{A}$, $x \succ_E^\sigma y$ implies $x \succ^\sigma y$.*

Non-Triviality states that every “part” of E is relevant, in the sense that it can potentially determine strict preference. Strict Determination states that, if x is strictly preferred to y conditional on E , then x is also unconditionally strictly preferred to y (regardless of the outcomes x and y may deliver outside of E). Observe that Non-Triviality ensures that Strict Determination does not hold vacuously. When the state space is finite, Non-Triviality holds if the LPS has full support.

BFK provide the following characterization for LCPSs.

Theorem 3.1 (BFK, Proposition A.2 and Lemma B.1) *Fix an LCPS σ . A set $E \subseteq \Omega$ is BFK-assumed under \succsim^σ if and only if it is Borel and there exists $j \in \{0, \dots, n-1\}$ such that*

$$(i) \mu_i(E) = 1 \text{ for all } i \leq j,$$

$$(ii) \mu_i(E) = 0 \text{ for all } i > j, \text{ and}$$

$$(iii) E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i.$$

The remainder of this section discusses how to modify Conditions (i)–(iii) in order to obtain an analog of Theorem 3.1, i.e., a characterization of BFK-assumption for LPSs. As discussed in the Introduction, an alternative is to modify the notion of BFK-assumption. We discuss this approach in Section 4.

3.2 From LCPSs to LPSs

We next explore the extent to which Theorem 3.1 holds for arbitrary LPSs. One direction of this theorem holds for all LPSs: If there is some j that satisfies conditions (i)—(iii), then \succsim^σ BFK-assumes E . This holds even if σ does *not* have disjoint supports. However, the conditions in Theorem 3.1 are not necessary for an event E to be BFK-assumed. In particular, we now argue that the problem arises from Condition (ii).

Example 3.1 illustrates that, for LPSs, conditions (i) and (iii) do not suffice for BFK-assumption. The remaining examples and observations in this subsection illustrate that, for LPSs, BFK-assumption implies that Conditions (i) and (iii) must hold for some j , but (ii) may fail for any such j . Thus (i) and (iii) are necessary but not sufficient. Since (i), (ii), and (iii) are jointly sufficient, we are led to weaken (ii). (With this in mind the reader interested in getting to the results can skip to Subsection 3.3.)

In this subsection we focus on finite state spaces, which is enough to illustrate these issues. For LCPSs on finite state spaces Condition (ii) is redundant: if (iii) holds for some j then so does (ii). Hence, for LCPSs, Condition (ii) only plays a role in infinite state spaces—that is, for LCPSs on finite state spaces (i) and (iii) are equivalent to BFK-assumption. Hence, for LCPSs, (ii) might be seen as a technicality. However, for LPSs, its weakening plays a substantive role even for finite state spaces.

We now provide an outline of the remainder of this subsection. Example 3.1 shows that (i) and (iii) are not sufficient for BFK-assumption; the next example (3.2) and observation (3.1) concern a case where (i) and (iii) hold for $j = 0$, and in particular illustrate that if E is BFK-assumed then $\mu_0(E) = 1$. This is essentially Lemma 5.1, the first step in the proof of our main result. The final example (3.3) and observation (3.2) illustrate more generally that (i) and (iii) must hold for some $j < n - 1$, mirroring Lemmas 5.2 and 5.5 .

Example 3.1 *Let $\Omega = \{\omega_0, \omega_1, \omega_2\}$ and consider the LPS $\sigma = (\mu_0, \mu_1, \mu_2)$ such that $\mu_0 = (\frac{1}{2}, \frac{1}{2}, 0)$, $\mu_1 = (0, 0, 1)$, and $\mu_2 = (0, 1, 0)$. Observe that σ has full support, so no state is \succsim^σ -null. Furthermore, σ is not an LCPS.*

The event $E = \{\omega_0, \omega_1\}$ is not BFK-assumed under \succsim^σ . Given acts $x = (0, 1, 0)$ and $y = (1, 0, 1)$, $x \succ_E^\sigma y$ and $y \succ^\sigma x$. This contradicts Strict Determination.

Nevertheless, Conditions (i) and (iii) do hold for $j = 0$, because $\mu_0(E) = 1$ and $\text{supp } \mu_0 = E$. Of course, Condition (ii) must fail for $j = 0$, and indeed it does: $\mu_2(E) > 0$.

Notice that, if Ω is finite, σ has full support, and $E \subsetneq \Omega$, then Condition (i) can only hold for some $j < n - 1$. (If $j = n - 1$ then, by (i), $\Omega = \bigcup_{i=0}^{n-1} \text{supp } \mu_i \subseteq E$, a contradiction.) So, if Conditions (i)—(iii) are to hold for *some* j , it must be the case that $j < n - 1$. However, we now provide examples where BFK-assumption holds but Condition (ii) *only* holds for

$j = n - 1$.¹⁰ So, as discussed, to obtain a characterization of BFK-assumption for all LPSs, we will relax Condition (ii).

Example 3.2 Take $\Omega = \{\omega_0, \omega_1\}$ and $E = \{\omega_0\}$. Consider an LPS $\sigma = (\mu_0, \mu_1)$ with $\mu_0 = (1, 0)$ and $\mu_1 = (\frac{1}{2}, \frac{1}{2})$. If $x \succ_{\{\omega_0\}}^\sigma y$, then $x(\omega_0) > y(\omega_0)$, and so $x \succ^\sigma y$. It follows that E is BFK-assumed under \succ^σ . Note that Condition (i) holds only for $j = 0$, Condition (ii) holds only for $j = 1$, and Condition (iii) holds for both $j = 0, 1$. Thus, there is no single j for which all three conditions hold. So one direction of BFK's characterization fails. ■

Observation 3.1 There is a natural generalization of Example 3.2, which is essentially the proof of Lemma 5.1. As above, take $\Omega = \{\omega_0, \omega_1\}$ and $E = \{\omega_0\}$. Then, for any LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ on Ω , if the event E is BFK-assumed under \succ^σ , then $\mu_0(E) = 1$. To prove this fact, suppose that E is BFK-assumed under \succ^σ , but $\mu_0(E) < 1$. By Non-Triviality, $E = \{\omega_0\}$ is not \succ^σ -null, so by Remark 2.1, $\mu_i(E) > 0$ for some $i = 0, \dots, n - 1$. Let $x = (x_0, 0)$, with $x_0 > 0$; let $y = (0, 1)$. Then $x \succ_E^\sigma y$, but, since $\mu_0(\{\omega_1\}) > 0$, for x_0 sufficiently small, $y \succ^\sigma x$. This yields a contradiction.

To sum up, if E is BFK-assumed under \succ^σ , Conditions (i) and (iii) hold for $j = 0$. However, Example 3.2 illustrated that Condition (ii) may fail to hold for $j = 0$; indeed it may only hold—and trivially hold—for $j = n - 1$. ■

In Example 3.2, BFK-assumption implies that Conditions (i) and (iii) hold for $j = 0$. The next example illustrates that, with more than two states, Conditions (i) and (iii) need not hold for $j = 0$. However, these conditions will hold for some $j > 0$.

Example 3.3 Let $\Omega = \{\omega_0, \omega_1, \omega_2\}$ and $E = \{\omega_0, \omega_1\}$. Consider the LPS $\sigma = (\mu_0, \mu_1, \mu_2)$ such that $\mu_0 = (1, 0, 0)$, $\mu_1 = (\frac{1}{2}, \frac{1}{2}, 0)$ and $\mu_2 = (0, \frac{1}{2}, \frac{1}{2})$. We claim that \succ^σ BFK-assumes E . Non-Triviality holds because Ω is finite and σ has full support. For Strict Determination, suppose $x \succ_E^\sigma y$. It must be the case that $x(\omega_0) \geq y(\omega_0)$; if not, $y \succ_E^\sigma x$. Hence there are two possibilities: either (i) $x(\omega_0) > y(\omega_0)$, or (ii) $x(\omega_0) = y(\omega_0)$ and $x(\omega_1) > y(\omega_1)$; if not, $y \succ_E^\sigma x$. In either case, $x \succ^\sigma y$, so \succ^σ BFK-assumes E . Notice that Conditions (i) and (iii) do not hold for $j = 0$, but do hold for $j = 1$. On the other hand, Condition (ii) holds only for $j = 2$. ■

We now extend this example to illustrate that BFK-assumption implies that (i) and (iii) must hold for some j .

¹⁰Regardless of the event E , conditions (ii) and (iii) always hold trivially for $j = n - 1$. For any non-null E , (ii) holds as well.

Observation 3.2 Consider the following generalization of Example 3.3. Again, $\Omega = \{\omega_0, \omega_1, \omega_2\}$ and $E = \{\omega_0, \omega_1\}$. We show that, for any full-support LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ on Ω , if the event E is BFK-assumed under \succsim^σ , then Conditions (i) and (iii) hold for some j .

To prove this claim, note first that the argument given in Observation 3.1 implies that $\mu_0(E) = 1$. Furthermore, since E is BFK-assumed, Non-Triviality holds and implies that $\mu_i(\{\omega_0\}) > 0$ for some $i = 0, \dots, n-1$, and similarly $\mu_\ell(\{\omega_1\}) > 0$ for some (possibly different) ℓ . We now argue that, if the support of μ_0 does not contain E —i.e., Condition (iii) fails for $j = 0$ —then $\mu_1(E) = 1$. For simplicity, let $\mu_0(\{\omega_0\}) = 1$. Now suppose that $\mu_1(E) < 1$. Consider acts x, y such that $x(\omega_0) = y(\omega_0) = 0$, $x(\omega_1) > 0 = y(\omega_1)$, and $x(\omega_2) = 0 < 1 = y(\omega_2)$. Then $x \succsim_E^\sigma y$; however, $\int_\Omega x d\mu_0 = x(\omega_0) = y(\omega_0) = \int_\Omega y d\mu_0$ but, for $x(\omega_1)$ sufficiently small, $\int_\Omega x d\mu_1 < \int_\Omega y d\mu_1$, and so $y \succ^\sigma x$. This violates Strict Determination. Hence, $\mu_1(E) = 1$. If $\mu_1(\{\omega_1\}) > 0$, then E is contained in the union of the supports of μ_0 and μ_1 , so Conditions (i) and (iii) both hold for $j = 1$. Otherwise, we can repeat the argument to conclude that $\mu_2(E) = 1$. And so on. Since, as noted, Non-Triviality implies that $\mu_\ell(\{\omega_1\}) > 0$ for some ℓ , we will eventually reach a j such that E is contained in $\bigcup_{i=0}^j \text{supp } \mu_i$. For this j , Conditions (i) and (iii) both hold. ■

To sum up, in Example 3.2, an event was assumed, but Condition (ii) failed. On the other hand, Example 3.1 illustrates that Condition (ii) cannot simply be dropped. This leads us to weaken Condition (ii) in order to characterize BFK-assumption.

3.3 LPS-based characterization

We provide our characterization of BFK-assumption for general LPSs. (The characterization refers to an LPS σ and not the preference relation \succsim^σ .)

Definition 3.2 Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. Say a set $E \subseteq \Omega$ is **BFK-assumed under σ at level j** if E is Borel and

$$(i) \quad \mu_i(E) = 1 \text{ for all } i \leq j,$$

$$(ii^*) \quad \text{for each } k > j, \text{ there exists } (\alpha_0^k, \dots, \alpha_j^k) \in \mathbb{R}^{j+1} \text{ so that, for each Borel } F \subseteq E, \\ \mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F),$$

$$(iii) \quad E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i.$$

Say a set $E \subseteq \Omega$ is **BFK-assumed under σ** if it is BFK-assumed under σ at some level j .

Conditions (i) and (iii) in Definition 3.2 are Conditions (i) and (iii) in Theorem 3.1; Condition (ii*) is the required weakening of Condition (ii) therein.

Observe that Condition (ii) in Theorem 3.1 implies Condition (ii*) by taking $(\alpha_0^k, \dots, \alpha_j^k) = (0, \dots, 0)$. Second, when σ is an LCPS, Condition (ii*) implies (ii); hence, for LCPSs, these conditions are equivalent. To see this, suppose that (ii) fails. Then there is $k > j$ such that $\mu_k(E) > 0$. Since σ is an LCPS, there is a Borel U_k such that $\mu_k(U_k) = 1$ and $\mu_i(U_k) = 0$ for all $i \neq k$. Therefore, $F \equiv E \cap U_k \subseteq E$ is Borel and $\mu_k(F) = \mu_k(E) > 0$. Moreover, $\mu_i(F) \leq \mu_i(U_k) = 0$ for all $i \neq k$, so (ii*) fails.

Theorem 3.2 *Fix an LPS σ . A set $E \subseteq \Omega$ is BFK-assumed under \succsim^σ if and only if it is BFK-assumed under σ .*

To understand the intuition, it will be useful to recall the following fact from Blume, Brandenburger and Dekel (1991a, Theorem 3.1): Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. If μ_i is a linear combination of $(\mu_0, \dots, \mu_{i-1})$, then $\sigma' = (\mu_0, \dots, \mu_{i-1}, \mu_{i+1}, \dots, \mu_{n-1})$ represents the same lexicographic preferences, i.e., $\succsim^\sigma = \succsim^{\sigma'}$. That is, μ_i is irrelevant for determining preference.

A similar idea applies to conditional preferences. Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ and consider the conditional preference \succsim_E^σ . Recall, $x \succsim_E^\sigma y$ if and only if $(x_E, z_{\Omega \setminus E}) \succsim^\sigma (y_E, z_{\Omega \setminus E})$. Thus, if the restriction of μ_i to (Borel sets in) E is a linear combination of the restrictions of μ_0, \dots, μ_{i-1} to (Borel sets in) E , then the measure does not affect the ranking of $x \succsim_E^\sigma y$, i.e., $\succsim_E^\sigma = \succsim_E^{\sigma'}$ where $\sigma' = (\mu_0, \dots, \mu_{i-1}, \mu_{i+1}, \dots, \mu_{n-1})$. This is precisely the content of Condition (ii*). Specifically, Condition (ii*) requires that, when $i > j$, the measure μ_i is irrelevant for determining the conditional preference given E .

We next apply this result to Examples 3.1 and 3.3.

Example 3.4 (Example 3.1, Continued) *Here, $E = \{\omega_0, \omega_1\}$ is not BFK-assumed under \succsim^σ . Conditions (i) and (iii) hold only for $j = 0$. However, Condition (ii*) fails for $j = 0$: since $\mu_0(\omega_0) = \mu_0(\omega_1) = \frac{1}{2}$ and $\mu_1(\omega_0) = \mu_1(\omega_1) = 0$, there are no α_0^2, α_1^2 such that $\mu_2(\omega_0) = 0 = \alpha_0^2 \mu_0(\omega_0) + \alpha_1^2 \mu_1(\omega_0)$ and $\mu_2(\omega_1) = 1 = \alpha_0^2 \mu_0(\omega_1) + \alpha_1^2 \mu_1(\omega_1)$. ■*

Example 3.5 (Example 3.3, Continued) *Here, \succsim^σ BFK-assumes $E = \{\omega_0, \omega_1\}$ and Conditions (i) and (iii) hold (only) for $j = 1$. To see that Condition (ii*) also holds for $j = 1$, take $(\alpha_0^2, \alpha_1^2) = (-\frac{1}{2}, 1)$ and notice that $\mu_2(F) = -\frac{1}{2} \mu_0(F) + \mu_1(F)$ for all $F \subseteq \{\omega_0, \omega_1\}$.*

Notice that this example also demonstrates that, in Definition 3.2 we cannot replace linear combinations with convex combinations. There is no $\alpha_0^2, \alpha_1^2 \geq 0$, so that $\mu_2(F) = \alpha_0^2 \mu_0(F) + \alpha_1^2 \mu_1(F)$ for all $F \subseteq \{\omega_0, \omega_1\}$. ■

There is no redundancy in our characterization result. Example 3.4 illustrates that Conditions (i) and (iii) alone do not imply BFK-assumption. Conditions (i) and (ii*) alone are also not sufficient. BFK’s Supplemental Appendix S.1 shows that, when Ω is uncountable, a full-support LCPS σ may satisfy these conditions for some event E even though the preference relation \succsim^σ does not assume E . Obviously, (ii*) and (iii) alone do not imply BFK-assumption either: consider a one-level LPS with full support on a finite Ω and any $E \subsetneq \Omega$.

4 Alternative notions of assumption

4.1 Weak dominance on infinite state spaces

Asheim and Søvik (2005) restrict attention to full support LCPSs on a finite state space and provide an alternate preference-based characterization of assumption. Their alternate characterization is based on what Asheim and Dufwenberg (2003) term “full belief.” The basic idea is condition (**) in the Introduction: whenever an act x weakly dominates y on E , the player prefers x to y unconditionally. To adapt this approach to arbitrary (i.e., uncountable) state spaces, we must specify what it means for x to weakly dominate y on E . We consider two possible definitions.

Definition 4.1 *Say x **P-weakly dominates** y on E if*

- (a) $x(\omega) \geq y(\omega)$ for all $\omega \in E$, and
- (b) there exists some Borel $F \subseteq E$ not \succsim^σ -null so that $x(\omega) > y(\omega)$ for all $\omega \in F$.

Definition 4.2 *Say x **T-weakly dominates** y on E if*

- (a) $x(\omega) \geq y(\omega)$ for all $\omega \in E$, and
- (b) there exists some open set U such that $U \cap E \neq \emptyset$ and $x(\omega) > y(\omega)$ for all $\omega \in U \cap E$.

If x P-weakly dominates y on E , we write $x \text{PWD}_E^\sigma y$; if x T-weakly dominates y on E , we write $x \text{TWD}_E y$.

In both definitions, condition (b) requires strict preference on a subset of E that is “significant.” P-weak dominance employs a Preference-based notion of significance, while T-weak dominance instead invokes a Topological notion. These two definitions coincide with the usual notion of weak dominance on finite state spaces.

Observation 4.1 Let Ω be finite and consider $E \subseteq \Omega$ and acts $x, y \in \mathcal{A}$.

- x T-weakly dominates y on E if and only if x weakly dominates y on E .
- If the support of σ contains E (in particular, if σ has full support), then x P-weakly dominates y on E if and only if x weakly dominates y on E .

To see this, recall that, if the state space is finite, every point is open. This immediately gives the first statement. Moreover, if every point of E is in the support of σ , then every such point is not \succsim^σ -null, and the second statement follows.

Definitions 4.1 and 4.2 do not coincide, as the next example shows.

Example 4.1 Let $\Omega = [0, 1]$ and $E = [0, 1)$, and consider the LCPS $\sigma = (\mu_0, \mu_1, \mu_2)$, where μ_0 is the uniform measure on $[0, 1]$, μ_1 is the Dirac measure on $\omega = 1$, and μ_2 is the Dirac measure on $\omega = 0$.

Consider the acts $x, y \in \mathcal{A}$ such that $x(0) = 1$, $x(\omega) = 0$ for all $\omega \in (0, 1]$, $y(1) = 1$, and $y(\omega) = 0$ for all $\omega \in [0, 1)$. Then x P-weakly dominates y on E : the event $F = \{0\} \subset [0, 1) = E$ is not \succsim^σ -null and $x(0) > y(0)$. However, x does not T-weakly dominate y , because for any open set U such that $U \cap E \neq \emptyset$, there is $\omega \in U \cap E$ such that $x(\omega) = y(\omega) = 0$.

P-weak dominance has a natural behavioral interpretation: x must be strictly better than y on a set F that is *subjectively* meaningful to the player. By way of contrast, T-weak dominance requires that x be strictly better than y on a *topologically* non-trivial set. On the other hand, we shall see that T-weak dominance leads to a version of assumption that admits a simpler characterization in terms of LPSs.

4.2 Weak-dominance assumption

We can now extend [Asheim and Dufwenberg \(2003\)](#)'s full belief to uncountable state spaces.

Definition 4.3 Say a set E is **P-weak dominance assumed (PWD-assumed)** under \succsim^σ if E is Borel, and Non-Triviality as well as the following hold:

PWD Determination: for all $x, y \in \mathcal{A}$, $x \text{ PWD}_E^\sigma y$ implies $x \succ^\sigma y$.

Definition 4.4 Say a set E is **T-weak-dominance assumed (TWD-assumed)** under \succsim^σ if E is Borel, and Non-Triviality as well as the following hold:

TWD Determination: for all $x, y \in \mathcal{A}$, $x \text{ TWD}_E y$ implies $x \succ^\sigma y$.

We view BFK-assumption, PWD-assumption, and TWD-assumption as different manifestations of the same intuitive idea. With this in mind, we use the term assumption when we discuss properties that hold for, or are implied by all three notions. The remainder of this subsection discusses the precise relationship among these three notions. It is summarized in Table 1.

General Case		\Rightarrow (Obs. 4.2) $\not\Leftarrow$ (Ex. 4.2)		\Rightarrow (Obs. 4.3) $\not\Leftarrow$ (Ex. 4.3)	
Finite Ω	BFK- assumption	\Rightarrow (Obs. 4.2) $\not\Leftarrow$ (Ex. 4.2)	PWD- assumption	\Leftrightarrow (Obs. 4.1)	TWD- assumption
LCPSs		\Leftrightarrow (Prop. 4.1)		\Rightarrow (Obs. 4.3) $\not\Leftarrow$ (Ex. 4.3)	

Table 1: Different notions of assumption

Observation 4.2 If E is BFK-assumed, it is also PWD-assumed. This follows because, if $x \text{ PWD}_E^\sigma y$, then surely $x \succ_E^\sigma y$; hence, if Strict Determination holds for E , PWD Determination holds as well.

However, the converse need not hold, as the following example demonstrates.

Example 4.2 (Example 3.1, continued) Recall that the event E is not BFK-assumed under \succ^σ . However, E is PWD-assumed under \succ^σ . Non-triviality holds because Ω is finite and σ has full support. Suppose that $x \text{ PWD}_E^\sigma y$. Then $x(\omega_0) \geq y(\omega_0)$ and $x(\omega_1) \geq y(\omega_1)$, with at least one strict inequality. In either case $x \succ^\sigma y$. Thus, PWD Determination holds.

Despite this difference, BFK-assumption and PWD-assumption coincide for LCPSs

Proposition 4.1 Fix an LCPS σ . An event E is BFK-assumed under \succ^σ if and only if it is PWD-assumed under \succ^σ .

This result follows from Theorems 3.2 and 4.1. It extends a result by [Asheim and S¸ovik, 2005](#) from finite to infinite state spaces. Since BFK restrict attention to LCPSs, this equivalence implies that their epistemic analyses and results have an alternate preference-based interpretation that builds upon PWD-assumption.¹¹

Now turn to PWD- and TWD-assumption.

¹¹We thank an anonymous referee for proposing this conjecture.

Observation 4.3 If E is PWD-assumed, then it is TWD-assumed. To see this, note that, if Non-Triviality holds, then $x \text{TWD}_E y$ implies $x \text{PWD}_E^\sigma y$. (Let U be an open set such that $U \cap E \neq \emptyset$ and $x(\omega) > y(\omega)$ for all $\omega \in U \cap E$, and observe that, by Non-Triviality, $U \cap E$ is not \succsim^σ -null.) Therefore, PWD Determination implies TWD Determination under Non-Triviality.

However, the converse need not hold, even for LCPS beliefs.

Example 4.3 (Continuation of Example 4.1) Recall that σ is an LCPS. We saw that $x \text{PWD}_E^\sigma y$; however, $x \prec^\sigma y$, so PWD Determination fails. We now show that, nonetheless, E satisfies TWD Determination. Consider arbitrary acts $x, y \in \mathcal{A}$ such that $x \text{TWD}_E y$. Then there is an open U such that $U \cap E \neq \emptyset$ and $x(\omega) > y(\omega)$ for all $\omega \in U \cap E$. Since U is open and $E = [0, 1)$, there is $\epsilon > 0$ and an $\omega \in U \cap E$ such that $V \equiv (\omega, \omega + \epsilon) \subseteq U \cap E$. Since V is open and $V \cap E \neq \emptyset$, by Non-Triviality and Remark 2.1, $\mu_i(V) > 0$ for some $i = 0, 1, 2$; but since $V \subseteq (0, 1)$, $i = 0$. Finally, since $V \subseteq U \cap E$, $x(\omega) > y(\omega)$ for all $\omega \in V$. This implies that $\int_\Omega (x - y) d\mu_0 = \int_E (x - y) d\mu_0 \geq \int_V (x - y) d\mu_0 > 0$, so $x \succ^\sigma y$. Thus, TWD Determination holds.

Nonetheless, when the state space is finite, PWD-assumption and TWD-assumption coincide. To see this, recall from Observation 4.1 that T-weak dominance on E coincides with P-weak dominance on E , provided that the support of the LPS contains E . The key is that, when Non-Triviality holds for E , the support of the LPS must contain E . Thus, under Non-Triviality, PWD and TWD Determination coincide and so PWD- and TWD-assumption coincide. However, Example 4.2 shows that, even with a finite state space, BFK-assumption is stronger than both.

By Proposition 4.1, the same Example shows that, even for LCPSs, BFK-assumption and TWD-assumption need *not* coincide in uncountable state spaces.

4.3 LPS-based characterizations

We provide characterizations of PWD-assumption and TWD-assumption in terms of LPSs.

Definition 4.5 Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. A set $E \subseteq \Omega$ is **PWD-assumed under σ at level j** if it is Borel and

$$(i) \quad \mu_i(E) = 1 \text{ for all } i \leq j,$$

(ii**) for each $k > j$ and each Borel $F \subseteq E$, if $\mu_k(F) > 0$, then there exists $i \leq j$ with $\mu_i(F) > 0$.

(iii) $E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i$.

A set $E \subseteq \Omega$ is **PWD-assumed under** σ if it is PWD-assumed under σ at some level j .

Like Condition (ii*), Condition (ii**) is an irrelevance requirement (cf. p. 12). However, the former dictates that the measures μ_k , for $k > j$ do not affect conditional preferences over *any* pair of acts. On the other hand, Condition (ii**) is an irrelevance requirement that only directly applies to specific pairs of “betting” acts. In particular, fix some Borel $F \subseteq E$ and (winning-losing) prizes $w, \ell \in [0, 1]$ such that $w > \ell$. Consider the act $(\vec{w}_F, \vec{\ell}_{\Omega \setminus F})$ and the constant act $\vec{\ell}$. Condition (ii**) requires that the ranking of these two acts is completely determined by the measures μ_0, \dots, μ_j . To see this, suppose first that $\mu_i(F) > 0$ for some $i \leq j$: then $(\vec{w}_F, \vec{\ell}_{\Omega \setminus F}) \succ^\sigma \vec{\ell}$. Suppose that instead $\mu_i(F) = 0$ for all $i \leq j$: then the Condition requires that also $\mu_k(F) = 0$ for $k > j$, and so $(\vec{w}_F, \vec{\ell}_{\Omega \setminus F}) \sim^\sigma \vec{\ell}$. Therefore, the measures $\mu_{j+1}, \dots, \mu_{n-1}$ are irrelevant for determining the ranking of $(\vec{w}_F, \vec{\ell}_{\Omega \setminus F})$ vs. $\vec{\ell}$. However, they may well be relevant for the ranking of other acts.

Definition 4.6 Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. A set $E \subseteq \Omega$ is **TWD-assumed under** σ **at level** j if E is Borel and

(i) $\mu_i(E) = 1$ for all $i \leq j$,

(iii) $E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i$.

A set $E \subseteq \Omega$ is **TWD-assumed under** σ if it is TWD-assumed under σ at some level j .

Definition 4.6 does not have an analog to Conditions (ii), (ii*), or (ii**). However, Conditions (i) and (iii) imply the following property:

(ii***) for each $k > j$ and each open U , if $\mu_k(U \cap E) > 0$, then there exists $i \leq j$ with $\mu_i(U \cap E) > 0$.

Thus, the difference between TWD-assumption and PWD-assumption under σ hinges on whether the set $F \subseteq E$ is Borel or relatively open. Indeed, the interpretation of (ii**) as an irrelevance property applies to (ii***) as well, but it is restricted to sets F that are relatively open, rather than just Borel.

Condition (ii*) immediately implies Condition (ii**), which trivially implies Condition (ii***). Our characterization results, Theorems 3.2 and 4.1, together with Examples 3.2, 4.2 and 4.3, show that the converse implications do not hold. That said, when σ is an LCPS, Conditions (ii*) and (ii**) are equivalent, and stronger than (ii***). The equivalence is established in the proof of Proposition 4.1. The fact that (ii*) and (ii**) are stronger than

(ii***) follows from our characterization results and Example 4.3. Also, Conditions (ii**) and (ii***) coincide when the state space is finite. This follows from Theorem 4.1 because in this case TWD- and PWD-assumption coincide.

Theorem 4.1 *Fix an LPS σ .*

- (A) *A set $E \subseteq \Omega$ is PWD-assumed under \succsim^σ if and only if it is PWD-assumed under σ .*
- (B) *A set $E \subseteq \Omega$ is TWD-assumed under \succsim^σ if and only if it is TWD-assumed under σ .*

There is no redundancy in Theorem 4.1. In Example 4.1, the event E is not PWD-assumed, even though it is immediate to verify that Conditions (i) and (iii) hold for $j = 0$. Conditions (i) and (ii**) are also not sufficient for PWD-assumption. Since PWD- and BFK-assumption coincide for LCPSs, our discussion at the end of Subsection 3.3 applies.

5 Proof of Theorems 3.2 and 4.1, and Proposition 4.1

We first prove that the LPS-based definitions are sufficient: if E is BFK-assumed (resp. PWD-assumed, TWD-assumed) under σ , then it is BFK-assumed (resp. PWD-assumed, TWD-assumed) under \succsim^σ . We need two preliminary results.

Remark 5.1 *Fix some LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ and a Borel set E . There is some i with $\mu_i(E) > 0$ if and only if there are $x, y \in \mathcal{A}$ with $x \succ_E^\sigma y$.*

Remark 5.1 is a corollary of Remark 2.1.

Remark 5.2 *Fix some LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ and a Borel $E \subseteq \Omega$. Suppose that Conditions (i) and (iii) hold for some j . Then, for each open set U with $E \cap U \neq \emptyset$, $\mu_i(U) = \mu_i(E \cap U) > 0$ for some $i \leq j$.*

Proof. Fix some open set U with $E \cap U \neq \emptyset$. By Condition (iii), for each $\omega \in E \cap U$, there is some $i \leq j$ with $\omega \in \text{supp } \mu_i$. Since U is an open neighborhood of ω , $\mu_i(U) > 0$. By Condition (i), $\mu_i(E \cap U) = \mu_i(U) > 0$. ■

Proof of Theorems 3.2 and 4.1, Sufficiency. If E is TWD-assumed under σ at level j (a fortiori, if it is PDW-assumed or BFK-assumed), then Non-triviality holds. This follows from Remark 5.2 and Remark 5.1. We show that, if E is TWD-assumed (resp. PWD-assumed, BFK-assumed) under σ , then TWD Determination (resp. PWD Determination, Strict Determination) holds.

The arguments for TWD and PWD Determination are similar; we present them concurrently. Consider acts $x, y \in \mathcal{A}$ such that $x(\omega) \geq y(\omega)$ for all $\omega \in E$, and suppose further that there exists a nonempty set $F \subseteq E$ such that $x(\omega) > y(\omega)$ for all $\omega \in F$. If E is TWD-assumed (a fortiori, if it is PWD-assumed) under σ at level j , then Condition (i) holds, and implies that

$$\int x d\mu_i = \int_E x d\mu_i \geq \int_E y d\mu_i = \int y d\mu_i$$

for all $i \leq j$.

If E is TWD-assumed and $\emptyset \neq F = U \cap E$ for some open U , then by Remark 5.2, there exists some $i \leq j$ such that $\mu_i(U \cap E) > 0$. If instead E is PWD-assumed and F is Borel and not \succsim^σ -null, then by Remark 5.1 there exists some i such that $\mu_i(F) > 0$; moreover, by (ii**), we can take $i \leq j$. Therefore, in either case, there is $i \leq j$ such that

$$\int x d\mu_i = \int_E x d\mu_i > \int_E y d\mu_i = \int y d\mu_i,$$

and so $x \succ^\sigma y$. This establishes TWD Determination and, respectively, PWD Determination.

Finally, suppose that E is BFK-assumed under σ . Consider acts $x, y \in \mathcal{A}$ such that $x \succ_E^\sigma y$. Then, there exists some $k = 0, \dots, n-1$ so that

(a) $\int_E (x - y) d\mu_i = 0$ for all $i \leq k-1$ and

(b) $\int_E (x - y) d\mu_k > 0$.

It suffices to show that $k \leq j$; if so, then by Condition (i), it follows that $x \succ^\sigma y$.

Suppose, contra hypothesis, $k > j$. Then, by Condition (ii*), there exists $(\alpha_0^k, \dots, \alpha_j^k) \in \mathbb{R}^{j+1}$ so that

$$\int_E (x - y) d\mu_k = \sum_{i=0}^j \alpha_i^k \int_E (x - y) d\mu_i = 0,$$

where the second equality follows from (a). But this contradicts (b). ■

We now turn to the proof of necessity: if E is TWD-assumed (resp. PWD-assumed, BFK-assumed) under \succsim^σ , then it is TWD-assumed (resp. PWD-assumed, BFK-assumed) under σ .

First, we show that, if E is TWD-assumed under \succsim^σ , Condition (i) must hold for $j = 0$. A fortiori, this conclusion holds if E is PWD-assumed or BFK-assumed under \succsim^σ . Second, we show that, if E is TWD-assumed, PWD-assumed, or BFK-assumed under \succsim^σ , then there exists a j such that Condition (i) holds, and in addition the measures $\mu_{j+1}, \dots, \mu_{n-1}$ are ‘redundant’ in the appropriate sense. (Refer to the discussion on pages 12 and 17.) Third,

we show that the weakest notion of redundancy—that implied by TWD-assumption—implies Condition (iii).

Several of these steps mirror the observations and examples above. The first is analogous to Observation 3.1.

Lemma 5.1 *Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. If E is TWD-assumed under \succsim^σ , then $\mu_0(E) = 1$.*

Proof. Suppose that E is TWD-assumed under \succsim^σ . By contradiction, suppose that $\mu_0(E) < 1$. Consider acts x and y so that $x(\omega) = \varepsilon \in (0, \mu_0(\Omega \setminus E))$ for all $\omega \in \Omega$, $y(\omega) = 0$ for all $\omega \in E$, and $y(\omega) = 1$ for all $\omega \notin E$. Since $x(\omega) > y(\omega)$ for all $\omega \in E = \Omega \cap E$, and Ω is open, $x \text{ TWD}_E y$. But,

$$\int_{\Omega} y d\mu_0 = \mu_0(\Omega \setminus E) > \varepsilon = \int_{\Omega} x d\mu_0,$$

contradicting TWD Determination. Thus, $\mu_0(E) = 1$. ■

The next two Lemmas state that assumption implies Condition (i) holds for some $j = 0, \dots, n-1$. Furthermore, the Lemmas show that, if an event E is TWD-assumed, PWD-assumed, or BFK-assumed under \succsim^σ , the measures $\mu_{j+1}, \dots, \mu_{n-1}$ satisfy suitable redundancy properties.

The cases of TWD-assumption and PWD-assumption can be handled concurrently.

Lemma 5.2 *Fix an LPS σ and an event E that is TWD-assumed or PWD-assumed under \succsim^σ . Then, there exists some j so that*

$$(i) \quad \mu_i(E) = 1 \text{ for all } i \leq j.$$

Furthermore, if E is TWD-assumed under \succsim^σ , then

$$(ii^{***}) \text{ for each } k > j \text{ and each open } U, \text{ if } \mu_k(U \cap E) > 0, \text{ then there exists } i \leq j \text{ so that } \mu_i(U \cap E) > 0;$$

and if E is PWD-assumed under \succsim^σ , then

$$(ii^{**}) \text{ for each } k > j \text{ and each } F \subseteq E \text{ Borel, if } \mu_k(F) > 0, \text{ then there exists } i \leq j \text{ so that } \mu_i(F) > 0.$$

Proof. By Lemma 5.1, Condition (i) holds for $j = 0$. Let j be the largest number satisfying Condition (i). We will show that, if j does not also satisfy Condition (ii^{***}) (resp. (ii^{**})), then E is not TWD-assumed (resp. PWD-assumed) under \succsim^σ .

If $j = n-1$, then Conditions (ii^{***}) and (ii^{**}) hold vacuously. Thus, suppose $j < n-1$. Since j is the largest number satisfying Condition (i), we have $\mu_{j+1}(\Omega \setminus E) > 0$. Fix a Borel

$F \subseteq E$ and consider acts x and y satisfying the following: $x(\omega) = \varepsilon \in (0, \mu_{j+1}(\Omega \setminus E))$ for $\omega \in F$, $x(\omega) = 0$ if $\omega \notin F$, $y(\omega) = 0$ if $\omega \in E$ and $y(\omega) = 1$ if $\omega \notin E$.

We now consider two cases. First, suppose that E is TWD-assumed, and there exists an open U such that $F = U \cap E$ and $\mu_k(U \cap E) > 0$ for some $k \geq j + 1$. Then $U \cap E \neq \emptyset$, so x TWD $_E y$, and TWD Determination implies that $x \succ^\sigma y$.

Second, suppose that E is PWD-assumed, and $\mu_k(F) > 0$ for some $k \geq j + 1$. By Remark 2.1, F is not \succsim^σ -null. Then x PWD $_E^\sigma y$ and so, by PWD Determination, $x \succ^\sigma y$.

To sum up, in either case $x \succsim^\sigma y$ and $\mu_k(F) > 0$ for some $k \geq j + 1$. Suppose that $\mu_i(F) = 0$ for all $i \leq j$. Then $\mu_i(E \setminus F) = 1$ for all $i \leq j$, so

$$\int_{\Omega} x d\mu_i = \int_E x d\mu_i = 0 = \int_E y d\mu_i = \int_{\Omega} y d\mu_i,$$

for all $i \leq j$. Moreover, since $x \succsim^\sigma y$,

$$\mu_{j+1}(F)\varepsilon = \int_{\Omega} x d\mu_{j+1} \geq \int_{\Omega} y d\mu_{j+1} = \mu_{j+1}(\Omega \setminus E).$$

But this contradicts the fact that $\mu_{j+1}(\Omega \setminus E) > \varepsilon > 0$ and $\mu_{j+1}(F) \leq 1$. Therefore, there must be $i \leq j$ such that $\mu_i(F) > 0$. Hence, if E is TWD-assumed (resp. PWD-assumed) under \succsim^σ , Condition (ii***) (resp. (ii**)) holds. ■

If U is open, then $U \cap E$ is Borel; thus, Condition (ii**) implies Condition (ii***).

Now consider BFK-assumption. We break up the argument into two Lemmas.

Lemma 5.3 *Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. If $E \subseteq \Omega$ is BFK-assumed under \succsim^σ , then there is some $j = 0, \dots, n - 1$ such that*

(i) $\mu_i(E) = 1$ for all $i \leq j$, and

(\diamond) if $\int_E (x - y) d\mu_i = 0$ for all $i \leq j$, then $\int_E (x - y) d\mu_i = 0$ for all $i = 0, \dots, n - 1$.

Proof. Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ on Ω such that E is BFK-assumed under \succsim^σ . By Lemma 5.1 and the fact that BFK-assumption implies TWD-assumption, $\mu_0(E) = 1$. Let $k = \max\{\ell = 0, \dots, n - 1 : \mu_i(E) = 1 \text{ for all } i \leq \ell\}$. If $k = n - 1$ then condition (\diamond) holds trivially. Suppose instead that $k < n - 1$ and condition (\diamond) fails. Then there are acts x, y and a number $l = k + 1, \dots, n - 1$ such that

- $\int_E (x - y) d\mu_i = 0$ for all $i \leq l - 1$, and
- $\int_E (x - y) d\mu_l > 0$.

We will use these acts and the fact that $\mu_{k+1}(E) < 1$, to construct acts \hat{x} and \hat{z} such that $\hat{x} \succ_E^\sigma \hat{z}$ and $\hat{z} \succ^\sigma \hat{x}$. This contradicts Strict Determination.

For each $\rho \in (0, 1)$, let $z[\rho]$ be the act with $z[\rho](\omega) = \rho x(\omega) + (1 - \rho)y(\omega)$ for all $\omega \in \Omega$. Note that for any $i = 0, \dots, n - 1$,

$$\int_E (x - z[\rho]) d\mu_i = (1 - \rho) \int_E (x - y) d\mu_i.$$

So, for each $\rho \in (0, 1)$, (i) $\int_E (x - z[\rho]) d\mu_i = 0$ for all $i \leq l - 1$, and (ii) $\int_E (x - z[\rho]) d\mu_l > 0$. It follows that, for each $\rho \in (0, 1)$, $x \succ_E^\sigma z[\rho]$.

Construct acts $\hat{x} = (x_E, \vec{0}_{\Omega \setminus E})$ and $\hat{z}[\rho] = (z[\rho]_E, \vec{1}_{\Omega \setminus E})$. Certainly, for each $\rho \in (0, 1)$, $\hat{x} \succ_E^\sigma \hat{z}[\rho]$. Moreover, since $\mu_i(E) = 1$ for all $i \leq k$, it follows that, for each $\rho \in (0, 1)$ and each $i \leq k$, $\int_\Omega (\hat{z}[\rho] - \hat{x}) d\mu_i = 0$. Next note that, for each $\rho \in (0, 1)$,

$$\int_\Omega (\hat{z}[\rho] - \hat{x}) d\mu_{k+1} = (1 - \rho) \int_E (y - x) d\mu_{k+1} + \mu_{k+1}(\Omega \setminus E).$$

Since $\mu_{k+1}(\Omega \setminus E) > 0$, there exists $\rho^* \in (0, 1)$ large enough so $\int_\Omega (\hat{z}[\rho^*] - \hat{x}) d\mu_{k+1} > 0$ and so $\hat{z}[\rho^*] \succ^\sigma \hat{x}$. ■

Lemma 5.4 Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. Suppose that $E \subseteq \Omega$ is Borel and, for some $j = 0, \dots, n - 1$,

$$(\diamond) \int_E (x - y) d\mu_i = 0 \text{ for all } i \leq j \implies \int_E (x - y) d\mu_i = 0 \text{ for all } i = 0, \dots, n - 1.$$

Then,

$$(ii^*) \text{ for each } k > j, \text{ there exists } (\alpha_0^k, \dots, \alpha_j^k) \in \mathbb{R}^{j+1} \text{ so that, for each Borel } F \subseteq E, \\ \mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F).$$

Proof. Take j so that condition (\diamond) holds. Fix some $k > j$. We will show that there exists $(\alpha_0^k, \dots, \alpha_j^k) \in \mathbb{R}^{j+1}$ so that, for any Borel $F \subseteq E$, $\mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F)$.

Let \mathcal{B} denote the vector space of bounded Borel-measurable functions $b : \Omega \rightarrow \mathbb{R}$. For each $i = 1, \dots, j, k$, define linear functionals T_1, \dots, T_j, T_k on \mathcal{B} by $T_i(b) = \int_E b d\mu_i$. By condition (\diamond) , if $x, y \in \mathcal{A}$ with $T_i(x - y) = 0$ for all $i \leq j$, then $T_k(x - y) = 0$. Now, note that \mathcal{B} is the set of all functions of the form $\gamma(x - y)$ for $\gamma \in \mathbb{R}_{++}$ and $x, y \in \mathcal{A}$. So, for each $b \in \mathcal{B}$, $T_i(b) = 0$ for all $i \leq j$ implies that $T_k(b) = 0$. Hence, by the Theorem of the Alternative (see [Aliprantis and Border, 2007](#), Corollary 5.92), there exists $(\alpha_0^k, \dots, \alpha_j^k) \in \mathbb{R}^{j+1}$ with $T_k = \sum_{i=0}^j \alpha_i^k T_i$.

For any $F \subseteq E$ Borel, it follows that

$$\mu_k(F) = \int_E (\vec{1}_F, \vec{0}_{\Omega \setminus F}) d\mu_k = \sum_{i=0}^j \alpha_i^k \int_E (\vec{1}_F, \vec{0}_{\Omega \setminus F}) d\mu_i = \sum_{i=0}^j \alpha_i^k \mu_i(F),$$

as desired. ■

Finally, we show that, under Non-Triviality, Condition (ii^{***}) implies Condition (iii) in Definition 4.6.

Lemma 5.5 *Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ and a Borel $E \subseteq \Omega$ for which Non-Triviality holds under \succsim^σ . If Condition (ii^{***}) holds for some j , then $E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i$, i.e., Condition (iii) holds.*

Proof. Let $U = \Omega \setminus (\bigcup_{i \leq j} \text{supp } \mu_i)$ and observe that U is open. Suppose, contra hypothesis, that $E \cap U \neq \emptyset$. Then, by Non-Triviality, $E \cap U$ is not \succsim^σ -null. By Remark 2.1, there exists k such that $\mu_k(E \cap U) > 0$. By the definition of U , for any ℓ with $\mu_\ell(E \cap U) > 0$, $\ell > j$. So there exists $k > j$ with $\mu_k(E \cap U) > 0$. But, $\mu_i(E \cap U) = 0$ for all $i \leq j$, which contradicts the fact that Condition (ii^{***}) must hold. ■

Proof of Theorems 3.2 and 4.1, Necessity. Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ and a Borel $E \subseteq \Omega$. Lemmas 5.2 and 5.5 show that, if E is TWD-assumed under \succsim^σ , then it is TWD-assumed under σ . Having established necessity for TWD-assumption, we turn to PWD- and BFK-assumption. First recall (see page 17) that if (ii^{*}) or (ii^{**}) hold for some j , then (ii^{***}) holds for that j . Therefore Lemmas 5.2 and 5.5 also show that, if E is PWD-assumed under \succsim^σ , then it is PWD-assumed under σ . Finally, Lemmas 5.3, 5.4, 5.5 show that, if E is BFK-assumed under \succsim^σ , then it is BFK-assumed under σ . ■

Theorems 3.2 and 4.1 readily imply Proposition 4.1, i.e., the equivalence of BFK-assumption and PWD-assumption for LCPSs.

Proof of Proposition 4.1. By Theorems 3.2 and 4.1, it is enough to show that Conditions (ii^{**}) and (ii^{*}) are equivalent. We argued on page 17 that (ii) in Theorem 3.1 implies (ii^{*}), which implies (ii^{**}). Thus, it is enough to show that (ii^{**}) implies (ii). To see this, suppose that (ii) fails. Then there is $k > j$ such that $\mu_k(E) > 0$. Since σ is an LCPS, there is a Borel U_k such that $\mu_k(U_k) = 1$ and $\mu_i(U_k) = 0$ for all $i \neq k$. Therefore, $E \cap U_k$ is Borel and $\mu_k(E \cap U_k) > 0$. Moreover, $\mu_i(E \cap U_k) \leq \mu_i(U_k) = 0$ for all $i \neq k$, so (ii^{**}) fails. ■

6 Application: SAS and IA

This section applies the LPS-based characterizations of assumption to BFK's game-theoretic analysis. We consider type structures where types map to arbitrary LPSs, rather than LCPSs. We formalize (lexicographic) rationality, assumption of rationality, etc., for the

three variants of assumption. We show that, independent of the variant of assumption, self-admissible sets capture the behavioral implications of rationality and common assumption of rationality across all type structure. Moreover, independent of the variant of assumption, iterated admissibility captures the behavioral implications of rationality and m^{th} -order assumption of rationality, in a complete type structure.

As in BFK, we restrict attention to two-player games. Fix a game $\langle S_a, S_b, \pi_a, \pi_b \rangle$ where S_a (resp. S_b) is a finite strategy set for Ann (resp. Bob) and π_a (resp. π_b) is a payoff function.

6.1 Solution Concepts

The following definitions are standard.

Definition 6.1 Fix $Y_a \times Y_b \subseteq S_a \times S_b$. A strategy $s_a \in Y_a$ is **weakly dominated with respect to** $Y_a \times Y_b$ if there exists $\mu_a \in \mathcal{P}(S_a)$, with $\mu_a(Y_a) = 1$, such that $\pi_a(\mu_a, s_b) \geq \pi_a(s_a, s_b)$ for every $s_b \in Y_b$, and $\pi_a(\mu_a, s_b) > \pi_a(s_a, s_b)$ for some $s_b \in Y_b$. Otherwise, say s_a is **admissible with respect to** $Y_a \times Y_b$. If s_a is admissible with respect to $S_a \times S_b$, simply say that s_a is **admissible**.

Definition 6.2 Set $S_a^0 = S_a$ and $S_b^0 = S_b$. Define inductively

$$S_a^{m+1} = \{s_a \in S_a^m : s_a \text{ is admissible with respect to } S_a^m \times S_b^m\};$$

and, likewise, define S_b^{m+1} . A strategy $s_a \in S_a^m$ is called **m -admissible**. A strategy $s_a \in \bigcap_{m=0}^{\infty} S_a^m$ is called **iteratively admissible (IA)**.

The following definitions are due to BFK.

Definition 6.3 Say r_a **supports** s_a if there exists some $\mu_a \in \mathcal{P}(S_a)$ with $r_a \in \text{supp } \mu_a$ and $\pi_a(\mu_a, s_b) = \pi_a(s_a, s_b)$ for all $s_b \in S_b$. Write $\text{su}(s_a)$ for the set of $r_a \in S_a$ that support s_a .

Definition 6.4 Fix $Q_a \times Q_b \subseteq S_a \times S_b$. The set $Q_a \times Q_b$ is a **self-admissible set (SAS)** if:

- (a) each $s_a \in Q_a$ is admissible,
- (b) each $s_a \in Q_a$ is admissible with respect to $S_a \times Q_b$,
- (c) for any $s_a \in Q_a$, if $r_a \in \text{su}(s_a)$ then $r_a \in Q_a$,

and likewise for each $s_b \in Q_b$.

6.2 Epistemic Analysis

For each $n \in \mathbb{N}$, write $\mathcal{N}_n(\Omega)$ for the set of LPSs of length n , $\sigma = (\mu_0, \dots, \mu_{n-1})$, and write $\mathcal{N}(\Omega) = \bigcup_{n \in \mathbb{N}} \mathcal{N}_n(\Omega)$ for the set of LPS. Write $\mathcal{N}^+(\Omega)$ for the set of $\sigma \in \mathcal{N}$ that have full support. Define a metric on $\mathcal{N}(\Omega)$ as follows: The distance between two sequences of measures $(\mu_0, \dots, \mu_{n-1})$ and $(\nu_0, \dots, \nu_{n-1})$ of the same length is the maximum of the Prohorov distances between μ_i and ν_i for all $i < n$. The distance between two sequences of measures of different lengths is 1. With this, $\mathcal{N}(\Omega)$ is a Polish space and, by Corollary C.1 in BFK, $\mathcal{N}^+(\Omega)$ is Borel.

Definition 6.5 An (S_a, S_b) -based type structure is a structure

$$\langle S_a, S_b, T_a, T_b, \lambda_a, \lambda_b \rangle,$$

where T_a and T_b are nonempty Polish **type spaces**, and $\lambda_a : T_a \rightarrow \mathcal{N}(S_b \times T_b)$ and $\lambda_b : T_b \rightarrow \mathcal{N}(S_a \times T_a)$ are Borel measurable **belief maps**.

Type structures are a basic representation of interactive LPS-based beliefs. Definition 6.5 differs from BFK's Definition 7.1 in that it does not require that types be mapped to LCPSs (or limits of LCPSs). A type structure induces a set of **states**, i.e., $S_a \times T_a \times S_b \times T_b$.

In the remainder of this subsection, we fix a (S_a, S_b) -based type structure $\langle S_a, S_b, T_a, T_b, \lambda_a, \lambda_b \rangle$. All definitions have counterparts with a and b reversed.

Definition 6.6 A strategy s_a is **optimal** under $\sigma = (\mu_0, \dots, \mu_{n-1})$ if $\sigma \in \mathcal{N}(S_b \times T_b)$ and

$$\left(\pi_a(s_a, \text{marg}_{S_b} \mu_i(s_b)) \right)_{i=0}^{n-1} \geq^L \left(\pi_a(r_a, \text{marg}_{S_b} \mu_i(s_b)) \right)_{i=0}^{n-1}$$

for all $r_a \in S_a$.

Here, $\text{marg}_{S_b} \mu_i$ denotes the marginal on S_b of the measure μ_i . In words, Ann will prefer strategy s_a to strategy r_a if the associated sequence of expected payoffs under s_a is lexicographically greater than the sequence under r_a . (If σ is a length-one LPS (μ_0) , we will sometimes say that s_a is optimal under the measure μ_0 if it is optimal under (μ_0) .)

We now formalize the epistemic conditions of interest as restrictions on strategy-type pairs.

Definition 6.7 A strategy-type pair $(s_a, t_a) \in S_a \times T_a$ is **rational** if $\lambda_a(t_a)$ is a full-support LPS and s_a is optimal under $\lambda_a(t_a)$.

Next, for $E \subseteq S_b \times T_b$, set

$$A_a^X(E) = \{t_a \in T_a : E \text{ is } X\text{-assumed under } \lambda_a(t_a)\}, \quad \text{for } X \in \{\text{BFK, PWD, TWD}\}.$$

In words, $A_a^{\text{BFK}}(E)$ is the set of types $t_a \in T_a$ such that the associated LPSs $\lambda_a(t_a)$ BFK-assumes the event $E \subseteq S_b \times T_b$. Likewise, $A_a^{\text{PWD}}(E)$ and $A_a^{\text{TWD}}(E)$ are the sets of types that PWD-assume and, respectively, TWD-assume E . We frequently refer to properties that hold for several variants of assumption. For brevity, we use X to denote these variants, as in the equation above. Henceforth, we drop explicit reference to the fact that X is an element of the set $\{\text{BFK, PWD, TWD}\}$; this fact should be taken as implicit (as in the next observation). Note, if $E \subseteq S_b \times T_b$ is not Borel, then $A_a^X(E) = \emptyset$

For finite m and any X , define the sets $R_a^{X,m}$ as follows. Let $R_a^{X,1}$ be the set of all rational $(s_a, t_a) \in S_a \times T_a$. Inductively, set

$$R_a^{X,m+1} = R_a^{X,m} \cap [S_a \times A_a^X(R_b^{X,m})].$$

If $(s_a, t_a, s_b, t_b) \in R_a^{X,m+1} \times R_b^{X,m+1}$, say there is **rationality and m th-order X -assumption of rationality** at this state. If $(s_a, t_a, s_b, t_b) \in \bigcap_{m=1}^{\infty} R_a^{X,m} \times \bigcap_{m=1}^{\infty} R_b^{X,m}$, say there is **rationality and common X -assumption of rationality (RCA^XR)** at this state.

Recall that BFK-assumption implies PWD-assumption which, in turn, implies TWD-assumption. In light of this, one might conjecture that, for any given type structure and any given m , $R_a^{\text{BFK},m} \times R_b^{\text{BFK},m} \subseteq R_a^{\text{PWD},m} \times R_b^{\text{PWD},m} \subseteq R_a^{\text{TWD},m} \times R_b^{\text{TWD},m}$. (And, if so, this would allow us to simplify proofs, taking BFK-assumption as a lower bound on behavior and TWD-assumption as an upper bound of behavior.) However, this is not the case because assumption—in all its forms—is not monotonic: we can have $E \subseteq F$, and E assumed, even though F is not assumed (see BFK, p. 323). Nonetheless, we will show two behavioral equivalence results; so, at some level, the differences between these variants of assumption will not be material for observed behavior.

Because there is no ranking of rationality and m^{th} -order assumption of rationality across the variants of assumption, we will need to establish the results separately, for each variant of assumption. However, the arguments take a similar structure to one another and all follow the line of argument in BFK (Theorems 8.1–10.1). In the Appendix, we discuss the required modifications.

6.3 Self-admissible sets

We begin by showing that SASs characterize RCA^XR , independent of the choice of $X \in \{\text{BFK}, \text{PWD}, \text{TWD}\}$.

Theorem 6.1

- (1) For every type structure, $\text{proj}_{S_a} \bigcap_m R_a^{X,m} \times \text{proj}_{S_b} \bigcap_m R_b^{X,m}$ is an SAS.
- (2) For every SAS $Q_a \times Q_b$, there exists a type structure such that $\text{proj}_{S_a} \bigcap_m R_a^{X,m} \times \text{proj}_{S_b} \bigcap_m R_b^{X,m} = Q_a \times Q_b$.

Part (2) is essentially Theorem 8.1(ii) in BFK. They show this by constructing a finite, LCPS-based type structure which is, a fortiori, an LPS-based type structure. In such a type structure, all three notions of assumption coincide.¹² Part (1) is an analogue of Theorem 8.1(i) in BFK. In contrast to BFK, we allow for arbitrary LPS-based type structures and employ our characterizations of assumption. As noted, in the Appendix we indicate how to adapt BFK’s proofs.

Within a given type structure, RCA^XR may have different—and potentially incompatible—behavioral implications for different admissibility concepts. That is, within a given type structure, the sets $\text{proj}_{S_a} \bigcap_m R_a^{X,m} \times \text{proj}_{S_b} \bigcap_m R_b^{X,m}$ may be disjoint for any pair of distinct $X \in \{\text{BFK}, \text{PWD}, \text{TWD}\}$. Nevertheless, Theorem 6.1 states that, if we quantify across all type structures, then RCA^XR has the same behavioral implications for all X . For instance, suppose that for a given type structure, $Q_a \times Q_b \subseteq S_a \times S_b$ is the projection on the strategy set of $\text{RCA}^{\text{BFK}}\text{R}$; then Theorem 6.1 says that, there exists a (potentially different) type structure so that $Q_a \times Q_b$ is also the projection of $\text{RCA}^{\text{TWD}}\text{R}$ in the different type structure. Thus, if the analyst can only observe behavior, then $\text{RCA}^{\text{BFK}}\text{R}$, $\text{RCA}^{\text{PWD}}\text{R}$, and $\text{RCA}^{\text{TWD}}\text{R}$ are indistinguishable. However, if the analyst also has information about both behavior and hierarchies of beliefs, then the three epistemic conditions are distinguishable.

6.4 Iterated Admissibility

BFK’s foundations for iterated admissibility focus on type structures that satisfy a particular property, known as completeness. Write $\text{range } \lambda_a$ for the range of the function λ_a .

Definition 6.8 A type structure is **complete** if $\mathcal{N}^+(S_b \times T_b) \subsetneq \text{range } \lambda_a$ and $\mathcal{N}^+(S_a \times T_a) \subsetneq \text{range } \lambda_b$.

¹²This is implied by Observation 4.1 and Proposition 4.1.

A complete type structure is one that is sufficiently rich: For every possible full-support LPS-based belief a player can hold, there is a type of the player that holds that belief. Moreover, there is (at least) some type of the player that has an LPS-based belief without full support.¹³ Section 2.4 in BFK illustrates why, from the perspective of providing foundations for iterated admissibility, it is important to include types without full support (i.e., why it is important to require that λ_a and λ_b be a strict superset of $\mathcal{N}^+(S_b \times T_b)$ and $\mathcal{N}^+(S_a \times T_a)$).

Again, within a given complete type structure, the sets $R_a^{X,m}$ may well be different for different values of X . Nonetheless, Theorem 6.2 shows that they all characterize m rounds of iterated admissibility.

Theorem 6.2 *Fix a complete type structure. For each m ,*

$$\text{proj}_{S_a} R_a^{X,m} \times \text{proj}_{S_b} R_b^{X,m} = S_a^m \times S_b^m.$$

Theorem 6.2 is an analogue of Theorem 9.1 in BFK. It states that, in a complete type structure, the strategies consistent with rationality and m th order BFK-assumption (resp. PWD-assumption, TWD-assumption) of rationality are precisely the ones that survive m rounds of iterated admissibility. Unlike BFK's result, Theorem 6.2 allows for an LPS-based notion of a complete type structure. It employs our characterizations of assumption to prove this result.

Finally, BFK show a negative result on the impossibility of RCAR in a complete type structure. Again, an analogous result holds in our setting. Say that player a is *not indifferent* if there exist $s_a, r_a \in S_a$ and $s_b \in S_b$ such that $\pi_a(r_a, s_b) \neq \pi_a(s_a, s_b)$.

Theorem 6.3 *Fix a complete type structure $(S_a, S_b, T_a, T_b, \lambda_a, \lambda_b)$ where λ_a and λ_b are continuous. If player a is not indifferent then, for any X , there is no state at which there is $RCA^X R$.*

Theorem 6.3 is an analogue of Theorem 10.1 in BFK. The result shows that BFK's impossibility of RCAR does not hinge on mutual singularity, or the choice between BFK-assumption, PWD-assumption, or TWD-assumption.

7 Discussion: Related Literature

In contemporaneous work, Lee (2013) extends the results of BFK to LPSs. His elegant approach is different from but complementary to ours. His starting point is that the same

¹³A type structure that is complete in the sense of BFK is complete according to this definition; the converse does not hold.

lexicographic preference relation may be represented by more than one LPS. (See [Blume, Brandenburger and Dekel, 1991a](#), page 66). He shows that a lexicographic preference relation \succsim assumes an event E if and only if Conditions (i)–(iii) in Theorem 3.1 hold for *some* LPS σ for which $\succsim^\sigma = \succsim$. That is, instead of providing conditions that a *given* LPS must satisfy for the corresponding preference relation to assume an event E , he provides conditions that must be satisfied by at least one of the many LPSs that represent the same preferences.

Theorem 7.1 ([Lee, 2013](#)) *Fix an LPS σ . A set $E \subseteq \Omega$ is BFK-assumed under \succsim^σ if and only if there is some LPS ρ satisfying Conditions (i)–(iii) in Theorem 3.1 such that $\succsim^\sigma = \succsim^\rho$.*

Lee’s result can also be derived from our Theorem 3.2. In fact, [Lee \(2015b\)](#) gives a self-contained proof using our characterization of assumption.

Appendix A Theorems 6.1-6.3

The proofs of Theorems 8.1 and 9.1 in BFK rely on three results concerning the properties that hold in LCPS-based type structures. (See Lemma D.1, Property 6.3, and Lemma C.4 in BFK.) The statement and proofs of these results rely on BFK’s characterization of assumption for LCPSs. (Proposition 4.1 implies that the same characterization holds for PWD-assumption.) We have seen that the characterization does not apply to BFK-assumption and PWD-assumption with arbitrary LPSs, and to TWD-assumption. To address this, we state and prove analogous properties in our setting. (See Lemmas A.1, A.2, and A.3.)

Lemma A.1 *Let $\lambda_a(t_a) = (\mu_0, \dots, \mu_{n-1})$ be a full-support LPS. Suppose t_a TWD-assumes $E \subseteq S_b \times T_b$. Then, there exist some j so that*

$$\bigcup_{i \leq j} \text{supp marg}_{S_b} \mu_i = \text{proj}_{S_b} E.$$

Note that if $S_b \times T_b$ is finite and $\lambda_a(t_a) = (\mu_0, \dots, \mu_{n-1})$ TWD-assumes (a fortiori, PWD-assumes or BFK-assumes) E at level j , then $E = \bigcup_{i \leq j} \text{supp } \mu_i$. If $S_b \times T_b$ is infinite, the same may not hold. Lemma A.1 shows that, if $\Omega = S_b \times T_b$ is infinite, a similar statement holds if we consider the marginal LPS $(\text{marg}_{S_b} \mu_0, \dots, \text{marg}_{S_b} \mu_{n-1})$ and, correspondingly, the projection of E on S_b .

Lemma A.2 *Fix a full-support LPS $\sigma \in \mathcal{N}^+(S_b \times T_b)$. If σ X-assumes $R_b^{X,1}, R_b^{X,2}, \dots$, then it X-assumes $\bigcap_m R_b^{X,m}$.*

Lemma A.2 will be a consequence of a conjunction property of X-assumption .

Lemma A.3 *The sets $R_a^{X,m}$ and $R_b^{X,m}$ are Borel.*

To prove part (1) of Theorem 6.1, it is enough to replace Lemma D.1 and Property 6.3 in BFK's proof with Lemma A.1 and Lemma A.2. To prove part (2) simply repeat BFK's proof, observing that BFK-assumption implies both PWD-assumption and TWD-assumption. (Observations 4.2-4.3.)

To prove Theorem 6.2, two changes to BFK's proof are needed. First, replace Lemma D.1 and Lemma C.4 in BFK with Lemma A.1 and Lemma A.3. Second, modify the proof of Lemma E.3 in BFK, for the case where $m \geq 2$: Skip the construction that ensures that $\mu_i(U) = 0$ for all i . (That particular construction does not work for arbitrary LPSs. Fortunately, it is not needed in our setting.)

We now prove Lemmas A.1, A.2, and A.3.

Proof of Lemma A.1. Suppose t_a TWD-assumes $E \subseteq S_b \times T_b$ at level j . If $s_b \in \text{proj}_{S_b} E$, then there exists $i \leq j$ such that $\mu_i(\{s_b\} \times T_b) > 0$. (See Remark 5.2.) It follows that, if $s_b \in \text{proj}_{S_b} E$, $s_b \in \text{supp marg}_{S_b} \mu_i$. Conversely, if $s_b \notin \text{proj}_{S_b} E$, then $E \cap (\{s_b\} \times T_b) = \emptyset$. Since each $\mu_i(E) = 1$ for $i \leq j$, it follows that $\mu_i(\{s_b\} \times T_b) = 0$ for $i \leq j$, i.e., $s_b \notin \bigcup_{i \leq j} \text{supp marg}_{S_b} \mu_i$. ■

Lemma A.4 *Fix Borel sets E_1, E_2, \dots , with $E_{m+1} \subseteq E_m$. If a full-support LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$ X-assumes each of E_1, E_2, \dots , then it X-assumes $\bigcap_m E_m$.*

Proof. For each $m = 1, 2, \dots$, there exists some $j[m] \in \{0, \dots, n-1\}$ so that σ X-assumes E_m at level $j[m]$. Let $j = \min\{j[m] : m \geq 0\}$. Let M be some m with $j = j[M]$. We show that $\bigcap_m E_m$ is X-assumed under σ at level $j = j[M]$.

For Condition (i), note that, for each $i \leq j$, $\mu_i(E_m) = 1$ for all m . So, by continuity, $\mu_i(\bigcap_m E_m) = 1$. For Condition (iii), note that $\bigcap_m E_m \subseteq E_M \subseteq \bigcup_{i \leq j} \text{supp } \mu_i$. Hence, $\bigcap_m E_m$ is TWD-assumed at level j . Now suppose that $X \in \{\text{BFK}, \text{PWD}\}$. For Condition (ii*) (resp. (ii**)), note that each Borel $F \subseteq \bigcap_m E_m$ is also a subset of E_M . Thus, Condition (ii*) (resp. (ii**)) applied to $\bigcap_m E_m$ follows from Condition (ii*) (resp. (ii**)) applied to E_M . ■

Proof of Lemma A.2. Immediate from Lemma A.4. ■

We now turn to the proof of Lemma A.3. We will break the proof into several Lemmas. The first Lemma is standard (and so the proof is omitted).

Lemma A.5 *Fix some strategy $s_a \in S_a$.*

(1) *The set of $\mu \in \mathcal{P}(S_b)$ so that s_a is optimal under μ is closed.*

(2) The set of $\mu \in \mathcal{P}(S_b)$ so that s_a is strictly optimal under μ is open.

Lemma A.6 The sets $R_a^{X,1}$ and $R_b^{X,1}$ are Borel.

Proof. For each $s_a \in S_a$, define $\mathbb{O}[s_a, n]$ to be

$$\mathbb{O}[s_a, n] = \{\sigma \in \mathcal{N}_n(S_b \times T_b) : s_a \text{ is optimal under } \sigma\}.$$

Note, that

$$R_a^{X,1} = \bigcup_{s_a \in S_a} \bigcup_{n \in \mathbb{N}^0} [\{s_a\} \times ((\lambda_a)^{-1}(\mathbb{O}[s_a, n]) \cap (\lambda_a)^{-1}(\mathcal{N}_n^+(S_b \times T_b)))] .$$

Since λ_a is measurable and $\mathcal{N}_n^+(S_b \times T_b)$ is Borel (Corollary C.1 in BFK), it suffices to show that each $\mathbb{O}[s_a, n]$ is measurable.

Write $O[s_a]$ for the set of $\mu \in \mathcal{P}(S_b)$ under which s_a is optimal, $O^s[s_a]$ for the set of $\mu \in \mathcal{P}(S_b)$ under which s_a is strictly optimal, and $O^w[s_a] = O[s_a] \setminus O^s[s_a]$. By Lemma A.5, $O^w[s_a]$, $O^s[s_a]$, and $O[s_a]$ are Borel. Note that

$$\mathbb{O}[s_a, n] = (O^s[s_a] \times \mathcal{N}_{n-1}(S_b \times T_b)) \cup (O^w[s_a] \times O^s[s_a] \times \mathcal{N}_{n-2}(S_b \times T_b)) \cup \dots \cup (O^w[s_a] \times O^w[s_a] \times \dots \times O[s_a]),$$

so that $\mathbb{O}[s_a, n]$ is Borel. ■

Given a Borel set $E \subseteq \Omega$, write \mathcal{S}_E for the set of $F \subseteq E$ that are Borel. Of course, $\mathcal{S}_E \subseteq \mathcal{S}$. Moreover, \mathcal{S}_E is the Borel σ -algebra on E . (See [Aliprantis and Border, 2007](#), Lemma 4.20.)

Lemma A.7 Fix $n \in \mathbb{N}^0$ and $j = 0, \dots, n-1$. If $E \in \mathcal{S}$, then

$$\{\sigma \in \mathcal{N}_n(\Omega) : E \text{ is } X\text{-assumed under } \sigma \text{ at level } j\},$$

is Borel.

A Corollary of Lemma A.7 is:

Corollary A.1 If $E \in \mathcal{S}$, then $\{\sigma \in \mathcal{N}(\Omega) : E \text{ is } X\text{-assumed under } \sigma\}$ is Borel.

To show Lemma A.7, define the sets

$$A[1, j, E] = \bigcap_{i=0}^j \{\sigma \in \mathcal{N}_n(\Omega) : \mu_i(E) = 1\}$$

and

$$A[3, j, E] = \{\sigma \in \mathcal{N}_n(\Omega) : E \subseteq \bigcup_{i \leq j} \text{supp } \mu_i\}.$$

Repeating the arguments in the proof of [Brandenburger, Friedenberg and Keisler's \(2008\)](#) Lemma C.3, for each j , the sets $A[1, j, E]$ and $A[3, j, E]$ are Borel. Now observe that

$$\{\sigma \in \mathcal{N}(\Omega) : E \text{ is TWD-assumed under } \sigma\} = \bigcup_{j=0}^{n-1} (A[1, j, E] \cap A[3, j, E]).$$

This establishes Lemma [A.7](#) for TWD-assumption.

To establish Lemma [A.7](#) for BFK- and PWD-assumption, we will need to define sets $A[\text{BFK}, j, E]$ and $A[\text{PWD}, j, E]$. For $j = 0, \dots, n-2$, let

$$A[\text{BFK}, j, E] = \bigcap_{k=j+1}^{n-1} \bigcup_{\alpha^k \in \mathbb{R}^{j+1}} \bigcap_{F \in \mathcal{S}_E} \{\sigma \in \mathcal{N}_n(\Omega) : \mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F)\}$$

and

$$A[\text{PWD}, j, E] = \bigcap_{F \in \mathcal{S}: F \subseteq E} \bigcap_{k=j+1}^{n-1} \{\sigma = (\mu_0, \dots, \mu_{n-1}) \in \mathcal{N}_n : \mu_k(F) > 0 \Rightarrow \exists i \leq j, \mu_i(F) > 0\}.$$

Let $A[\text{BFK}, n-1, E] = A[\text{PWD}, n-1, E] = \mathcal{N}_n(\Omega)$.¹⁴ Observe that, for $X \in \{\text{BFK}, \text{PWD}\}$

$$\{\sigma \in \mathcal{N}(\Omega) : E \text{ is X-assumed under } \sigma\} = \bigcup_{j=0}^{n-1} (A[1, j, E] \cap A[X, j, E] \cap A[3, j, E]).$$

Thus, to show Lemma [A.7](#) for BFK- and PWD-assumption, it suffices to show that the sets $A[\text{BFK}, j, E]$ and $A[\text{PWD}, j, E]$ are Borel. This is immediate for $j = n-1$. So we focus on the case of $j = 0, \dots, n-2$.

Lemma A.8 *Fix a Borel $E \subseteq \Omega$. There exists a countable algebra \mathcal{F}_E on E that generates \mathcal{S}_E .*

Proof. Since E is a subset of a second countable space, it is second countable. Thus, there exists a countable subbase $\{U^1, U^2, \dots\}$ that generates \mathcal{S}_E . Let \mathcal{F}_E be the algebra generated by $\{U^1, U^2, \dots\}$. By [Rao and Rao \(1983, Corollary 1.1.14\)](#), \mathcal{F}_E is countable. Moreover, it generates \mathcal{S}_E . ■

In what follows, we write \mathcal{F}_E for a countable algebra on E that generates \mathcal{S}_E .

Lemma A.9 *Fix an LPS $\sigma = (\mu_0, \dots, \mu_{n-1})$. Fix also some $j = 0, \dots, n-2$ and $k > j$. Then, the following are equivalent:*

- (1) *There exists $\alpha \in \mathbb{R}^{j+1}$ with $\mu_k(F) = \sum_{i=0}^j \alpha_i \mu_i(F)$ for all $F \in \mathcal{S}_E$.*

¹⁴For $A[X, n-1, E]$ is independent of E . But, for $j = 0, \dots, n-2$, $A[X, j, E]$ depends on E .

(2) There exists an integer $M \geq 1$ such that, for all integers $m \geq 1$, there exists $\beta^m = (\beta_0^m, \dots, \beta_j^m) \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}$ with $|\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)| \leq \frac{1}{m}$ for all $F \in \mathcal{S}_E$.

(3) There exists an integer $M \geq 1$ such that, for all integers $m \geq 1$, there exists $\beta^m = (\beta_0^m, \dots, \beta_j^m) \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}$ with $|\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)| \leq \frac{1}{m}$ for all $F \in \mathcal{F}_E$.

Proof. Suppose part (1) holds. If $\sum_{i=0}^j \mu_i(E) = 0$, then $\mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F) = 0$ for every $F \subseteq E$ Borel. In this case, take $M = 1$ and $\beta = (0, \dots, 0) \in \mathbb{Q}^{j+1} \cap [-1, 1]^{j+1}$.

Thus, we focus on the case where $\sum_{i=0}^j \mu_i(E) > 0$. In this case, for each $m \geq 1$, we can choose $\varepsilon^m \in (0, \frac{1}{m \sum_{i=0}^j \mu_i(E)}]$ and $\beta^m \in \mathbb{Q}^{j+1}$ such that $\max_i |\beta_i^m - \alpha_i| \leq \varepsilon^m$. By construction, $\beta^m \rightarrow \alpha$, and so the sequence $(\beta^m)_m$ is bounded. This implies that there exists $M \geq 0$ such that $\beta^m \in [-M, M]^{j+1}$ for all m . Moreover, for each $m \geq 1$ and each $F \subseteq E$ Borel,

$$\begin{aligned} \left| \mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F) \right| &= \left| \sum_{i=0}^j \alpha_i \mu_i(F) - \sum_{i=0}^j \beta_i^m \mu_i(F) \right| \\ &= \left| \sum_{i=0}^j (\alpha_i - \beta_i^m) \mu_i(F) \right| \\ &\leq |\alpha_i - \beta_i^m| \sum_{i=0}^j \mu_i(F) \\ &\leq \varepsilon^m \sum_{i=0}^j \mu_i(E) \\ &\leq \frac{1}{m}. \end{aligned}$$

This establishes part (2), which in turn establishes part (3).

Next, suppose part (3) holds, i.e., there exist an integer $M \geq 1$ and a sequence $(\beta^m)_m$ such that, for every $m \geq 1$, $\beta^m \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}$ and $|\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)| \leq \frac{1}{m}$ for all $F \in \mathcal{F}_E$. Let \mathcal{M} be the collection of all $F \in \mathcal{S}_E$ for which $|\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)| \leq \frac{1}{m}$ holds for all $m \geq 1$. We will show that $\mathcal{S}_E \subseteq \mathcal{M}$, thereby establishing part (2).

By Lemma A.8, \mathcal{S}_E is the σ -algebra generated by \mathcal{F}_E . So, by the Monotone Class Lemma [Aliprantis and Border \(2007, Lemma 4.13\)](#), \mathcal{S}_E is the smallest monotone class containing \mathcal{F}_E . As such, to show $\mathcal{S}_E \subseteq \mathcal{M}$, it suffices to show that \mathcal{M} is a monotone class containing \mathcal{F}_E .

The fact that \mathcal{M} contains \mathcal{F}_E follows from part (3). To see that \mathcal{M} is a monotone class, consider a monotonically increasing (resp. decreasing) sequence (F^n) of elements of \mathcal{M} . Then $F \equiv \bigcup_n F^n$ (resp. $F \equiv \bigcap_n F^n$) are Borel and, by continuity of the measures $\mu_0, \dots, \mu_j, \mu_k$, $\lim_{n \rightarrow \infty} \mu_i(F^n) = \mu_i(F)$ for $i = 0, \dots, j, k$. Therefore, $\lim_{n \rightarrow \infty} |\mu_k(F^n) - \sum_{i=0}^j \beta_i^m \mu_i(F^n)| = |\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)|$, and so $|\mu_k(F) - \sum_{i=0}^j \beta_i^m \mu_i(F)| \leq \frac{1}{m}$. Thus, \mathcal{M} is a monotone class containing \mathcal{F}_E .

Finally, suppose part (2) holds. Since $\beta^m \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}$, there exists a convergent subsequence $(\beta^{m(\ell)})_\ell$; let $\beta = (\beta_0, \dots, \beta_j)$ be its limit. By construction, for each $m(\ell) \geq 1$

and each $F \in \mathcal{S}_E$ $|\mu_k(F) - \sum_{i=0}^j \beta_i^{m(\ell)} \mu_i(F)| \leq \frac{1}{m(\ell)}$. It follows that, for each $F \in \mathcal{S}_E$, $|\mu_k(F) - \sum_{i=0}^j \beta_i \mu_i(F)| = 0$. This establishes (1). ■

Lemma A.10 *Fix some $n \in \mathbb{N}^0$ and some $j = 0, \dots, n-2$. If $E \in \mathcal{S}$, then $A[\text{BFK}, j, E]$ is Borel.*

Proof. It suffices to show that the set

$$A^k[\text{BFK}, j, E] := \bigcup_{\alpha^k \in \mathbb{R}^{j+1}} \bigcap_{F \in \mathcal{S}_E} \{\sigma \in \mathcal{N}_n(\Omega) : \mu_k(F) = \sum_{i=0}^j \alpha_i^k \mu_i(F)\}$$

is Borel. Note, by Lemma A.9, $A^k[\text{BFK}, j, E] = Y^k$ where

$$Y^k := \bigcup_{M \in \mathbb{N}} \bigcap_{m \in \mathbb{N}} \bigcup_{\alpha \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}} \bigcap_{F \in \mathcal{F}_E} \{\sigma \in \mathcal{N}_n(\Omega) : |\mu_k(F) - \sum_{i=0}^j \alpha_i \mu_i(F)| \leq \frac{1}{m}\}.$$

Note that, in the definition of Y^k , each of the unions and intersections are taken over countable sets. (Use Lemma A.8 to conclude that \mathcal{F}_E is countable.) Thus, to show that the set $A^k[\text{BFK}, j, E]$ is Borel, it suffices to show that, for each $M \geq 1$, $\alpha \in \mathbb{Q}^{j+1} \cap [-M, M]^{j+1}$, $m \in \mathbb{N}$, and $F \in \mathcal{F}_E$ the set

$$\{\sigma \in \mathcal{N}_n(\Omega) : |\mu_k(F) - \sum_{i=0}^j \alpha_i \mu_i(F)| \leq \frac{1}{m}\}$$

is Borel. To show this set is Borel, it suffices to show that the map $F : \mathcal{N}_n(\Omega) \rightarrow \mathbb{R}$ defined by

$$F(\mu_0, \dots, \mu_n) = |\mu_k(F) - \sum_{i=0}^j \alpha_i \mu_i(F)|$$

is measurable.

Note that F is measurable if and only if G is measurable, where

$$G(\mu_0, \dots, \mu_n) = \mu_k(F) - \sum_{i=0}^j \alpha_i \mu_i(F).$$

(See Aliprantis and Border, 2007, Theorem 4.27.) Define maps $g_i : \mathcal{N}_n(\Omega) \rightarrow \mathbb{R}$ where $g_i(\mu_0, \dots, \mu_n) = \mu_i(F)$. For each i , g_i is measurable. (See Aliprantis and Border (2007, Lemma 15.16).) With this $G = g_k - \sum_{i=0}^j \alpha_i g_i$ is measurable Aliprantis and Border (2007, Theorem 4.27), as desired. ■

Fix some $n \in \mathbb{N}^0$ and $j = 0, \dots, n-2$. For any $k > j$ and Borel $F \in \mathcal{S}$, define

$$A^k[\text{PWD}, j, E](F) = \{\sigma = (\mu_0, \dots, \mu_{n-1}) \in \mathcal{N}_n : \mu_k(F) > 0 \Rightarrow \exists i \leq j, \mu_i(F) > 0\}.$$

Then, set

$$A^k[\text{PWD}, j, E] \equiv \bigcap_{F \in \mathcal{S}: F \subseteq E} A^k[\text{PWD}, j, E](F)$$

and observe that $A[\text{PWD}, j, E] = \bigcap_{k=j+1}^{n-1} A^k[\text{PWD}, j, E]$.

Lemma A.11 *Fix some $n \in \mathbb{N}^0$ and $j = 0, \dots, n-2$.*

(i) *For any $k > j$ and Borel $F \in \mathcal{S}$, $A^k[\text{PWD}, j, E](F)$ is Borel.*

(ii) *If $E \in \mathcal{S}$, then $A^k[\text{PWD}, j, E]$ is Borel.*

(iii) *If $E \in \mathcal{S}$, then $A[\text{PWD}, j, E]$ is Borel.*

Before we prove Lemma A.11, it will be convenient to introduce a background lemma. Observe that, since Ω is second-countable, its topology admits a countable base \mathcal{O} . Let $\mathcal{C} = \{\Omega \setminus O : O \in \mathcal{O}\}$. Then \mathcal{C} is countable, and every closed set C is an intersection of elements of \mathcal{C} .

Lemma A.12 *If $F \in \mathcal{S}$, then $\mu(F) = \sup\{\mu(C) : C \in \mathcal{C}\}$.*

Proof. Since Ω is Polish, every Borel measure on Ω is regular (Aliprantis and Border, 2007, Theorem 12.7) and hence inner regular (Aliprantis and Border, 2007, Definition 12.2 and Lemma 12.3): that is, for every Borel $F \in \mathcal{S}$, $\mu(F) = \sup\{\mu(C) : C \subseteq F, C \text{ closed}\}$. Since any closed set is an intersection of elements of \mathcal{C} , this is equivalent to $\mu(F) = \sup\{\mu(C) : C \in \mathcal{C}\}$. ■

Proof of Lemma A.11. Let $F \in \mathcal{S}$ and define

$$P_i(F) = \{(\mu_0, \dots, \mu_{n-1}) \in \mathbb{N}_n : \mu_i(F) > 0\} \quad \text{and} \quad Z_i(F) = \{(\mu_0, \dots, \mu_{n-1}) \in \mathbb{N}_n : \mu_i(F) = 0\}$$

for each $i = 0, \dots, n-1$. By Lemma 15.16 in Aliprantis and Border (2007), the sets $P_i(F)$ and $Z_i(F)$ are Borel.

Fix $k > j$. Observe that $\sigma = (\mu_0, \dots, \mu_{n-1}) \in A^k[\text{PWD}, j, E](F)$ if either $\mu_k(F) = 0$, or $\mu_i(F) > 0$ for some $i \leq j$. Therefore,

$$A^k[\text{PWD}, j, E](F) = Z_k(F) \cup \left(\bigcup_{i=0}^j P_i(F) \right).$$

This is a finite union of Borel sets, and so it is Borel. This establishes part (i).

To complete the proof, it suffices to show part (ii). (Part (iii) follows immediately from part (iii).) Set

$$\tilde{A}^k[\text{PWD}, j] = \bigcap_{C \in \mathcal{C}: C \subseteq E} A^k[\text{PWD}, j, E](C).$$

First observe that, by part (i) and the fact that \mathcal{C} is countable, $\tilde{A}^k[\text{PWD}, j]$ is Borel. Thus, it suffices to show that $\tilde{A}^k[\text{PWD}, j] = A^k[\text{PWD}, j, E]$.

Observe that $\mathcal{C} \subset \mathcal{S}$ and so $A^k[\text{PWD}, j, E] \subseteq \tilde{A}^k[\text{PWD}, j]$. We show $\tilde{A}^k[\text{PWD}, j] \subseteq A^k[\text{PWD}, j, E]$: Fix some $\sigma = (\mu_0, \dots, \mu_{n-1}) \in \tilde{A}^k[\text{PWD}, j]$. Consider a Borel $F \in \mathcal{S}$. Suppose that $\mu_k(F) > 0$. By Lemma A.12, there is $C \in \mathcal{C}$ such that $\mu_k(C) > 0$ and $C \subseteq F \subseteq E$. Since $\sigma \in \tilde{A}^k[\text{PWD}, j]$, there is $i \leq j$ such that $\mu_i(C) > 0$. Hence $\mu_i(F) > 0$. Thus $\sigma \in A^k[\text{PWD}, j, E]$. ■

Proof of Lemma A.7. Immediate from the earlier argument and Lemmas A.10-A.11. ■

Finally, we conclude with Theorem 6.3. The proof of Theorem 10.1 in BFK relies on Property 6.2 and Lemmas F.1 and F.2 therein. Property 6.2 holds for all LPSs (repeat the proof in BFK). The proofs of the two Lemmas begin by fixing a type t^a which maps to a full-support LCPS. BFK then use this type to construct a continuum of related types u^a that satisfy certain properties. Although in our setting types may correspond to general LPSs, we can still find a type t^a that maps to a (full-support) LCPS. In that case, the proof of Lemma F.1 follows verbatim. Likewise, the proof of Lemma F.2 applies verbatim for BFK-assumption and PWD-assumption, since the concepts coincide for LCPSs (Proposition 4.1).¹⁵

Because TWD-assumption does not coincide with BFK-assumption for LCPSs, we must provide a separate proof of Lemma F.2 for arbitrary LPSs. The key step is to replace Lemma E.2 with an analogue for LPSs. With that, we can repeat the argument in BFK (modulo making the correction discussed in Footnote 15). Parts (i) and (ii) of Lemma E.2 follow immediately for arbitrary LPSs. Part (iv) follows immediately from Part (iii). Thus, it suffices to show part (iii) for arbitrary LPSs and TWD-assumption. We conclude with that argument.

Say two LPSs on $S_b \times T_b$, viz. σ and ρ , are *equivalent* if $\sigma = (\mu_0, \dots, \mu_{n-1})$, $\rho = (\nu_0, \dots, \nu_{n-1})$ and for each i , (a) $\text{marg}_{S_b} \mu_i = \text{marg}_{S_b} \nu_i$ and (b) μ_i and ν_i have the same null sets. The following is the analogue of Lemma E.2(iii) for TWD-assumption:

Lemma A.13 *If $\sigma = (\mu_0, \dots, \mu_{n-1})$ and $\rho = (\nu_0, \dots, \nu_{n-1})$ are equivalent and σ TWD-assumes a Borel set $E \subseteq S_b \times T_b$, the ρ TWD-assumes E .*

Proof. Since σ TWD-assumes E , there exists some $j = 0, \dots, n-1$ so that conditions (i) and (iii) hold. Since, for each i , μ_i and ν_i have the same null sets, it follows that

¹⁵ That said, there is a correctable error in BFK's proof of Lemma F.2. The first paragraph must be amended to say the following: There exists some type t_a with $\lambda_a(t_a) = (\mu_0, \dots, \mu_{n-1})$ and a finite set $U \subseteq R_b^{m-1} \setminus R_b^m$ so that R_b^m is assumed at level 0, $\text{proj}_{S_b} U = \text{proj}_{S_b} R_b^{m-1}$, and $\mu_i(E) = 0$ for all $i < m$. The ability to choose U is crucial for establishing the base case, but not important for the inductive step.

$\mu_i((S_b \times T_b) \setminus E) = \nu_i((S_b \times T_b) \setminus E) = 0$ for $i \leq j$; thus, $\nu_i(E) = 1$ for all $i \leq j$. We next show that $E \subseteq \bigcup_{i \leq j} \text{supp } \nu_i$.

Suppose not, i.e., E is not contained in $\bigcup_{i \leq j} \text{supp } \nu_i$. Then $U = (S_b \times T_b) \setminus \bigcup_{i \leq j} \text{supp } \nu_i$ is an open set with $E \cap U \neq \emptyset$. Observe that $E \cap U \subseteq \bigcup_{i \leq j} \text{supp } \mu_i$. So, for each $(s_b, t_b) \in (E \cap U)$, there exists $i \leq j$ with $(s_b, t_b) \in \text{supp } \mu_i$. Since U is an open neighborhood of (s_b, t_b) , $\mu_i(U) > 0$. And since $\mu_i(E) = 1$, $\mu_i(E \cap U) > 0$. But then since μ_i and ν_i have the same null sets, $\nu_i(E \cap U) > 0$, contradicting the definition of U . ■

References

- Aliprantis, C.D. and K.C. Border. 2007. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Verlag.
- Asheim, Geir B. 2002. "Proper Rationalizability in Lexicographic Beliefs." *International Journal of Game Theory* 30(4):453–478.
- Asheim, Geir B and Martin Dufwenberg. 2003. "Admissibility and common belief." *Games and Economic Behavior* 42(2):208–234.
- Asheim, Geir B and Ylva Søvik. 2005. "Preference-based belief operators." *Mathematical Social Sciences* 50(1):61–82.
- Blume, L., A. Brandenburger and E. Dekel. 1991a. "Lexicographic probabilities and choice under uncertainty." *Econometrica* 59(1):61–79.
- Blume, L., A. Brandenburger and E. Dekel. 1991b. "Lexicographic Probabilities and Equilibrium Refinements." *Econometrica* pp. 81–98.
- Brandenburger, A. 1992. Lexicographic Probabilities and Iterated Admissibility. In *Analysis of Markets and Games*, ed. P. Dasgupta, D. Gale, O. Hart and E. Maskin. MIT Press.
- Brandenburger, A., A. Friedenberg and H.J. Keisler. 2008. "Admissibility in Games." *Econometrica* 76(2):307.
- Catonini, E. and N. De Vito. 2014. "Common Assumption of Cautious Rationality and Iterated Admissibility."
- Govindan, Srihari and Tilman Klumpp. 2003. "Perfect Equilibrium and Lexicographic Beliefs." *International Journal of Game Theory* 31(2):229–243.
- Keisler, H.J. and B.S. Lee. 2010. "Common Assumption of Rationality."

- Lee, Byung Soo. 2013. “Conditional Beliefs and Higher-Order Preferences.”.
- Lee, Byung Soo. 2015a. “Admissibility and Assumption.”. (previously titled “Conditional Beliefs and Higher-Order Preferences”).
URL: <http://individual.utoronto.ca/byungsoolee/papers/cbhop.pdf>
- Lee, Byung Soo. 2015b. “Strict Determination is Mutual Singularity.”.
- Mailath, George J, Larry Samuelson and Jeroen M Swinkels. 1997. “How Proper is Sequential Equilibrium?” *Games and Economic Behavior* 18(2):193–218.
- Morris, Stephen. 1997. Alternative definitions of knowledge. In *Epistemic logic and the theory of games and decisions*. Springer pp. 217–233.
- Rajan, U. 1998. “Trembles in the Bayesian Foundations of Solution Concepts of Games.” *Journal of Economic Theory* 82(1):248–266.
- Rao, KPS Bhaskara and M Bhaskara Rao. 1983. *Theory of charges: a study of finitely additive measures*. Vol. 109.
- Stahl, D.O. 1995. “Lexicographic rationalizability and iterated admissibility.” *Economics Letters* 47(2):155–159.
- Yang, C. 2013. “Weak Assumption and Iterated Admissibility.”.