Admissibility and common belief

Geir B. Asheim [a,*] and Martin Dufwenberg [b]

[a] *Department of Economics, University of Oslo, PO Box 1095 Blindern, N-0317 Oslo, Norway*
[b] *Department of Economics, Stockholm University, SE-10691 Stockholm, Sweden*

Abstract

The concept of 'fully permissible sets' is defined by an algorithm that eliminates strategy subsets. It is characterized as choice sets when there is common certain belief of the event that each player prefer one strategy to another if and only if the former weakly dominates the latter on the set of all opponent strategies or on the union of the choice sets that are deemed possible for the opponent. The concept refines the Dekel–Fudenberg procedure and captures aspects of forward induction.
© 2003 Elsevier Science (USA). All rights reserved.

JEL classification: C72

1. Introduction

Iterated (maximal) elimination of weakly dominated strategies (IEWDS) has a long history and some intuitive appeal, yet it does not have as clear an interpretation as iterated elimination of strongly dominated strategies (IESDS). IESDS is known to be equivalent to common belief of rational choice (cf. Tan and Werlang, 1988). IEWDS would appear simply to add a requirement of *admissibility*, i.e., that one strategy should be preferred to another if the former weakly dominates the latter on a set of strategies that the opponent "may choose." However, numerous authors—in particular, Samuelson (1992)—have noted that it is not clear that we can interpret IEWDS this way. To see this, consider the following two examples.

Figure 1 shows the pure strategy reduced strategic form of the 'Battle-of-the-sexes-with-an-outside-option' game. Here IEWDS works by eliminating D, R, and U, leading to the forward induction outcome (M, L). This prediction appears consistent: if 2 believes that 1

---

* Corresponding author.
  E-mail addresses: g.b.asheim@econ.uio.no (G.B. Asheim), md@ne.su.se (M. Dufwenberg).

	L	R
U	2, 2	2, 2
M	3, 1	0, 0
D	0, 0	1, 3

Fig. 1. Battle-of-the-sexes-with-an-outside-option (G_1).

	L	R
U	1, 1	1, 1
M	0, 1	2, 0
D	1, 0	0, 1

Fig. 2. Game where IEWDS is problematic (G_2).

will choose M, then she will prefer L to R as 2's preference over her strategies depends only on the relative likelihood of M and D.

The situation is different in G_2 of Fig. 2, where IEWDS works by eliminating D, R, and M, leading to (U, L). Since 2 is indifferent at the predicted outcome, we must here appeal to admissibility on a superset of $\{U\}$, namely $\{U, M\}$, to justify the statement that 2 *must* play L. However, it is not clear that this is reasonable. Admissibility on $\{U, M\}$ means that 2's preferences respect weak dominance on this set and implies that M is deemed infinitely more likely (in the sense of Blume et al., 1991a, Definition 5.1; see also Appendix A) than D. However, why should 2 deem M more likely than D? If 2 believes that 1 believes in the prediction that 2 plays L (as IEWDS suggests), then it seems odd to assume that 2 believes that 1 considers D to be a less attractive choice than M.

A sense in which D is "less rational" than M is simply that it was eliminated first. This hardly seems a justification for insisting on the belief that D is much less likely than M. Still, Stahl (1995) has shown that IEWDS effectively assumes this: a strategy survives IEWDS if and only it is a best response to a belief where one strategy is infinitely less likely than another if the former is eliminated at an earlier round than the latter. Thus, IEWDS adds extraneous and hard-to-justify restrictions on beliefs, and may not appear to correspond to the most natural formalization of deductive reasoning under admissibility. So what does?

This paper proposes the concept of 'fully permissible sets' as an answer. In G_1 this concept agrees with the *prediction* of IEWDS, as seems natural. The *procedure* leading to this prediction is quite different, though, as is its *interpretation*. In G_2, however, full permissibility predicts that 1's set of rational choices is *either* $\{U\}$ *or* $\{U, M\}$, while 2's set of rational choices is *either* $\{L\}$ *or* $\{L, R\}$. This has appealing features. If 2 is certain that 1's set is $\{U\}$, then—absent extraneous restrictions on beliefs—one cannot conclude that 2 prefers L to R or v.v. On the other hand, if 2 considers it possible that 1's set is $\{U, M\}$, then L weakly dominates R on this set and justifies $\{L\}$ as 2's set of rational choices. Similarly, one can justify that U is preferred to M if and only if 1 considers it impossible that 2's set is $\{L, R\}$. Thus, full permissibility tells a consistent story of deductive reasoning under admissibility, *without* adding extraneous restrictions on beliefs.

Our paper is organized as follows. Section 2 illustrates the key features of the requirement—called 'full admissible consistency'—that is imposed on players to arrive at full permissibility. Section 3 formally defines the concept of fully permissible sets through an algorithm that eliminates strategy sets under full admissible consistency. General existence as well as other properties are shown. Section 4 establishes epistemic conditions for the concept of fully permissible sets, while Section 5 checks that these conditions are indeed necessary and thereby relates full permissibility to other concepts. Section 6 investigates examples, showing how forward induction is promoted and how multiple permissible sets may arise. Section 7 compares our epistemic conditions to those provided in related literature. Some technical material (including the proofs) are contained in three appendices. For ease of presentation, the analysis will be limited to 2-player games, but everything can essentially be generalized to n-player games (with $n > 2$).

2. Illustrating the key features

Our modeling captures three key features:

Caution.   A player should prefer one strategy to another if the former weakly dominates the latter. Such admissibility of a player's preferences on the set of *all* opponent strategies is defended by, e.g., Luce and Raiffa (1957, Chapter 13) and is implicit in procedures that start out by eliminating all weakly dominated strategies.

Full belief of opponent rationality.   A player should deem any opponent strategy that is a rational choice infinitely more likely than any opponent strategy not having this property. This is equivalent to preferring one strategy to another if the former weakly dominates the latter on the set of rational choices for the opponent. Such admissibility of a player's preferences on a *particular subset* of opponent strategies is an ingredient of the analyses of weak dominance by Samuelson (1992) and Börgers and Samuelson (1992), and is essentially satisfied by 'extensive form rationalizability' (EFR; cf. Pearce, 1984 and Battigalli, 1996, 1997) and IEWDS.

No extraneous restrictions on beliefs.   A player should prefer one strategy to another *only if* the former weakly dominates the latter on the set of all opponent strategies or on the set of rational choices for the opponent. Such equal treatment of opponent strategies that are all rational—or all irrational—have in principle been argued by Samuelson (1992, p. 311), Gul (1997), and Mariotti (1997).

These features are combined as follows. A player's preferences over his own strategies, which depend *both* on his payoff function *and* on his beliefs about opponent choice, leads to a *choice set* (i.e., a set of maximal strategies; cf. Section 4.3). A player's preferences are said to be *fully admissibly consistent* with the opponent's preferences if one strategy is preferred to another if and only if the former weakly dominates the latter

- on the set of all opponent strategies, or
- on the union of the choice sets that are deemed possible for the opponent.

	L	R
U	1, 1	1, 1
M	1, 1	1, 0
D	1, 0	0, 1

Fig. 3. Game that illustrates the key features (G_3).

A subset of strategies is a *fully permissible set* if and only if it can be a choice set when there is common certain belief of full admissible consistency, where an event is 'certainly believed' if the complement is deemed impossible (i.e., Savage-null; cf. Section 4.5). Hence, the analysis yields a solution concept that determines a collection of choice sets for each player. This collection can be found via a simple algorithm, introduced in the next section.

We use G_3 of Fig. 3 to illustrate the consequences of imposing 'caution' and 'full belief of opponent rationality.' Since 'caution' means that each player takes all opponent strategies into account, it follows that player 1's preferences over his strategies will be $U \sim M \succ D$ (where $\sim$ and $\succ$ denote indifference and preference, respectively). Player 1 must prefer each of the strategies U and M to the strategy D, because the former strategies weakly dominate D. Hence, U and M are maximal, implying that 1's choice set is $\{U, M\}$.

The requirement of 'full belief in opponent rationality' comes into effect when considering the preferences of player 2. Suppose that 2 certainly believes that 1 is cautious and therefore (as indicated above) certainly believes that $\{U, M\}$ is 1's choice set. Our assumption that 2 has full belief of 1's rationality captures that 2 deems *each* element of $\{U, M\}$ infinitely more likely than D. Thus, 2's preferences respect weak dominance on 1's choice set $\{U, M\}$, regardless of what happens if 1 chooses D. Hence, 2's preferences over her strategies will be $L \succ R$.

Summing up, we get to the following solution for G_3:

1's preferences: $U \sim M \succ D$,

2's preferences: $L \succ R$.

Hence, $\{U, M\}$ and $\{L\}$ are the players' fully permissible sets.

The third feature of full admissible consistency—'no extraneous restrictions on beliefs'—means in G_3 that 2 does not assess the relative likelihood of 1's maximal strategies U and M. This does not have any bearing on the analysis of G_3, but is essential for capturing forward induction in G_1. In this case the issue is not whether a player assesses the relative likelihood of different maximal strategies, but rather whether a player assesses the relative likelihood of different non-maximal strategies. To see the significance in G_1, assume that 1 deems R infinitely more likely than L, while 2 deems U infinitely more likely than D and D infinitely more likely than M. Then the players rank their strategies as follows:

1's preferences: $U \succ D \succ M$,

2's preferences: $R \succ L$.

Both 'caution' and 'full belief of opponent rationality' are satisfied and still the forward induction outcome (M, L) is not promoted. However, the requirement of 'no extraneous restrictions on beliefs' is not satisfied since the preferences of 2 introduce extraneous restrictions on beliefs by deeming one of 1's non-maximal strategies, D, infinitely more likely than another non-maximal strategy, M. When we return to G_1 in Sections 3.3 and 6.1 we show how the additional imposition of 'no extraneous restrictions on beliefs' leads to (M, L) in this game.

Several concepts with natural epistemic foundations fail to match these predictions in G_1 and G_3. In the case of rationalizability (Bernheim, 1984; Pearce, 1984) this is perhaps not so surprising since this concept in 2-player games corresponds to IESDS. It can be understood as a consequence of common belief of rational choice without imposing caution, so there is no guarantee that a player prefers one strategy to another if the former weakly dominates the latter. In G_3, for example, all strategies are rationalizable.

It is more surprising that the concept of 'permissibility' does not match our solution of G_3. Permissibility can be given rigorous epistemic foundations in models with cautious players (cf. Börgers, 1994, and Brandenburger, 1992, who coined the term 'permissible'; see also Ben-Porath, 1997, and Gul, 1997). In these models players take into account all opponent strategies, while assigning more weight to a *subset* of those deemed to be rational choices. Permissibility corresponds to the DF procedure (after Dekel and Fudenberg, 1990) where one round of elimination of (all) weakly dominated strategies is followed by iterated elimination of strongly dominated strategies. In G_3, this means that 1 cannot choose his weakly dominated strategy D. However, while 2 prefers L to R in our solution, permissibility allows that 2 chooses R. To exemplify using Brandenburger's (1992) approach, this will be the case if 2 deems U to be infinitely more likely than D which in turn is deemed infinitely more likely than M. The problem is that 'full belief of opponent rationality' is not satisfied: Player 2 deems D more likely than M even though M is in 1's choice set, while D is not. In Section 3.4 we establish (Proposition 3.2) that the concept of fully permissible sets refines the DF procedure.

3. The algorithm

We present in this section an algorithm—'iterated elimination of choice sets under full admissible consistency' (IECFA)—leading to the concept of 'fully permissible sets.' This concept will in turn be given an epistemic characterization in Section 4 by imposing common certain belief of full admissible consistency. We present the algorithm before the epistemic characterization for different reasons:

- IECFA is fairly accessible. By defining it early, we can apply it early, and offer early indications of the nature of the solution concept we wish to promote.
- By defining IECFA, we point to a parallel to the concepts of rationalizable strategies and permissible strategies. These concepts are motivated by epistemic assumptions, but turn out to be identical in 2-player games to the set of strategies surviving simple algorithms: respectively, IESDS and the DF procedure.

- Just like IESDS and the DF procedure, IECFA is easier to use than the corresponding epistemic characterizations. The algorithm should be handy for applied economists, independently of the foundational issues discussed in Sections 4 and 5.

IESDS and the DF procedure iteratively eliminate dominated strategies. In the corresponding epistemic models, these strategies in turn cannot be rational choices, cannot be rational choices given that other players do not use strategies that cannot be rational choices, etc. IECFA is also an elimination procedure. However, the interpretation of the basic item thrown out is not that of a strategy that cannot be a rational choice, but rather that of *a set* of strategies that cannot be a choice set for any preferences that are in a given sense consistent with the preferences of the opponent. The specific kind of consistency involved in IECFA—which will be defined in Section 4.6 and referred to as 'full admissible consistency'—requires that a player's preferences are characterized by the properties of 'caution,' 'full belief of opponent rationality' and 'no extraneous restrictions on beliefs.' Thus, IECFA does not start with each player's strategy set and then iteratively eliminates strategies. Rather, IECFA starts with each player's collection of non-empty subsets of his strategy set and then iteratively eliminates subsets from this collection.

3.1. A strategic game

With $I = \{1, 2\}$ as the set of *players*, let, for each i, S_i denote player i's finite set of *pure strategies* and $u_i : S \to \Re$ be a vNM utility function that assigns payoff to any strategy vector, where $S = S_1 \times S_2$ is the set of strategy vectors. Then $G = (S_1, S_2, u_1, u_2)$ is a finite *strategic* two-player *game*. Write p_i, r_i, and s_i ($\in S_i$) for pure strategies and x_i and y_i ($\in \Delta(S_i)$) for mixed strategies. We may extend u_i to mixed strategies: $u_i(x_i, s_j) = \sum_{s_i \in S_i} x_i(s_i) u_i(s_i, s_j)$.

3.2. Definition

Let S'_j ($\subseteq S_j$) be a set of opponent strategies. Say that x_i *weakly dominates* y_i on S'_j if, $\forall s_j \in S'_j$, $u_i(x_i, s_j) \geqslant u_i(y_i, s_j)$, with strict inequality for some $s_j \in S'_j$. Interpret Q_j ($\subseteq S_j$) as the set of strategies that player i deems to be the set of rational choices for his opponent. Let i's choice set be equal to $S_i \backslash D_i(Q_j)$, where, for any ($\emptyset \neq$) $Q_j \subseteq S_j$,

$$D_i(Q_j) := \left\{ s_i \in S_i \mid \exists x_i \in \Delta(S_i) \text{ s.t. } x_i \text{ weakly dom. } s_i \text{ on } Q_j \text{ or } S_j \right\}.$$

Hence, i's choice set consists of pure strategies that are not weakly dominated by any mixed strategy on Q_j or S_j. In Section 4 we show how this corresponds to a set of maximal strategies given the player's preferences over his own strategies.

Let $\Sigma = \Sigma_1 \times \Sigma_2$, where $\Sigma_i := 2^{S_i} \backslash \{\emptyset\}$ denotes the collection of non-empty subsets of S_i. Write π_i, ρ_i, and σ_i ($\in \Sigma_i$) for subsets of pure strategies. For any ($\emptyset \neq$) $\Xi = \Xi_1 \times \Xi_2 \subseteq \Sigma$, write $\alpha(\Xi) := \alpha_1(\Xi_2) \times \alpha_2(\Xi_1)$, where

$$\alpha_i(\Xi_j) := \left\{ \pi_i \in \Sigma_i \,\middle|\, \exists (\emptyset \neq) \Psi_j \subseteq \Xi_j \text{ s.t. } \pi_i = S_i \backslash D_i \left(\bigcup_{\sigma_j \in \Psi_j} \sigma_j \right) \right\}.$$

Hence, $\alpha_i(\Xi_j)$ is the collection of strategy subsets that can be choice sets for player i if he associates Q_j—the set of rational choices for his opponent—with the union of the strategy subsets in a non-empty subcollection of Ξ_j.

We can now define the concept of a fully permissible set.

Definition 3.1. Consider the sequence defined by $\Xi(0) = \Sigma$ and, $\forall g \geqslant 1$, $\Xi(g) = \alpha(\Xi(g-1))$. A non-empty strategy set π_i is said to be a *fully permissible set* for player i if $\pi_i \in \bigcap_{g=0}^{\infty} \Xi_i(g)$.

Let $\Pi = \Pi_1 \times \Pi_2$ denote the *collection* of vectors of fully permissible sets. Since $\emptyset \neq \alpha_i(\Xi_j') \subseteq \alpha_i(\Xi_j'') \subseteq \alpha_i(\Sigma_j)$ whenever $\emptyset \neq \Xi_j' \subseteq \Xi_j'' \subseteq \Sigma_j$ and since the game is finite, $\Xi(g)$ is a monotone sequence that converges to Π in a finite number of iterations. IECFA is the procedure that in round g eliminates sets in $\Xi(g-1) \backslash \Xi(g)$ as possible choice sets. As defined in Definition 3.1 IECFA eliminates maximally in each round in the sense that, $\forall g \geqslant 1$, $\Xi(g) = \alpha(\Xi(g-1))$. However, it follows from the monotonicity of α_i that any non-maximal procedure, where $\exists g \geqslant 1$ such that $\Xi(g-1) \supset \Xi(g) \supset \alpha(\Xi(g-1))$, will also converge to Π.

A strategy subset survives elimination round g if it can be a choice set when the set of rational choices for his opponent is associated with the union of some (or all) of opponent sets that have survived the procedure up till round $g - 1$. A fully permissible set is a set that survives in this way for any g. The analysis of Section 4 justifies that strategy subsets that this algorithm has not eliminated by round g be interpreted as choice sets compatible with $g - 1$ order of mutual certain belief of full admissible consistency.

3.3. Applications

We illustrate IECFA by applying it. Consider G_3 of Section 2. We get:

$$\Xi(0) = \Sigma_1 \times \Sigma_2,$$
$$\Xi(1) = \big\{\{U, M\}\big\} \times \Sigma_2,$$
$$\Pi = \Xi(2) = \big\{\{U, M\}\big\} \times \big\{\{L\}\big\}.$$

Independently of Q_2, $S_1 \backslash D_1(Q_2) = \{U, M\}$, so for 1 only $\{U, M\}$ can survive the first elimination round, while $S_2 \backslash D_2(\{U, M\}) = \{L\}$, $S_2 \backslash D_2(\{D\}) = \{R\}$ and $S_2 \backslash D_2(\{U\}) = \{L, R\}$, so that no elimination is possible for player 2. However, in the second round only $\{L\}$ survives since L weakly dominates R on $\{U, M\}$, implying that $S_2 \backslash D_2(\{U, M\}) = \{L\}$.

Next, consider G_1 of the introduction. Applying IECFA we get:

$$\Xi(0) = \Sigma_1 \times \Sigma_2,$$
$$\Xi(1) = \big\{\{U\}, \{M\}, \{U, M\}\big\} \times \Sigma_2,$$
$$\Xi(2) = \big\{\{U\}, \{M\}, \{U, M\}\big\} \times \big\{\{L\}, \{L, R\}\big\},$$
$$\Xi(3) = \big\{\{M\}, \{U, M\}\big\} \times \big\{\{L\}, \{L, R\}\big\},$$
$$\Xi(4) = \big\{\{M\}, \{U, M\}\big\} \times \big\{\{L\}\big\},$$
$$\Pi = \Xi(5) = \big\{\{M\}\big\} \times \big\{\{L\}\big\}.$$

Again the algorithm yields a unique fully permissible set for each player.

Finally, apply IECFA to G_2 of the introduction:

$$\Xi(0) = \Sigma_1 \times \Sigma_2,$$

$$\Xi(1) = \big\{\{U\}, \{M\}, \{U, M\}\big\} \times \Sigma_2,$$

$$\Xi(2) = \big\{\{U\}, \{M\}, \{U, M\}\big\} \times \big\{\{L\}, \{L, R\}\big\},$$

$$\Pi = \Xi(3) = \big\{\{U\}, \{U, M\}\big\} \times \big\{\{L\}, \{L, R\}\big\}.$$

Here we are left with two fully permissible sets for each player. There is no further elimination, as $\{U\} = S_1 \backslash D_1(\{L\})$, $\{U, M\} = S_1 \backslash D_1(\{L, R\})$, $\{L\} = S_2 \backslash D_2(\{U, M\})$, and $\{L, R\} = S_2 \backslash D_2(\{U\})$.

The elimination process for G_1 and G_2 is explained and interpreted in Section 6.

3.4. Results

The following proposition characterizes the strategy subsets that survive IECFA and thus are fully permissible.

Proposition 3.1. (i) $\forall i \in I$, $\Pi_i \neq \emptyset$.

(ii) $\Pi = \alpha(\Pi)$.

(iii) $\forall i \in I$, $\pi_i \in \Pi_i$ if and only if there exists $\Xi = \Xi_1 \times \Xi_2$ with $\pi_i \in \Xi_i$ such that $\Xi \subseteq \alpha(\Xi)$.

Proposition 3.1(i) establishes existence, but not uniqueness, of each player's fully permissible set(s). In addition to G_2, games with multiple strict Nash equilibria illustrate the possibility of such multiplicity; by Proposition 3.1(iii), any strict Nash equilibrium corresponds to a vector of fully permissible sets. Proposition 3.1(ii) means that Π is a fixed point in terms of a collection of vectors of strategy sets as illustrated by G_2 above. By Proposition 3.1(iii) it is the largest such fixed point.

We close this section by recording some connections between IECFA on the one hand, and IESDS, the DF-procedure, and IEWDS on the other. First, we note through the following Proposition 3.2 that IECFA has more bite than the DF procedure. Both G_1 and G_3 illustrate that this refinement may be strict.

Proposition 3.2. *A pure strategy p_i is permissible (i.e., survives the DF procedure) if there exists a fully permissible set π_i such that $p_i \in \pi_i$.*

It is a corollary that IECFA has more cutting power also than IESDS.

However, neither of IECFA and IEWDS has more bite than the other, as demonstrated by the game G_4 of Fig. 4. It is straightforward to verify that a and b for player 1, and e for player 2 survive IEWDS, while $\{a\}$ for 1 and $\{e, f\}$ for 2 survive IECFA and are thus the fully permissible sets, as shown below:

	e	f	g
a	1, 1	1, 1	0, 0
b	1, 1	0, 1	1, 0
c	0, 1	0, 0	2, 0
d	0, 0	0, 1	0, 2

Fig. 4. The relation between IECFA and IEWDS (G_4).

$$\Xi(0) = \Sigma_1 \times \Sigma_2,$$
$$\Xi(1) = \big\{\{a\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\big\}$$
$$\times \big\{\{e\}, \{g\}, \{e,f\}, \{e,g\}, \{f,g\}, \{e,f,g\}\big\},$$
$$\Xi(2) = \big\{\{a\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\big\} \times \big\{\{e\}, \{e,f\}\big\},$$
$$\Xi(3) = \big\{\{a\}, \{a,b\}\big\} \times \big\{\{e\}, \{e,f\}\big\},$$
$$\Xi(4) = \big\{\{a\}, \{a,b\}\big\} \times \big\{\{e,f\}\big\},$$
$$\Pi = \Xi(5) = \big\{\{a\}\big\} \times \big\{\{e,f\}\big\}.$$

Strategy b survives IEWDS but does not appear in any fully permissible set. Strategy f appears in a fully permissible set but does not survive IEWDS.

4. Epistemic conditions for fully permissible sets

When justifying rationalizable and permissible strategies through epistemic conditions, players are usually modeled as decision makers under uncertainty. Tan and Werlang (1988) characterize rationalizable strategies by common belief (with probability 1) of the event that each player chooses a maximal strategy given preferences that are represented by a subjective probability distribution. Hence, preferences are both *complete* and *continuous*. Brandenburger (1992) characterizes permissible strategies by common belief (with primary probability 1) of the event that each player chooses a maximal strategy given preferences that are represented by a *lexicographic probability system* (LPS; cf. Blume et al., 1991a) with full support on the set of opponent strategies. Hence, preferences are still complete, but not continuous due to the full support requirement. Since preferences are complete and representable by a probability distribution or an LPS, these epistemic justifications differ significantly from the corresponding algorithms, IESDS and the DF procedure, neither of which makes reference to subjective probabilities.

When doing analogously for fully permissible sets, not only must continuity of preferences be relaxed to allow for 'caution' and 'full belief of opponent rationality,' as discussed in Section 2. One must also relax completeness of preferences to accommodate 'no extraneous restrictions on beliefs,' which is a requirement of minimal completeness and implies that preferences are expressed solely in terms of admissibility on nested sets. Hence, preferences are not in general representable by subjective probabilities (except through treating incomplete preferences as a *set* of complete preferences; cf. Aumann, 1962). This means that epistemic operators must be derived directly from the underlying

preferences (as observed by Morris, 1997) since there is no probability distribution or LPS that represents the preferences. It also entails that the resulting characterization (Proposition 4.1) must be closely related to the algorithm used in the definition of fully permissible sets.

There is another fundamental difference. When characterizing rationalizable and permissible strategies, the event that is made subject to interactive epistemology is often defined by requiring that each player's *strategy choice* is an element of his choice set (i.e., his set of maximal strategies) given his belief about the opponent's strategy choice. In contrast, in the characterization of Proposition 4.1, the event that is made subject to interactive epistemology is defined by imposing requirements on how each player's *choice set* is related to his belief about the opponent's choice set. Since a player's choice set equals the set of undominated strategies given the ranking that the player has over his strategies, the imposed requirements relate a player's ranking over his strategies to the opponent's ranking.

To prepare for the characterization result we first present a framework for strategic games where each player is modeled as a decision-maker under uncertainty and introduce the belief operator that is used for the interactive epistemology. We then move to the characterization of Proposition 4.1, which is interpreted in the subsequent Section 5 in terms of the requirements of 'caution,' 'full belief of opponent rationality,' and 'no extraneous restrictions on beliefs.'

4.1. A strategic game form

Let $z: S \rightarrow Z$ map strategy vectors into *outcomes*, where Z is the set of outcomes. Then (S_1, S_2, z) is a finite *strategic* two-player *game form*.

4.2. States and types

The uncertainty faced by a player i in a strategic game form concerns the strategy choice of his opponent j, j's belief about i's strategy choice, and so on (cf. Tan and Werlang, 1988). A type of a player i corresponds to a vNM utility function and a belief about j's strategy choice, a belief about j's belief about i's strategy choice, and so on. Models of such infinite hierarchies of beliefs (Böge and Eisele, 1979; Mertens and Zamir, 1985; Brandenburger and Dekel, 1993; Epstein and Wang, 1996) yield $S \times T$ as the complete state space, where $T = T_1 \times T_2$ is the set of all feasible type vectors. Furthermore, for each i, there is a homeomorphism between T_i and the set of beliefs on $S \times T_j$. For each type of any player i, the type's decision problem is to choose one of i's strategies. For the modeling of this problem, the type's belief about i's strategy choice is not relevant and can be ignored.[1] Hence, in the setting of a strategic game form the beliefs can be restricted to the set of opponent strategy-type pairs, $S_j \times T_j$. Combined with a vNM utility function, the set of beliefs on $S_j \times T_j$ corresponds to a set of binary relations on the set of acts on

---

[1] This does not mean that i is not aware of i's strategy choice. It signifies that i's belief about j's belief about j's choice plays no role in the analysis. Note that Tan and Werlang (1988, Sections 2 and 3) characterize rationalizable strategies without specifying beliefs about one's own choice.

$S_j \times T_j$, where an *act* on $S_j \times T_j$ is a function that to any element of $S_j \times T_j$ assigns an objective randomization on Z (cf. Appendix A).

In conformity with the literature on infinite hierarchies of beliefs, let

- the set of *states of the world* (or simply *states*) be $\Omega := S \times T$,
- each *type* t_i of any player i correspond to a binary relation $\succeq^{t_i}$ on the set of acts on $S_j \times T_j$.

However, as the above results on infinite hierarchies of beliefs are not applicable in the present setting,[2] we instead consider an implicit model—with a finite type set T_i for each player i—from which infinite hierarchies of beliefs can be constructed. Moreover, since completeness and continuity of preferences are not imposed, the conditions on $\succeq^{t_i}$ are specified as follows.

Assumption 4.1. The binary relation $\succeq^{t_i}$ satisfies *reflexivity*, *transitivity*, *objective independence*, *nontriviality*, *conditional completeness*, *conditional continuity*, and *non-null state independence*.

This means that $\succeq^{t_i}$ is conditionally represented by a vNM utility function $\upsilon_i^{t_i} : Z \to \Re$ that assigns a payoff to any outcome (cf. Proposition A.1).[3] Being a vNM utility function, $\upsilon_i^{t_i}$ can be extended to objective randomizations on Z. Since $\succeq^{t_i}$ is conditionally represented, it follows that strong and weak dominance are well-defined. The construction is summarized by the following definition.

Definition 4.1. A *belief system* for a game form (S_1, S_2, z) consists of

- for each player i, a finite set of types T_i,
- for each type t_i of any player i, a binary relation $\succeq^{t_i}$ (t_i's *preferences*) on the set of acts on $S_j \times T_j$, where $\succeq^{t_i}$ satisfies Assumption 4.1, and where $\upsilon_i^{t_i}$ denotes the vNM utility function that conditionally represents $\succeq^{t_i}$.

4.3. Choice sets and rationality

Let $\succeq^{t_i}_{S_j}$ denote the *marginal* of $\succeq^{t_i}$ on S_j (cf. Appendix A). A pure strategy $s_i \in S_i$ can be viewed as an act $\mathbf{x}_{S_j}$ on S_j that assigns $z(s_i, s_j)$ to any $s_j \in S_j$. A *mixed* strategy $x_i \in \Delta(S_i)$ can be viewed as an act $\mathbf{x}_{S_j}$ on S_j that assigns $z(x_i, s_j)$ to any $s_j \in S_j$. Hence,

---

[2] Brandenburger (1998) and Brandenburger and Keisler (1999) have shown that a complete state space may not exist if beliefs are not based on subjective probabilities. In contrast to Brandenburger and Keisler's (2002) and Battigalli and Siniscalchi's (2002) epistemic foundations for IEWDS and EFR respectively, a complete state space is not needed for the present analysis.

[3] If conditional completeness is strengthened to completeness, then it follows from Blume et al. (1991a) that $\succeq^{t_i}$ is represented by $\upsilon_i^{t_i}$ and an LPS $\lambda^{t_i} = (\mu_1^{t_i}, \ldots, \mu_L^{t_i}) \in \mathbf{L}\Delta(S_j \times T_j)$. If, in addition, conditional continuity is strengthened to continuity, then $\succeq^{t_i}$ is represented by $\upsilon_i^{t_i}$ and a subjective probability distribution $\mu^{t_i} \in \Delta(S_j \times T_j)$.

$\succeq_{S_j}^{t_i}$ is a binary relation also on the subset of acts on S_j that correspond to i's mixed strategies. Thus, $\succeq_{S_j}^{t_i}$ can be referred to as t_i's *preferences over i's mixed strategies*. The set of mixed strategies $\Delta(S_i)$ is the set of acts that are at t_i's actual disposal. Since $\succeq^{t_i}$ is reflexive and transitive and satisfies objective independence, $\succeq_{S_j}^{t_i}$ shares these properties, and t_i's *choice set*,

$$C_i^{t_i} := \left\{ s_i \in S_i \mid s_i \text{ is maximal w.r.t. } \succeq_{S_j}^{t_i} \text{ in } \Delta(S_i) \right\},$$

is non-empty and supports any maximal mixed strategy.

The event that i is *rational* is defined as

$$[rat_i] := \left\{ (s_1, s_2, t_1, t_2) \in \Omega \mid s_i \in C_i^{t_i} \right\}.$$

4.4. Playing the game

The event that i plays the game $G = (S_1, S_2, u_1, u_2)$ is given by

$$[u_i] := \left\{ (s_1, s_2, t_1, t_2) \in \Omega \mid v_i^{t_i} \circ z \text{ is a positive affine transformation of } u_i \right\},$$

while $[u_1] \cap [u_2]$ is the event that both players play G.

We allow for the possibility that $\Omega \setminus [u_i] \neq \emptyset$ merely to ensure that strategy subsets not eliminated by round 1 in the algorithm of Definition 3.1 correspond to choice sets compatible with full admissible consistency in the epistemic characterization of Section 4.6. This requires that, for each i and for any $\sigma_i \in \Sigma_i$, there exists a belief system and some state $(s_1, s_2, t_1, t_2) \in \Omega$ such that $C_i^{t_i} = \sigma_i$.

4.5. Epistemic operators

As illustrated by the analysis of G_2 in Section 6.2, the interactive epistemology must allow a player type to deem an opponent type impossible even if it is the true type of the opponent. This requires a subjective epistemic operator that does not satisfy the truth axiom. Such subjective operators are usually based on the subjective probability distribution $\mu^{t_i} \in \Delta(S_j \times T_j)$ or the LPS $\lambda^{t_i} = (\mu_1^{t_i}, \ldots, \mu_L^{t_i}) \in \mathbf{L}\Delta(S_j \times T_j)$ that represents the preferences.

In the present setting where incomplete preferences entail that no such representation is available, operators can instead be based directly the underlying preferences. For this purpose, say that $\succeq^{t_i}$ is *admissible on* β_j, where $\emptyset \neq \beta_j \subseteq S_j \times T_j$, if $\mathbf{x} \succ^{t_i} \mathbf{y}$ whenever $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$, and note that

- with $\mu^{t_i} \in \Delta(S_j \times T_j)$ as a representation, $\mathrm{supp}\,\mu^{t_i}$ is the unique set of opponent strategy-type pairs on which $\succeq^{t_i}$ is admissible, while
- with $\lambda^{t_i} \in \mathbf{L}\Delta(S_j \times T_j)$ as a representation,

$$\left\{ \mathrm{supp}\,\mu_1^{t_i}, \; \bigcup_{\ell=1}^{2} \mathrm{supp}\,\mu_\ell^{t_i}, \; \ldots, \; \bigcup_{\ell=1}^{L} \mathrm{supp}\,\mu_\ell^{t_i} \right\}$$

is the collection of nested sets on which $\succeq^{t_i}$ is admissible.

Such a collection of nested sets exists also if preferences are incomplete. Let

$$\kappa_j^{t_i} := \left\{ (s_j, t_j) \in S_j \times T_j \mid (s_j, t_j) \text{ is not Savage-null acc. to } \succeq^{t_i} \right\}$$

denote the set of opponent strategy-type pairs that t_i deems possible. If $\succeq^{t_i}$ is admissible on β_j, then $\beta_j \subseteq \kappa_j^{t_i}$ and any $(s_j', t_j') \in \beta_j$ is deemed infinitely more likely than any $(s_j'', t_j'') \in (S_j \times T_j) \backslash \beta_j$. Since (s_j', t_j') being infinitely more likely than (s_j'', t_j'') implies that (s_j'', t_j'') is *not* infinitely more likely than (s_j', t_j'), it follows that $\beta_j' \subseteq \beta_j''$ or $\beta_j' \supseteq \beta_j''$ whenever $\succeq^{t_i}$ is admissible on both β_j' and β_j''. Since $\succeq^{t_i}$ is admissible on $\kappa_j^{t_i}$, it follows that there exists a unique smallest (w.r.t. set inclusion) non-empty set on which $\succeq^{t_i}$ is admissible; let this set be denoted $\beta_j^{t_i}$. Hence, there is a collection of sets, $\{\beta_j^{t_i}, \ldots, \kappa_j^{t_i}\}$, on which $\succeq^{t_i}$ is admissible, where $\beta_j^{t_i} \subseteq \cdots \subseteq \kappa_j^{t_i}$, and where $\beta_j^{t_i} \neq \kappa_j^{t_i}$ only if $\succeq^{t_i}$ is not continuous.

Since we will be concerned with what types of j that i deems possible, what types of i that the possible types of j deem possible, and so on, the operator used for the interactive epistemology will be based on $\kappa_j^{t_i}$. To state this operator—which will be referred to as 'certain belief'—let, for each player i and each state $\omega \in \Omega$, $t_i(\omega)$ denote the projection of ω on T_i. It will follow that at ω player i 'certainly believes' the event that his type is $t_i(\omega)$. As it is unnecessary to specify at ω player i's belief about his own strategy choice, we consider only events $E \subseteq \Omega$ satisfying that $E = S_i \times \text{proj}_{S_j \times T_i \times T_j} E$. For any such event E, let

$$E_j^{t_i} := \left\{ (s_j, t_j) \in S_j \times T_j \mid (s_j, t_i, t_j) \in \text{proj}_{S_j \times T_i \times T_j} E \right\}$$

denote the set of opponent strategy-type pairs that are consistent with $\omega \in E$ and $t_i(\omega) = t_i$. If $E \subseteq \Omega$ satisfies that $E = S_i \times \text{proj}_{S_j \times T_i \times T_j} E$, then say that at ω player i *certainly believes* the event E if $\omega \in K_i E$ where

$$K_i E := \left\{ \omega \in \Omega \mid \kappa_j^{t_i(\omega)} \subseteq E_j^{t_i(\omega)} \right\}.[4]$$

Hence, at ω player i certainly believes E (where $E = S_i \times \text{proj}_{S_j \times T_i \times T_j} E$) if the complement of $E_j^{t_i(\omega)}$, $(S_j \times T_j) \backslash E_j^{t_i(\omega)}$, is deemed impossible.

If $E \subseteq \Omega$ satisfies that $E = S_i \times S_j \times \text{proj}_{T_i \times T_j} E$, then say that there is *mutual certain belief* of E at ω if $\omega \in KE$, where $KE := K_1 E \cap K_2 E$. Say that there is *common certain belief* of E at ω if $\omega \in CKE$, where $CKE := KE \cap KKE \cap KKKE \cap \cdots$.

4.6. Full admissible consistency

To characterize the concept of fully permissible sets, consider for each i,

$$[cau_i] := \left\{ \omega \in \Omega \mid \kappa_j^{t_i(\omega)} = S_j \times T_j^{t_i(\omega)} \right\},$$

$$\overline{B}_i^0[rat_j] := \Big\{ \omega \in \Omega \mid \beta_j^{t_i(\omega)} = [rat_j]_j \cap \kappa_j^{t_i(\omega)}, \text{ and } \mathbf{x} \succ^{t_i(\omega)} \mathbf{y} \text{ only if }$$
$$\mathbf{x}_{\beta_j} \text{ weakly dom. } \mathbf{y}_{\beta_j} \text{ for } \beta_j = \beta_j^{t_i(\omega)} \text{ or } \beta_j = \kappa_j^{t_i(\omega)} \Big\},$$

---

[4] 'Certain belief' corresponds to what Morris (1997) calls 'Savage-belief.' This operator and those introduced in Section 5—'full belief' and 'belief'—all coincide with 'belief with probability 1' in the case where preferences are complete and continuous. The properties of the operators are summarized in Appendix B.

where $T_j^{t_i} := \text{proj}_{T_j} \kappa_j^{t_i}$ denotes the set of opponent types that t_i deems possible, and where $[rat_j]_j := \text{proj}_{S_j \times T_j}[rat_j] = \{(s_j, t_j) \mid s_j \in C_j^{t_j}\}$. Note that if $\omega \in [cau_i] \cap \overline{B}_i^0[rat_j]$, then $t_i(\omega)$'s preferences are determined by admissibility on two particular nested set of opponent strategy-type pairs.

Say that at ω player i is *fully admissibly consistent* (with the game G and the preferences of his opponent) if $\omega \in \overline{A}_i^0$, where

$$\overline{A}_i^0 := [u_i] \cap [cau_i] \cap \overline{B}_i^0[rat_j].$$

Refer to $\overline{A}^0 := \overline{A}_1^0 \cap \overline{A}_2^0$ as the event of *full admissible consistency*. We can now characterize the concept of fully permissible sets as sets of maximal strategies in states where there is common certain belief of full admissible consistency.

Proposition 4.1. *A non-empty strategy set π_i for player i is fully permissible in a finite strategic game G if and only if there exists a belief system with $\pi_i = C_i^{t_i(\omega)}$ for some $\omega \in CK\overline{A}^0$.*

5. Are the conditions necessary?

In this section we first show how the event used to characterize fully permissible sets—full admissible consistency—can be interpreted in terms of the requirements of 'caution,' 'full belief of opponent rationality,' and 'no extraneous restrictions on beliefs.' Furthermore, following a common procedure of the axiomatic method, we verify that these requirements are indeed necessary for the characterization in Proposition 4.1 by investigating the consequences of relaxing one requirement at a time. These exercises contribute to the understanding of fully permissible sets by showing that our concept is related to proper equilibrium as well as permissible and rationalizable strategies in the following manner:

- When allowing extraneous restrictions on beliefs, we open for any strategy that can be played with positive probability in a proper equilibrium, implying that forward induction is no longer promoted in G_1.
- When weakening 'full belief of opponent rationality' to 'belief of opponent rationality,' we characterize the concept of permissible strategies independently of whether a requirement of 'no extraneous restrictions on beliefs' is retained.
- When removing 'caution,' we characterize the concept of rationalizable strategies independently of whether extraneous restrictions on beliefs are allowed and full belief of opponent rationality is weakened.

To relax 'no extraneous restrictions on beliefs' we need a model—as the one introduced in Section 4—that is versatile enough to allow for preferences that are more complete than being determined by admissibility on two nested sets.

5.1. Interpreting full admissible consistency

It is clear that $[cau_1] \cap [cau_2]$ corresponds to 'caution.' If $\omega \in [cau_i]$, then (s_j, t_j) is deemed possible according to $\succeq^{t_i(\omega)}$ whenever t_j is deemed possible. Hence, $\forall(s_j, t_j) \in S_j \times T_j^{t_i(\omega)}$, $\omega \notin K_i\{(s'_1, s'_2, t'_1, t'_2) \in \Omega \mid (s'_j, t'_j) \neq (s_j, t_j)\}$ (cf. Dekel and Gul's, 1997, definition of caution). It implies that the marginal of $\succeq^{t_i(\omega)}$ on S_j (i.e., $t_i(\omega)$'s preferences over $\Delta(S_i)$, $\succeq^{t_i(\omega)}_{S_j}$) is admissible on S_j.

To interpret full admissible consistency it remains to split $\overline{B}_1^0[rat_2] \cap \overline{B}_2^0[rat_1]$ into 'full belief of opponent rationality' and 'no extraneous restrictions on beliefs.'

To state the condition of 'full belief of opponent rationality' we need to introduce 'full belief' as an epistemic operator. Also this operator will be defined on the class of events $E \subseteq \Omega$ satisfying that $E = S_i \times \text{proj}_{S_j \times T_i \times T_j} E$. For any E in this class, say that at ω player i *fully believes* the event E if $\omega \in B_i^0 E$ where

$$B_i^0 E := \left\{ \omega \in \Omega \mid \succeq^{t_i(\omega)} \text{ is admissible on } E_j^{t_i(\omega)} \cap \kappa_j^{t_i(\omega)} \right\}.$$

If $\omega \in B_i^0[rat_j]$, then at ω i fully believes that j is rational. This means that any (s'_j, t'_j) that is deemed possible and where s'_j *is* a rational choice by t'_j is considered infinitely more likely than any (s''_j, t''_j) where s''_j *is not* a rational choice by t''_j.

As $\omega \in \overline{B}_i^0[rat_j]$ implies that $\succeq^{t_i(\omega)}$ is admissible on $\beta_j^{t_i(\omega)} = [rat_j]_j \cap \kappa_j^{t_i(\omega)}$, it follows that $\overline{B}_i^0[rat_j] \subseteq B_i^0[rat_j]$. Hence, relative to $B_1^0[rat_2] \cap B_2^0[rat_1]$, $\overline{B}_1^0[rat_2] \cap \overline{B}_2^0[rat_1]$ is obtained by imposing minimal completeness, which in this context yields the requirement of 'no extraneous restrictions on beliefs.'

As pointed out in Appendix B, the operator B_i^0 does not satisfy monotonicity since $E \subseteq F$ does not imply $B_i^0 E \subseteq B_i^0 F$. Such non-monotonic operators arise also in other contributions that provide epistemic conditions for forward induction. In particular, Brandenburger and Keisler (2002) use essentially the same operator, which they refer to as 'assumption,' and Battigalli and Siniscalchi (2002) use a non-monotonic operator that they call 'strong belief.' However, in contrast to the use of non-monotonic operators in these contributions, our non-monotonic operator B_i^0 is used *only* to interpret 'full admissible consistency,' while the monotonic operator K_i is used for the *interactive* epistemology. The importance of this will be discussed in Section 7. There we also comment on how the present requirement of 'no extraneous restrictions on beliefs' is related to Brandenburger and Keisler's (2002) and Battigalli and Siniscalchi's (2002) use of a complete epistemic model.

5.2. Allowing extraneous restrictions on beliefs

In view of the previous subsection, we allow extraneous restrictions on beliefs by replacing, for each i, $\overline{B}_i^0[rat_j]$ by $B_i^0[rat_j]$. Hence, let for each i,

$$A_i^0 := [u_i] \cap [cau_i] \cap B_i^0[rat_j],$$

and write $A^0 := A_1^0 \cap A_2^0$. The following result shows that any proper equilibrium is consistent with common certain belief of A^0.[5]

Proposition 5.1. *If $(x_1, x_2) \in \Delta(S_1) \times \Delta(S_2)$ is a proper equilibrium in a finite strategic game G, then, for each i and any $s_i \in \operatorname{supp} x_i$, there exists a belief system with $s_i \in C_i^{t_i(\omega)}$ for some $\omega \in CKA^0$.*

Note that (U, R) is a proper equilibrium in the 'Battle-of-the-sexes-with-an-outside-option' (G_1), while neither U nor R is consistent with common certain belief of full admissible consistency. This demonstrates that 'no extraneous restrictions on beliefs' is necessary for the characterization in Proposition 4.1 of the concept of fully permissible sets, which in G_1 promotes only the forward induction outcome (M, L) (cf. the analysis of G_1 in Sections 3.3 and 6.1).

5.3. Weakening full belief of opponent rationality

To weaken 'full belief' to 'belief,' recall that $\beta_j^{t_i}$ denotes the smallest non-empty set on which $\succeq^{t_i}$ is admissible. If $E \subseteq \Omega$ satisfies that $E = S_i \times \operatorname{proj}_{S_j \times T_i \times T_j} E$, then say that at ω player i *believes* the event E if $\omega \in B_i E$ where

$$B_i E := \left\{ \omega \in \Omega \mid \beta_j^{t_i(\omega)} \subseteq E_j^{t_i(\omega)} \right\}.^{[6]}$$

As shown in Appendix B, we have that $K_i E \subseteq B_i^0 E \subseteq B_i E$. In particular, 'belief' is implied by 'full belief.'

We can now weaken $B_1^0[rat_2] \cap B_2^0[rat_1]$ (i.e., 'full belief of opponent rationality') to $B_1[rat_2] \cap B_2[rat_1]$ (i.e., 'belief of opponent rationality') and weaken $\overline{B}_1^0[rat_2] \cap \overline{B}_2^0[rat_1]$ to $\overline{B}_1[rat_2] \cap \overline{B}_2[rat_1]$, where for each i,

$$\overline{B}_i[rat_j] := \left\{ \omega \in \Omega \mid \beta_j^{t_i(\omega)} \subseteq [rat_j]_j, \text{ and } \mathbf{x} \succ^{t_i(\omega)} \mathbf{y} \text{ only if} \right.$$
$$\left. \mathbf{x}_{\beta_j} \text{ weakly dom. } \mathbf{y}_{\beta_j} \text{ for } \beta_j = \beta_j^{t_i(\omega)} \text{ or } \beta_j = \kappa_j^{t_i(\omega)} \right\}.$$

Relative to $B_1[rat_2] \cap B_2[rat_1]$, $\overline{B}_1[rat_2] \cap \overline{B}_2[rat_1]$ is obtained by imposing minimal completeness, which in the context of 'belief of opponent rationality' yields a requirement of 'no extraneous restrictions on beliefs.'

To impose 'caution' and 'belief of opponent rationality,' let for each i,

$$A_i := [u_i] \cap [cau_i] \cap B_i[rat_j].$$

To add 'no extraneous restrictions on beliefs,' consider for each i,

---

$$\overline{A}_i := [u_i] \cap [cau_i] \cap \overline{B}_i[rat_j],$$

where $\overline{A}_i \subseteq A_i$. Write $A := A_1 \cap A_2$ and $\overline{A} := \overline{A}_1 \cap \overline{A}_2$. Since $\overline{A} \subseteq A$, the following proposition implies that the DF procedure is characterized if 'full belief of opponent rationality' is weakened to 'belief of opponent rationality,' independently of whether a requirement of 'no extraneous restrictions on beliefs' is retained. This shows that 'full belief of opponent rationality' is necessary for the characterization in Proposition 4.1 of the concept of fully permissible sets.

Proposition 5.2. *Consider a finite strategic game G. If a pure strategy p_i for player i is permissible, then there exists a belief system with $p_i \in C_i^{t_i(\omega)}$ for some $\omega \in CK\overline{A}$. A pure strategy p_i for player i is permissible if there exists a belief system with $p_i \in C_i^{t_i(\omega)}$ for some $\omega \in CKA$.*

5.4. Removing caution

Since, for each i, $\overline{B}_i^0[rat_j] \subseteq B_i[rat_j]$, the following result means that the removal of 'caution' leads to a characterization of IESDS, independently of whether extraneous restrictions on beliefs are allowed and full belief of opponent rationality is weakened. Thus, 'caution' is necessary for the characterization in Proposition 4.1.

Proposition 5.3. *Consider a finite strategic game G. If a pure strategy r_i for player i is rationalizable, then there exists a belief system with $r_i \in C_i^{t_i(\omega)}$ for some $\omega \in CK([u_1] \cap [u_2] \cap \overline{B}_1^0[rat_2] \cap \overline{B}_2^0[rat_1])$. A pure strategy r_i for player i is rationalizable if there exists a belief system with $r_i \in C_i^{t_i(\omega)}$ for some $\omega \in CK([u_1] \cap [u_2] \cap B_1[rat_2] \cap B_2[rat_1])$.*

6. Investigating examples

The present section illustrates the concept of fully permissible sets by returning to the previously discussed games G_1 and G_2, as well as by considering one new example. Of the three examples, G_2 will be used to interpret the occurrence of multiple fully permissible sets, while the others will be used to show how our concept captures aspects of forward induction.

All three examples will be used to shed light on the differences between, on the one hand, the approach suggested here and, on the other hand, IEWDS as characterized by Stahl (1995): A strategy survives IEWDS if and only if it is a best response to a belief where one strategy is infinitely less likely than another if the former is eliminated at an earlier round than the latter.[7]

---

[7] Cf. Brandenburger and Keisler (2002, Theorem 1), Battigalli (1996), and Rajan (1998). See also Bicchieri and Schulte (1997), who give conceptually related interpretations of IEWDS.

6.1. Forward induction

Reconsider G_1 of the introduction, and apply our algorithm IECFA to this 'Battle-of-the-sexes-with-an-outside-option' game. Since D is a dominated strategy, D cannot be an element of 1's choice set. This *does not* imply, as in the procedure of IEWDS (given Stahl's, 1995, characterization), that 2 deems M infinitely more likely than D. However, 2 certainly believes that only $\{U\}$, $\{M\}$, and $\{U, M\}$ are candidates for 1's choice set. This excludes $\{R\}$ as 2's choice set, since $\{R\}$ is 2's choice set only if 2 deems $\{D\}$ or $\{U, D\}$ possible. This in turn means that 1 certainly believes that only $\{L\}$ and $\{L, R\}$ are candidates for 2's choice set, implying that $\{U\}$ cannot be 1's choice set. Certainly believing that only $\{M\}$ and $\{U, M\}$ are candidates for 1's choice set *does* imply that 2 deems M infinitely more likely than D. Hence, 2's choice set is $\{L\}$ and, therefore, 1's choice set $\{M\}$. Thus, the forward induction outcome (M, L) is promoted.

To show how common certain belief of the event $\overline{A}^0$ is consistent with the fully permissible sets $\{M\}$ and $\{L\}$—and thus illustrate Proposition 4.1—consider a belief system with only one type of each player; i.e., $T_1 \times T_2 = \{t_1\} \times \{t_2\}$. Let, for each i, $\succeq_i^{t_i}$ satisfy that $\upsilon_i^{t_i} \circ z = u_i$. Furthermore, let

$$\beta_2^{t_1} = \{L\} \times \{t_2\}, \qquad \kappa_2^{t_1} = S_2 \times \{t_2\},$$
$$\beta_1^{t_2} = \{M\} \times \{t_1\}, \qquad \kappa_1^{t_2} = S_1 \times \{t_1\}.$$

Finally, let for each i, $\mathbf{x} \succ^{t_i} \mathbf{y}$ if and only if $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i}$ or $\beta_j = \kappa_j^{t_i}$. Then

$$C_1^{t_1} = \{M\}, \qquad C_2^{t_2} = \{L\}.$$

Inspection will verify that $CK\overline{A}^0 = \overline{A}^0 = \Omega = S \times T_1 \times T_2$.

Turn now to the 'Burning money' game due to van Damme (1989) and Ben-Porath and Dekel (1992). G_5 of Fig. 5 is the pure strategy reduced strategic form of a 'Battle-of-the-sexes' (B-o-s) game with the addition that 1 can publicly destroy 1 unit of payoff before the B-o-s game starts. BU (NU) is the strategy where 1 burns (does not burn), and then plays U, etc., while LR is the strategy where 2 responds with L conditional on 1 not burning and R conditional on 1 burning, etc. The forward induction outcome (supported, e.g., by IEWDS) involves implementation of 1's preferred B-o-s outcome, with *no payoff being burnt*.

One might be skeptical to the use of IEWDS in the 'Burning money' game, because it effectively requires 2 to infer that BU is infinitely more likely than BD based on the sole

	LL	LR	RL	RR
NU	3, 1	3, 1	0, 0	0, 0
ND	0, 0	0, 0	1, 3	1, 3
BU	2, 1	-1, 0	2, 1	-1, 0
BD	-1, 0	0, 3	-1, 0	0, 3

Fig. 5. Burning money (G_5).

Table 1
Applying IECFA to 'Burning money'

$$\begin{aligned}
\varXi(0) &= \Sigma_1 \times \Sigma_2, \\
\varXi(1) &= \big\{\{NU\}, \{ND\}, \{BU\}, \{NU, ND\}, \{ND, BU\}, \{NU, BU\}, \{NU, ND, BU\}\big\} \times \Sigma_2, \\
\varXi(2) &= \big\{\{NU\}, \{ND\}, \{BU\}, \{NU, ND\}, \{ND, BU\}, \{NU, BU\}, \{NU, ND, BU\}\big\} \\
&\quad \times \big\{\{LL\}, \{RL\}, \{LL, LR\}, \{RL, RR\}, \{LL, RL\}, \{LL, LR, RL, RR\}\big\}, \\
\varXi(3) &= \big\{\{NU\}, \{BU\}, \{ND, BU\}, \{NU, BU\}, \{NU, ND, BU\}\big\} \\
&\quad \times \big\{\{LL\}, \{RL\}, \{LL, LR\}, \{RL, RR\}, \{LL, RL\}, \{LL, LR, RL, RR\}\big\}, \\
\varXi(4) &= \big\{\{NU\}, \{BU\}, \{ND, BU\}, \{NU, BU\}, \{NU, ND, BU\}\big\} \\
&\quad \times \big\{\{LL\}, \{RL\}, \{LL, LR\}, \{LL, RL\}\big\}, \\
\varXi(5) &= \big\{\{NU\}, \{BU\}, \{NU, BU\}\big\} \times \big\{\{LL\}, \{RL\}, \{LL, LR\}, \{LL, RL\}\big\}, \\
\varXi(6) &= \big\{\{NU\}, \{BU\}, \{NU, BU\}\big\} \times \big\{\{LL\}, \{LL, LR\}, \{LL, RL\}\big\}, \\
\varXi(7) &= \big\{\{NU\}, \{NU, BU\}\big\} \times \big\{\{LL\}, \{LL, LR\}, \{LL, RL\}\big\}, \\
\varXi(8) &= \big\{\{NU\}, \{NU, BU\}\big\} \times \big\{\{LL\}, \{LL, LR\}\big\}, \\
\varXi(9) &= \big\{\{NU\}\big\} \times \big\{\{LL\}, \{LL, LR\}\big\}, \\
\varPi = \varXi(10) &= \big\{\{NU\}\big\} \times \big\{\{LL, LR\}\big\}
\end{aligned}$$

premise that BD is eliminated before BU, even though all strategies involving burning (i.e., both BU and BD) are eventually eliminated by the procedure. On the basis of this premise such an inference seems at best to be questionable. As shown in Table 1, the application of our algorithm IECFA yields an iteration where at no stage need 2 deem BU infinitely more likely than BD, since $\{NU\}$ is always included as a candidate for 1's choice set. The procedure uniquely determines $\{NU\}$ as 1's fully permissible set and $\{LL, LR\}$ as 2's fully permissible set. Even though the forward induction *outcome* is obtained, 2 does not have any assessment concerning the relative likelihood of opponent strategies conditional on burning; hence, she need not interpret burning as a signal that 1 will play according with his preferred B-o-s outcome.[8]

We can conclude that the concept of fully permissible sets yields the forward induction outcome in G_1 and G_5. Furthermore, the concept promotes forward induction for different reasons than does the procedure of IEWDS (and the concept of EFR, which in the extensive form of these games works like IEWDS).

6.2. Multiple fully permissible sets

Let us also return to G_2 of the introduction, where IEWDS eliminates D in the first round, R in the second round, and M in the third round, so that U and L survive. Stahl's (1995) characterization of IEWDS entails that 2 deems *each* of U and M infinitely more likely than D. Hence, the procedure forces 2 to deem M infinitely more likely than D for the sole reason that D is eliminated before M, even though both M and D are eventually eliminated by the procedure.

Applying our algorithm IECFA yields the following result. Since D is a weakly dominated strategy, D cannot be an element of 1's choice set. Hence, 2 certainly believes that only $\{U\}$, $\{M\}$, and $\{U, M\}$ are candidates for 1's choice set. This excludes $\{R\}$ as

---

[8] Also Battigalli (1991), Asheim (1994), and Dufwenberg (1994) (as well as Hurkens (1996), in a different context) argue that (NU, LR) in addition to (NU, LL) is viable in 'Burning money.'

2's choice set, since $\{R\}$ is 2's choice set only if 2 deems $\{D\}$ or $\{U, D\}$ possible. This in turn means that 1 certainly believes that only $\{L\}$ and $\{L, R\}$ are candidates for 2's choice set, implying that $\{M\}$ cannot be 1's choice set. There is no further elimination. This means that 1's collection of fully permissible sets is $\{\{U\}, \{U, M\}\}$ and 2's collection of fully permissible sets is $\{\{L\}, \{L, R\}\}$. Thus, common certain belief of full admissible consistency implies that 2 deems U infinitely more likely than D since U (respectively, D) is an element of any (respectively, no) fully permissible set for 1. However, whether 2 deems M infinitely more likely than D depends on the type of player 2.

To show how common certain belief of the event $\overline{A}^0$ is consistent with the collections of fully permissible sets $\{\{U\}, \{U, M\}\}$ and $\{\{L\}, \{L, R\}\}$—and thus illustrate Proposition 4.1 also in the case of G_2—consider a belief system with two types of each player; i.e., $T_1 \times T_2 = \{t_1', t_1''\} \times \{t_2', t_2''\}$. Let, for each type t_i of any player i, $\succeq_i^{t_i}$ satisfy that $\upsilon_i^{t_i} \circ z = u_i$. Furthermore, let

$$\beta_2^{t_1'} = \{L\} \times \{t_2'\}, \qquad\qquad \kappa_2^{t_1'} = S_2 \times \{t_2'\},$$
$$\beta_2^{t_1''} = \{(L, t_2'), (L, t_2''), (R, t_2'')\}, \qquad \kappa_2^{t_1''} = S_2 \times T_2,$$
$$\beta_1^{t_2'} = \{(U, t_1'), (U, t_1''), (M, t_1'')\}, \qquad \kappa_1^{t_2'} = S_1 \times T_1,$$
$$\beta_1^{t_2''} = \{U\} \times \{t_1'\}, \qquad\qquad \kappa_1^{t_2''} = S_1 \times \{t_1'\}.$$

Finally, let for each type t_i of any player i, $\mathbf{x} \succ^{t_i} \mathbf{y}$ if and only if $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i}$ or $\beta_j = \kappa_j^{t_i}$. Then

$$C_1^{t_1'} = \{U\}, \qquad C_1^{t_1''} = \{U, M\}, \qquad C_2^{t_2'} = \{L\}, \qquad C_2^{t_2''} = \{L, R\}.$$

Inspection will verify that $CK\overline{A}^0 = \overline{A}^0 = \Omega = S \times T_1 \times T_2$.

Our analysis of G_2 allows a player to deem an opponent choice set to be impossible even when it is the true choice set of the opponent. E.g., at $\omega = (s, t_1, t_2)$ with $t_1 = t_1'$ and $t_2 = t_2''$, player 1 deems it impossible that player 2's choice set is $\{L, R\}$ even though this is the true choice set of player 2. Likewise, at $\omega = (s, t_1, t_2)$ with $t_1 = t_1''$ and $t_2 = t_2''$, player 2 deems it impossible that player 1's choice set is $\{U, M\}$ even though this is the true choice set of player 1. This is an unavoidable feature of this game as there exists no pair of non-empty strategy subsets (Q_1, Q_2) such that $Q_1 = S_1 \backslash D_1(Q_2)$ and $Q_2 = S_2 \backslash D_2(Q_1)$. It implies that under full admissible consistency we cannot have in G_2 that each player is certain of the true choice set of the opponent.

Multiplicity of fully permissible sets arises also in the strategic form of certain extensive games in which the application of backward induction is controversial, e.g., the 'Centipede' game. For more on this, see Asheim and Dufwenberg (2001) where the concept of fully permissible sets is used to analyze extensive games.

7. Related literature

It is instructive to explain how our analysis differs from the epistemic foundations of IEWDS and EFR provided by Brandenburger and Keisler (2002) (BK) and Battigalli and

Siniscalchi (2002) (BS), respectively. It is of minor importance for the comparison that EFR makes use of the extensive form, while the present analysis is performed in the strategic form. The reason is that, by 'caution,' a rational choice in the whole game implies a rational choice at all information sets that are not precluded from being reached by the player's own strategy.

To capture forward induction players must essentially deem any opponent strategy that is a rational choice infinitely more likely than any opponent strategy not having this property. An analysis incorporating this feature must involve a non-monotonic epistemic operator, which is called 'full belief' in the present analysis (cf. Section 5.1), while the corresponding operators are called 'assumption' and 'strong belief' by BK and BS, respectively.

We use 'full belief' only to define the event that the preferences of each player is 'fully admissibly consistent' with the preferences of his opponent, while the monotonic 'certain belief' operator is used for the interactive epistemology:

- each player certainly believes (in the sense of deeming the complement impossible) that the preferences of his opponent are fully admissibly consistent,
- each player certainly believes that his opponent certainly believes that he himself has preferences that are fully admissibly consistent, etc.

As the examples of Section 6 illustrate, it is here a central question what opponent types (choice sets) a player deems possible (i.e., not Savage-null). Consequently, the 'certain belief' operator is appropriate for the interactive epistemology.

In contrast, BK and BS use their non-monotonic operators for the interactive epistemology. In the process of defining higher order beliefs both BK and BS impose that lower order beliefs are maintained. This is precisely how BK obtain Stahl's (1995) characterization which—e.g., in G_2 of the introduction—seems to correspond to extraneous and hard-to-justify restrictions on beliefs.

Stahl's characterization provides an interpretation of IEWDS where strategies eliminated in the first round are completely irrational, while strategies eliminated in later rounds are at intermediate degrees of rationality. Likewise, Battigalli (1996) has shown how EFR corresponds to the 'best rationalization principle,' entailing that some opponent strategies are neither completely rational nor completely irrational. The present analysis, in contrast, differentiates only between whether a strategy is maximal (i.e., a rational choice) or not. As the examples of Section 6 illustrate, although a strategy that is weakly dominated on the set of all opponent strategies is a "stupid" choice, it need not be "more stupid" than any remaining admissible strategy, as this depends on the interactive analysis of the game.

The fact that a non-monotonic epistemic operator is involved when capturing forward induction also means that the analysis must ensure that *all* rational choices for the opponent are included in the epistemic model. BK and BS ensure this by employing *complete* epistemic models, where all possible epistemic types for each player are represented. Instead, the present analysis achieves this by requiring 'no extraneous restrictions on beliefs,' meaning that the preferences are minimally complete (cf. Section 5.1). Since an ordinary monotonic operator is used for the interactive epistemology, there is no more need

for a complete epistemic model here than in usual epistemic analyses of rationalizability and permissibility.

Our paper has a predecessor in Samuelson (1992), who also presents an epistemic analysis of admissibility that leads to a collection of sets for each player, called a 'generalized consistent pair.' Samuelson requires that a player's choice set equals the set of strategies that are not weakly dominated on the union of choice sets that are deemed possible for the opponent; this implies our requirements of 'full belief of opponent rationality' and 'no extraneous restrictions on beliefs' (cf. Samuelson, 1992, p. 311). However, he does not require that each player deems no opponent strategy impossible, as implied by our requirement of 'caution.' Hence, his analysis does not yield $\{\{U, M\}\} \times \{\{L\}\}$ in G_3 of Section 2. Furthermore, he defines possibility relative to a knowledge operator that satisfies the truth axiom, while our analysis—as illustrated by the discussion of G_2 in Section 6.2—allows a player to deem an opponent choice set to be impossible (or more precisely, Savage-null) even when it is the true choice set of the opponent. This explains why we in contrast to Samuelson obtain general existence (cf. Proposition 3.1(i)).

If each player is certain of the true choice set of the opponent, one obtains a 'consistent pair' (cf. Börgers and Samuelson, 1992), a concept that need not exist even when a generalized consistent pair exists. Ewerhart (1998) modifies the concept of a consistent pair by adding 'caution.' However, since he allows extraneous restrictions on beliefs to ensure general existence, his concept of a 'modified consistent pair' does not promote forward induction in G_1. Basu and Weibull's (1991) 'tight curb* set' is another variant of a consistent pair that ensures existence without yielding forward induction in G_1, as they impose 'caution' but weaken 'full belief of opponent rationality' to 'belief of opponent rationality.' In particular, the set of permissible strategy vectors is tight curb*.

'Caution' and 'full belief of opponent rationality' are admissibility requirements on the preferences (or beliefs) of players. Moreover, by imposing 'no extraneous restrictions on beliefs' as a requirement of minimal completeness, preferences are not in general representable by subjective probabilities. By not employing subjective probabilities, the analysis is related to the filter model of belief presented by Brandenburger (1998). By imposing requirements on the preferences of players rather than their choice, our paper follows a tradition in equilibrium analysis where concepts are characterized as equilibria in conjectures (cf. Blume et al., 1991b). This approach can also be used to characterize backward induction and to define rationalizability analogs to sequential, quasi-perfect and proper equilibrium (cf. Asheim, 2001, 2002, and Asheim and Perea, 2002).

Appendix A. The decision-theoretic framework

The purpose of this appendix is to present the decision-theoretic terminology, notation and results utilized and referred to in the main text.

Consider a decision-maker, and let F be a finite set of states. The decision-maker is uncertain about what state in F will be realized. Let Z be a finite set of outcomes. In the tradition of Anscombe and Aumann (1963), the decision-maker is endowed with a binary relation over all functions that to each element of F assigns an objective randomization on Z. Any such function $\mathbf{x}_F : F \to \Delta(Z)$ is called an *act* on F. Write $\mathbf{x}_F$ and $\mathbf{y}_F$ for acts on F. A *reflexive* and *transitive* binary relation on the set of acts on F is denoted by $\succeq_F$, where $\mathbf{x}_F \succeq_F \mathbf{y}_F$ means that $\mathbf{x}_F$ is *preferred* or *indifferent* to $\mathbf{y}_F$. As usual, let $\succ_F$ (*preferred to*) and $\sim_F$ (*indifferent to*) denote the asymmetric and symmetric parts of $\succeq_F$.

Assume that $\succeq_F$ satisfies *objective independence*, *nontriviality*, *conditional continuity* (or "the conditional Archimedean property"), and *non-null state independence*, where these terms are defined in Blume et al. (1991a, pp. 64–65). Let, for any $E \subseteq F$, $\succeq_E$ denote the *conditional* binary relation. Weaken *completeness* to *conditional completeness*: $\forall f \in F$, $\succeq_{\{f\}}$ is complete.

Say that the state $f \in F$ is *Savage-null* if $\mathbf{x}_F \sim_{\{f\}} \mathbf{y}_F$ for all acts $\mathbf{x}_F$ and $\mathbf{y}_F$ on F. If $e, f \in F$ and $\succeq_F$ is conditionally complete, then e is deemed *infinitely more likely* than f ($e \gg f$) if e is not Savage-null and $\mathbf{x}_F \succ_{\{e\}} \mathbf{y}_F$ implies $(\mathbf{x}_{-\{f\}}, \mathbf{x}'_{\{f\}}) \succ_{\{e,f\}} (\mathbf{y}_{-\{f\}}, \mathbf{y}'_{\{f\}})$ for all $\mathbf{x}'_F, \mathbf{y}'_F$. Say that $\mathbf{y}_F$ is *maximal* w.r.t. $\succeq_E$ if there is no $\mathbf{x}_F$ such that $\mathbf{x}_F \succ_E \mathbf{y}_F$.

If $\upsilon : Z \to \Re$ is a vNM utility function, abuse notation slightly by writing $\upsilon(x) = \sum_{z \in Z} x(z)\upsilon(z)$ whenever $x \in \Delta(Z)$ is an objective randomization. Say that $\mathbf{x}_E$ *strongly dominates* $\mathbf{y}_E$ w.r.t. υ if, $\forall f \in E$, $\upsilon(\mathbf{x}_E(f)) > \upsilon(\mathbf{y}_E(f))$. Say that $\mathbf{x}_E$ *weakly dominates* $\mathbf{y}_E$ w.r.t. υ if, $\forall f \in E$, $\upsilon(\mathbf{x}_E(f)) \geqslant \upsilon(\mathbf{y}_E(f))$, with strict inequality for some $e \in E$. Say that $\succeq_F$ is *admissible* on E ($\neq \emptyset$) if $\mathbf{x}_F \succ_F \mathbf{y}_F$ whenever $\mathbf{x}_E$ weakly dominates $\mathbf{y}_E$.

The following representation result—which follows directly from the von Neumann–Morgenstern theorem on expected utility representation—requires the notion of conditional representation: Say that $\succeq_F$ is *conditionally represented* by υ if (a) $\succeq_F$ is nontrivial and (b) $\mathbf{x}_F \succeq_{\{f\}} \mathbf{y}_F$ iff $\upsilon(\mathbf{x}_F(f)) \geqslant \upsilon(\mathbf{y}_F(f))$ whenever f is not Savage-null.

Proposition A.1. *If $\succeq_F$ is reflexive and transitive, and satisfies objective independence, nontriviality, conditional completeness, conditional continuity, and non-null state independence, then there exists a vNM utility function $\upsilon : Z \to \Re$ such that $\succeq_F$ is conditionally represented by υ.*

If $F = F_1 \times F_2$ and $\succeq_F$ is a binary relation on the set of acts on F, then say that $\succeq_{F_1}$ is the *marginal* of $\succeq_F$ on F_1 if, $\mathbf{x}_{F_1} \succeq_{F_1} \mathbf{y}_{F_1}$ iff $\mathbf{x}_F \succeq_F \mathbf{y}_F$ whenever $\mathbf{x}_{F_1}(f_1) = \mathbf{x}_F(f_1, f_2)$ and $\mathbf{y}_{F_1}(f_1) = \mathbf{y}_F(f_1, f_2)$ for all (f_1, f_2).

Appendix B. Properties of epistemic operators

With $\{\beta_j^{t_i}, \ldots, \kappa_j^{t_i}\}$ as the collection of nested sets on which $\succeq^{t_i}$ is admissible, we have in Sections 4.5, 5.1, and 5.3 defined the K_i ('certain belief'), B_i^0 ('full belief'), and B_i ('belief') operators on $\{E \subseteq \Omega \mid E = S_i \times \text{proj}_{S_j \times T_i \times T_j} E\}$. We have that, in general, $K_i E \subseteq B_i^0 E \subseteq B_i E$ since

- $\kappa_j^{t_i(\omega)} \subseteq E_j^{t_i(\omega)}$ implies that $\succeq^{t_i(\omega)}$ is admissible on $\kappa_j^{t_i(\omega)} = E_j^{t_i(\omega)} \cap \kappa_j^{t_i(\omega)}$,
- $\succeq^{t_i(\omega)}$ is admissible on $E_j^{t_i(\omega)} \cap \kappa_j^{t_i(\omega)}$ implies that $\beta_j^{t_i(\omega)} \subseteq E_j^{t_i(\omega)}$.

Hence, 'certain belief' implies 'full belief,' which in turn implies 'belief.'

If $\succeq^{t_i}$ in addition to Assumption 4.1 also satisfies completeness and continuity—so that $\succeq^{t_i}$ is represented by $\upsilon_i^{t_i}$ and a subjective probability distribution $\mu^{t_i} \in \Delta(S_j \times T_j)$—then supp μ^{t_i} is the unique set on which $\succeq^{t_i}$ is admissible. In this special case we have that the three operators coincide: $K_i E = B_i^0 E = B_i E$. Hence, they are all generalizations of 'belief with probability 1.'

For this reason, it follows that none of the three operators satisfy the truth axiom (i.e., $K_i E \subseteq E$, $B_i^0 E \subseteq E$, and $B_i E \subseteq E$ need not hold).

It is straightforward to verify that, in general, K_i and B_i correspond to KD45 systems. However, even though the operator B_i^0 satisfies $B_i^0 \emptyset = \emptyset$, $B_i^0 \Omega = \Omega$, and $B_i^0 E \cap B_i^0 F \subseteq B_i^0 (E \cap F)$ as well as positive and negative introspection, it does not satisfy monotonicity: $E \subseteq F$ does not imply $B_i^0 E \subseteq B_i^0 F$. To see this, consider G_3 of Section 2: If 2 prefers any strategy that (weakly) dominates another on $\{U\}$, regardless of what happen outside $\{U\}$, then it does not follow that 2 prefer any strategy that weakly dominates another on $\{U, M\}$, regardless of what happens outside $\{U, M\}$, since weak dominance on $\{U, M\}$ does not imply (weak) dominance on $\{U\}$. This is illustrated by L and R: $\omega \in B_2^0(\{(s_1, s_2, t_1, t_2) \mid s_1 = U\})$ does *not* imply that $t_2(\omega)$ prefers L to R, while $\omega \in B_2^0(\{(s_1, s_2, t_1, t_2) \mid s_1 \in \{U, M\}\})$ *does* imply that $t_2(\omega)$ prefers L to R.

For the proofs of Propositions 4.1 and 5.2, we need to establish some properties of iterated mutual certain belief. With $KE = K_1 E \cap K_2 E$ as the mutual certain belief operator defined on $\{E \subseteq \Omega \mid E = S_i \times S_j \times \text{proj}_{T_i \times T_j} E\}$, write $K^0 E := E$ and, for each $g \geqslant 1$, $K^g E := KK^{g-1}E$. Since $K_i(E \cap F) = K_i E \cap K_i F$, and since $K_i \emptyset = \emptyset$, conjunction, and positive and negative introspection imply $K_i K_i E = K_i E$, it follows $\forall g \geqslant 2$, $K^g E = K_1 K^{g-1} E \cap K_2 K^{g-1} E \subseteq K_1 K_1 K^{g-2} E \cap K_2 K_2 K^{g-2} E = K_1 K^{g-2} E \cap K_2 K^{g-2} E = K^{g-1} E$. Even though the truth axiom ($K_i E \subseteq E$) is not satisfied, the present paper considers certain belief only of events $E \subseteq \Omega$ that can be written as $E = E_1 \cap E_2$ where, for each i, $E_i = S_i \times S_j \times \text{proj}_{T_i} E \times T_j$. Mutual certain belief of any such event E implies that E is true: $KE = K_1 E \cap K_2 E \subseteq K_1 E_1 \cap K_2 E_2 = E_1 \cap E_2 = E$ since, for each i, $K_i E_i = E_i$. Hence,

(i) $\forall g \geqslant 1$, $K^g E \subseteq K^{g-1} E$, and
(ii) $\exists g' \geqslant 0$ such that $K^g E = KE \cap KKE \cap KKKE \cap \cdots = CKE$ for $g \geqslant g'$ since Ω is finite, implying $CKE = KCKE$.

Appendix C. Proofs

Proof of Proposition 3.1. Standard results given that Σ is finite, and, $\forall i \in I$, α_i is monotone: $\emptyset \neq \Xi'_j \subseteq \Xi''_j \subseteq \Sigma_j$ implies $\emptyset \neq \alpha_i(\Xi'_j) \subseteq \alpha_i(\Xi''_j) \subseteq \alpha_i(\Sigma_j)$. $\square$

To define the concept of permissible strategies, we use the equivalent Dekel–Fudenberg procedure as the primitive definition. For any $(\emptyset \neq) X = X_1 \times X_2 \subseteq S$, write $\tilde{a}(X) := \tilde{a}_1(X_2) \times \tilde{a}_2(X_1)$, where

$$\tilde{a}_i(X_j) := S_i \setminus \{s_i \in S_i \mid \exists x_i \in \Delta(S_i) \text{ s.t. } x_i \text{ strongly dom. } s_i \text{ on } X_j \text{ or } x_i \text{ weakly dom. } s_i \text{ on } S_j\}.$$

Definition C.1. Consider the sequence defined by $X(0) = S$ and, $\forall g \geqslant 1$, $X(g) = \tilde{a}(X(g-1))$. A pure strategy p_i for player i is said to be *permissible* if $p_i \in \bigcap_{g=0}^{\infty} X_i(g)$.

Let $P = P_1 \times P_2$ denote the set of permissible strategy vectors. To characterize P, write for any $(\emptyset \neq) X = X_1 \times X_2 \subseteq S$, $a(X) := a_1(X_2) \times a_2(X_1)$, where

$$a_i(X_j) := \{p_i \in S_i \mid \exists (\emptyset \neq) Q_j \subseteq X_j \text{ s.t. } p_i \in S_i \setminus D_i(Q_j)\}.$$

Lemma C.1. *For any* $(\emptyset \neq) X_j \subseteq S_j$, $a_i(X_j) = \tilde{a}_i(X_j)$.

Proof. The proof is available on request from the authors. $\square$

Proposition C.1. (i) *The sequence defined by* $X(0) = S$ *and,* $\forall g \geqslant 1$, $X(g) = a(X(g-1))$ *converges to* P *in a finite number of iterations.*
   (ii) $\forall i \in I$, $P_i \neq \emptyset$.
   (iii) $P = a(P)$.
   (iv) $\forall i \in I$, $p_i \in P_i$ *if and only if there exists* $X = X_1 \times X_2$ *with* $p_i \in X_i$ *such that* $X \subseteq a(X)$.

Proof. Part (i) follows from Lemma C.1 given that S is finite, and, $\forall i \in I$, a_i is monotone: $\emptyset \neq X'_j \subseteq X''_j \subseteq S_j$ implies $\emptyset \neq a_i(X'_j) \subseteq a_i(X''_j) \subseteq a_i(S_j)$. Parts (ii)–(iv) are then standard results. $\square$

Proof of Proposition 3.2. Using Proposition 3.1(ii), the definitions of $\alpha(\cdot)$ and $a(\cdot)$ imply, $\forall i \in I$, $P_i^0 := \bigcup_{\sigma_i \in \Pi_i} \sigma_i = \bigcup_{\sigma_i \in \alpha_i(\Pi_j)} \sigma_i \subseteq a_i(P_j^0)$. Since $P^0 \subseteq a(P^0)$ implies $P^0 \subseteq P$ (by Proposition C.1(iv)), it follows that, $\forall i \in I$, $\bigcup_{\sigma_i \in \Pi_i} \sigma_i \subseteq P_i$. $\square$

Proof of Proposition 4.1. *Part 1: If π_i is fully permissible, then there exists a belief system with $\pi_i = C_i^{t_i(\omega)}$ for some $\omega \in CK\overline{A}^0$.* By Proposition 3.1(i) it is sufficient to construct a belief system with $\overline{A}^0 = \Omega = S \times T_1 \times T_2$ such that, $\forall i \in I$, $\forall \pi_i \in \Pi_i$, there exists $t_i \in T_i$ with $\pi_i = C_i^{t_i}$. Construct a belief system with, $\forall i \in I$, a bijection $\sigma_i : T_i \to \Pi_i$ from the set of types to the collection of fully permissible sets. By Proposition 3.1(ii) we have that, $\forall i \in I$, $\forall t_i \in T_i$, $\exists \Psi_j^{t_i} \subseteq \Pi_j$ such that $\sigma_i(t_i) = S_i \backslash D_i(Q_j^{t_i})$, where $Q_j^{t_i} := \{s_j \in S_j \mid \exists \sigma_j \in \Psi_j^{t_i} \text{ s.t. } s_j \in \sigma_j\}$. Determine the set of opponent types that t_i does not deem Savage-null as follows: $T_j^{t_i} = \{t_j \in T_j \mid \sigma_j(t_j) \in \Psi_j^{t_i}\}$. Let, $\forall i \in I$, $\forall t_i \in T_i$, $\succeq^{t_i}$ satisfy

(1) $\upsilon_i^{t_i} \circ z = u_i$, which means that, $\forall i \in I$, $\Omega \subseteq [u_i]$, and
(2) $\mathbf{x} \succ^{t_i} \mathbf{y}$ iff $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i} = \{(s_j, t_j) | s_j \in \sigma_j(t_j) \text{ and } t_j \in T_j^{t_i}\}$ or $\beta_j = \kappa_j^{t_i} = S_j \times T_j^{t_i}$, which implies, $\forall i \in I$, $\Omega \subseteq [cau_i]$.

This means that $C_i^{t_i} = S_i \backslash D_i(Q_j^{t_i}) = \sigma_i(t_i)$ since $\mathbf{x}_{S_j} \succ_{S_j}^{t_i} \mathbf{y}_{S_j}$ iff $\mathbf{x}_{S'_j}$ weakly dominates $\mathbf{y}_{S'_j}$ for $S'_j = Q_j^{t_i}$ or $S'_j = S_j$. This in turn implies, $\forall i \in I$, $\forall t_i \in T_i$,

(3) $\beta_j^{t_i} = [rat_j]_j \cap \kappa_j^{t_i}$, which combined with (2) yields, $\forall i \in I$, $\Omega \subseteq \overline{B}_j^0[rat_j]$.

It follows that, $\forall i \in I$, $\Omega \subseteq [u_i] \cap [cau_i] \cap \overline{B}_i^0[rat_j] = \overline{A}_i^0$, implying $\overline{A}^0 = \Omega$.

  Part 2: If there exists a belief system with $\pi_i = C_i^{t_i(\omega)}$ for some $\omega \in CK\overline{A}^0$, then π_i is fully permissible. Consider any belief system for which $CK\overline{A}^0 \neq \emptyset$. Let, $\forall i \in I$, $T'_i := \{t_i(\omega) \mid \omega \in CK\overline{A}^0\}$ and $\Xi_i := \{C_i^{t_i} \mid t_i \in T'_i\}$. Note that, $\forall i \in I$ and $\forall t_i \in T'_i$, (s_j, t_j) is Savage-null acc. to $\succeq^{t_i}$ if $t_j \in T_j \backslash T'_j$ since $CK\overline{A}^0 = KCK\overline{A}^0 \subseteq K_i CK\overline{A}^0$, implying $T_j^{t_i} \subseteq T'_j$. Since, $\forall i \in I$ and $t_i \in T'_i$, $\mathbf{x} \succ^{t_i} \mathbf{y}$ iff $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i} = [rat_j]_j \cap \kappa_j^{t_i}$ or $\beta_j = \kappa_j^{t_i} = S_j \times T_j^{t_i}$, it follows that $\mathbf{x}_{S_j} \succ_{S_j}^{t_i} \mathbf{y}_{S_j}$ iff $\mathbf{x}_{S'_j}$ weakly dominates $\mathbf{y}_{S'_j}$ for $S'_j = Q_j^{t_i}$ or $S'_j = S_j$, where $Q_j^{t_i} := \{s_j \in S_j \mid \exists \sigma_j \in \Psi_j^{t_i} \text{ s.t. } s_j \in \sigma_j\}$ and $\Psi_j^{t_i} := \{C_j^{t_j} \mid t_j \in T_j^{t_i}\} \subseteq \Xi_j$. This implies, $\forall i \in I$ and $t_i \in T'_i$, $C_i^{t_i} = S_i \backslash D_i(Q_j^{t_i})$, and $\Xi \subseteq \alpha(\Xi)$. Hence, by Proposition 3.1(iii), $\pi_i \in \Pi_i$ if there exists a belief system with $\pi_i = C_i^{t(\omega)}$ for some $\omega \in CK\overline{A}^0$. $\square$

Proof of Proposition 5.1. It is sufficient to show that one can construct a belief system with $A^0 = \Omega = S \times \{t_1\} \times \{t_2\}$ such that, $\forall i \in I$, $\text{supp}\, x_i \subseteq C_i^{t_i}$, whenever (x_1, x_2) is a proper equilibrium. Let (x_1, x_2) be a proper equilibrium. By Blume et al. (1991b, Proposition 5), there exists a pair of preferences, $\succeq^{t_1}$ and $\succeq^{t_2}$, that are represented by $\upsilon_1^{t_1}$ and $\lambda^{t_1} = (\mu_1^{t_1}, \dots) \in \mathbf{L}\Delta(S_2 \times \{t_2\})$, and $\upsilon_2^{t_2}$ and $\lambda^{t_2} = (\mu_1^{t_2}, \dots) \in \mathbf{L}\Delta(S_1 \times \{t_1\})$, respectively—with $\upsilon_1^{t_1} \circ z = u_1$ and, $\forall s_2 \in S_2$, $\mu_1^{t_1}(s_2, t_2) = x_2(s_2)$, and $\upsilon_2^{t_2} \circ z = u_2$ and, $\forall s_1 \in S_1$, $\mu_1^{t_2}(s_1, t_1) = x_1(s_1)$—satisfying, $\forall i \in I$,

(i) $\text{supp}\, x_i \subseteq C_i^{t_i}$,
(ii) $\kappa_j^{t_i} = S_j \times \{t_j\}$, and
(iii) $(r_j, t_j) \gg (s_j, t_j)$ whenever $r_j \succ^{t_j} s_j$.

Properties (ii) and (iii) imply that $\succeq^{t_i}$ is admissible on $C_j^{t_j} \times \{t_j\} = [rat_j]_j \cap \kappa_j^{t_i}$. By letting $\Omega = S \times \{t_1\} \times \{t_2\}$, it follows that, $\forall i \in I$, $\Omega \subseteq [u_i] \cap [cau_i] \cap B_i^0[rat_j] = A_i^0$. Hence, $A^0 = \Omega$ and, by property (i), $\forall i \in I$, $\text{supp}\, x_i \subseteq C_i^{t_i}$. $\square$

Proof of Proposition 5.2. *Part* 1: *If* p_i *is permissible, then there exists a belief system with* $p_i \in C_i^{t_i(\omega)}$ *for some* $\omega \in CK\overline{A}$. By Proposition C.1(ii) it is sufficient to construct a belief system with $\overline{A} = \Omega = S \times T_1 \times T_2$ such that, $\forall i \in I$, $\forall p_i \in P_i$, there exists $t_i \in T_i$ with $p_i \in C_i^{t_i}$. Construct a belief system with, $\forall i \in I$, a bijection $\mathbf{s}_i : T_i \to P_i$ from the set of types to the set of permissible strategies. By Proposition C.1(iii) we have that, $\forall i \in I$, $\forall t_i \in T_i$, $\exists Q_j^{t_i} \subseteq P_j$ such that $\mathbf{s}_i(t_i) \in S_i \backslash D_i(Q_j^{t_i})$. Determine the set of opponent types that t_i does not deem Savage-null as follows: $T_j^{t_i} = \{t_j \in T_j \mid \mathbf{s}_j(t_j) \in Q_j^{t_i}\}$. Let, $\forall i \in I$, $\forall t_i \in T_i$, $\succeq^{t_i}$ satisfy

(1) $v_i^{t_i} \circ z = u_i$, which means that, $\forall i \in I$, $\Omega \subseteq [u_i]$, and
(2) $\mathbf{x} \succ^{t_i} \mathbf{y}$ iff $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i} = \{(s_j, t_j) \mid s_j = \mathbf{s}_j(t_j)$ and $t_j \in T_j^{t_i}\}$ or $\beta_j = \kappa_j^{t_i} = S_j \times T_j^{t_i}$, which implies, $\forall i \in I$, $\Omega \subseteq [cau_i]$.

This means that $C_i^{t_i} = S_i \backslash D_i(Q_j^{t_i}) \ni \mathbf{s}_i(t_i)$ since $\mathbf{x}_{S_j} \succ_{S_j}^{t_i} \mathbf{y}_{S_j}$ iff $\mathbf{x}_{S_j'}$ weakly dominates $\mathbf{y}_{S_j'}$ for $S_j' = Q_j^{t_i}$ or $S_j' = S_j$. This in turn implies, $\forall i \in I$, $\forall t_i \in T_i$,

(3) $\beta_j^{t_i} \subseteq [rat_j]_j$, which combined with (2) yields, $\forall i \in I$, $\Omega \subseteq \overline{B}_i[rat_j]$.

It follows that, $\forall i \in I$, $\Omega \subseteq [u_i] \cap [cau_i] \cap \overline{B}_i[rat_j] = \overline{A}_i$, implying $\overline{A} = \Omega$.

Part 2: *If there exists a belief system with* $p_i \in C_i^{t_i(\omega)}$ *for some* $\omega \in CKA$, *then* p_i *is permissible.* Consider any belief system for which $CKA \neq \emptyset$. Let, $\forall i \in I$, $T_i' := \{t_i(\omega) \mid \omega \in CKA\}$ and $X_i := \bigcup_{t_i \in T_i'} C_i^{t_i}$. Note that, $\forall i \in I$ and $\forall t_i \in T_i'$, (s_j, t_j) is Savage-null acc. to $\succeq^{t_i}$ if $t_j \in T_j \backslash T_j'$ since $CKA = KCKA \subseteq K_iCKA$, implying $T_j^{t_i} \subseteq T_j'$. Since, $\forall i \in I$ and $t_i \in T_i'$, $\mathbf{x} \succ^{t_i} \mathbf{y}$ if $\mathbf{x}_{\beta_j}$ weakly dominates $\mathbf{y}_{\beta_j}$ for $\beta_j = \beta_j^{t_i} \subseteq [rat_j]_j \cap \kappa_j^{t_i}$ or $\beta_j = \kappa_j^{t_i} = S_j \times T_j^{t_i}$, it follows that $\mathbf{x}_{S_j} \succ_{S_j}^{t_i} \mathbf{y}_{S_j}$ if $\mathbf{x}_{S_j'}$ weakly dominates $\mathbf{y}_{S_j'}$ for $S_j' = Q_j^{t_i}$ or $S_j' = S_j$, where $Q_j^{t_i} := \mathrm{proj}_{S_j} \beta_j^{t_i} \subseteq X_j$. This implies, $\forall i \in I$ and $t_i \in T_i'$, $C_i^{t_i} \subseteq S_i \backslash D_i(Q_j^{t_i})$, and $X \subseteq a(X)$. Hence, by Proposition C.1(iv), $p_i \in P_i$ if there exists a belief system with $p_i \in C_i^{t(\omega)}$ for some $\omega \in CKA$. □

Proof of Proposition 5.3. The proof is available on request from the authors. □

References

Anscombe, F.J., Aumann, R., 1963. A definition of subjective probability. Ann. of Math. Stat. 34, 199–205.

Asheim, G.B., 1994, Defining rationalizability in 2-player extensive games. Memorandum No. 25/94, Dept. of Economics, University of Oslo.

Asheim, G.B., 2001. Proper rationalizability in lexicographic beliefs. Int. J. Game Theory 30, 453–478.

Asheim, G.B., 2002. On the epistemic foundation for backward induction. Math. Soc. Sci. 44, 121–144.

Asheim, G.B., Dufwenberg, M., 2001. Deductive reasoning in extensive games, Econ. J. Forthcoming.

Asheim, G.B., Perea, A., 2002. Sequential and quasi-perfect rationalizability in extensive games. University of Oslo and University of Maastricht.

Aumann, R., 1962. Utility theory without the completeness axiom. Econometrica 30, 445–462.

Basu, K., Weibull, J.W., 1991. Strategy subsets closed under rational behavior. Econ. Lett. 36, 141–146.

Battigalli, P., 1991. Algorithmic solutions for extensive games. In: Ricci, G. (Ed.), Decision Processes in Economics. Springer-Verlag, Berlin.

Battigalli, P., 1996. Strategic rationality orderings and the best rationalization principle. Games Econ. Behav. 13, 178–200.

Battigalli, P., 1997. On rationalizability in extensive games. J. Econ. Theory 74, 40–61.

Battigalli, P., Siniscalchi, M., 2002. Strong belief and forward induction reasoning. J. Econ. Theory 106, 356–391.

Ben-Porath, E., 1997. Rationality, Nash equilibrium, and backwards induction in perfect information games. Rev. Econ. Stud. 64, 23–46.

Ben-Porath, E., Dekel, E., 1992. Coordination and the potential for self-sacrifice. J. Econ. Theory 57, 36–51.

Bernheim, D., 1984. Rationalizable strategic behavior. Econometrica 52, 1007–1028.

Bicchieri, C., Schulte, O., 1997. Common reasoning about admissibility. Erkenntnis 45, 299–325.

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

Blume, L., Brandenburger, A., Dekel, E., 1991b. Lexicographic probabilities and equilibrium refinements. Econometrica 59, 81–98.

Böge, W., Eisele, T., 1979. On the solutions of Bayesian games. Int. J. Game Theory 8, 193–215.

Börgers, T., 1994. Weak dominance and approximate common knowledge. J. Econ. Theory 64, 265–276.

Börgers, T., Samuelson, L., 1992. 'Cautious' utility maximization and iterated weak dominance. Int. J. Game Theory 21, 13–25.

Brandenburger, A., 1992. Lexicographic probabilities and iterated admissibility. In: Dasgupta, P., Gale, D., Hart, O., Maskin, E. (Eds.), Economic Analysis of Markets and Games. MIT Press, Cambridge, MA, pp. 282–290.

Brandenburger, A., 1998. On the existence of a 'complete' belief model. Harvard Business School Working Paper 99-056.

Brandenburger, A., Dekel, E., 1993. Hierarchies of beliefs and common knowledge. J. Econ. Theory 59, 189–198.

Brandenburger, A., Keisler, H.J., 1999. An impossibility theorem on beliefs in games. Harvard Business School Working Paper 00-010.

Brandenburger, A., Keisler, H.J., 2002. Epistemic conditions for iterated admissibility. Harvard Business School.

Dekel, E., Fudenberg, D., 1990. Rational behavior with payoff uncertainty. J. Econ. Theory 52, 243–267.

Dekel, E., Gul, F., 1997. Rationality and knowledge. In: Kreps, D.M., Wallis, K.F. (Eds.), Advances in Economics and Econometrics: Theory and Applications. Cambridge Univ. Press, Cambridge, UK, pp. 87–172.

Dufwenberg, M., 1994. Tie-break rationality and tie-break rationalizability. Working Paper 1994:29, Dept. of Economics, Uppsala University.

Epstein, L.G., Wang, T., 1996. 'Beliefs about beliefs' without probabilities. Econometrica 64, 1343–1373.

Ewerhart, C., 1998. Rationality and the definition of consistent pairs. Int. J. Game Theory 27, 49–59.

Gul, F., 1997. Rationality and coherent theories of strategic behavior. J. Econ. Theory 70, 1–31.

Hurkens, S., 1996. Multi-sided pre-play communication by burning money. J. Econ. Theory 69, 186–197.

Luce, D., Raiffa, H., 1957. Games and Decisions. Wiley, New York, NY.

Mariotti, M., 1997. Decisions in games: why there should be a special exemption from Bayesian rationality. J. Econ. Meth. 4, 43–60.

Mertens, J.-M., Zamir, S., 1985. Formulation of Bayesian analysis for games of incomplete information. Int. J. Game Theory 14, 1–29.

Morris, S., 1997. Alternative notions of belief. In: Bacharach, M., Gérard-Varet, L.-A., Mongin, Ph., Shin, H. (Eds.), Epistemic Logic and the Theory of Games and Decisions. Kluwer, Dordrecht, pp. 217–234.

Pearce, D.G., 1984. Rationalizable strategic behavior and the problem of perfection. Econometrica 52, 1029–1050.

Rajan, U., 1998. Trembles in the Bayesian foundations of solution concepts of games. J. Econ. Theory 82, 248–266.

Samuelson, L., 1992. Dominated strategies and common knowledge. Games Econ. Behav. 4, 284–313.

Stahl, D., 1995. Lexicographic rationality, common knowledge, and iterated admissibility. Econ. Lett. 47, 155–159.

Tan, T., Werlang, S.R.C., 1988. The Bayesian foundations of solution concepts of games. J. Econ. Theory 45, 370–391.

van Damme, E., 1989. Stable equilibria and forward induction. J. Econ. Theory 48, 476–496.