Common belief of weak-dominance rationality in strategic-form games: a qualitative analysis*

Giacomo Bonanno†
University of California, Davis

Elias Tsakas‡
Maastricht University

August 22, 2018

Abstract

We study common belief of weak-dominance rationality in strategic-form games with ordinal utilities, employing a qualitative model of beliefs. We characterize two standard solution concepts for such games: the Iterated Deletion of Börgers-dominated Strategies (IDBS) and the Iterated Deletion of Inferior Strategy Profiles (IDIP). We do so by imposing nested restrictions on the doxastic models: namely, the respective epistemic conditions differ in the fact that IDIP requires the truth axiom whereas IDBS does not. Hence, IDIP refines IDBS.

*We are indebted to two anonymous referees, an associate editor and the audience of the LOFT 2018 Conference for their helpful comments and suggestions. Elias thanks the Department of Economics at UC Davis for its hospitality while working on this project. The paper previously circulated under the title "Qualitative analysis of common belief of rationality in strategic-form games".

†Department of Economics, University of California, Davis, CA 95616-8578, USA; Homepage: http://faculty.econ.ucdavis.edu/faculty/bonanno/; E-mail: gfbonanno@ucdavis.edu

‡Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Homepage: http://www.elias-tsakas.com/home.html; E-mail: e.tsakas@maastrichtuniversity.nl

1. Introduction

Traditionally, game-theoretic analysis has been based on the assumption that the game under consideration is common knowledge among the players. That is, besides postulating that the rules of the game (i.e., the set of players, the set of strategies and the outcome which is associated with each strategy profile) are commonly known, we typically assume that *the players have vNM preferences* and that *these preferences are also commonly known*.[1] Under these assumptions, common belief of rationality characterizes correlated rationalizability; that is, the strategy profiles that survive the Iterated Deletion of Strictly Dominated Strategies are exactly those that can be rationally played under common belief of (Bayesian) rationality (e.g., see Brandenburger and Dekel, 1987; Tan and Werlang, 1988).

While it is certainly reasonable to assume that the rules of the game are commonly known, the last two assumptions seem harder to justify at the outset. The issue with the preferences being commonly known has already been addressed by Harsanyi (1967-68) and the extensive literature on incomplete information games that followed his seminal contribution. Within Harsanyi's extended model, rationality and common belief of (Bayesian) rationality characterizes interim correlated rationalizability (e.g., see Dekel et al., 2007; Ely and Peski, 2006). However, in Harsanyi's program, preferences are still assumed to be vNM and therefore the utilities of the game outcomes remain cardinal.

Relaxing this assumption can be motivated not only from a theoretical, but also from an applied point of view, given that in lab experiments we typically test predictions obtained by employing solution concepts for games with ordinal utilities. From a theoretical standpoint, the main consequence of postulating only ordinal utilities is that we have to replace the usual models of probabilistic beliefs with Kripke structures, and thus abandon the standard notion of Bayesian rationality. The two usual alternatives are a weak notion of rationality which is typically employed to characterize Iterated Deletion of Strictly Dominated Strategies (see Section 4) and a strong notion of rationality, often called weak-dominance rationality (Hillas and Samet, 2014). In this paper we investigate the content of the notion of *common belief of (weak-dominance) rationality in strategic-form games* with ordinal payoffs.

[1] These assumptions are consistent with extending Savage's standard decision-theoretic framework to an interactive setting (e.g., see Epstein and Wang, 1996).

In particular, we consider qualitative doxastic models, which consist, for each player, of a belief operator (which is represented by a Kripke structure) and a specification of the set of the opponents' strategy profiles deemed possible at each state. Then, a strategy is said to be weak-dominance rational (henceforth, simply called *rational*) at some state whenever it is *not weakly dominated relative to the opponents' strategy profiles that are deemed possible*. Within this model, we provide a full characterization of two standard solution concepts for strategic-form games with ordinal utilities, viz., Iterated Deletion of Börgers-dominated Strategies (Börgers, 1993, henceforth IDBS) and Iterated Deletion of Inferior Strategy Profiles (the pure-strategy version of strong rationalizability à la Stalnaker, 1994, henceforth IDIP). In particular, we show that IDBS is characterized by common belief in rationality within the class of models that satisfy Consistency (Theorem 1), while IDIP is characterized by common belief in rationality within the sub-class of models that satisfy Truth (Theorem 2). The previous results complete the epistemic analysis of strategic-form games with ordinal utilities, by providing not only sufficient but also necessary conditions.[2]

With the above-mentioned results, not only do we manage to put under the same umbrella two main solution concepts for games with ordinal utilities, but we also manage to prove that they monotonically refine each other (Corollary 1): our results imply that IDIP refines IDBS.[3] As we will point out, while one can prove the same result algorithmically via the corresponding procedures, our approach allows us to epistemically pin down the reason for the aforementioned refinement. In particular, the difference between the two concepts is attributed to the Truth Axiom.

The paper is structured as follows: In Section 2 we introduce the notion qualitative doxastic model of a strategic-form game; in Section 3 we define our notion of rationality and prove our characterization results; in Section 4 we present some additional results; Section 5 concludes. All proof are relegated to the Appendix.

[2]Earlier results of Apt and Zvesper (2010) and Hillas and Samet (2014) provided (merely) sufficient conditions in similar frameworks.

[3]It is trivial to show that, in turn, IDBS refines IDSDS: the Iterated Deletion of Strictly Dominated Strategies.

2. Qualitative models of ordinal games

2.1. The underlying ordinal game

A *finite strategic-form game with ordinal payoffs* is a quintuple $G = \langle I, (S_i)_{i\in I}, O, z, (\succeq_i)_{i\in I}\rangle$, where $I = \{1,\dots,n\}$ is a finite set of *players*, S_i is a finite set of *strategies* of player $i \in I$ with $S = S_1 \times \cdots \times S_n$ being the set of strategy profiles, O is a finite set of *outcomes*, $z : S \to O$ is a function that associates with every strategy profile $s = (s_1,\dots,s_n) \in S$ an outcome $z(s) \in O$, $\succeq_i$ is player i's *ordinal ranking* of the outcomes, i.e., a binary relation on O which is complete (for all $o, o' \in O$, $o \succeq_i o'$ or $o' \succeq_i o$) and transitive (for all $o, o', o'' \in O$, if $o \succeq_i o'$ and $o' \succeq_i o''$ then $o \succeq_i o''$). The interpretation of $o \succeq_i o'$ is that player i considers outcome o to be at least as good as outcome o'.

Games are often represented in *reduced form* by replacing the triple $\langle O, z, (\succeq_i)_{i\in I}\rangle$ with a list $(\pi_i)_{i\in I}$ of *payoff functions*, where $\pi_i : S \to \mathbb{R}$ is any real-valued function that satisfies the property that, for all $s, s' \in S$, $\pi_i(s) \geq \pi_i(s')$ if and only if $z(s) \succeq_i z(s')$. In the following we will adopt this more succinct representation of strategic-form games. It is important to note, however, that the payoff functions are taken to be purely ordinal and one could replace π_i with any other function obtained by composing π_i with an arbitrary strictly increasing function on the set of real numbers.[4]

A strategic-form game provides only a partial description of an interactive situation, since it does not specify what choices the players make, nor what beliefs they have about their opponents' choices. A specification of these missing elements is obtained by introducing the notion of a "model of the game" (Section 2.3), which represents a possible context in which the game is played.

2.2. Qualitative beliefs

2.2.1. Belief operators

The players' beliefs are represented by means of a finite model $\langle \Omega, (\mathbb{B}_i)_{i\in I}\rangle$, where Ω is a finite set of *states* (or possible worlds). As usual, 2^Ω denotes the collection of all subsets of Ω (i.e., events), while $\neg E := \Omega \setminus E$ denotes the

[4] This is in contrast to von Neumann-Morgenstern utility functions which are invariant only with respect to positive affine transformations.

complement of E for each event $E \subseteq \Omega$. Moreover, for every player $i \in I$, $\mathbb{B}_i : 2^\Omega \to 2^\Omega$ is the *belief operator* that associates with each $E \subseteq \Omega$ the set of states $\mathbb{B}_i E$ where E is believed by $i \in I$. The belief operator is assumed to satisfy the following property: for every $E \subseteq \Omega$ and every $i \in I$,

(D) CONSISTENCY: $\mathbb{B}_i E \subseteq \neg \mathbb{B}_i \neg E$.

Consistency rules out the possibility that a player may simultaneously believe an event E and its complement $\neg E$.

Later in the paper we restrict attention to belief operators that satisfy the stronger property that beliefs cannot be erroneous: for every $E \subseteq \Omega$ and every $i \in I$,

(T) TRUTH: $\mathbb{B}_i E \subseteq E$.

Remark 1. *It is customary in the literature to further restrict the belief operators to satisfy the following properties: for every $E \subseteq \Omega$ and every $i \in I$,*

(4) POSITIVE INTROSPECTION: $\mathbb{B}_i E \subseteq \mathbb{B}_i \mathbb{B}_i E$,

(5) NEGATIVE INTROSPECTION: $\neg \mathbb{B}_i E \subseteq \mathbb{B}_i \neg \mathbb{B}_i E$.

Since our results do not require such restrictions, we will adopt a "minimalistic" approach and refrain from imposing Positive and/or Negative Introspection of beliefs.

2.2.2. Semantic characterization of the belief operators

It is common in the literature to characterize belief operators by means of binary relations in Kripke frames. A Kripke frame is a tuple $\langle \Omega, (\mathcal{B}_i)_{i \in I} \rangle$ where $\mathcal{B}_i \subseteq \Omega \times \Omega$ is a binary relation on Ω that describes i's doxastic accessibility at each state: the interpretation of $\omega \mathcal{B}_i \omega'$ is that at state ω player i considers state ω' possible. In the game-theoretic literature, it is more common to view $\mathcal{B}_i$ as a function that associates with every state $\omega \in \Omega$ the set of states $\mathcal{B}_i(\omega) \subseteq \Omega$ that player i considers possible at ω and to call such a function a *possibility correspondence* or *information correspondence* (e.g., Brandenburger and Keisler, 2006). Of course, the two views (binary relation and possibility correspondence) are equivalent: given a relation $\mathcal{B}_i \subseteq \Omega \times \Omega$, define $\mathcal{B}_i(\omega) = \{\omega' \in \Omega : \omega \mathcal{B}_i \omega'\}$; conversely, given a function $\mathcal{B}_i : \Omega \to 2^\Omega$

define $\omega \mathcal{B}_i \omega'$ if and only if $\omega' \in \mathcal{B}_i(\omega)$.[5] A belief operator $\mathbb{B}_i$ is characterized by a binary relation $\mathcal{B}_i$ if, for every $E \subseteq \Omega$,

$$\mathbb{B}_i E = \{\omega \in \Omega : \mathcal{B}_i(\omega) \subseteq E\}.$$

A Kripke frame is said to be a

- ◇ *KD frame* if the relation $\mathcal{B}_i$ is *serial*: for all $\omega \in \Omega$, $\mathcal{B}_i(\omega) \neq \varnothing$.
- ◇ *KT frame* if the relation $\mathcal{B}_i$ is *reflexive*: $\omega \in \mathcal{B}_i(\omega)$ for all $\omega \in \Omega$.[6]

It is well known that seriality characterizes Consistency (Property (D) of the belief operator), and reflexivity characterizes Truth (Property (T) of the belief operator).[7]

2.3. Doxastic models of games

So far we have introduced the notion of a frame rather abstractly, viz., we have not assigned a meaning to the states and thus to the possible events. We now give an interpretation to the states by introducing a *strategy function* $\sigma_i : \Omega \to S_i$ for each player $i \in I$. Thus, each state $\omega \in \Omega$ is associated with the strategy profile $\sigma(\omega) = \big(\sigma_1(\omega), \ldots, \sigma_n(\omega)\big)$. We denote by $\sigma_{-i}(\omega)$ the profile of strategies played, at ω, by the players other than i, that is, $\sigma_{-i}(\omega) = \big(\sigma_1(\omega), \ldots, \sigma_{i-1}(\omega), \sigma_{i+1}(\omega), \ldots, \sigma_n(\omega)\big)$; thus the entire profile, $\sigma(\omega)$, can also be denoted by $\big(\sigma_i(\omega), \sigma_{-i}(\omega)\big)$.

[5] For more details on Kripke frames see, e.g., Aumann (1999); Battigalli and Bonanno (1999); Chellas (1980); van Ditmarsch et al. (2015); Fagin et al. (1995); Hughes and Cresswell (1968); Kripke (1959).

[6] For completeness we mention that a Kripke frame is said to be a

- ◇ *K4 frame* if the relation $\mathcal{B}_i$ is *transitive*: if $\omega' \in \mathcal{B}_i(\omega)$ then $\mathcal{B}_i(\omega') \subseteq \mathcal{B}_i(\omega)$.
- ◇ *K5 frame* if the relation $\mathcal{B}_i$ is *euclidean*: if $\omega' \in \mathcal{B}_i(\omega)$ then $\mathcal{B}_i(\omega) \subseteq \mathcal{B}_i(\omega')$.

[7] Furthermore, transitivity characterizes Positive Introspection (Property (4) of the belief operator) and euclideanness characterizes Negative Introspection (Property (5) of the belief operator): see Remark 1. A Kripke frame is $KD45$ if it satisfies seriality, trasitivity and euclideanness; by transitivity and euclideanness, we obtain that $KD45$ Kripke frames satisfy the property that $\mathcal{B}_i(\omega') = \mathcal{B}_i(\omega)$ for every $\omega' \in \mathcal{B}_i(\omega)$. A Kripke frame is $S5$ or $KT5$ if it satisfies euclideanness and reflexivity (note that these two properties together imply both seriality and transitivity); in this case, we typically use the term "knowledge" instead of "belief". It is straightforward to see that in an $S5$ Kripke frame $\mathcal{B}_i$ is an equivalence relation.

For an arbitrary $s_i \in S_i$, we define the event $||s_i|| := \{\omega \in \Omega : \sigma_i(\omega) = s_i\}$. We impose the following standard (measurability) property, for every $i \in I$ and every $s_i \in S_i$:

(Σ_0) KNOWLEDGE OF OWN STRATEGY: $||s_i|| = \mathbb{B}_i||s_i||$.

That is, each player knows her own strategy: at every state player i plays s_i if and only if she believes that she plays s_i.

Definition 1. Given a strategic-form game with ordinal payoffs $G = \langle I, (S_i, \pi_i)_{i\in I}\rangle$ a *qualitative doxastic model of* G is a tuple $M = \langle \Omega, (\mathbb{B}_i)_{i\in I}, (\sigma_i)_{i\in I}\rangle$, where

- Ω is finite,
- $\mathbb{B}_i$ is the belief operator of player i, and
- σ_i is a strategy function that satisfies Knowledge-of-own-strategy (property (Σ_0)).

Then, we define:

$\mathcal{M}_D$: the class of models where each belief operator satisfies Consistency (Property (D)).

$\mathcal{M}_T$: the class of models where each belief operator satisfies Truth (Property (T)).

Note that, since reflexivity implies seriality, $\mathcal{M}_T \subset \mathcal{M}_D$.

Remark 2. *We briefly remark on how to relate the "orthodox" approach, based on probabilistic beliefs, and our more general, qualitative, approach. When beliefs are represented by probability distributions, one defines a function* $p_i : \Omega \to \Delta(\Omega)$ *(with* $\Delta(\Omega)$ *being the set of probability distributions over* Ω*) where* $p_{i,\omega}$ *(we use the notation* $p_{i,\omega}$ *rather than* $p_i(\omega)$*) are the probabilistic beliefs of player i at state* ω*. For the probabilistic case, our* $\mathcal{B}_i(\omega)$ *coincides with the support of* $p_{i,\omega}$*, that is,* $\mathcal{B}_i(\omega) = \{\omega' \in \Omega : p_{i,\omega}(\omega') > 0\}$ *so that the expression "player i believes event E" is interpreted as "player i attaches probability 1 to E".*[8]

[8] In probabilistic models it is customary to impose the restriction that

$$\text{if } \; p_{i,\omega}(\omega') > 0 \; \text{ then } \; p_{i,\omega'} = p_{i,\omega}$$

3. Common belief of rationality

Fix a player i and two strategies $a, b \in S_i$ of player i. We denote by

$$\|b \succeq a\| = \{\omega \in \Omega : \pi_i(b, \sigma_{-i}(\omega)) \geq \pi_i(a, \sigma_{-i}(\omega))\}$$

the event that strategy b yields at least as high a payoff for player i as strategy a. Similarly, $\|b \succ a\| = \{\omega \in \Omega : \pi_i(b, \sigma_{-i}(\omega)) > \pi_i(a, \sigma_{-i}(\omega))\}$ is the event that b is strictly preferred to a by i. Finally, $\|b \sim a\| = \|b \succeq a\| \setminus \|b \succ a\| = \{\omega \in \Omega : \pi_i(b, \sigma_{-i}(\omega)) = \pi_i(a, \sigma_{-i}(\omega))\}$ denotes the event that i is indifferent between a and b.

Definition 2. Player i is *(weak-dominance) rational* at state ω whenever, for all $b \in S_i$,

$$\text{if } \omega \in \mathbb{B}_i\|b \succeq \sigma_i(\omega)\| \text{ then } \omega \in \mathbb{B}_i\|b \sim \sigma_i(\omega)\|. \tag{1}$$

Let $\mathbf{R}_i \subseteq \Omega$ be the event that player i is rational and $\mathbf{R} = \bigcap_{i \in I} \mathbf{R}_i$ be the event that all players are rational.

Intuitively, if at ω player i believes that b yields at least as high a payoff as the chosen strategy $\sigma_i(\omega)$, then he does not consider it possible that b yields a *strictly higher* payoff than $\sigma_i(\omega)$, i.e., $\sigma_i(\omega)$ is not weakly dominated by any $b \in S_i$ *given the strategy profiles that are played by the opponents at the doxastically accessible states.* For a formal definition of weak dominance relative to a subset of the opponents' strategy profiles, see Section 3.1.

We want to investigate the implications of common belief of rationality. Given an event E, let $\mathbb{B}_I E = \bigcap_{i \in I} \mathbb{B}_i E$ denote the event that all the players believe E. Then the event that E is commonly believed, denoted by $\mathbb{CB}E$, is defined as the infinite intersection $\mathbb{CB}E = \mathbb{B}_I E \cap \mathbb{B}_I\mathbb{B}_I E \cap \mathbb{B}_I\mathbb{B}_I\mathbb{B}_I E \cap \cdots$, that is, the event that everybody believes E, and everybody believes that everybody believes E, and everybody believes that everybody believes that everybody believes E, and so on. It is well-known that, for every state ω and every event E, $\omega \in \mathbb{CB}E$ if and only if $\mathcal{B}^*(\omega) \subseteq E$, where $\mathcal{B}^*(\omega)$ is the transitive closure of $\bigcup_{i \in I} \mathcal{B}_i(\omega)$.[9] We are interested in the event that there is

---

to capture the fact that the player knows her own beliefs. It follows from this restriction that $\mathcal{B}_i$ satisfies seriality, transitivity and euclideanness, that is, the corresponding Kripke frame is a $KD45$ frame.

[9] $\mathcal{B}^*$ is thus defined as follows: $\omega' \in \mathcal{B}^*(\omega)$ if and only if there is a sequence $\{\omega_1, \ldots, \omega_m\}$ in Ω and a sequence $\{i_1, \ldots, i_{m-1}\}$ in I such that (1) $\omega_1 = \omega$, (2) $\omega_m = \omega'$, and (3) for every $j = 1, \ldots, m-1$, $\omega_{j+1} \in \mathcal{B}_{i_j}(\omega_j)$.

common belief of rationality, henceforth denoted by $\mathbb{CB}\mathbf{R}$. In particular, we ask the question: *which strategy profiles are compatible with states in* $\mathbb{CB}\mathbf{R}$*?*

Definition 3. We say that common belief of rationality in a class of models $\mathcal{M}$ *(epistemically) characterizes* the set $S^* \subseteq S$ of strategy profiles whenever the following two conditions hold:

(A) in every model $M \in \mathcal{M}$, if $\omega \in \mathbb{CB}\mathbf{R}$ then $\sigma(\omega) \in S^*$,

(B) for every $s \in S^*$, there exists a model $M \in \mathcal{M}$ and a state ω in that model such that $\sigma(\omega) = s$ and $\omega \in \mathbb{CB}\mathbf{R}$.

In the following sections, we will epistemically characterize two well-known solution concepts for ordinal strategic-form games by means of common belief of rationality, by successively imposing stronger properties on the models of qualitative beliefs. That way, (i) we will place these different solution concepts under the same umbrella of common belief of rationality, and (ii) we will formally order the solution concepts in terms of the strategy profiles that they predict.

3.1. Iterated Deletion of Börgers-dominated Strategies

Börgers (1993) introduced a notion of pure-strategy dominance which is stronger than strict dominance. Let $a, b \in S_i$ be two pure strategies of player i, and let $X_{-i} \subseteq S_{-i}$ be a non-empty set of strategy-profiles of the players other than i. We say that b *weakly dominates* a *relative to* X_{-i} whenever:
(1) $\pi_i(b, x_{-i}) \geq \pi_i(a, x_{-i})$ for all $x_{-i} \in X_{-i}$, and
(2) there exists some $\hat{x}_{-i} \in X_{-i}$ such that $\pi_i(b, \hat{x}_{-i}) > \pi_i(a, \hat{x}_{-i})$.
Then, a pure strategy $a \in S_i$ is *Börgers-dominated* (henceforth *B-dominated*) if for every non-empty subset $X_{-i} \subseteq S_{-i}$ there exists a strategy $b \in S_i$ (which is allowed to vary with X_{-i}) such that b weakly dominates a relative to X_{-i}. The Iterated Deletion of B-dominated Strategies (IDBS) is the following algorithm: reduce the game by deleting, for each player, all the strategies that are B-dominated and then repeat the procedure in the reduced game, and so on, until there are no B-dominated strategies left.

Definition 4. Given a strategic-form game with ordinal payoffs $G = \langle I, (S_i, \pi_i)_{i \in I} \rangle$, recursively define the sequence of reduced games $\{G^0, G^1, \ldots, G^m, \ldots\}$ as follows: for each $i \in I$, (4.1) let $B_i^0 = S_i$, and let $E_i^0 \subsetneq B_i^0$ be the set of i's strategies that are B-dominated in $G^0 = G$;

(4.2) for each $m \geq 1$ let G^{m-1} be the reduced game with strategy sets B_i^{m-1}, and define $B_i^m = B_i^{m-1} \setminus E_i^{m-1}$, where $E_i^{m-1} \subsetneq B_i^{m-1}$ is the set of i's strategies that are B-dominated in G^{m-1}.

Let $B_i^\infty = \bigcap_{m=0}^\infty B_i^m$. The strategy profiles in $B^\infty = B_1^\infty \times \cdots \times B_n^\infty$ are those surviving IDBS.

Since the strategy sets are finite, there exists an integer r such that $B^\infty = B^k$ for every $k \geq r$, that is, the procedure terminates after finitely many steps. Furthermore, it is straightforward to verify that $B^\infty \neq \varnothing$.

For example, in the game of Figure 1, strategy a of Player 1 is B-dominated. Indeed, a is weakly dominated by b relative to $\{d\}$ and also relative to $\{d, e\}$, and it is weakly dominated by c relative to $\{e\}$. Eliminating a for Player 1, we are left with a reduced game where for Player 2 d is strictly dominated by e (thus B-dominated). In the game remaining after the deletion of d, for Player 1 b is strictly dominated by c (thus B-dominated). Hence in this game $B^\infty = \{(c, e)\}$.

		Player 2: d	Player 2: e
Player 1	a	1 , 1	1 , 0
	b	2 , 0	1 , 1
	c	0 , 0	2 , 1

Figure 1: $B^\infty = \{(c, e)\}$.

Theorem 1 (Characterization of IDBS)**.** *In the class of models $\mathcal{M}_D$, common belief of rationality characterizes IDBS. Formally,*

(A_1) *in every model $M \in \mathcal{M}_D$, if $\omega \in \mathbb{CB}\mathbf{R}$ then $\sigma(\omega) \in B^\infty$,*

(B_1) *for every $s \in B^\infty$, there exists a model $M \in \mathcal{M}_D$ and a state ω in that model such that $\sigma(\omega) = s$ and $\omega \in \mathbb{CB}\mathbf{R}$.*

Intuitively, the proof of the result relies on the observation that a strategy is B-dominated (in the full game) if and only if it is not (weak-dominance) rational at any state of any consistent frame. A similar intuition carries for the later iterations.[10]

A weaker version of (A_1) follows from Apt and Zvesper (2010). In particular, they show that in every $KD45$ frame, common belief in rationality leads to strategy profiles consistent with IDBS. The latter follows from the fact that B-dominance is monotonic (without actually using any of the introspection axioms), i.e., if a is B-dominated relative to some $X_{-i} \subseteq S_{-i}$ then it is also B-dominated relative to every nonempty subset of X_{-i}.

Now turning to (B_1) which is arguably the more challenging and innovative part of our previous result, note that in order to "rationalize" a strategy profile in B^∞, *it may be necessary for a player to have erroneous beliefs.* To see this, consider the game in Figure 2, where $B^\infty = S$, that is, IDBS does not eliminate any strategy; in particular, $(a, d) \in B^\infty$.[11] Consider an arbitrary $\mathcal{M}_D$-model of this game and a state ω_0 such that $\sigma(\omega_0) = (a, d)$. Since, for every $s_2 \in \{c, d\}$; $\pi_1(b, s_2) \geq \pi_1(a, s_2)$, $\|b \succeq a\| = \Omega$; thus $\mathcal{B}_1(\omega_0) \subseteq \|b \succeq a\|$, that is, $\omega_0 \in \mathbb{B}_1\|b \succeq a\|$. Hence, if Player 1 is rational at ω_0 (according to Definition 2) then $\|b \succ a\| \cap \mathcal{B}_1(\omega_0) = \varnothing$. Thus, $\sigma_2(\omega) = c$ for all $\omega \in \mathcal{B}_1(\omega_0)$. In particular, it must be that $\omega_0 \notin \mathcal{B}_1(\omega_0)$. Thus at state ω_0 Player 2 actually plays d but Player 1 – who plays a – must erroneously believe that Player 2 is playing c.

In the next section we investigate the consequences of ruling out false beliefs.

3.2. Iterated Deletion of Inferior Strategy Profiles

The following algorithm is the pure-strategy version of a procedure first introduced by Stalnaker (1994) and further studied in Bonanno (2008); Bonanno and Nehring (1998); Hillas and Samet (2014); Trost (2013). Unlike the IDBS

[10] Dekel and Siniscalchi (2015) provide a characterization of IDBS in a "quantitative" framework: they assume cardinal utilities (vNM preferences) and use the notion of Bayesian rationality (expected utility maximization) but employ an incomplete information model where players have uncertainty (probabilistic beliefs) over the risk preferences of their opponents, while the ordinal rankings are commonly known.

[11] For Player 1, a is weakly dominated by b relative to $\{d\}$ and $\{c, d\}$ but not relative to $\{c\}$; for Player 2, d is weakly dominated by c relative to $\{a\}$ but not relative to $\{b\}$ or $\{a, b\}$ and c is weakly dominated by d relative to $\{b\}$ but not relative to $\{a\}$ or $\{a, b\}$.

		Player 2	
		c	d
Player 1	a	1 , 1	1 , 0
	b	1 , 0	2 , 2

Figure 2: Rationalization of (a, d) requires erroneous beliefs.

procedure considered above, this procedure merely deletes strategy profiles, rather than individual strategies. In particular, let $X \subseteq S$ be a set of strategy profiles (not necessarily having a product structure). A strategy profile $x \in X$ is *inferior relative to* X if there exist a player i and a strategy $s_i \in S_i$ of player i (with s_i not necessarily belonging to the projection of X onto S_i) such that
(1) $\pi_i(s_i, x_{-i}) > \pi_i(x_i, x_{-i})$, and
(2) for all $s_{-i} \in S_{-i}$, either $(x_i, s_{-i}) \notin X$ or $\pi_i(s_i, s_{-i}) \geq \pi_i(x_i, s_{-i})$.
The Iterated Deletion of Inferior Profiles (IDIP) is the following algorithm: reduce the game by deleting all the inferior strategy profiles and then repeat the procedure by eliminating inferior profiles relative to the strategy profiles that have not been eliminated so far, until there are no inferior profiles left. Formally, the algorithm is defined as follows:

Definition 5. Given a strategic-form game with ordinal payoffs $G = \langle I, (S_i, \pi_i)_{i \in I} \rangle$, recursively define the sequence of sets of strategy profiles $\{U^0, U^1, \ldots, U^m, \ldots\}$ as follows:

(5.1) let $U^0 = S$, and let $F^0 \subsetneq U^0$ be the set of inferior strategy profiles relative to U^0;

(5.2) for each $m \geq 1$ let $U^m = U^{m-1} \setminus F^{m-1}$, where $F^{m-1} \subsetneq U^{m-1}$ is the set of strategy profiles in U^{m-1} that are inferior relative to U^{m-1}.

Then $U^\infty = \bigcap_{m=0}^{\infty} U^m$ denotes the strategy profiles surviving IDIP.

Once again, since the strategy sets are finite, there exists an integer r such that $U^\infty = U^k$ for every $k \geq r$, i.e., the procedure terminates after finitely many steps. Besides, it is straightforward to verify that $U^\infty \neq \varnothing$.

		Player 2	
		c	d
Player 1	a	1 , 1	1 , 2
	b	1 , 1	2 , 0

Figure 3: IDIP.

As an illustration of this procedure, consider the game in Figure 3. In this game (a, d) is inferior relative to $U^0 = S$ since $\pi_1(b, d) > \pi_1(a, d)$ and $\pi_1(b, c) = \pi_1(a, c)$ (and $(a, c) \in S$). No other strategy profile is inferior relative to U^0 and thus $F^0 = \{(a, d)\}$ so that $U^1 = \{(a, c), (b, c), (b, d)\}$. Now (b, d) is inferior relative to U^1 since $\pi_2(b, c) > \pi_2(b, d)$ and $(a, d) \notin U^1$. No other strategy profile is inferior relative to U^1 and thus $F^1 = \{(b, d)\}$ so that $U^2 = \{(a, c), (b, c)\}$. Since no strategy profile is inferior relative to U^2, we have that $U^\infty = U^2$.

We now turn to investigating the consequences of ruling out false beliefs. At state ω player i has correct beliefs if ω is one of the states that player i considers possible at ω, that is, if $\omega \in \mathcal{B}_i(\omega)$. Let $\mathbf{T}_i = \{\omega \in \Omega : \omega \in \mathcal{B}_i(\omega)\}$ be the event that player i has correct beliefs. Imposing the Truth condition (Property (T)) on the belief operator of player i amounts to requiring that $\mathbf{T}_i = \Omega$. Recall from Definition 1 that $\mathcal{M}_T \subsetneq \mathcal{M}_D$ denotes the class of finite qualitative doxastic models that rule out false beliefs.

Theorem 2 (Characterization of IDIP)**.** *When erroneous beliefs are ruled out, common belief of rationality characterizes IDIP. Formally,*

(A_2) *in every model $M \in \mathcal{M}_T$, if $\omega \in \mathbb{CBR}$ then $\sigma(\omega) \in U^\infty$,*

(B_2) *for every $s \in U^\infty$, there exists a model $M \in \mathcal{M}_T$ and a state ω in that model such that $\sigma(\omega) = s$ and $\omega \in \mathbb{CBR}$.*

The intuition behind the previous result is as follows: a strategy profile is not inferior (in the full game) if and only if it is (weak-dominance) rational for all player to play according to this profile at some state deemed possible by all players where this profile is played. The same logic applies to later iterations.

Property (T) says that no player can have false beliefs. This is a stronger condition than simply requiring that it is commonly believed that every player has correct beliefs. In fact, in order to get a characterization of the set U^∞, common belief that all players have correct beliefs is not sufficient: it only yields *common belief* that only strategy profiles in U^∞ are played (see Section 4.2).

A weaker version of (A_2) follows from Hillas and Samet (2014). In particular, they show that in every $S5$ frame, common belief in rationality yields strategy profiles that survive the Stalnaker procedure (i.e., the mixed strategy version of IDIP), taking an approach similar to the one in Apt and Zvesper (2010). In this respect, our paper completes the picture by providing a full epistemic characterization of IDIP, viz., in particular by proving (B_2).

4. Discussion

4.1. Monotonicity result

We have characterized two well-known solution concepts for games with ordinal payoffs (viz., IDBS and IDIP) by means of restrictions imposed on the belief operator.[12] A direct implication of our results (Theorems 1 and 2) is the following (monotonicity) result.

Corollary 1 (Monotonicity result). $U^\infty \subseteq B^\infty$.

The proof follows directly from the fact that $\mathcal{M}_T \subsetneq \mathcal{M}_D$. Without invoking Theorems 1 and 2 this monotonicity result would not be straightforward to establish, since there are games where $B^m \subsetneq U^m$ for some $m > 0$, as illustrated in the game in Figure 4.
In this game, c is B-dominated, while no other strategy is subsequently eliminated. That is, $B^\infty = B^1 = \{a, b\} \times \{d, e\}$. On the other hand, the only inferior strategy profile relative to the entire game is (c, e), and therefore $U^1 \supsetneq B^1$. However, (c, d) is inferior relative to U^1, thus implying that $B^\infty = B^2 = U^2 = U^\infty$, consistently with the conclusions of Corollary 1.

[12] A weaker notion of rationality than the one given in Definition 2 would yield a characterization of the Iterated Deletion of Strictly Dominated Strategies within the class of models $\mathcal{M}_D$, namely the following definition: player i is rational at state ω if for all $b \in S_i$, $\omega \notin \mathbb{B}_i \| b \succ \sigma_i(\omega) \|$.

		Player 2	
		d	e
	a	1 , 1	2 , 1
Player 1	b	2 , 1	0 , 1
	c	1 , 1	1 , 1

Figure 4: Monotonicity.

Note that we are not claiming that ours is the only way to prove Corollary 1. In fact, using the fact that IDBS and IDIP are order-independent, one can show that IDIP refines IDBS. However, our approach (via the two characterization Theorems, 1 and 2), not only proves the aforementioned refinement, but also manages to attribute it to the Truth Axiom.

4.2. Correct beliefs

As we have already mentioned, common belief in correct beliefs does not suffice for a strategy that survives IDIP to be played. Recall that $\mathbf{T}_i = \{\omega \in \Omega : \omega \in \mathcal{B}_i(\omega)\}$ is the event that player i has correct beliefs and that imposing the Truth condition (Property (T)) on the belief operator of player i amounts to requiring that $\mathbf{T}_i = \Omega$. A weaker condition is – possibly erroneous – common belief that every player has correct beliefs. This can be expressed as the event $\mathbb{CB}\mathbf{T}$ where $\mathbf{T} = \bigcap_{i \in I} \mathbf{T}_i$ is the event that all players have correct beliefs.

The following example shows that it is possible that $\omega \in \mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T}$ and yet the strategy profile played at ω does not survive IDIP.[13] Consider the following model of the game shown in Figure 5: $\Omega = \{\omega_1, \omega_2, \omega_3\}, \mathcal{B}_1(\omega_1) = \{\omega_1\}, \mathcal{B}_1(\omega_2) = \mathcal{B}_1(\omega_3) = \{\omega_3\}, \mathcal{B}_2(\omega_1) = \mathcal{B}_2(\omega_2) = \{\omega_1\}, \mathcal{B}_2(\omega_3) = \{\omega_3\}$, $\sigma_1(\omega_1) = b, \sigma_1(\omega_2) = \sigma_1(\omega_3) = a, \sigma_2(\omega_1) = \sigma_2(\omega_2) = d$ and $\sigma_2(\omega_3) = c$. Then $\sigma(\omega_2) = (a, d) \notin U^\infty$ and yet $\omega_2 \in \mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T}$ (in fact, $\mathcal{B}^*(\omega_2) =$

[13]Note that, in this example, the belief relations of both players are serial, transitive and euclidean, that is, the belief operators satisfy Consistency, Positive Introspection and Negative Introspection.

$\{\omega_1, \omega_3\}, \mathbf{T} = \{\omega_1, \omega_3\}$ and $\mathbf{R} = \Omega$). Note that, in this model, at state ω_2 both players have false beliefs; in particular, although it is common belief at ω_2 that only strategy profiles in U^∞ are played, the strategy profile *actually* played does not belong to U^∞.

		Player 2	
		c	d
Player 1	a	1 , 1	1 , 0
	b	1 , 0	2 , 2

Figure 5: Common belief in correct beliefs is not sufficient for U^∞.

Although common belief in correct beliefs does not suffice for IDIP, it does guarantee common belief in the event that only strategy profiles in U^∞ are played. Let $\mathbf{U}^\infty = \{\omega \in \Omega : \sigma(\omega) \in U^\infty\}$.

Proposition 1. *In every model $M \in \mathcal{M}_D$, $\mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T} \subseteq \mathbb{CB}\mathbf{U}^\infty$.*

The condition that there is common belief that all players have correct beliefs ($\omega \in \mathbb{CB}\mathbf{T}$) is necessary for Proposition 1. To see this, consider the game shown in Figure 6, where $U^\infty = \{(a, c), (b, c)\}$.

		Player 2	
		c	d
Player 1	a	1 , 1	1 , 1
	b	1 , 1	2 , 0

Figure 6: Correct beliefs.

Consider the following model of this game: $\Omega = \{\omega_1, \omega_2\}, \mathcal{B}_1(\omega_1) = \mathcal{B}_1(\omega_2) = \{\omega_2\}, \mathcal{B}_2(\omega_1) = \{\omega_1\}, \mathcal{B}_2(\omega_2) = \{\omega_2\}, \sigma_1(\omega_1) = \sigma_1(\omega_2) = a, \sigma_2(\omega_1) = d, \sigma_2(\omega_2) = c$. Then $\mathbf{R} = \mathbb{CB}\mathbf{R} = \Omega$, while $\mathbf{U}^\infty = \{\omega_2\}$ (since $\sigma(\omega_1) = (a, d) \notin U^\infty$).

Since $\mathcal{B}^*(\omega_1) = \{\omega_1, \omega_2\}$, $\omega_1 \in \mathbb{CB}\mathbf{R}$ but $\omega_1 \notin \mathbb{CB}\mathbf{U}^\infty$. In this model, at state ω_1 Player 1 has false beliefs ($\mathbf{T}_1 = \{\omega_2\}$) and thus $\omega_1 \notin \mathbb{CB}\mathbf{T}$.[14]

The following Corollary shows that if, to the hypotheses of Proposition 1, we add the further hypothesis that *at least one player* does not have false beliefs, then it follows that the strategy profile *actually played* also belongs to U^∞. Let $\mathbf{T}_\cup = \bigcup_{i \in I} \mathbf{T}_i$ be the event that at least one player has correct beliefs.

Corollary 2. *In every model $M \in \mathcal{M}_D$, $\mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T} \cap \mathbf{T}_\cup \subseteq \mathbf{U}^\infty$. That is, common belief of rationality, common belief of correct beliefs and correct beliefs of at least one player imply IDIP.*

5. Conclusion

We have studied the behavioral implications of common belief of rationality in strategic-form games with ordinal utilities, using qualitative beliefs. Focusing on ordinal utilities is relevant both theoretically (as we implicitly relax the admittedly unrealistic assumption of commonly known vNM preferences), as well as empirically (as experimental economists typically use solution concepts for games with ordinal payoffs for their benchmark theoretical predictions).

Our main contribution is twofold. Firstly, we provide a full characterization of two well-known solution concepts for games with ordinal payoffs in terms of common belief of rationality, by gradually strengthening the properties of the doxastic model that we use. Then, as a consequence of our characterization results, we prove that the aforementioned solution concepts monotonically refine each other, that is., IDIP refines IDBS.

The qualitative doxastic model that we have used is quite permissive, in that it does not specify the relative likelihood between any two events. Thus our model weakens earlier models on qualitative beliefs, which typically rely on likelihood relations.[15]

[14] Note that this is a KD45 frame (the beliefs of both players satisfy Consistency, Positive Introspection and Negative Introspection: see Footnote 2.2.2) and yet the common belief relation is not euclidean (that is, the common belief operator does not satisfy Negative Introspection). Bonanno and Nehring (2000) have shown that Negative Introspection of common belief is equivalent to the property that whenever a player believes that an event E is commonly believed then E is indeed commonly believed.

[15] Qualitative beliefs have been extensively studied in the literature since the early con-

Our qualitative doxastic model is also related to the literature on possibility models (e.g., Brandenburger and Keisler, 2006) and knowledge-belief models (e.g., Meier, 2008). In particular, using our terminology, in its simplest form a possibility model is essentially a Kripke structure that merely employs a belief operator, whereas a knowledge-belief model is also endowed with a probabilistic belief over the set of states that are deemed possible.

There are also qualitative models of knowledge and beliefs (e.g., Lorini, 2016) that employ two Kripke relations: an equivalence relation representing knowledge and a sub-relation of the equivalence relation representing beliefs. For our results there is no need to employ a two-level epistemic structure: belief is all that is needed.

A. Proofs

Proof of Theorem 1. (A_1) Fix a strategic-form game with ordinal payoffs, a model $M \in \mathcal{M}_D$ and a state ω_1 in M. Suppose that $\omega_1 \in \mathbb{CB}\mathbf{R}$, that is, $\mathcal{B}^*(\omega_1) \subseteq \mathbf{R}$. We want to show that $\sigma(\omega_1) \in B^\infty$. The proof is by induction.

Base Step. First we show (by contradiction) that, for every player $i \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \notin E_i^0$ (see Definition 4). Suppose not. Then there exist a player i and an $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma_i(\omega_2) \in E_i^0$, that is, strategy $\sigma_i(\omega_2)$ of player i is B-dominated relative to S_{-i}, that is, for every non-empty $X_{-i} \subseteq S_{-i}$ there exists a strategy $s_i \in S_i$ such that:
(I) for all $x_{-i} \in X_{-i}$, $\pi_i(s_i, x_{-i}) \geq \pi_i(\sigma_i(\omega_2), x_{-i})$, and
(II) there exists an $\hat{x}_{-i} \in X_{-i}$ such that $\pi_i(s_i, \hat{x}_{-i}) > \pi_i(\sigma_i(\omega_2), \hat{x}_{-i})$.
Let $X_{-i} = \sigma_{-i}\big(\mathcal{B}_i(\omega_2)\big) = \{s_{-i} \in S_{-i} : s_{-i} = \sigma_{-i}(\omega) \text{ for some } \omega \in \mathcal{B}_i(\omega_2)\}$. Note that, by Consistency (Property (D)), $\mathcal{B}_i(\omega_2) \neq \varnothing$ and thus $X_{-i} \neq \varnothing$. Let $s_i \in S_i$ and $\hat{x}_{-i} \in X_{-i}$ satisfy (I) and (II) above and let $\hat{\omega} \in \mathcal{B}_i(\omega_2)$ be

tributions of de Finetti (1949) and Koopman (1940). Most papers in the literature have focused on whether a qualitative likelihood relation can be represented by a probability measure (Kraft et al., 1959; Mackenzie, 2017; Scott, 1964; Scott and Suppes, 1958; Villegas, 1967) and on the respective logical foundations (Gärdenfors, 1975; Segerberg, 1971; van der Hoek, 1996). For an early overview of qualitative beliefs see Fishburn (1986). To the best of our knowledge there has not been any attempt to embed qualitative probability in a game-theoretic model. Notice that there is a duality between the approach adopted in these models and the one taken by preference-based models of beliefs in games (e.g., Di Tillio, 2008), which starts with a preference relation over acts and derives the collection of Savage-null events.

such that $\sigma_{-i}(\hat{\omega}) = \hat{x}_{-i}$. Then, by (I), $\mathcal{B}_i(\omega_2) \subseteq \|s_i \succeq \sigma_i(\omega_2)\|$, that is,

$$\omega_2 \in \mathbb{B}_i\|s_i \succeq \sigma_i(\omega_2)\|$$

and, by (II), $\mathcal{B}_i(\omega_2) \cap \|s_i \succ \sigma_i(\omega_2)\| \supseteq \{\hat{\omega}\} \neq \varnothing$, that is (recall that, by Property (Σ_0), $\sigma_i(\hat{\omega}) = \sigma_i(\omega_2)$),

$$\omega_2 \notin \mathbb{B}_i\|s_i \sim \sigma_i(\omega_2)\|.$$

Hence $\omega_2 \notin \mathbf{R}_i$ (Definition 2). Since $\mathbf{R}_i \subseteq \mathbf{R}$, it follows that $\omega_2 \notin \mathbf{R}$, contradicting our hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1)$ and $\mathcal{B}^*(\omega_1) \subseteq \mathbf{R}$. Thus we have shown that

$$\text{for every } \ \omega \in \mathcal{B}^*(\omega_1), \ \sigma_i(\omega) \in S_i \backslash E_i^0 = B_i^1.$$

Inductive Step. Fix an integer $m \geq 1$ and suppose that, for every player $j \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_j(\omega) \in B_j^m$, that is, $\mathcal{B}^*(\omega_1) \subseteq B^m$. We want to show (by contradiction) that, for every player $i \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \notin E_i^m$. Suppose not. Then there exist a player i and an $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma_i(\omega_2) \in E_i^m$, that is, strategy $\sigma_i(\omega_2)$ of player i is B-dominated relative to B_{-i}^m: for every $X_{-i} \subseteq B_{-i}^m$ there exists a strategy $s_i \in S_i$ such that:
(I') for all $x_{-i} \in X_{-i}$, $\pi_i(s_i, x_{-i}) \geq \pi_i(\sigma_i(\omega_2), x_{-i})$ and
(II') there exists an $\hat{x}_{-i} \in X_{-i}$ such that $\pi_i(s_i, \hat{x}_{-i}) > \pi_i(\sigma_i(\omega_2), \hat{x}_{-i})$.
Let $X_{-i} = \sigma_{-i}\big(\mathcal{B}_i(\omega_2)\big) = \{s_{-i} \in S_{-i} : s_{-i} = \sigma_{-i}(\omega) \text{ for some } \omega \in \mathcal{B}_i(\omega_2)\}$. Note, again, that, by Consistency (Property (D)), $\mathcal{B}_i(\omega_2) \neq \varnothing$ and thus $X_{-i} \neq \varnothing$. By the induction hypothesis and the fact that $\mathcal{B}_i(\omega_2) \subseteq \mathcal{B}^*(\omega_2) \subseteq \mathcal{B}^*(\omega_1)$ (the latter inclusion follows from transitivity of $\mathcal{B}^*$), $X_{-i} \subseteq B_{-i}^m$. Let $s_i \in S_i$ and $\hat{x}_{-i} \in X_{-i}$ satisfy (I') and (II') above and let $\hat{\omega} \in \mathcal{B}_i(\omega_2)$ be such that $\sigma_{-i}(\hat{\omega}) = \hat{x}_{-i}$. Then, by (I'),

$$\omega_2 \in \mathbb{B}_i\|s_i \succeq \sigma_i(\omega_2)\|$$

and, by (II'), $\|s_i \succ \sigma_i(\omega_2)\| \cap \mathcal{B}_i(\omega_2) \neq \varnothing$, that is,

$$\omega_2 \notin \mathbb{B}_i\|s_i \sim \sigma_i(\omega_2)\|.$$

Hence $\omega_2 \notin \mathbf{R}_i$ (Definition 2). Since $\mathbf{R}_i \subseteq \mathbf{R}$, it follows that $\omega_2 \notin \mathbf{R}$, contradicting our hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1)$ and $\mathcal{B}^*(\omega_1) \subseteq \mathbf{R}$.

Thus we have shown that, for every player $i \in I$ and for every state $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \in \bigcap_{m=1}^{\infty} B_i^m = B_i^\infty$.

It only remains to show that $\sigma_i(\omega_1) \in B_i^\infty$. Take any $\omega_2 \in \mathcal{B}_i(\omega_1)$. Since $\mathcal{B}_i(\omega_1) \subseteq \mathcal{B}^*(\omega_1)$, $\omega_2 \in \mathcal{B}^*(\omega_1)$. Thus $\sigma_i(\omega_2) \in B_i^\infty$. By (Σ_0), since $\omega_2 \in \mathcal{B}_i(\omega_1)$, $\sigma_i(\omega_2) = \sigma_i(\omega_1)$. Thus $\sigma_i(\omega_1) \in B_i^\infty$.

(B_1) Given a game G construct the following model $M \in \mathcal{M}_D$: $\Omega = B^\infty = B_1^\infty \times \cdots \times B_n^\infty$; for every player i and for every $s \in B^\infty$, $\sigma_i : B^\infty \to S_i$ is defined by $\sigma_i(s) = s_i$ (that is, $\sigma_i(s)$ is the i^{th} coordinate of s). To define $\mathcal{B}_i$ first note that, by Definition of B^∞, every $s_i \in B_i^\infty$ is not B-dominated relative to B_{-i}^∞, that is, there exists an $X_{-i}^{s_i} \subseteq B_{-i}^\infty$ (note that this set may vary with s_i, hence the superscript 's_i') such that, for all $s_i' \in S_i$, there exists an $\hat{x}_{-i} \in X_{-i}^{s_i}$ such that either

$$\pi_i(s_i', \hat{x}_{-i}) < \pi_i(s_i, \hat{x}_{-i}) \tag{A.1}$$

or, for all $x_{-i} \in X_{-i}^{s_i}$,

$$\pi_i(s_i', \hat{x}_{-i}) \leq \pi_i(s_i, \hat{x}_{-i}). \tag{A.2}$$

For every $s_i \in B_i^\infty$ fix one such set $X_{-i}^{s_i}$ (there may be several) and define $\mathcal{B}_i(s_i, s_{-i}') = \{s_i\} \times X_{-i}^{s_i}$. By construction, $(s_i, \hat{x}_{-i}) \in \mathcal{B}_i(s)$ and thus, either, by (A.1), $s \notin \mathbb{B}_i \|s_i' \succeq s_i\|$ or, by (A.2), $\|s_i' \succ s_i\| \cap \mathcal{B}_i(s) = \varnothing$. It follows that, for every $i \in I$ and for every $s \in B^\infty$, $s \in \mathbf{R}_i$ and thus $B^\infty = \mathbf{R} = \mathbb{CB}\mathbf{R}$. $\square$

The proof of Theorem 2 makes use of Proposition 1 and Corollary 2, which are proved below.

Proof of Theorem 2. (A_2) Given a game, consider a model $M \in \mathcal{M}_T$. Then $\mathbf{T} = \mathbb{CB}\mathbf{T} = \mathbf{T}_\cup = \Omega$ (so that $\mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T} \cap \mathbf{T}_\cup = \mathbb{CB}\mathbf{R}$). Let $\omega \in \mathbb{CB}\mathbf{R}$. Then, by Corollary 2 (Section 4.2), $\omega \in \mathbf{U}^\infty$.

(B_2) Given a game construct the following model of it: $\Omega = U^\infty$; for every player i and for every $s \in U^\infty$, $\mathcal{B}_i(s) = \{s' \in U^\infty : s_i' = s_i\}$ (that is, $s' \in \mathcal{B}_i(s)$ if and only if both s and s' belong to U^∞ and player i's strategy is the same in s and s'); $\sigma_i : U^\infty \to S_i$ is defined by $\sigma_i(s) = s_i$ (that is, $\sigma_i(s)$ is the i^{th} coordinate of s). Note that each relation $\mathcal{B}_i$ is an equivalence relation (in particular it satisfies reflexivity). Fix an arbitrary state $s \in U^\infty$ and an arbitrary player i and suppose that, for some $s_i' \in S_i$, $\pi_i(s_i', s_{-i}) > \pi_i(s_i, s_{-i})$, that is, $s \in \|s_i' \succ s_i\|$, so that $\|s_i' \succ s_i\| \cap \mathcal{B}_i(s) \supseteq \{s\} \neq \varnothing$. Then, by definition of U^∞, there exists an $\hat{s}_{-i} \in S_{-i}$ such that $(s_i, \hat{s}_{-i}) \in U^\infty$ and

$\pi_i(s_i', \hat{s}_{-i}) < \pi_i(s_i, \hat{s}_{-i})$; by construction, $(s_i, \hat{s}_{-i}) \in \mathcal{B}_i(s)$ so that $s \notin \mathbb{B}_i \| s_i' \succeq s_i \|$. Thus, by Definition 2, player i is rational at state s, that is, $s \in \mathbf{R}_i$. Since i and s were chosen arbitrarily, it follows that $\mathbf{R} = U^\infty$. □

Proof of Corollary 1. Fix an arbitrary $s \in U^\infty$. Then, by Theorem 2, there exists some model in $M \in \mathcal{M}_T$ such that for some state ω (in this model), $\sigma(\omega) = s$ and $\omega \in \mathbb{CB}\mathbf{R}$. Since $M_T \subseteq M_D$ it follows that $M \in \mathcal{M}_D$, and therefore, by Theorem 1, $s \in B^\infty$, thus proving that $U^\infty \subseteq B^\infty$. □

Proof of Proposition 1. Fix a strategic-form game and a model $M \in \mathcal{M}_D$. Suppose that $\omega_1 \in \mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T}$, i.e., $\mathcal{B}^*(\omega_1) \subseteq \mathbf{R} \cap \mathbf{T}$. We want to show that $\sigma(\omega_1) \in U^\infty$. As before, the proof is by induction.

Initial Step. First we show (by contradiction) that, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma(\omega) \notin F^0$ (see Definition 5). Suppose, that there exists an $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma(\omega_2) \in F^0$, that is, $\sigma(\beta)$ is inferior relative to the entire set of strategy profiles S. Then there exists a player i and a strategy $\hat{s}_i \in S_i$ such that

$$\begin{aligned} \pi_i(\hat{s}_i, s_{-i}) &\geq \pi_i(\sigma_i(\omega_2), s_{-i}), \text{ for all } s_{-i} \in S_{-i}, && \text{(A.3)} \\ \pi_i(\hat{s}_i, \sigma_{-i}(\omega_2)) &> \pi_i(\sigma_i(\omega_2), \sigma_{-i}(\omega_2)). && \text{(A.4)} \end{aligned}$$

Hence, for every $\omega \in \mathcal{B}_i(\omega_2)$, $\pi_i(\hat{s}_i, \sigma_{-i}(\omega)) \geq \pi_i(\sigma_i(\omega_2), \sigma_{-i}(\omega))$, that is, $\omega_2 \in \mathbb{B}_i \| \hat{s}_i \succeq \sigma_i(\omega_2) \|$. Furthermore, since $\mathcal{B}^*(\omega_1) \subseteq \mathbf{T} \subseteq \mathbf{T}_i$ and $\omega_2 \in \mathcal{B}^*(\omega_1)$, $\omega_2 \in \mathcal{B}_i(\omega_2)$. Hence, by (A.3), $\| \hat{s}_i \succ \sigma_i(\omega_2) \| \cap \mathcal{B}_i(\omega_2) \neq \varnothing$, so that, by Definition 2, player i is not rational at state ω_2, contradicting the hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1)$ and $\omega_1 \in \mathbb{CB}\mathbf{R}$. Thus, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma(\omega) \in U^0 \backslash F^0 = U^1$ (recall that $U^0 = S$).

Inductive Step. Fix an integer $m \geq 1$ and suppose that, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma(\omega) \in U^m$. We want to show that, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma(\omega) \notin F^m$. Suppose, by contradiction, that there exists an $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma(\omega_2) \in F^m$, that is, $\sigma(\omega_2)$ is inferior relative to U^m. Then there exist a player i and a strategy $\tilde{s}_i \in S_i$ such that

$$\begin{aligned} \pi_i(\tilde{s}_i, \sigma_{-i}(\omega_2)) &> \pi_i(\sigma_i(\omega_2), \sigma_{-i}(\omega_2)), && \text{(A.5)} \\ \pi_i(\tilde{s}_i, s_{-i}) &\geq \pi_i(\sigma_i(\omega_2), s_{-i}), && \text{(A.6)} \\ & \text{for all } s_{-i} \in S_{-i} \text{ such that } (\sigma_i(\omega_2), s_{-i}) \in U^m. \end{aligned}$$

By the induction hypothesis, for every $\omega \in \mathcal{B}^*(\omega_1)$, $(\sigma_i(\omega), \sigma_{-i}(\omega)) \in U^m$. Thus, since $\mathcal{B}_i(\omega_2) \subseteq \mathcal{B}^*(\omega_2) \subseteq \mathcal{B}^*(\omega_1)$ (the latter inclusion follows from transitivity of $\mathcal{B}^*$), we have that, for every $\omega \in \mathcal{B}_i(\omega_2)$, $(\sigma_i(\omega_2), \sigma_{-i}(\omega)) \in U^m$ (recall that, by (Σ_0), if $\omega \in \mathcal{B}_i(\omega_2)$ then $\sigma_i(\omega) = \sigma_i(\omega_2)$). Since $\mathcal{B}^*(\omega_1) \subseteq \mathbf{T} \subseteq \mathbf{T}_i$ and $\omega_2 \in \mathcal{B}^*(\omega_1)$, $\omega_2 \in \mathcal{B}_i(\omega_2)$. Hence $\|\hat{s}_i \succ \sigma_i(\omega_2)\| \cap \mathcal{B}_i(\omega_2) \neq \varnothing$ so that, by Definition 2, player i is not rational at state ω_2, contradicting the hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1)$ and $\omega_1 \in \mathbb{CB}\mathbf{R}$. Thus, we have shown that, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma(\omega) \in \bigcap_{m=1}^{\infty} U^m = U^\infty$, that is, $\omega_1 \in \mathbb{CB}\mathbf{U}^\infty$. □

Proof of Corollary 2. Fix a strategic-form game with ordinal payoffs and a model $M \in \mathcal{M}_D$. Suppose that $\omega_0 \in \mathbb{CB}\mathbf{R} \cap \mathbb{CB}\mathbf{T} \cap \mathbf{T}_\cup$. Since $\omega_0 \in \mathbf{T}_\cup$, there exists a player $i \in I$ such that $\omega_0 \in \mathbf{T}_i$, that is, $\omega_0 \in \mathcal{B}_i(\omega_0)$. Hence, by definition of $\mathcal{B}^*$, $\omega_0 \in \mathcal{B}^*(\omega_0)$. By Proposition 1, $\omega_0 \in \mathbb{CB}\mathbf{U}^\infty$, that is, for every $\omega \in \mathcal{B}^*(\omega_0)$, $\omega \in \mathbf{U}^\infty$. Hence $\omega_0 \in \mathbf{U}^\infty$. □

References

Apt, K.R. & Zvesper, J.A. (2010). The role of monotonicity in the epistemic analysis of strategic games. *Games* 1, 381–394.

Aumann, R.J. (1999). Interactive epistemology I: Knowledge. *International Journal of Game Theory* 28, 263–300.

Battigalli, P. & Bonanno, G. (1999). Recent results on belief, knowledge and the epistemic foundations of game theory. *Research in Economics* 53, 149–225.

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

Bonanno, G. (2008). A syntactic approach to rationality in games with ordinal payoffs. *Logic and the Foundations of Game and Decision Theory (LOFT7)* (Edited by G. Bonanno, W. van der Hoek & M. Wooldridge), Vol. 3 of *Texts in Logic and Games*, Amsterdam University Press, 59–86.

Bonanno, G. & Nehring, K. (1998). On Stalnaker's notion of strong rationalizability and Nash equilibrium in perfect information games. *Theory and Decision* 45, 291–295.

Bonanno, G. & Nehring, K. (2000). Common belief with the logic of individual belief. *Mathematical Logic Quarterly* 46, 49–52.

Börgers, T. (1993). Pure strategy dominance. *Econometrica* 61, 423–430.

Brandenburger, A. & Dekel, E. (1987). Rationalizability and correlated equilibria. *Econometrica* 55, 1391–1402.

Brandenburger, A. & Keisler, J. (2006). An impossibility theorem on beliefs in games. *Studia Logica* 84, 211–240.

Chateauneuf, A., Eichberger, J. & Grant, S. (2007). Choice under uncertainty with the best and worst in mind: Neo-additive capacities. *Journal of Economic Theory* 137, 538–567.

Chellas, B.F. (1980). *Modal logic: an introduction.* Cambridge University Press, Cambridge.

Chen, Y.C., Long, N.V. & Luo, X. (2007). Iterated strict dominance in general games. *Games and Economic Behavior* 61, 299-351.

de Finetti, B. (1949). La 'logica del plausible' secondo la concezione di Polya. *Atti della XLII Riunione, Societa Italiana per il Progresso delle Scienze*, 227–236.

Dekel, E., Fudenberg, D. & Morris, S. (2007). Interim correlated rationalizability. *Theoretical Economics* 2, 15–40.

Dekel, E. & Siniscalchi, M. (2015). Epistemic game theory. In: Peyton Young, H. & Zamir, S. (Editors), *Handbook of Game Theory with Economic Applications, Vol. 4*, Elsevier, 2015, Chapter 12, 619–702.

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

Ely, J. & Peski, M. (2006). Hierarchies of belief and interim rationalizability. *Theoretical Economics* 1, 19–65.

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

FAGIN, R., HALPERN, J.Y., MOSES, Y. & VARDI, M.Y. (1995). *Reasoning about knowledge.* MIT press.

FISHBURN, P. (1986). The axioms of subjective probability. *Statistical Science* 1, 335–358.

GÄRDENFORS, P. (1975). Qualitative probability as an intentional logic. *Journal of Philosophical Logic* 4, 171–185.

HARSANYI, J. (1967-68). Games with incomplete information played by Bayesian players, I-III. *Management Science* 14, 159–182, 320–334, 486–502.

HILLAS, J. & SAMET, D. (2014). Weak dominance: a mystery cracked. *Technical Report, Tel Aviv University.*

HUGHES, G.E. & CRESSWELL, M.J. (1968). *An introduction to modal logic.* Methuen, London, UK.

KOOPMAN, B. (1940). The axioms and algebra of intuitive probability. *Annals of Mathematics* 41, 269–292.

KRAFT, C., PRATT, J.W. & SEIDENBERG, A. (1959). Intuitive probability on finite sets. *Annal of Mathematical Statistics* 30, 408–419.

KRIPKE, S. (1959). A completeness theorem in modal logic. *Journal of Symbolic Logic* 24, 1–14.

LORINI, E. (2016). A minimal logic for interactive epistemology. *Synthese* 193, 725–755.

MACKENZIE, A. (2017). A foundation for probabilistic beliefs with or without atoms. *Working Paper, University of Rochester.*

MEIER, M. (2008). Universal knowledge-belief structures. *Games and Economic Behavior* 62, 53–66.

PEARCE, D.G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52, 1029–1050.

SCOTT, D. (1964). Measurement structures and linear inequalities. *Journal of Mathematical Psychology* 1, 233–247.

Scott, D. & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic* 23, 113–128.

Segerberg, K. (1971). Qualitative probability in a modal setting. *Proceedings of the Second Scandinavian Logic Symposium* 63, 341–352.

Stalnaker, R. (1994). On the evaluation of solution concepts of games. *Theory and Decision* 37, 49–74.

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

Trost, M. (2013). Epistemic characterization of iterated deletion of inferior strategy profiles in preference-based type spaces. *International Journal of Game Theory* 42, 755–776.

van der Hoek, W. (1996). Qualitative modalities. *Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 4, 45–59.

van Ditmarsch, H., Halpern, J., van der Hoek, W. & Kooi, B. (2015). *Handbook of Epistemic Logic*. College Publications.

Villegas, C. (1967). On qualitative probability. *American Mathematical Monthly* 74, 661–669.