Bonanno, Giacomo; Tsakas, Elias

Working Paper
Qualitative analysis of common belief of rationality in strategic-form games

Working Paper, No. 17-5

Provided in Cooperation with:
University of California Davis, Department of Economics

Suggested Citation: Bonanno, Giacomo; Tsakas, Elias (2017) : Qualitative analysis of common belief of rationality in strategic-form games, Working Paper, No. 17-5, University of California, Department of Economics, Davis, CA

This Version is available at:
http://hdl.handle.net/10419/187495

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
We study common belief of rationality in strategic-form games with ordinal utilities, employing a model of qualitative beliefs. We characterize the three main solution concepts for such games, viz., Iterated Deletion of Strictly Dominated Strategies (IDSDS), Iterated Deletion of Boergers-dominated Strategies (IDBS) and Iterated Deletion of Inferior Strategy Profiles (IDIP), by means of gradually restrictive properties imposed on the models of qualitative beliefs. As a corollary, we prove that IDIP refines IDBS, which refines IDSDS.
Qualitative analysis of common belief of rationality in strategic-form games

GIACOMO BONANNO

University of California, Davis

Elias Tsakas

Maastricht University

May 11, 2017

Abstract

In this paper we study common belief in rationality in strategic-form games with ordinal utilities, employing a model of qualitative beliefs. We characterize the three main solution concepts for such games, viz., Iterated Deletion of Strictly Dominated Strategies (IDSDS), Iterated Deletion of Börgers-dominated Strategies (IDBS) and Iterated Deletion of Inferior Strategy Profiles (IDIP), by means of gradually restrictive properties imposed on the models of qualitative beliefs. As a corollary, we prove that IDIP refines IDBS, which refines IDSDS.

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 asking that the rules of the game (i.e., the set of players, the set of strategies and the set of outcomes for 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, rationality and common belief of rationality characterizes correlated rationalizability, i.e., the strategy profiles that survive Iterated Deletion of Strictly Dominated Strategies are exactly those that can be rationally played under common belief of 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 rationality characterizes interim correlated rationalizability

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

There have been several attempts to relax this last assumption by considering ordinal utilities. 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 made by solution concepts for games with ordinal utilities, such as pure-strategy Nash Equilibrium, or pure-strategy Iterated Deletion of Strictly (resp., Weakly) Dominated Strategies. From a theoretical standpoint, the main consequence of sticking to 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. Depending on the notion of rationality that we adopt, there are various solution concepts characterized by rationality and common belief in rationality, e.g., Iterated Deletion of Strictly Dominated Strategies (the pure strategy version of rationalizability à la Bernheim, 1984; Pearce, 1984, henceforth IDSDS), Iterated Deletion of Börgers-dominated Strategies (Börgers, 1993, henceforth IDBS), Iterated Deletion of Inferior Strategy Profiles (the pure strategy version of strong rationalizability à la Stalnaker, 1994, henceforth IDIP).

In this paper we investigate the content of the notion of common belief of rationality in strategic-form games with ordinal utilities within a qualitative context. In particular, following the literature on qualitative beliefs (e.g., see de Finetti, 1949; Koopman, 1940), we endow the standard KD45 Kripke structures with a qualitative likelihood relation for each player at each state. This additional structure allows us to characterize each of the aforementioned solution concepts within our framework in terms of restrictions on the qualitative beliefs, and without needing to vary the notion of rationality that we employ. In particular, we prove that IDSDS is characterized by common belief in rationality in a very broad class of models (Theorem 1); IDBS is characterized by common belief in rationality if we restrict attention to full-support beliefs (Theorem 2); and finally, IDIP is characterized by common belief in rationality if we further restrict attention to correct full-support beliefs (Theorem 3).

With our three main results at hand, not only do we manage to put under the same umbrella the three main solution concepts for games with ordinal utilities that have been studied in the literature, but we also manage to prove that they monotonically refine each other (see Corollary 1). In particular, the fact that we impose stronger and stronger restrictions on our epistemic characterizations, implies that IDIP refines IDBS, while IDBS refines IDSDS. In fact, while the second relationship (between IDSDS and IDBS) seems to be perhaps straightforward, the one between IDBS and IDIP is far from trivial.

Qualitative beliefs have been extensively studied in the literature since the early contributions 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 on 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.

The paper is structured as follows: In Section 2 we introduce qualitative likelihood relations to a game-theoretic framework; in Section 3 we define our notion of rationality and we prove our three characterization results; in Section 4 we present some secondary results; Section 5 concludes; all proof are relegated to the Appendices.
2. Qualitative models of ordinal games

A finite strategic-form game with ordinal payoffs is a quintuple $G = (I, (S_i)_{i \in I}, O, z, (\geq_i)_{i \in I})$, where $I = \{1, 2, \ldots, n\}$ is a finite set of players, $S_i$ is a finite set of strategies (or actions) 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, $z : S \rightarrow O$ is a function that associates with every strategy profile $s = (s_1, \ldots, s_n) \in S$ an outcome $z(s) \in O$, $\geq_i$ is player $i$’s ordinal ranking of the outcomes, i.e., a binary relation on $O$ which is complete (i.e., for all $o, o' \in O$, $o \geq_i o'$ or $o' \geq_i o$) and transitive (i.e., for all $o, o', o'' \in O$, if $o \geq_i o'$ and $o' \geq_i o''$ then $o \geq_i o''$). The interpretation of $o \geq_i o'$ 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 $\langle O, z, (\geq_i)_{i \in I} \rangle$ with a list $(\pi_i)_{i \in I}$ of payoff functions, where $\pi_i : S \rightarrow \mathbb{R}$ 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 $z(s) \geq z(s')$. 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 $\pi_i$ with any other function obtained by composing $\pi_i$ with an arbitrary strictly increasing function on the set of real numbers.\(^2\)

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”, which represents a possible context in which the game is played. The players’ beliefs are represented by means of a finite $KD45$ Kripke frame $\langle \Omega, (\mathcal{B}_i)_{i \in I} \rangle$, where $\Omega$ is a finite set of states (or possible worlds) and for every player $i \in I$, $\mathcal{B}_i$ is a binary relation on $\Omega$. We denote by $\mathcal{B}_i(\omega)$ the set of states that are reachable from $\omega$ by $\mathcal{B}_i$, that is, $\mathcal{B}_i(\omega) = \{ \omega' \in \Omega : \omega \mathcal{B}_i \omega' \}$. The relation $\mathcal{B}_i$ is assumed to be serial (i.e., for all $\omega \in \Omega$, $\mathcal{B}_i(\omega) \neq \emptyset$), transitive (i.e., if $\omega' \in \mathcal{B}_i(\omega)$ then $\mathcal{B}_i(\omega') \subseteq \mathcal{B}_i(\omega)$) and euclidean (i.e., if $\omega' \in \mathcal{B}_i(\omega)$ then $\mathcal{B}_i(\omega) \subseteq \mathcal{B}_i(\omega')$). Obviously, by transitivity and euclideanness, we obtain that for every $\omega' \in \mathcal{B}_i(\omega)$, $\mathcal{B}_i(\omega') = \mathcal{B}_i(\omega)$.\(^3\)

As usual, $\mathcal{B}_i(\omega)$ is interpreted as the set of states that are doxastically accessible to player $i$ at state $\omega$, that is, the states that she considers possible according to her beliefs. At state $\omega$, player $i$ is said to believe an event $E \subseteq \Omega$ if and only if $\mathcal{B}_i(\omega) \subseteq E$, i.e., if $E$ is true at every state that she considers possible at $\omega$. Thus we can define a belief operator $\mathcal{B}_i : \Omega \rightarrow 2^\Omega$ (where $2^\Omega$ denotes the collection of subsets of $\Omega$) as follows: for $E \subseteq \Omega$, we denote by $\mathcal{B}_i E$ the event that player $i$ believes event $E$, i.e., formally, $\mathcal{B}_i E = \{ \omega \in \Omega : \mathcal{B}_i(\omega) \subseteq E \}$. Seriality of the accessibility relation $\mathcal{B}_i$ guarantees that the player’s beliefs are consistent (i.e., it is not the case that she believes $E$ and also $\neg E$: $\mathcal{B}_i E \subseteq \neg \mathcal{B}_i \neg E$), while transitivity corresponds to positive introspection (i.e., if the player believes $E$ then she believes that she believes $E$: for all $E \subseteq \Omega$, $\mathcal{B}_i E \subseteq \mathcal{B}_i \mathcal{B}_i E$) and euclideanness corresponds to negative introspection (i.e., if the player does not believe $E$ then she believes that she does not believe $E$: for all $E \subseteq \Omega$, $\neg \mathcal{B}_i E \subseteq \mathcal{B}_i \neg \mathcal{B}_i E$).\(^4\) Note that erroneous beliefs are not ruled out: it is possible for a player to believe $E$ even though $E$ is actually false, that is, it may be the case that $\omega \in \mathcal{B}_i E$ even though $\omega \notin E$. Erroneous beliefs are ruled out if one imposes the restriction that $\mathcal{B}_i$ be reflexive (i.e., $\omega \in \mathcal{B}_i(\omega)$ for all $\omega \in \Omega$). If reflexivity is added to the above assumptions, then $\mathcal{B}_i$ becomes an equivalence relation and thus gives rise to a partition of $\Omega$; in such a case it is

\(^2\)This is in contrast to von Neumann-Morgenstern utility functions whose properties (e.g., risk attitudes) are preserved only under positive affine transformations.

\(^3\)In the game-theoretic literature, it is more common to view $\mathcal{B}_i$ as a function that associates with every state $\omega \in \Omega$ a set of states $\mathcal{B}_i(\omega) \subseteq \Omega$ and to call such a function a possibility correspondence or information correspondence. Of course, the two views (binary relation and possibility correspondence) are equivalent. 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).

\(^4\)For every event $F \subseteq \Omega$, $\neg F$ denotes the complement of $F$, that is, $\neg F = \Omega \setminus F$. 

3
common to use the term “knowledge” rather than “belief”.

While the relation $\mathcal{B}_i$ captures what player $i$ believes or is certain of, a second (complete) relation $\succeq^\omega_i$ on the algebra of subsets of $\mathcal{B}_i(\omega)$ captures the relative likelihood judgments of player $i$ on the events contained in $\mathcal{B}_i(\omega)$: formally, for each $E, F \subseteq \mathcal{B}_i(\omega)$, $E \succeq^\omega_i F$ if, at state $\omega$, player $i$ considers event $E$ to be at least as likely as event $F$. Define $E \succ^\omega_i F$ as $E \succeq^\omega_i F$ and $F \nprec^\omega_i E$, i.e., the interpretation of $E \succ^\omega_i F$ is that, at state $\omega$, player $i$ considers event $E$ to be strictly more likely than event $F$. Similarly define $E \sim^\omega_i F$ as $E \succeq^\omega_i F$ and $F \succeq^\omega_i E$, i.e., the interpretation of $E \sim^\omega_i F$ is that, at state $\omega$, player $i$ considers event $E$ to be as likely as event $F$. An event $E \subseteq \mathcal{B}_i(\omega)$ is null for player $i$ at state $\omega$ if $E \sim^\omega_i \varnothing$, and it is non-null otherwise.

The following minimal properties on the qualitative likelihood relation $\succeq^\omega_i$ are imposed: for every $i \in I$, every $\omega \in \Omega$ and every $E, F \subseteq \mathcal{B}_i(\omega)$,

$$(L_1) \mathcal{B}_i(\omega) \succeq^\omega_i \varnothing,$$

$$(L_2) \text{ if } E \subseteq F \text{ then (a) } F \succeq^\omega_i E, \text{ and (b) if } E \succ^\omega_i \varnothing \text{ then } F \succ^\omega_i \varnothing,$$

$$(L_3) \text{ if } E \sim^\omega_i \varnothing, \text{ } F \sim^\omega_i \varnothing \text{ and } E \cup F \neq \mathcal{B}_i(\omega) \text{ then } E \cup F \sim^\omega_i \varnothing.$$

Intuitively, $(L_1)$ guarantees that not every event is deemed null; $(L_2)$ says that whenever $E$ implies $F$, player $i$ must deem the consequence $F$ at least as likely as its cause $E$; $(L_3)$ says that the union of any two null events is also null, unless this union covers the states deemed possible by the player. As usual, we require each player to know her own qualitative likelihood judgement, i.e., formally, the following condition is imposed: for every $i \in I$ and every $\omega \in \Omega$,

$$(L_4) \text{ if } \omega' \in \mathcal{B}_i(\omega) \text{ then } \succeq^\omega_i = \succeq^{\omega'}_i.$$

Further properties will be introduced as needed.\(^5\)

So far we have introduced the notion of a frame rather abstractly, viz., we have not assigned a meaning to each event $E \subseteq \Omega$. Let us do so, by introducing a strategy function $\sigma_i : \Omega \to S_i$ for each player $i \in I$. Then, each state $\omega \in \Omega$ is associated with the strategy profile $\sigma(\omega) = (\sigma_1(\omega), \ldots, \sigma_n(\omega))$. Moreover, we denote by $\sigma_{-i}(\omega)$ the profile of strategies played, at $\omega$, by the players other than $i$, that is, $\sigma_{-i}(\omega) = (\sigma_1(\omega), \ldots, \sigma_{i-1}(\omega), \sigma_{i+1}(\omega), \ldots, \sigma_n(\omega))$; thus the entire profile, $\sigma(\omega)$, can also be denoted by $(\sigma_i(\omega), \sigma_{-i}(\omega))$. Then we impose the following standard property: for every $i \in I$ and every $\omega \in \Omega$,

$$(\Sigma_0) \text{ if } \omega' \in \mathcal{B}_i(\omega) \text{ then } \sigma_i(\omega') = \sigma_i(\omega').$$

That is, intuitively each player knows her own strategy at every state.

**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 $M = \langle \Omega, (\mathcal{B}_i)_{i \in I}, (\succeq^\omega_i)_{\omega \in \Omega}, (\sigma_i)_{i \in I} \rangle$, where $\langle \Omega, (\mathcal{B}_i)_{i \in I} \rangle$ is a $KD45$ Kripke frame, $\succeq^\omega_i$ is a qualitative likelihood relation satisfying $(L_1) - (L_4)$ and $\sigma_i : \Omega \to S_i$ is a strategy function satisfying $(\Sigma_0)$. Let $\mathcal{M}_0$ be the class of all finite qualitative doxastic models of $G$.

\(^5\)As we have already mentioned, early contributions in this literature focused on the problem of “representing a qualitative likelihood relation with a probability measure” (e.g., see Fishburn, 1986). It is not difficult to verify that $(L_1) - (L_3)$ do not suffice for a probability-measure representation to be obtained, e.g., the well-known example of Kraft et al. (1959) satisfies our three basic properties and yet the likelihood relation cannot be probabilistically represented. In fact, even in the presence additional properties – that we will impose in the upcoming sections – our likelihood relations will not always be represented by a probability measure.
3. Common belief of rationality

Fix a player $i$ and two strategies $a, b \in S_i$ of player $i$, and denote by $\|b \geq a\|$ the event that strategy $b$ yields at least as high a payoff for player $i$ as strategy $a$, that is, $\|b \geq a\| = \{ \omega \in \Omega : \pi_i(b, \sigma_{-i}(\omega)) \geq \pi_i(a, \sigma_{-i}(\omega)) \}$. Similarly, $\|b > a\| = \{ \omega \in \Omega : \pi_i(b, \sigma_{-i}(\omega)) > \pi_i(a, \sigma_{-i}(\omega)) \}$ is the event that strategy $b$ yields a strictly higher payoff for player $i$ than strategy $a$.

**Definition 2.** Player $i$ is rational at state $\omega$ whenever, for all $b \in S_i$,

$$\text{if } \omega \in B_i \|b \geq \sigma_i(\omega)\| \text{ then } \|b > \sigma_i(\omega)\| \cap B_i(\omega) \sim_i^\omega \emptyset.$$  \hspace{1cm} (1)

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

Intuitively, if at $\omega$ player $i$ believes that $b$ yields at least as high a payoff as the chosen strategy $\sigma_i(\omega)$ at every state deemed possible, then the event that $b$ yields a strictly higher payoff than $\sigma_i(\omega)$ is a null event for player $i$ at $\omega$.

We want to investigate the implications of common belief of rationality. Given an event $E$, let $B_i E = \bigcap_{i \in I} B_i E$ denote the event that all the players believe $E$. Then the event that $E$ is commonly believed, denoted by $\mathbb{C}BE$, is defined as the infinite intersection $\mathbb{C}BE = B_i E \cap B_i B_j E \cap B_i B_j B_k E \cap \cdots$, that is, the event that everybody believes $E$, and everybody believes that everybody believes $E$, and everybody believes that everybody believes $E$, and so on. It is well-known that, for every state $\omega$ and every event $E$, $\omega \in \mathbb{C}BE$ if and only if $B^*(\omega) \subseteq E$, where $B^*(\omega)$ is the transitive closure of $\bigcup_{i \in I} B_i(\omega)$. We are interested in the event that there is common belief of rationality, henceforth denoted by $\mathbb{C}BR$. In particular, we ask the question: *which strategy profiles are compatible with states in $\mathbb{C}BR$?*

**Definition 3.** We say that common belief of rationality in a class of models $\mathcal{M} \subseteq \mathcal{M}_0$ (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 $\omega \in \mathbb{C}BR$ then $\sigma(\omega) \in S^*$,

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

In the following sections, we will epistemically characterize three well-known solution concepts for ordinal strategic-form games (viz., Iterated Deletion of Strictly Dominated Strategies, Iterated Deletion of Börgers-dominated Strategies, and Iterated Deletion of Inferior Profiles) 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 Strictly Dominated Strategies

A strategy $a \in S_i$ of player $i$ is strictly dominated if there is another strategy $b \in S_i$ such that $\pi(b, s_{-i}) > \pi(a, s_{-i})$ for every strategy profile $s_{-i} \in S_{-i}$ of the players other than $i$, where as usual $S_{-i} = S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$. Iterated Deletion of Strictly Dominated Strategies (IDSDS) is the following algorithm: reduce the game by deleting, for each player, all the strategies that are strictly dominated and then repeat the procedure in the reduced game, and so on, until there are no strictly dominated strategies left. Formally, the procedure is defined as follows:

---

$B^*$ is thus defined as follows: $\omega' \in B^*(\omega)$ if and only if there is a sequence $\{\omega_1, \ldots, \omega_m\}$ in $\Omega$ 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 $j = 1, \ldots, m-1$, $\omega_{j+1} \in B_{i_j}(\omega_j)$.
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 $S^0_i = S_i$, and let $D^0_i \subsetneq S^0_i$ be the set of $i$'s strategies that are strictly dominated in $G^0 = G$;

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

Let $S^\infty = \bigcap_{m=0}^\infty S^m_i$. The strategy profiles in $S^\infty = S^\infty_1 \times \cdots \times S^\infty_n$ are those surviving IDSDS.

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

This is the pure-strategy version of the procedure that was first introduced by Bernheim (1984); Brandenburger and Dekel (1987); Tan and Werlang (1988). Recall that we have restricted attention to ordinal payoffs and thus a pure strategy of player $i$ can be strictly dominated only by another pure strategy; in other words, domination by a mixed strategy is not meaningful in this context.

Theorem 1 (Characterization of IDSDS). If no restrictions (besides (L1) – (L4)) are imposed on the relative likelihood relations $(\succeq_i)_{i \in I}$ then common belief of rationality characterizes IDSDS. Formally,

$$(A_0) \text{ in every model } M \in \mathcal{M}_0, \text{ if } \omega \in \mathbb{CB}, \text{ then } \sigma(\omega) \in S^\infty,$$

$$(B_0) \text{ for every } s \in S^\infty, \text{ there exists a model } M \in \mathcal{M}_0 \text{ and a state } \omega \text{ in that model such that } \sigma(\omega) = s \text{ and } \omega \in \mathbb{CB}. \text{ }$$

However, it is possible that a strategy that survives IDSDS is compatible with common belief of rationality only at states $\omega$ where the only non-null event is $B_i(\omega)$ itself.\(^7\) To see this, consider the game in Figure 1, where no strategy is strictly dominated and thus IDSDS does not eliminate any strategy, i.e., $S^\infty = \{a, b, c\} \times \{d, e\}$. Consider an arbitrary model $M \in \mathcal{M}_0$ and a state $\omega_0 \in \Omega$

\begin{tabular}{c|cc}
  & $d$ & $e$ \\
\hline
$a$ & 1, 1 & 1, 0 \\
$b$ & 2, 0 & 1, 1 \\
$c$ & 1, 1 & 2, 0 \\
\end{tabular}

Figure 1: IDSDS.

where Player 1 is rational and plays $a$. That is, $\omega_0 \in \mathbb{R}_1$ and $\sigma_1(\omega_0) = a$ (thus implying $\sigma_1(\omega) = a$ for every $\omega \in B_1(\omega_0)$ by ($\Sigma_0$)). By seriality, $B_1(\omega_0) \neq \emptyset$. Note that $B_1(\omega_0) = D \cup E$, where $D = \{\omega \in B_1(\omega_0) : \sigma_2(\omega) = d\}$ and $E = \{\omega \in B_1(\omega_0) : \sigma_2(\omega) = e\}$. If $E = \emptyset$ then $B_1(\omega_0) \subset ||b > a||$

\(^7\)Obviously such likelihood relation cannot be represented by a probability measure. Instead it could be represented by a capacity, and in particular by a (Hurwicz) capacity that assigns weight 0 to every strict subset of $B_i(\omega)$ and weight 1 to $B_i(\omega)$ itself (e.g., see Chateauneuf et al., 2007).
and thus, since $B_1(\omega_0) \succ i_{\omega_0} \emptyset$, it follows that Player 1 is not rational at state $\omega_0$, contradicting our hypothesis that $\omega_0 \in R_1$. Similarly, if $D = \emptyset$ then $B_1(\omega_0) \subseteq \|c > a\|$ and thus, since $B_1(\omega_0) \succ i_{\omega_0} \emptyset$, it follows that Player 1 is not rational at state $\omega_0$, yielding a contradiction. Suppose now that $D \neq \emptyset$ and $E \neq \emptyset$. Then, since $\omega_0 \in B_1(\|b \geq a\| \cap \|c \geq a\|)$ and Player 1 is rational at state $\omega_0$, it must be that $D \sim i_{\omega_0} \emptyset$ and $E \sim i_{\omega_0} \emptyset$. It follows that every event $F \subseteq B_1(\omega_0)$ with $F \neq B_1(\omega_0)$ is such that $F \sim i_{\omega_0} \emptyset$.

3.2. Iterated Deletion of Börners-dominated Strategies

Börners (1993) introduced a refined notion of pure-strategy dominance. In particular, 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 $i$ (note that $X_{-i}$ need not have a product structure). 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 Börners-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 $X_{-i}$. 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 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 reduced games $\{G^0, G^1, \ldots, G^n, \ldots\}$ as follows: for each $i \in I$,

(5.1) let $B_i^0 = S_i$, and let $E_i^0 \subseteq B_i^0$ be the set of $i$'s strategies that are $B$-dominated in $G^0 = G$;

(5.2) 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 $E_i^{m-1} \subseteq B_i^{m-1}$ 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.

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

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 $\{e\}$. However, as we saw above, unless additional restrictions (besides $(L_1) - (L_4)$) are imposed on the relative likelihood relations, there exists a model and a state within this model such that strategy $a$ is rational according to Definition 2. In what follows we shall restrict attention to models with cautious players, i.e., with players who have full-support beliefs.

Definition 6. A finite qualitative doxastic model of a strategic-form game with ordinal payoffs has full support if, for every $i \in I$ and every $\omega \in \Omega$,

$$(L_5) \{\omega'\} \succ i_{\omega} \emptyset \text{ for all } \omega' \in B_i(\omega).$$

Let $M_1 \subseteq M_0$ denote the class of finite full-support qualitative doxastic models.

Theorem 2 (Characterization of IDBS). If the relative likelihood relations $(\succ i_{\omega})_{i \in I, \omega \in \Omega}$ are full-support then common belief of rationality characterizes IDBS. Formally,

\[\text{By Property } (L_2) \text{ of the relative likelihood relation, every subset of } D \text{ is null and so is every subset of } E \text{ and by property } (L_3) \text{ the union of two such sets is null, unless it is equal to } B_i(\omega_0).\]
(A₁) in every model \( M \in \mathcal{M}_1 \), if \( \omega \in \mathbb{C} \mathbb{R} \) then \( \sigma(\omega) \in B^\infty \),

(B₁) for every \( s \in B^\infty \), there exists a model \( M \in \mathcal{M}_1 \) and a state \( \omega \) in that model such that \( \sigma(\omega) = s \) and \( \omega \in \mathbb{C} \mathbb{R} \).

Note that, in order to “rationalize” a strategy profile in \( B^\infty \), 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, \( (a, d) \in B^\infty \).⁹ Consider an arbitrary full-support model of this game and a state \( \omega_0 \) such that \( \sigma(\omega_0) = (a, d) \). Since, for every \( s_2 \in \{c, d\} \), \( \pi_1(b, s_2) \geq \pi_1(a, s_2) \), \( ||b \geq a|| = \Omega \) and thus \( B_1(\omega_0) \subseteq ||b \geq a|| \). That is, \( \omega_0 \in B_1 ||b \geq a|| \). Hence, if Player 1 is rational at \( \omega_0 \) (according to Definition 2) then \( ||b > a|| \cap B_1(\omega_0) = \emptyset \).¹⁰ Thus, \( \sigma_2(\omega) = c \) for all \( \omega \in B_1(\omega_0) \). In particular, it must be that \( \omega_0 \notin B_1(\omega_0) \). Thus at state \( \omega_0 \) 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

![Figure 2: IDBS.](image)

the consequences of ruling out false beliefs, while maintaining caution, i.e. full-support beliefs.

### 3.3. 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 procedures considered above (viz., IDSDS and IDBS), this procedure deletes entire 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 \( i \) (i.e., \( s_i \) need not belong 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 \( \pi_i(s_i, s_{-i}) \geq \pi_i(x_i, s_{-i}) \). 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 7.** 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 \( \{T_0, T_1, \ldots, T^m, \ldots\} \) as follows:

(7.1) let \( T^0 = S \), and let \( I^0 \subseteq T^0 \) be the set of inferior strategy profiles relative to \( T^0 \);

---

⁹ For Player 1, \( a \) is weakly dominated by \( b \) relative to \( \{d\} \) and \( \{c, d\} \) but not relative to \( \{c\} \) and for Player 2 \( d \) is weakly dominated by \( c \) relative to \( \{a\} \) but not relative to \( \{b\} \) or \( \{a, b\} \).

¹⁰ Suppose that \( \omega_0 \in R_1 \). Then, by Definition 2, since \( \omega_0 \in B_1 ||b \geq a|| \), it must be that \( ||b > a|| \cap B_1(\omega_0) \sim^{\omega_0} \emptyset \). Fix an arbitrary \( \omega \in B_1(\omega_0) \) and suppose that \( \omega \notin ||b > a|| \) so that \( ||b > a|| \cap B_1(\omega_0) \supseteq \{\omega\} \), noting that, by \( (\Sigma_0) \), \( \sigma_1(\omega) = \sigma_1(\omega_0) = a \). By the assumption of full support, \( \{\omega\} \sim^{\omega_0} \emptyset \) and thus, by \( (L_2) \), \( ||b > a|| \cap B_1(\omega_0) \sim^{\omega_0} \emptyset \), yielding a contradiction.
for each \( m \geq 1 \) let \( T^m = T^{m-1} \setminus I^{m-1} \), where \( I^{m-1} \subset T^{m-1} \) is the set of strategy profiles in \( T^{m-1} \) that are inferior relative to \( T^{m-1} \).

Then \( T^\infty = \bigcap_{m=0}^{\infty} T^m \) denotes the strategy profiles surviving IDIP.

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

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

We now turn to investigating the consequences of ruling out false beliefs. At state \( \omega \) player \( i \) has correct beliefs if \( \omega \) is one of the states that player \( i \) considers possible at \( \omega \), that is, if \( \omega \in B_i(\omega) \).

**Definition 8.** A finite qualitative doxastic model of a strategic-form game with ordinal payoffs rules out false beliefs if, for every \( i \in I \) and every \( \omega \in \Omega \),

\[
(K_0) \quad \omega \in B_i(\omega).
\]

Let \( \mathcal{M}_2 \subset \mathcal{M}_1 \) denote the class of finite full-support qualitative doxastic models that rule out false beliefs.

**Theorem 3** (Characterization of IDIP). If the relative likelihood relations \((\succeq^\omega_i)_{\omega \in \Omega}\) are full-support and the binary relations \((B_i)_{i \in I}\) are reflexive then common belief of rationality characterizes IDIP. Formally,

\[
(A_2) \quad \text{in every model } M \in \mathcal{M}_2, \text{ if } \omega \in \mathcal{CB} \text{ then } \sigma(\omega) \in T^\infty,
\]

\[
(B_2) \quad \text{for every } s \in T^\infty, \text{ there exists a model } M \in \mathcal{M}_2 \text{ and a state } \omega \text{ in that model such that } \sigma(\omega) = s \text{ and } \omega \in \mathcal{CB}.
\]

Property \((K_0)\) says that no player can have false beliefs. This is actually stronger 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 \( T^\infty \), common belief that all players have correct beliefs is not sufficient (see Section 4.2).
4. Discussion

4.1. Monotonicity result

A direct implication of our three theorems is the following (monotonicity) result, which proves that IDIP is a refinement of IDBS, which is a refinement of IDSDS.

**Corollary 1 (Monotonicity result).** \( T^\infty \subseteq B^\infty \subseteq R^\infty \).

The proof follows directly from \( \mathcal{M}_2 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}_0 \). The second part of the result is not very surprising and can also be proven directly, viz., it can be shown that for every \( m \geq 0 \) it is the case that \( B^m \subseteq R^m \). However, this is not the case with the first part of our monotonicity result, which is far from trivial. The difficulty of proving the result stems from the fact that there exist games where \( B^m \subsetneq T^m \) for some \( m > 0 \), as illustrated in the game in Figure 4.

![Figure 4: Monotonicity.](image)

In this game, \( c \) is B-dominated, while no other strategy is subsequently eliminated. That is, formally \( B^\infty = B^1 = \{a, b\} \times \{d, e\} \). On the other hand, the only inferior strategy profile relative to the entire game is \( (c, e) \), and therefore \( T^1 \supseteq B^1 \). But then, \( (c, d) \) is inferior relative to \( T^1 \), thus implying that \( B^\infty = B^2 = T^2 = T^\infty \), consistently with the conclusions of our Corollary 1.

4.2. Correct beliefs

As we have already mentioned above, common belief in correct beliefs does not suffice for a strategy that survives IDIP to be played. Formally, let \( C_i = \{ \omega \in \Omega : \omega \in B_i(\omega) \} \) be the event that player \( i \) has correct beliefs, and let \( C = \bigcup_{i \in I} C_i \) be the event that at least one player has correct beliefs and \( C = \bigcap_{i \in I} C_i \) the event that all players have correct beliefs. As the following example shows, it is possible that \( \omega \in \text{CBR} \cap \text{CBC} \) and yet the strategy profile played at \( \omega \) does not survive IDIP.

![Figure 5: Common belief in correct beliefs.](image)
Consider the following model of the game in Figure 5: \( \Omega = \{ \omega_1, \omega_2, \omega_3 \}, B_1(\omega_1) = \{ \omega_1 \}, B_1(\omega_2) = B_1(\omega_3) = \{ \omega_3 \}, B_2(\omega_1) = B_2(\omega_3) = \{ \omega_1 \}, B_2(\omega_2) = \{ \omega_3 \}, \sigma_1(\omega_1) = b, \sigma_1(\omega_2) = \sigma_1(\omega_3) = a, \sigma_2(\omega_1) = \sigma_2(\omega_2) = c \) and \( \sigma_2(\omega_3) = d \). Then \( \sigma(\omega_2) = (a, d) \notin T^\infty \) and yet \( \omega_2 \in \text{CBR} \cap \text{CBC} \) (in fact, \( B^*(\omega_2) = \{ \omega_1, \omega_3 \}, C = \{ \omega_1, \omega_3 \} \) and \( R = \Omega \)). Note that, in this model, at state \( \omega_2 \) both players have false beliefs. This is because, although it is common belief at \( \omega_2 \) that only strategy profiles in \( T^\infty \) are played, the strategy profile \textit{actually} played does not belong to \( T^\infty \).

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

**Proposition 1.** If the relative likelihood relations \( (\succeq_i)_{i \in I}^{\omega \in \Omega} \) are full-support then common belief of rationality and common belief of correct beliefs imply common belief in IDIP. Formally, \( \text{CBR} \cap \text{CBC} \subseteq \text{CBT}^\infty \) in every model \( M \in \mathcal{M}_1 \).

The condition that there is common belief that all players have correct beliefs \( (\omega \in \text{CBC}) \) is necessary for Proposition 1. To see this, consider the game in Figure 6, where \( T^\infty = \{ (a, c), (b, c) \} \). Consider the following model of this game: \( \Omega = \{ \omega_1, \omega_2 \}, B_1(\omega_1) = B_1(\omega_2) = \{ \omega_2 \}, B_2(\omega_1) = \{ \omega_1 \} \).

\[
\begin{array}{c|cc}
\text{Player 1} & c & d \\
\hline
a & 1, 1 & 1, 1 \\
b & 1, 1 & 2, 0
\end{array}
\]

Figure 6: Correct beliefs.

\( \{ \omega_1 \}, B_2(\omega_2) = \{ \omega_2 \}, \sigma_1(\omega_1) = \sigma_1(\omega_2) = a, \sigma_2(\omega_1) = d, \sigma_2(\omega_2) = c \). Then \( R = \text{CBR} = \Omega \), while \( T^\infty = \{ \omega_2 \} \) (since \( \sigma(\omega_1) = (a, d) \notin T^\infty \)). Since \( B^*(\omega_1) = \{ \omega_1, \omega_2 \}, \omega_1 \in \text{CBR} \) but \( \omega_1 \notin \text{CBT}^\infty \). In this model, at state \( \omega_1 \) Player 1 has false beliefs \( (C_1 = \{ \omega_2 \}) \) and thus \( \omega_1 \notin \text{CBC} \).

The following Corollary shows that if, to the hypotheses of Proposition 1, we add the further hypothesis that \textit{at least one player} does not have false beliefs, then it follows that the strategy profile \textit{actually played} also belongs to \( T^\infty \). Recall that \( C_u = \bigcup_{i \in I} C_i \) is the event that at least one player has correct beliefs.

**Corollary 2.** If the relative likelihood relations \( (\succeq_i)_{i \in I}^{\omega \in \Omega} \) are full-support then common belief of rationality, common belief of correct beliefs and correct beliefs of at least one player imply IDIP. Formally, \( \text{CBR} \cap \text{CBC} \cap C_u \subseteq T^\infty \) in every model \( M \in \mathcal{M}_1 \).

5. Conclusion

In this paper 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 contribution is threefold. Firstly, this is the first paper in the literature to embed qualitative likelihood relations into a game-theoretic model. Secondly, we manage to characterize three well-known solution concepts for games with ordinal payoffs in terms of common belief of rationality,
without needing to vary the notion of rationality that we employ, but rather by gradually strengthening the properties of the model of qualitative beliefs that we use. Finally, as a consequence of our characterization results, we prove that the aforementioned solution concepts monotonically refine each other, viz., IDIP refines IDBS, which in turn refines IDSDS. Notably, the first refinement result is far from trivial to prove.

A. Proofs of Section 3

Proof of Theorem 1. (A0) Fix a strategic-form game with ordinal payoffs and a model $M \in \mathcal{M}_0$ of the game. Suppose that $\omega_1 \in \mathbb{CBR}$. That is, $\mathcal{B}^*(\omega_1) \subseteq \mathbb{R}$. We want to show that $\sigma(\omega_1) \in S^\infty$. The proof is by induction.

Initial Step. First we show (by contradiction) that, for every player $i \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \notin D_{i}^0$ (see Definition 4). Suppose not. Then there exist a player $i$ and an $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma_i(\omega_2) \in D_{i}^0$, that is, strategy $\sigma_i(\omega_2)$ is strictly dominated in $G$ by some other strategy $\tilde{s}_i \in S_i$: for every $s_{-i} \in S_{-i}$, $\pi_i(\tilde{s}_i, s_{-i}) > \pi_i(\sigma_i(\omega_2), s_{-i})$. Thus, for every $\omega \in \mathcal{B}_i(\omega_2)$, $\pi_i(\tilde{s}_i, \sigma_{i-}(\omega)) > \pi_i(\sigma_i(\omega_2), \sigma_{i-}(\omega))$, that is, $\|\tilde{s}_i > \sigma_i(\omega_2)\| \cap \mathcal{B}_i(\omega_2) = \mathcal{B}_i(\omega_2)$. Since $\mathcal{B}_i(\omega_2) \succ^\omega_i \emptyset$, it follows from Definition 2 that $\omega_2 \notin \mathcal{R}_i$, thus contradicting the hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1) \subseteq \mathbb{R}$ (recall that $\mathbb{R} \subseteq \mathcal{R}_i$). Thus, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \in S_i \setminus D_{i}^0 = S_i^1$.

Inductive Step. Fix an integer $m \geq 1$ and suppose that, for every player $j \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_j(\omega) \in S_j^m$. We want to show (again by contradiction) that, for every player $i \in I$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \notin D_{i}^m$. Suppose not. Then there exist a player $i$ and a $\omega_2 \in \mathcal{B}^*(\omega_1)$ such that $\sigma_i(\omega_2) \in D_{i}^m$, that is, strategy $\sigma_i(\omega_2)$ is strictly dominated in $G^m$ by some other strategy $\tilde{s}_i \in S_i^m$. Since, by hypothesis, for every player $j$ and for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_j(\omega) \in S_j^m$, it follows – since $\mathcal{B}_i(\omega_2) \subseteq \mathcal{B}^*(\omega_2) \subseteq \mathcal{B}^*(\omega_1)$ (the latter inclusion follows from transitivity of $\mathcal{B}^*$) – that, for every $\omega \in \mathcal{B}_i(\omega_2)$, $\pi_i(\tilde{s}_i, \sigma_{i-}(\omega)) > \pi_i(\sigma_i(\omega_2), \sigma_{i-}(\omega))$, that is, $\|\tilde{s}_i > \sigma_i(\omega_2)\| \cap \mathcal{B}_i(\omega_2) = \mathcal{B}_i(\omega_2)$. Since $\mathcal{B}_i(\omega_2) \succ^\omega_i \emptyset$, it follows from Definition 2 that $\omega_2 \notin \mathcal{R}_i$, contradicting the hypothesis that $\omega_2 \in \mathcal{B}^*(\omega_1) \subseteq \mathbb{R}$. Thus, for every $\omega \in \mathcal{B}^*(\omega_1)$, $\sigma_i(\omega) \in \bigcap_{m=1}^\infty S_i^m = S_i^\infty$.

It only remains to show that $\sigma_i(\omega_1) \in S_i^\infty$. Fix an arbitrary $\omega_2 \in \mathcal{B}_i(\omega_1)$. Since $\mathcal{B}_i(\omega_1) \subseteq \mathcal{B}^*(\omega_1)$, $\omega_2 \in \mathcal{B}^*(\omega_1)$. Thus $\sigma_i(\omega_2) \in S_i^\infty$. By $(\Sigma_0)$, since $\omega_2 \in \mathcal{B}_i(\omega_1)$, $\sigma_i(\omega_2) = \sigma_i(\omega_1)$. Thus $\sigma_i(\omega_1) \in S_i^\infty$. □

Proof of Theorem 2. (A1) Fix a strategic-form game with ordinal payoffs and a model $M \in \mathcal{M}_1$. Suppose that $\omega_1 \in \mathbb{CBR}$ (that is, $\mathcal{B}^*(\omega_1) \subseteq \mathbb{R}$). We want to show that $\sigma(\omega_1) \in B^\infty$. The proof is by induction.

Initial 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 5). Suppose not. Then there exist a player $i$ and a $\omega_2 \in \mathcal{B}^*(\omega_1)$ such

Proof of Theorem 2. (A1) Fix a strategic-form game with ordinal payoffs and a model $M \in \mathcal{M}_1$. Suppose that $\omega_1 \in \mathbb{CBR}$ (that is, $\mathcal{B}^*(\omega_1) \subseteq \mathbb{R}$). We want to show that $\sigma(\omega_1) \in B^\infty$. The proof is by induction.

Initial 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 5). Suppose not. Then there exist a player $i$ and a $\omega_2 \in \mathcal{B}^*(\omega_1)$ such
that \( \sigma_i(\omega_2) \in E^0_1 \), that is, strategy \( \sigma_i(\omega_2) \) of player \( i \) is B-dominated relative to \( S_{-i} \), i.e., for every \( X_{-i} \subseteq S_{-i} \) there exists a strategy \( s_i \in S_i \) such that (1) for all \( x_{-i} \in X_{-i} \), \( \pi_i(s_i, x_{-i}) \geq \pi_i(\sigma_i(\omega_2), x_{-i}) \), and (2) 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 \( X_{-i} = \sigma_{-i}(B_i(\omega_2)) = \{ s_{-i} \in S_{-i} : s_{-i} = \sigma_{-i}(\omega) \text{ for some } \omega \in B_i(\omega_2) \} \). Let \( s_i \in S_i \) and \( \hat{x}_{-i} \in X_{-i} \) satisfy (1) and (2) and let \( \omega \in B_i(\omega_2) \) be such that \( \sigma_{-i}(\omega) = \hat{x}_{-i} \). Then, by (1), \( \omega_2 \in \mathcal{B}_i \| s_i \geq \sigma_i(\omega_2) \) and by (2) \( ||s_i > \sigma_i(\omega_2)|| \cap B_i(\omega_2) \supseteq \{ \hat{x} \} \). By \((L5)_\omega \), \( \hat{x} \in \mathcal{B}_i \| \mathcal{B}_i(\omega_2) \supseteq \{ \hat{x} \} \). By \((L2)_\omega \), \( ||s_i > \sigma_i(\omega_2)|| \cap B_i(\omega_2) \supseteq \{ \hat{x} \} \); hence \( \omega_2 \notin \mathcal{R}_i \) (see Definition 2), contradicting the hypothesis that \( \omega_1 \in \mathcal{C} \mathcal{B} \mathcal{R} \) and \( \omega_2 \in \mathcal{B}^*(\omega_1) \) (which implies that \( \omega_2 \in \mathcal{R} \subseteq \mathcal{R}_i \)). Thus we have shown that, for every \( \omega \in \mathcal{B}^*(\omega_1) \), \( \sigma_i(\omega) \in S_i \backslash E^0_1 = B^1_1 \).

**Inductive Step.** Fix an integer \( m \geq 1 \) and suppose that, for every player \( j \in I \) and for every \( \omega \in \mathcal{B}^*(\omega_1) \), \( \sigma_j(\omega) \in B^m_j \), 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 \( \omega \in \mathcal{B}^*(\omega_1) \), \( \sigma_i(\omega) \notin E^m_i \). Suppose not. Then there exist a player \( i \) and an \( \omega_2 \in \mathcal{B}^*(\omega_1) \) such that \( \sigma_i(\omega_2) \in E^m_i \), that is, strategy \( \sigma_i(\omega_2) \) of player \( i \) is B-dominated relative to \( B^m_i \); for every \( X_{-i} \subseteq B^m_i \) there exists a strategy \( s_i \in S_i \) such that (1) for all \( x_{-i} \in X_{-i} \), \( \pi_i(s_i, x_{-i}) \geq \pi_i(\sigma_i(\omega_2), x_{-i}) \) and (2) 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 \( X_{-i} = \sigma_{-i}(B_i(\omega_2)) = \{ s_{-i} \in S_{-i} : s_{-i} = \sigma_{-i}(\omega) \text{ for some } \omega \in B_i(\omega_2) \} \). By the induction hypothesis and the fact that \( \mathcal{B}_i(s_2) \subseteq \mathcal{B}^*(\omega_1) \) (the latter inclusion follows from transitivity of \( \mathcal{B}^* \)), \( X_{-i} \subseteq B^m_i \). Let \( s_i \in S_i \) and \( \hat{x}_{-i} \in X_{-i} \) satisfy (1) and (2) and let \( \omega \in B_i(\omega_2) \) be such that \( \sigma_{-i}(\omega) = \hat{x}_{-i} \). Then, by (1), \( \omega_2 \in \mathcal{B}_i \| s_i \geq \sigma_i(\omega_2) \) and by (2) \( ||s_i > \sigma_i(\omega_2)|| \cap B_i(\omega_2) \supseteq \{ \hat{x} \} \). By \((L5)_\omega \), \( \hat{x} \in \mathcal{B}_i \| \mathcal{B}_i(\omega_2) \supseteq \{ \hat{x} \} \). By \((L2)_\omega \), \( ||s_i > \sigma_i(\omega_2)|| \cap B_i(\omega_2) \supseteq \{ \hat{x} \} \); hence \( \omega_2 \notin \mathcal{R}_i \) (see Definition 2), contradicting the hypothesis that \( \omega_1 \in \mathcal{C} \mathcal{B} \mathcal{R} \) and \( \omega_2 \in \mathcal{B}^*(\omega_1) \) (which implies that \( \omega_2 \in \mathcal{R} \subseteq \mathcal{R}_i \)). Thus, for every player \( i \in I \) and for every \( \omega \in \mathcal{B}^*(\omega_1) \), \( \sigma_i(\omega) \in \bigcap_{m=1}^{\infty} B^m_i = B^\infty_i \).

It only remains to show that \( \sigma_i(\omega_1) \in B^\infty_i \). Take \( \omega_2 \in B_i(\omega_1) \). Since \( B_i(\omega_1) \subseteq \mathcal{B}^*(\omega_1) \), \( \omega_2 \in \mathcal{B}^*(\omega_1) \). Thus \( \sigma_i(\omega_1) \in B^\infty_i \).

\((B_1)\) Given a game \( G \) construct the following model \( M \in \mathcal{M}_1 \); \( \Omega = B^\infty = B^1_1 \times \cdots \times B^n_\infty \); for every player \( i \) and for every \( s \in B^\infty \), \( \sigma_i : B^\infty \rightarrow S_i \) 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 \( B^\infty \), every \( s_i \in B_i^\infty \) is not B-dominated relative to \( B_i^\infty \); that is, there exists an \( X^i_{-i} \subseteq B^\infty_{-i} \) (note this set may vary with \( s_i \), hence the superscript “\( s_i \)” such that, for all \( s'_i \in S_i \), either there exists an \( \hat{x}_{-i} \in X^i_{-i} \) such that:

\[
\pi_i(s'_i, \hat{x}_{-i}) < \pi_i(s_i, \hat{x}_{-i}) \tag{A.1}
\]

or for all \( x_{-i} \in X^i_{-i} \),

\[
\pi_i(s'_i, \hat{x}_{-i}) \leq \pi_i(s_i, \hat{x}_{-i}). \tag{A.2}
\]

For every \( s_i \in B^\infty_i \) fix one such set \( X^i_{-i} \) (there may be several) and define \( \mathcal{B}_i(s_i, s'_{-i}) = \{ s_i \} \times X^i_{-i} \). By construction, \( (s_i, \hat{x}_{-i}) \in \mathcal{B}_i(s) \) and thus, either, by \((A.1)\), \( s \notin \mathcal{B}_i \| s'_i \geq s_i \) or, by \((A.2)\), \( s'_i > s_i \| \mathcal{B}_i(s) = \emptyset \). It follows that, for every \( i \in I \) and for every \( s \in B^\infty \), \( s \in \mathcal{R}_i \) and thus \( B^\infty = \mathcal{R} = \mathcal{C} \mathcal{B} \mathcal{R} \). Note that for completeness – although strictly speaking this is not needed – we can add the condition that, for every \( i \in I \) and \( s \in B^\infty \), \( \{ s'_i \} > s_i \| \mathcal{B}_i(s) \).}

The proof of Theorem 3 uses as intermediate results the ones stated in Section 4.2 and proved in Appendix B.

**Proof of Theorem 3.** \((A2)\) Given a game, consider a model \( M \in \mathcal{M}_2 \). Then \( C = \mathcal{C} \mathcal{B} \mathcal{C} = C_{ij} = \Omega \). Let \( \omega \in \mathcal{C} \mathcal{B} \mathcal{R} \). Then, by Corollary 2 in Section 4.2, \( \omega \in T^\infty \).

\((B_2)\) Given a game construct the following model of it: \( \Omega = T^\infty \); for every player \( i \) and for every \( s \in T^\infty \), \( \mathcal{B}_i(s) = \{ s'_i \in T^\infty : s'_i = s_i \} \) (that is, \( s'_i \in \mathcal{B}_i(s) \) if and only if both \( s \) and \( s' \) belong to
$T^\infty$ and player $i$'s strategy is the same in $s$ and $s'$; $\sigma_i : T^\infty \rightarrow S_i$ is defined by $\sigma_i(s) = s_i$ (that is, $\sigma_i(s)$ is the $i^{th}$ coordinate of $s$); finally, for all $i \in I$, $s \in T^\infty$ and $s' \in B_i(s), \{s'\} \not\succ_i \emptyset$. Note that each relation $B_i$ is an equivalence relation. Fix an arbitrary state $s \in T^\infty$ 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, $s \in \|s'_i > s_i\|$, so that $\|s'_i > s_i\| \cap B_i(s) \supseteq\{s\} \not\succ_i \emptyset$. Then, by definition of $T^\infty$, there exists an $s_{-i} \in S_{-i}$ such that $(s_i, s_{-i}) \in T^\infty$ and $\pi_i(s_i, s_{-i}) < \pi_i(s_i, s_{-i})$; by construction, $(s_i, s_{-i}) \in B_i(s)$ so that $s \not\in B_i\|s'_i \geq s_i\|$. Thus, by Definition 2, player $i$ is rational at state $s$, that is, $s \in R_i$. Since $i$ and $s$ were chosen arbitrarily, it follows that $R = T^\infty$. \hfill \square

B. Proofs of Section 4

Proof of Corollary 1. Fix an arbitrary $s \in T^\infty$. Then, by Theorem 3, there exists some model in $M \in M_2$ such that for some state $\omega$ (in this model), $\sigma(\omega) = s$ and $\omega \in CR$. Since $M_2 \subseteq M_1$ it follows that $M \in M_1$, and therefore, by Theorem 2, $s \in B^\infty$, thus proving $T^\infty \subseteq B^\infty$.

Likewise, fix an arbitrary $s' \in B^\infty$. Then, by Theorem 2, there exists some model $M' \in M_1$ such that for some state $\omega'$ (in this model), $\sigma(\omega') = s'$ and $\omega' \in CR$. Since $M_1 \subseteq M_0$ it follows that $M' \in M_0$, and therefore, by Theorem 1, $s' \in S^\infty$, thus proving $B^\infty \subseteq S^\infty$. \hfill \square

Proof of Proposition 1. Fix a strategic-form game and a model $M \in M_1$. Suppose that $\omega_1 \in CR \cap CB$, i.e., $B^*(\omega_1) \subseteq R \cap C$. We want to show that $\sigma(\omega_1) \in T^\infty$. As before, the proof is by induction.

Initial Step. First we show (by contradiction) that, for every $\omega \in B^*(\omega_1)$, $\sigma(\omega) \not\in I^0$ (see Definition 7). Suppose, that there exists an $\omega_2 \in B^*(\omega_1)$ such that $\sigma(\omega_2) \in I^0$, that is, $\sigma(\beta)$ is inferior relative to the entire set of strategy profiles $S$. Then there exists a player $i$ and a strategy $s_i \in S_i$ such that

$$\pi_i(s_i, s_{-i}) > \pi_i(\sigma_i(\omega_2), s_{-i}), \quad \text{for all } s_{-i} \in S_{-i}, \quad \text{(B.1)}$$

Hence, for every $\omega \in B_i(\omega_2)$, $\pi_i(s_i, s_{-i}) \geq \pi_i(\sigma_i(\omega_2), s_{-i})$, that is, $\omega_2 \in B_i\|s_i \geq \sigma_i(\omega_2)\|$. Furthermore, since $B^*(\omega_1) \subseteq C \subseteq C_i$ and $\omega_2 \in B^*(\omega_1)$, $\omega_2 \in B_i(\omega_2)$. Since the model has full support, $\{\omega_2\} \not\succ_i \emptyset$ and thus, by (B.1), $\|\hat{s}_i > \sigma_i(\omega_2)\| \cap B_i(\omega_2) \not\succ_i \emptyset$ (appealing to (L2) with $F = \|\hat{s}_i > \sigma_i(\omega_2)\|$ and $E = \{\omega_2\}$), so that, by Definition 2, player $i$ is not rational at state $\omega_2$, contradicting the hypothesis that $\omega_2 \in B^*(\omega_1)$ and $\omega_1 \in CR$. Thus, for every $\omega \in B^*(\omega_1)$, $\sigma(\omega) \in T^0 \setminus I^0 = T^1$ (recall that $T^0 = S$).

Inductive Step. Fix an integer $m \geq 1$ and suppose that, for every $\omega \in B^*(\omega_1)$, $\sigma(\omega) \in T^m$. We want to show that, for every $\omega \in B^*(\omega_1)$, $\sigma(\omega) \not\in I^m$. Suppose, by contradiction, that there exists a $\omega_2 \in B^*(\omega_1)$ such that $\sigma(\omega_2) \in I^m$, that is, $\sigma(\omega_2)$ is inferior relative to $T^m$. Then there exist a player $i$ and a strategy $s_i \in S_i$ such that

$$\pi_i(s_i, s_{-i}(\omega_2)) > \pi_i(\sigma_i(\omega_2), s_{-i}(\omega_2)), \quad \text{(B.3)}$$

$$\pi_i(s_i, s_{-i}) \geq \pi_i(\sigma_i(\omega_2), s_{-i}), \quad \text{for all } s_{-i} \in S_{-i} \text{ such that } (\sigma_i(\omega_2), s_{-i}) \in T^m. \quad \text{(B.4)}$$

By the induction hypothesis, for every $\omega \in B^*(\omega_1)$, $(\sigma_i(\omega), s_{-i}(\omega)) \in T^m$. Thus, since $B_i(\omega_2) \subseteq B^*(\omega_2) \subseteq B^*(\omega_1)$ (the latter inclusion follows from transitivity of $B^*$), we have that, for every $\omega \in B_i(\omega_2)$, $(\sigma_i(\omega_2), s_{-i}(\omega)) \in T^m$ (recall that, by (Σ0), if $\omega \in B_i(\omega_2)$ then $\sigma_i(\omega) = \sigma_i(\omega_2)$). Since $B^*(\omega_1) \subseteq C \subseteq C_i$ and $\omega_2 \in B^*(\omega_1)$, $\omega_2 \in B_i(\omega_2)$. Since the model has full support, $\{\omega_2\} \not\succ_i \emptyset$ and thus $\|\hat{s}_i > \sigma_i(\omega_2)\| \cap B_i(\omega_2) \not\succ_i \emptyset$ so that, by (B.3) and Definition 2, player $i$ is not rational at state
ω₂, contradicting the hypothesis that ω₂ ∈ B*(ω₁) and ω₁ ∈ CBR. Thus, we have shown that, for every ω ∈ B*(ω₁), σ(ω) ∈ ∩ₘₙ₌₁ Tᵐ = T∞, that is, ω₁ ∈ CBₜ∞.

Proof of Corollary 2. Fix a strategic-form game with ordinal payoffs and a model M ∈ M₁. Suppose that ω₀ ∈ CBR ∩ CBₜ C ∩ Cₜ. Since ω₀ ∈ Cₜ, there exists a player i ∈ I such that ω₀ ∈ Cᵢ, that is, ω₀ ∈ Bᵢ(ω₀). Hence, by definition of B*, ω₀ ∈ B*(ω₀). By Proposition 1, ω₀ ∈ CBₜ∞, that is, for every ω ∈ B*(ω₀), ω ∈ T∞. Hence ω₀ ∈ T∞.

References


