



Limited focus in dynamic games

Andrés Perea¹ · Elias Tsakas²

Accepted: 29 September 2018 / Published online: 5 October 2018

© The Author(s) 2018

Abstract

In this paper we introduce a novel framework that allows us to model games with players who reason about *the opponents' rationality only in some part of the game tree*. We refer to this type of bounded rationality as *limited focus*. In particular, players try to rationalize their opponents' moves only at the histories they focus on, i.e., formally, they strongly believe in their opponents' rationality in these particular histories only. Our main result characterizes the strategy profiles that can be played under rationality and common strong belief in rationality by means of a simple elimination procedure, for every specification of the players' focus. Finally, we present several special cases and applications of our framework, and we discuss how it differs from other forms of bounded perception such as unawareness.

Keywords Limited focus · Epistemic game theory · Bounded rationality · Forward induction · Backward induction

1 Introduction

The bounded rationality literature can be classified into two broad categories. On the one hand, we have models that explicitly embed specific mechanisms that are responsible for some cognitive limitation, e.g., Mullainathan (2002) introduces a model of

We are indebted to Giacomo Bonanno and the audiences in Maastricht, Bilbao, Brisbane, GAMES (Maastricht), LogiCIC Workshop (Amsterdam), LSB Workshop (Zurich), SAET (Paris), LOFT (Bergen) and LOFT (Maastricht) for valuable comments and useful discussions. Previous versions of this paper circulated under the title “Local reasoning in dynamic games”.

✉ Elias Tsakas
e.tsakas@maastrichtuniversity.nl

Andrés Perea
a.perea@maastrichtuniversity.nl

¹ EpiCenter, Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

² EpiCenter, Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200, MD Maastricht, The Netherlands

limited memory motivated by—empirically documented—psychological and biological mechanisms. On the other hand, there are models that primarily focus on the behavioral consequences of such limitations, rather than on their underlying mechanisms, e.g., the literature on automata studies repeated games with players who avoid complex strategies, without attributing this tendency to a single mechanism (for an early overview of this literature, see Kalai 1990).

This paper belongs to this second stream. In particular, we study the behavioral implications of a family of psychological traits and cognitive constraints, whose common denominator is that players fail to reason about the opponents' rationality in certain parts of the game tree, either because they cannot, or because they find it cognitively costly, or even because they simply do not deem it important to do so. We refer to this phenomenon as *limited focus*. Such restriction is particularly interesting, for instance, in the context of signalling future actions, where it reduces the signalling power of the sender (see our leading example in Sect. 2). The reason for classifying our work in the second stream within the bounded rationality literature is that we are not primarily interested in identifying the source of the players' failure to focus on certain histories, but rather in understanding the behavioral consequences of this failure.

Our main result characterizes (by means of an elimination procedure) the strategies that can be rationally played under common strong belief in rationality when the players focus only on a subset of the game tree (Theorem 2).¹ The underlying idea is that each player is endowed with a function that maps each history (where this player is active) to a subset of the non-terminal histories, viz., to *the player's focus at this history*. Then, the player tries to rationalize the opponents' moves *only at the histories that she focuses on*, in contrast to the standard forward induction solution concepts which postulate that players try to rationalize the opponents' actions *at all histories* (e.g., Battigalli and Siniscalchi 2002; Battigalli and Friedenberg 2012). In particular, we say that the player strongly believes in the opponents' rationality *given her focus* whenever the following holds: she will keep believing that her opponents will behave rationally *at the histories that she focuses on*, if they already behaved rationally *at the (past) histories that she focuses on*. In other words, the opponents' behavior *at the past histories that the player focuses on* constitutes the only evidence that the player can use against the opponents' rationality; on the other hand, *the future histories that the player focuses on* describes the degree to which the player believes in the opponents' future rationality (given that there is no evidence against it). Then, common strong belief of rationality naturally iterates this reasoning.

Our framework is flexible enough to accommodate as special cases several cognitive constraints and psychological traits that lead to limited focus. The most obvious examples of such constraints would be limited memory and limited foresight. But moreover, through our framework, we also identify new interesting cognitive constraints and psychological traits, such as for instance our notion of *focal histories*

¹ Recall that strong belief in an event E postulates that E is believed as long as it has not been contradicted by some observed evidence. In particular, a player is said to strongly believe in the opponents' rationality whenever it is the case that her conditional beliefs attach probability 1 to her opponents being rational at all histories that are consistent with none of her opponents having played an irrational strategy (Battigalli and Siniscalchi 2002). For instance, in the game in Fig. 1 below, the column player at both h_1 and h_2 maintains the belief in his opponent's rationality, as neither L nor R is irrational.

(Sect. 3.3). Finally, we show that standard solution concepts—such as common strong belief in rationality and common belief in future rationality—emerge as natural special cases of our model.

While presenting our framework's special cases, we also show that different cognitive constraints or psychological traits may lead to the same focus domain, and therefore to the same behavioral predictions. Thus, it is often impossible to identify the underlying constraint/trait from choice data or even from the players' beliefs (Remark 3). As an example, take backward induction which implicitly postulates that the players focus on their opponents' rationality only at present and future histories (e.g., Perea 2014). In this case, we cannot identify whether the players have forgotten the strategic incentives at past histories—and therefore they cannot reason about the opponents' past rationality—or whether they consciously disregard them. But then again, we are not mainly interested in identifying the precise cause of their failure to focus on past histories. Rather we are primarily interested in the behavioral consequences of this failure, and the latter only depends on *what* they focus on, and not on *why* they focus on these histories.

The paper is structured as follows: In Sect. 2 we present our leading example, which we use throughout the paper to illustrate both our framework and our results. In Sect. 3 we introduce our formal framework and we present different special cases and applications of our model. In Sect. 4 we introduce our epistemic model. In Sects. 5 and 6 we introduce our solution concept and our elimination procedure, respectively. In Sect. 7 we present the main characterization results. In Sect. 8 we embed standard solution concepts into our framework. Section 9 addresses some concluding remarks. All proofs are relegated to the Appendices. Since we touch upon several different issues, the relationship to the existing literature is discussed throughout the paper.

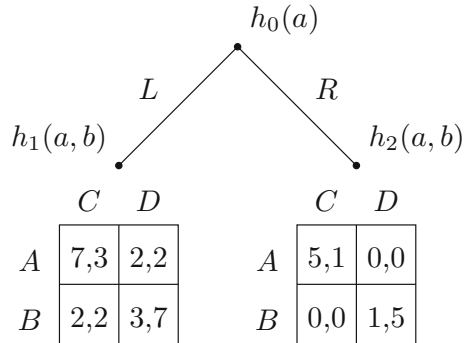
2 Heuristic treatment

Consider the standard Battle of the Sexes (BoS) game with the option to burn money, depicted in Fig. 1. The story goes as follows: Ann and Bob are scheduled to play the left hand side BoS game, but before doing so Ann has the option to burn 2 \$/utils for each player, thus choosing to play the right hand side BoS game instead. This class of games has been extensively used in the literature to study costly signaling of future actions (e.g., van Damme 1989; Ben-Porath and Dekel 1992).

Now, assume that Bob focuses on negative events and disregards neutral events, viz., he tries to rationalize Ann's choice at h_0 if and only if she has actually chosen to burn money. In other words, at h_2 Bob focuses on the entire game, while at h_1 he only focuses on the current history, without reasoning about how we ended up there. Note that there is empirical evidence supporting asymmetries in the responses to positive and negative news (e.g., Soroka 2006). Which are our predictions for this game?

Let us begin by observing that at h_0 Ann will never plan to choose B after having burned money, as this strategy would be strictly dominated by any plan involving first not burning money. Now, at h_2 Bob focuses on Ann's behavior at h_0 , and therefore he realizes that upon reaching h_2 Ann will definitely choose A , thus implying that it is optimal for him to choose C . Then, at h_0 Ann anticipates that if she moves Right, she

Fig. 1 The burning-money game



will definitely receive a utility of 5, and therefore she will never plan to choose B after having moved Left, which would give her at most a utility of 3. Now the crucial part of our analysis relies on the fact that at h_1 Bob does not focus on h_0 , and therefore he will not figure out that Ann has ruled out the strategy LB . Thus, any action of Bob at h_1 remains possible, and as a consequence we predict that one of the strategy profiles in $\{LA, RA\} \times \{CC, DC\}$ will be played.

The latter is in contrast with forward induction—viz., common strong belief in rationality—that predicts $\{LA\} \times \{CC\}$, as well as with backward induction—viz., common belief in future rationality—that yields $\{LA, LB, RA\} \times \{CC, CD, DC, DD\}$. The reason for these deviations is that Ann can only use some—but not her entire—signalling power, as Bob anyway does not pay attention to her in some instances. For instance, the fact that our analysis—contrary to backward induction—rules out Bob’s strategies CD and DD , is due to Bob focusing on h_0 while being at h_2 . On the other hand, the fact that our analysis—contrary to forward induction—permits Ann’s strategy RA is attributed to Ann believing at h_0 that Bob at h_1 will not focus on her choice at h_0 . Later in the paper we revisit this example using the formal machinery that we are going to introduce.

3 Basic framework

3.1 Dynamic games with observable actions

We consider dynamic games with observable actions and simultaneous moves. Our results can be extended to arbitrary dynamic games with perfect recall. Formally, the game structure is described by the following components:

Players Let I denote the finite set of players, with typical elements i and j . Throughout the paper, we often consider examples with the set of players being $I = \{\text{Ann } (a), \text{ Bob } (b)\}$.

Histories For each $i \in I$, let H_i denote the set of histories where player i moves. We permit more than one player to move at the same history, i.e., $H_i \cap H_j$ may be non-empty. For instance, in Fig. 1 we have $H_a = \{h_0, h_1, h_2\}$ and $H_b = \{h_1, h_2\}$, and we write $h_0(a)$ and $h_1(a, b)$ to signify that “only Ann moves at h_0 ” and that “both Ann

and Bob move at h_1 ” respectively. Let $H := \bigcup_{i \in I} H_i$ be the set of all non-terminal histories, and $H_{-i} := \bigcup_{j \neq i} H_j$ be the set of non-terminal histories where at least one player other than i moves. Moreover, Z denotes the set of terminal histories. Finally, for an arbitrary $h \in H$ (resp., for an arbitrary $z \in Z$), let $\text{Pr}(h)$ (resp., $\text{Pr}(z)$) denote the set of histories that weakly precede h (resp., z). Likewise, $\text{Fut}(h)$ denotes the set of histories that weakly follow h . That is, for every $h' \in \text{Pr}(h)$ (and likewise for every $h' \in \text{Fut}(h)$) there exists a path of play going through both h and h' .

Moves and strategies The finite set of moves (also called actions) from which player i chooses one at some history $h \in H_i$ is denoted by $A_i(h)$. Player i ’s strategy space is denoted by S_i with typical element s_i , e.g., in Fig. 1 we have $S_a = \{LA, LB, RA, RB\}$ and $S_b = \{CC, CD, DC, DD\}$. Notice that we define strategies as plans of actions, and not as elements of $\times_{h \in H_i} A_i(h)$. In either case our analysis would still hold under the alternative definition of a strategy that often appears in the literature (cf., Rubinstein 1991). As usual, $S := \times_{i \in I} S_i$ denotes the set of strategy profiles with typical element s , and $S_{-i} := \times_{j \neq i} S_j$ denotes the strategy profiles of all players other than i with typical element s_{-i} . We define player i ’s set of conditional strategies at some history h as the set of strategies that are consistent with h being reached, and we denote it by $S_i(h)$. Then, $S_{-i}(h)$ denotes the profiles $s_{-i} \in S_{-i}$ that are consistent with h being reached. For each $s_i \in S_i$ we define $H_i(s_i) := \{h \in H_i : s_i \in S_i(h)\}$, and naturally we let $H(s_i) := \{h \in H : s_i \in S_i(h)\}$ and $H_{-i}(s_i) := \{h \in H_{-i} : s_i \in S_i(h)\}$. There exists a function $\zeta : S \rightarrow Z$, mapping each strategy profile $s \in S$ to a unique terminal history. Each strategy profile induces a path of play, which contains the set of histories that are reached if s is played. Formally, this path contains the non-terminal histories $H(s) := \bigcap_{i \in I} H_i(s_i)$ and the terminal history $\zeta(s)$.

Utilities Player i has preferences over the terminal histories, represented by a mapping $v_i : Z \rightarrow \mathbb{R}$. Recall that each strategy profile s leads to a unique terminal history $\zeta(s)$. Thus, we obtain the utility function $u_i : S \rightarrow \mathbb{R}$, defined as the composition $u_i := v_i \circ \zeta$, that represents i ’s preferences over S .²

3.2 Focus structures

In this section we introduce the notion of a player’s focus at some history. This concept is new to the literature and we will use it throughout the paper to model dynamic games with players who reason about the opponents’ rationality in some part of the game only, while disregarding the remaining histories. This tool allows us to systematically study the behavioral consequences of several interesting—some existing and some new—cognitive constraints and psychological traits. Moreover, it provides a unifying framework, within which we can study standard solution concepts—such as backward induction and forward induction—which correspond to special cases of our generalized solution concept, as formally shown later in the paper.

Formally, for an arbitrary player i , the function

$$F_i : H_i \rightarrow 2^H \tag{1}$$

² As usual, we assume that i has vNM preferences over $\Delta(Z)$, and consequently also over $\Delta(S)$. Thus, u_i can be seen as the vNM representation of these preferences.

specifies a subset of histories that player i focuses on, upon reaching the history $h \in H_i$. Throughout the paper, we refer to F_i as i 's *focus function*, with $F_i(h) \subseteq H$ being i 's focus at h . The interpretation is that at h player i tries to rationalize the opponents' moves only at the histories in $F_i(h)$.

The general idea is that players do not necessarily focus on all histories in the game, either because they cannot, or because it is cognitively very costly for them, or even because they simply do not find it important to do so. In fact, we are not mainly interested in why the players fail to reason about certain histories, but rather in the consequences of this failure. Nevertheless, later in this section we provide several concrete examples of cognitive constraints and psychological traits that induce limited focus. In either case, it is important to stress that in our model, players are aware of all histories, including those outside their focus.³

Now, let \mathcal{F}_i denote the set of i 's focus functions, and let $\mathcal{F} := \times_{i \in I} \mathcal{F}_i$ be the set of respective function profiles. Following Harsanyi (1967-68) and Battigalli and Siniscalchi (1999), we model interactive uncertainty about the players' focus using an (\mathcal{F} -based) *focus structure*

$$\mathfrak{F} = ((\Theta_i)_{i \in I}, (f_i)_{i \in I}, (g_i)_{i \in I}),$$

where Θ_i is a finite set of focus types, also called Θ_i -types,⁴ $f_i : \Theta_i \rightarrow \mathcal{F}_i$ associates each focus type of player i with a focus function, and $g_i : \Theta_i \times H_i \rightarrow \Delta(\Theta_{-i})$ is a function mapping each $\theta_i \in \Theta_i$ at each history $h \in H_i$ to a conditional belief $g_i^h(\theta_i) \in \Delta(\Theta_{-i})$, with $\Theta_{-i} := \times_{j \neq i} \Theta_j$. For notation simplicity, throughout the paper we write $F_{\theta_i} := f_i(\theta_i)$. Thus, each θ_i is a full description of i 's focus, as well as her beliefs and higher-order beliefs about every player's focus.

Throughout the paper, we assume that players do not focus indirectly on a history, unless they also focus directly on it. This idea is formalized by the following assumption.

Assumption 1 For every $i \in I$, every $h \in H_i$ and every $\theta_i \in \Theta_i$, if $g_i^h(\theta_i)((\theta_j)_{j \neq i}) > 0$ then $F_{\theta_j}(h') \subseteq F_{\theta_j}(h)$ for all $h' \in F_{\theta_i}(h) \cap H_j$ and for all $j \neq i$. □

In fact, our analysis does not depend on Assumption 1, thus implying that it could be in principle dispensed with. Still, we find it to be intuitively appealing, and as such throughout the paper we consider structures \mathfrak{F} that satisfy the assumption.

Example 1 Recall the burning-money game from Sect. 2, and consider a focus structure with $\Theta_a = \{\theta_a, \theta'_a\}$ and $\Theta_b = \{\theta_b\}$. Moreover, let the corresponding focus functions and belief mappings be the ones presented in Table 1.

Notice that this focus structure satisfies Assumption 1. In fact, in order for the assumption to be satisfied, we need to enrich our focus structure with the additional focus type θ'_a for Ann. The reason is that at h_1 , Bob does not focus on h_0 , but does focus on h_1 itself. But, then he cannot consider Ann's type θ_a , since in order to do so,

³ In Sect. 9.1 we discuss the relationship of this framework to the literature on dynamic games with unawareness.

⁴ In Sect. 9.2 we discuss the case of infinite Θ_i 's.

Table 1 Focusing on negative events in the burning-money game

		h_0	h_1	h_2
Ann	θ_a	$\{h_0, h_1, h_2\}$ $(1 \otimes \theta_b)$	$\{h_0, h_1, h_2\}$ $(1 \otimes \theta_b)$	$\{h_0, h_1, h_2\}$ $(1 \otimes \theta_b)$
	θ'_a	$\{h_0, h_1, h_2\}$ $(1 \otimes \theta_b)$	$\{h_1\}$ $(1 \otimes \theta_b)$	$\{h_0, h_1, h_2\}$ $(1 \otimes \theta_b)$
Bob	θ_b		$\{h_1\}$	$\{h_0, h_1, h_2\}$
			$(1 \otimes \theta'_a)$	$(1 \otimes \theta_a)$

he would need to reason indirectly about h_0 . This highlights the need for incomplete information about the players' focus. □

Remark 1 In order to simplify the presentation of our results, we henceforth restrict attention to cases where the support of $g_i^h(\theta_i)$ is a singleton. Still, our analysis can be directly extended to any structure \mathfrak{F} . Finally, throughout the paper, we are often interested in cases where the focus function of each player is transparent across all players. Formally, this is the case whenever each Θ_i is a singleton. In this case, we identify the unique $\theta_i \in \Theta_i$ with i , thus simply writing F_i for F_{θ_i} , and we say that $F \in \mathcal{F}$ is commonly known. □

Remark 2 Throughout the paper we consider only complete information games. However, our framework can easily accommodate incomplete information games, by simply introducing for each $i \in I$ a second function $\delta_i : \Theta_i \rightarrow \mathcal{U}_i$, where \mathcal{U}_i is the set of i 's utility functions. □

3.3 Applications

As we have already mentioned, the primary aim of this paper is not to identify which psychological traits (or cognitive constraints) are responsible for the players' failure to focus on certain histories, but rather to understand the behavioral implications of this failure. Nevertheless, in order to convince the reader that our framework is useful, we present some specific traits that yield limited focus.

Remark 3 Note that the behavioral implications of these traits can be understood and studied within our framework, even without identifying the traits themselves, as long as we manage to identify the induced focus structures. Obviously the second identification problem appears to be easier—compared to the first one—as the players' focus is a well-defined parameter, in contrast to the underlying constraints/traits which are often vaguely defined. □

Games with limited memory Dynamic games with limited memory have been extensively studied in the literature.⁵ Some of these papers consider players who are literally

⁵ There is a rather large literature on how to model limited memory in dynamic games/decision problems and its behavioral consequences (e.g., Dow 1991; Rubinstein 1991, 1998; Piccione and Rubinstein 1997; Mullainathan 2002; Kline 2002; Bonanno 2004; Aumann and Sorin 1989; Lehrer 1988; Cole and Kocherlakota 2005; Barlo et al. 2009).

unable to remember certain past events (e.g., Mullainathan 2002), whereas others postulate that players choose what to remember (e.g., Dow 1991). Irrespective of the source of their failure to remember, what matters is that players fail to focus on certain past histories, and consequently they may fail to reason about their opponents' actions in these histories. In this sense, limited memory induces a particular form of limited focus, which can be modelled within our framework. In other words, within our framework, players do not try to rationalize opponents' moves that took place in the distant past.

Games with limited foresight Dynamic games with limited foresight have also been extensively studied in the literature.⁶ The most common motivation—though not the only one—for assuming players with limited foresight is that “*looking ahead is computationally expensive and unnatural because it means reasoning about events that probably will not occur*” as Johnson et al. (2002) elaborately put it. Thus, limited foresight can be seen as a special form of limited focus, and therefore captured by our framework. That is, within our framework, players do not reason about the opponents' rationality in the distant future.

Games with focal histories Suppose that due to some inherent psychological trait, a player's attitude towards different events changes depending on whether each of these events is relevant or not. In particular, consider the following traits:

- Assume that a player has the tendency to disregard positive and/or neutral news, like for instance in our motivating example in Sect. 2. In fact, there is empirical evidence suggesting that in certain situations this is indeed the case (e.g., Soroka 2006; Taylor 1991; Peeters and Czapinski 1990).
- A second such trait has been observed by psychologists and biologists, and it is called associativeness. Accordingly, players focus on histories where they encountered similar decisions as the one current one (e.g., Mullainathan 2002).
- A third such trait is exhibited by players who pay attention only to high-stakes decisions, while disregarding the low-stakes ones.
- A fourth trait is exhibited by people who have a tendency to disregard—at least some—counterfactual events (for a review of this literature, see Roese 1997). In fact there is evidence that counterfactual events are often taken into account when the actually realized event had negative consequences. In this sense, this psychological trait partially overlaps with the first one.

The common denominator of these traits is that the set of histories that a player focuses on (at an arbitrary h), depends on the relevance of the decision being taken at every history. Then, we say that these are the player's *focal histories* at h . Notice that similarly to the way focal points stand out in static games, focal histories stand out in dynamic games.

⁶ In fact, limited foresight has been studied in the context of repeated games (e.g., Jehiel 2001; Maenner 2008), learning (e.g., Mengel 2014), behavioral and experimental economics (e.g., Johnson et al. 2002) and computer science (e.g., Grossi and Turrini 2012; Turrini 2015).

3.4 Special cases: forward and backward induction reasoning

The two standard families of solution concepts for dynamic games are backward induction (BI) and forward induction (FI). While each of these families contains several concepts—which differ in certain aspects—they all share a common principle. In particular, all BI concepts postulate that players believe in their opponents’ rationality at all future histories, irrespective of what has been played up to that point, while most FI concepts postulate that players try to justify their opponents’ past and future actions, and consequently they will believe in their opponents’ future rationality whenever this is not contradicted by what has been chosen up to that point. Hence, it seems natural to associate BI with players who focus only on the future while disregarding the past histories, and FI with players who focus on all histories in the game. That is, at first glance BI and FI appear to implicitly postulate focus structures with $F_i(h) = \text{Fut}(h)$ and $F_i(h) = H$ respectively.

The question that arises then is whether our predictions under these particular focus structures will coincide with the standard BI and FI predictions. Our formal answer to this question is postponed till Sect. 8, as we first need to introduce our solution concept, which we do in Sect. 5. Still we can already mention that—as expected—our predictions under a focus structure with $F_i(h) = \text{Fut}(h)$ (resp., with $F_i(h) = H$) coincide with the predictions of the BI concepts of common belief in future rationality and backward rationalizability (resp., with the FI concept of extensive form rationalizability). Therefore, our framework allows us to unify under one umbrella the two standard families of solution concepts.

4 Subjective beliefs and rationality

In this section we formally introduce the epistemic model that will allow us to formalize the implications of limited focus in dynamic games.

4.1 Conditional beliefs

Using a variant of the standard framework of Battigalli and Siniscalchi (1999, 2002), we model conditional belief hierarchies by means of a type structure. Let us begin by fixing a focus structure \mathfrak{F} . Then, we consider the tuple

$$\mathcal{T}_{\mathfrak{F}} = ((T_i)_{i \in I}, (\phi_i)_{i \in I}, (\lambda_i)_{i \in I}),$$

where T_i is a compact metrizable space of player i ’s belief types with typical element t_i ,⁷ $\phi_i : T_i \rightarrow \Theta_i$ is a surjective Borel function endowing each T_i -type with a Θ_i -type, and $\lambda_i : T_i \times H_i \rightarrow \Delta(S_{-i} \times T_{-i})$ is a Borel function associating each type $t_i \in T_i$ at each history $h \in H_i$ with a Borel probability measure $\lambda_i^h(t_i) \in \Delta(S_{-i}(h) \times T_{-i})$,

⁷ Throughout the paper we often refer to belief types as T_i -types, in contrast to the previously introduced Θ_i -types.

where $T_{-i} := \times_{j \neq i} T_j$.⁸ Henceforth, we refer to the measure $\lambda_i^h(t_i)$ as t_i 's conditional beliefs (or simply beliefs) at a history h . The subset

$$T_{\theta_i} := \phi_i^{-1}(\theta_i) \quad (2)$$

contains the T_i -types with $\phi_i(t_i) = \theta_i$. Observe that $\{T_{\theta_i} | \theta_i \in \Theta_i\}$ is a partition of T_i . Obviously, it is the trivial partition whenever Θ_i is a singleton. Whenever $t_i \in T_{\theta_i}$, we naturally require the conditional belief $\lambda_i^h(t_i)$ to agree with $g_i^h(\theta_i)$. This restriction is formally imposed by the following assumption.

Assumption 2 For every $i \in I$, every $h \in H_i$, every $\theta_i \in \Theta_i$, every $t_i \in T_{\theta_i}$, every $(\theta_j)_{j \neq i} \in \Theta_{-i}$, it is the case that $\lambda_i^h(t_i)(S_{-i} \times (\times_{j \neq i} T_{\theta_j})) = g_i^h(\theta_i)((\theta_j)_{j \neq i})$. \square

Before moving forward, observe that $S_{-i} \times (\times_{j \neq i} T_{\theta_j})$ is a Borel event in $S_{-i} \times T_{-i}$, and therefore the probability $\lambda_i^h(t_i)(S_{-i} \times (\times_{j \neq i} T_{\theta_j}))$ is well-defined. This follows directly from ϕ_j being Borel measurable for every $j \in I$.

A type structure $\mathcal{T}_{\mathcal{F}}$ induces a conditional belief hierarchy for every $t_i \in T_i$. In particular, t_i holds a conditional belief at each $h \in H_i$ about the opponents' strategies (first order conditional beliefs), a conditional belief at each $h \in H_i$ about the opponents' strategies and first order conditional beliefs (second order conditional beliefs), and so on. Throughout the paper, we denote t_i 's first order conditional belief at h by

$$b_i^h(t_i) := \text{marg}_{S_{-i}} \lambda_i^h(t_i).$$

Remark 4 Notice that by the previous construction, it is for instance the case that each $t_i \in T_{\theta_i}$ forms a belief at $h \in H_i$ about j 's actions and beliefs at each $h' \in H_j$, even if $h' \notin F_{\theta_i}(h)$. At first sight this might seem to intuitively contradict our idea of limited focus. However, recall that our general idea is that θ_i at h can see all histories in H but only reasons about those histories that belong to $F_{\theta_i}(h)$. This distinction is formally captured by the fact that in our model θ_i 's beliefs about j 's beliefs at $h' \notin F_{\theta_i}(h)$ are arbitrary, i.e., $t_i \in T_{\theta_i}$ will form her beliefs without applying any form of reasoning about these histories. This rather subtle point will become clearer in Sect. 5, where we introduce a specific form of reasoning. \square

Definition 1 A type structure $\mathcal{T}_{\mathcal{F}}$ is said to be complete if for every $i \in I$, every $\theta_i \in \Theta_i$, every $h \in H_i$, every $(\theta_j)_{j \neq i} \in \Theta_{-i}$ and every $(\mu_i^h)_{h \in H_i}$ with $\mu_i^h \in \Delta(S_{-i}(h) \times T_{-i})$ and $\mu_i^h(S_{-i} \times (\times_{j \neq i} T_{\theta_j})) = g_i^h(\theta_i)((\theta_j)_{j \neq i})$ for all $h \in H_i$, there exists some $t_i \in T_{\theta_i}$ such that $\lambda_i^h(t_i) = \mu_i^h$ for all $h \in H_i$.

Before moving forward, notice that the standard notion of completeness is a special case of our definition for cases where F is common knowledge. In particular, if F is commonly known, completeness postulates that the function λ_i is surjective, i.e.,

⁸ The assumption that "upon reaching a history $h \in H_i$ every type t_i assigns probability 1 to $S_{-i}(h) \times T_{-i}$ " corresponds to the standard Condition 1 in (Battigalli and Siniscalchi 2002, Def. 1). Note that in their paper they further restrict beliefs to satisfy Bayesian updating whenever possible (see their Condition 3), thus implicitly assuming that the collection of conditional beliefs forms a conditional probability system, as originally defined by R enyi (1955).

for every collection of conditional beliefs $(\mu_i^h)_{h \in H_i}$ there is some type t_i such that $\lambda_i^h(t_i) = \mu_i^h$ for all $h \in H_i$. To see that the standard notion of completeness agrees with our definition whenever F is commonly known, observe that in this case $T_j = T_{\theta_j}$ for the unique $\theta_j \in \Theta_j$, and therefore $\mu_i^h(S_{-i} \times (\times_{j \neq i} T_{\theta_j})) = g_i^h(\theta_i)((\theta_j)_{j \neq i}) = 1$ is trivially satisfied. Thus, our definition reduces to the usual one.

The fact that our notion of completeness is in general different from the usual one justifies the need to distinguish between Θ_i -types and T_i -types. The idea is that the type structure $\mathcal{T}_{\mathfrak{F}}$ must be rich enough so that all conditional beliefs are included in T_i , but not too rich so that Assumption 2 is violated. In other words, the surjectiveness of λ_i is not a global property anymore, i.e., we do not require that every conditional belief over $S_{-i} \times T_{-i}$ is induced by some T_i -type; instead, we ask that every conditional belief over $S_{-i} \times (\times_{j \neq i} T_{\theta_j}) \subseteq S_{-i} \times T_{-i}$ is induced by some T_i -type only if there is some $\theta_i \in \Theta_i$ such $g_i^h(\theta_i)((\theta_j)_{j \neq i}) > 0$.

Battigalli and Siniscalchi (1999) showed the existence of a complete type structure (with commonly known F).⁹ It turns out that this result can be extended to arbitrary type structures, i.e., a complete type structure $\mathcal{T}_{\mathfrak{F}}$ exists for every focus structure \mathfrak{F} , even if F is not commonly known. This claim is formally proven in Appendix A. Throughout the paper, unless explicitly stated otherwise, we work with complete type structures. Finite type structures that we often consider in our examples can be seen as belief-closed subspaces of a complete type structure.

At some $h \in H_i$ a type t_i of player i is said to believe in some event $E \subseteq S_{-i} \times T_{-i}$ whenever $\lambda_i^h(t_i)(E) = 1$. Then, the types of i that believe in E at h are those in

$$B_i^h(E) := \{ t_i \in T_i : \lambda_i^h(t_i)(E) = 1 \} .$$

For instance, as we have already mentioned, it is trivially the case that $t_i \in B_i^h(S_{-i}(h) \times T_{-i})$ for all $t_i \in T_i$. Moreover, we say that a type believes in E whenever it belongs to

$$B_i(E) := \bigcap_{h \in H_i} B_i^h(E) .$$

Recall that throughout the paper, we restrict attention to structures \mathfrak{F} with the property that each $\theta_i \in \Theta_i$ puts positive probability to a unique $\theta_{-i} \in \Theta_{-i}$ at each $h \in H_i$ (see Remark 1). Thus, at an arbitrary $h \in H_i$, a type $t_i \in T_{\theta_i}$ is said to strongly believe in some event $E \subseteq S_{-i} \times T_{-i}$ whenever it belongs to

$$SB_{\theta_i}^h(E) := \{ t_i \in T_{\theta_i} : \text{if } (S_{-i}(h) \times T_{-\theta_i}) \cap E \neq \emptyset \text{ then } t_i \in B_i^h(E) \} \quad (3)$$

⁹ In fact, they proved existence of a complete type structure under Bayesian updating, but their result can be easily generalized to type structures without Bayesian updating. In a more recent paper, Friedenberg (2010) showed that for standard belief hierarchies a complete type structure that satisfies certain mild topological conditions induces all belief hierarchies, i.e., for every belief hierarchy of each player there exists a type associated with this hierarchy. Moreover, she conjectured—without formally proving it—that the same applies to conditional belief hierarchies that satisfy Bayesian updating. Finally, notice that her result is directly extended to conditional beliefs without Bayesian updating.

where $T_{-\theta_i} := \times_{j \neq i} T_{\theta_j}$, with $(\theta_j)_{j \neq i}$ being such that $g_i^h(\theta_i)((\theta_j)_{j \neq i}) = 1$. Notice that, unlike the standard notion of strong belief à la Battigalli and Siniscalchi (2002), here we require $(S_{-i}(h) \times T_{-\theta_i}) \cap E \neq \emptyset$ rather than $(S_{-i}(h) \times T_{-i}) \cap E \neq \emptyset$, the reason being that from θ_i 's point of view the only T_j -types in the model are those in T_{θ_j} , i.e., those corresponding to Θ_j -types that θ_i deems possible in the structure \mathfrak{F} . The set of all types strongly believing in E at h are those in

$$SB_i^h(E) := \bigcup_{\theta_i \in \Theta_i} SB_{\theta_i}^h(E).$$

Obviously, if F is commonly known, then $SB_{\theta_i}^h(E) = SB_i^h(E)$, as it is the case that $T_{-\theta_i} = T_{-i}$.

Of course, it is straightforward to verify that every $t_i \in T_{\theta_i}$ strongly believes in $S_{-i}(h) \times T_{-\theta_i}$ as well as in $S_{-i}(h) \times T_{-i}$ at $h \in H_i$. Finally, we say that E is strongly believed by θ_i , whenever it is strongly believed at every $h \in H_i$, i.e., formally

$$SB_{\theta_i}(E) := \bigcap_{h \in H_i} SB_{\theta_i}^h(E),$$

and at the same time a T_i -type strongly believes in E whenever it belongs to

$$SB_i(E) := \bigcup_{\theta_i \in \Theta_i} SB_{\theta_i}(E).$$

Once again, if F is commonly known, then $SB_{\theta_i}(E) = SB_i(E)$, as it is the case that $T_{-\theta_i} = T_{-i}$.

4.2 Subjective expected utility and rationality

For an arbitrary conditional belief $\beta_i^h \in \Delta(S_{-i}(h))$ and a strategy $s_i \in S_i(h)$, we define i 's (subjective) expected utility at $h \in H_i$ in the usual way, i.e., $U_i^h(s_i, \beta_i^h) := \sum_{s_{-i} \in S_{-i}} \beta_i^h(s_{-i}) \cdot u_i(s_i, s_{-i})$. Then, we define the expected utility of a strategy type pair $(s_i, t_i) \in S_i(h) \times T_i$ at a history $h \in H_i$ by

$$U_i^h(s_i, t_i) := U_i^h(s_i, b_i^h(t_i)). \tag{4}$$

Player's rationality at a history The event that a player is rational at some history $h \in H_i$ is given by

$$R_i^h := \{ (s_i, t_i) \in S_i(h) \times T_i : U_i^h(s_i, t_i) \geq U_i^h(s'_i, t_i) \text{ for all } s'_i \in S_i(h) \}. \tag{5}$$

If it is indeed the case that $(s_i, t_i) \in R_i^h$, we say that the strategy s_i is rational given (the first order beliefs induced by) t_i at h .

Opponents' rationality at a history Now, let

$$\begin{aligned}
 R_{-i}^h &:= \bigtimes_{j \neq i} \{ (s_j, t_j) \in S_j(h) \times T_j : \text{if } h \in H_j \text{ then } (s_j, t_j) \in R_j^h \} \\
 &= \bigtimes_{j \neq i} R_j^h.
 \end{aligned}
 \tag{6}$$

denote the event that every player other than i —who is active at h —is rational at h .¹⁰ *Player's rationality in a set of histories* Now, consider an arbitrary collection $G \subseteq H$ of histories. Then, a strategy-type combination (s_i, t_i) is rational in G whenever it is rational at all histories which (i) are consistent with s_i , and (ii) belong to G . Formally, the event

$$R_i^G := \{ (s_i, t_i) \in S_i \times T_i : (s_i, t_i) \in R_i^h \text{ for all } h \in H_i(s_i) \cap G \}
 \tag{7}$$

contains the strategy-type pairs that are rational in G . Notice that in general $R_i^{(h)}$ may differ from R_i^h , as $R_i^h \subseteq S_i(h) \times T_i$ considers only strategies that reach h , whereas $R_i^{(h)} \subseteq S_i \times T_i$ also allows for strategies that are not consistent with h . Because of this, R_i^G does not necessarily coincide with $\bigcap_{h \in G} R_i^h$. Finally, note that the standard notion of rationality corresponds to the event $R_i := R_i^{H_i}$, i.e., a strategy-type combination (s_i, t_i) is rational whenever it is rational at all histories $h \in H_i(s_i)$ given the respective conditional first order belief $b_i^h(t_i)$.

Opponents' rationality in a set of histories Now, let

$$\begin{aligned}
 R_{-i}^G &:= \bigtimes_{j \neq i} \{ (s_j, t_j) \in S_j \times T_j : (s_j, t_j) \in R_j^h \text{ for all } h \in H_j(s_j) \cap G \} \\
 &= \bigtimes_{j \neq i} R_j^G
 \end{aligned}
 \tag{8}$$

contain i 's opponents' strategy-type combinations that are rational in G . Then, the usual event of every player other than i being rational corresponds to $R_{-i} := R_{-i}^{H_{-i}}$.

5 Solution concept

So far, we have defined a general framework which allows us to model players who focus on only some histories in the game. However, we have not specified yet *how players reason about the particular histories that they focus on*. In this section, we formalize a reasoning process, based on Battigalli and Siniscalchi's (2002) notion of strong belief in rationality, and we introduce a solution concept that incorporates this reasoning process. This will be an epistemic concept, implying that it is defined by means of a sequence of restrictions on the players' types, and therefore it gives the

¹⁰ In order to obtain $R_{-i}^h = \bigtimes_{j \neq i} R_j^h$ we make use of the convention that $R_j^h = S_j(h) \times T_j$ for all $j \neq i$ with $h \notin H_j$.

set of types (for each player) that are consistent with the particular form of reasoning that we postulate. Then, we provide a simple procedure which yields the strategies that can be rationally played given the types that satisfy the restrictions imposed by the concept.

Let us begin by defining our notion of “a player focusing on the opponents’ rationality at some histories only”. Fix an arbitrary focus structure \mathfrak{F} . Then, for an arbitrary $\theta_i \in \Theta_i$, let

$$SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)}) := \{ t_i \in T_{\theta_i} : \text{if } (S_{-i}(h) \times T_{-\theta_i}) \cap R_{-i}^{F_{\theta_i}(h)} \neq \emptyset \text{ then } t_i \in B_i^h(R_{-i}^{F_{\theta_i}(h)}) \}. \tag{9}$$

The underlying idea is that, upon finding herself at history h , each $\theta_i \in \Theta_i$ tries to rationalize the opponents’ moves at every history in $F_{\theta_i}(h)$. If them being rational at every $h' \in F_{\theta_i}(h)$ does not contradict reaching h , then θ_i will believe at h that they are indeed rational at every history $h' \in F_{\theta_i}(h)$. This type of reasoning will be henceforth called *strong belief in rationality (with respect to \mathfrak{F}) at h* .¹¹

Let us now iterate the reasoning of strong belief in rationality to obtain our solution concept of common strong belief in rationality. Keeping the focus structure fixed, take an arbitrary $\theta_i \in \Theta_i$ and an arbitrary history $h \in H_i$. Then, we define the following sequences of subsets of $T_{\theta_i} \subseteq T_i$:

$$\begin{aligned} T_{\theta_i}^1(h) &:= SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)}) \\ T_{\theta_i}^2(h) &:= T_{\theta_i}^1(h) \cap SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^1(F_{\theta_i}(h)))) \\ &\vdots \\ T_{\theta_i}^k(h) &:= T_{\theta_i}^{k-1}(h) \cap SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h)))) \\ &\vdots \end{aligned}$$

where, for each $k > 1$,

$$T_{-\theta_i}^{k-1}(F_{\theta_i}(h)) := \times_{j \neq i} \{ t_j \in T_{\theta_j} : t_j \in T_{\theta_j}^{k-1}(h') \text{ for all } h' \in F_{\theta_i}(h) \cap H_j \},$$

with $(\theta_j)_{j \neq i}$ being the unique element of Θ_{-i} receiving positive probability by θ_i at h in \mathfrak{F} .

Obviously, $T_{\theta_i}^1(h)$ contains the types in T_{θ_i} that strongly believe at h that the opponents are rational at every $h' \in H_j \cap F_{\theta_i}(h)$. Throughout the paper, we refer to the types in

$$T_i^1(h) := \bigcup_{\theta_i \in \Theta_i} T_{\theta_i}^1(h)$$

¹¹ Interestingly, when θ_i focuses only on future histories, the strong belief operator reduces to a standard belief operator. We further elaborate on this issue in Sect. 8.2.

as those satisfying *1-fold strong belief in rationality at h*. Likewise, $T_{\theta_i}^2(h)$ contains those types in $T_{\theta_i}^1(h)$ that strongly believe at h that every opponent j (i) is rational at every $h' \in H_j \cap F_{\theta_i}(h)$, and (ii) strongly believes at every $h' \in H_j \cap F_{\theta_i}(h)$ that every opponent $k \neq j$ is rational at every $h'' \in H_k \cap F_{\theta_j}(h')$. The reason for explicitly requiring every type in $T_{\theta_i}^2(h)$ to belong to $T_{\theta_i}^1(h)$ is that the strong belief operator is not monotonic, thus implying that $SB_{\theta_i}^h(E \cap F)$ is not necessarily equal to $SB_{\theta_i}^h(E) \cap SB_{\theta_i}^h(F)$.¹² Therefore $SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)})$ does not follow directly from $SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)}) \cap (S_{-i} \times T_{-\theta_i}^1(F_{\theta_i}(h)))$. Throughout the paper, we refer to the types in

$$T_i^2(h) := \bigcup_{\theta_i \in \Theta_i} T_{\theta_i}^2(h)$$

as those satisfying *up to 2-fold strong belief in rationality at h*. The reason we add the term “up to” is that, by construction, $T_i^2(h) \subseteq T_i^1(h)$, as we have already discussed above.

Continuing inductively we define the set of types that satisfy *up to k-fold strong belief in rationality at h*. Those are the types in $T_i^k(h)$. Then, the types that satisfy *common strong belief in rationality at h* are those in

$$T_i^{\mathfrak{F}}(h) := \bigcap_{k=1}^{\infty} T_i^k(h). \tag{10}$$

The types that satisfy *common strong belief in rationality* (\mathfrak{F} -CSBR) are those in

$$T_i^{\mathfrak{F}} := \bigcap_{h \in H_i} T_i^{\mathfrak{F}}(h). \tag{11}$$

Observe that in order to obtain the types that satisfy \mathfrak{F} -CSBR, we need to take two intersections. In particular, first we find, for each $h \in H_i$, the types that that satisfy the (infinitely many) restrictions that \mathfrak{F} -CSBR imposes at h (see Eq. (10)), and then we select those types that satisfy all these restrictions at every $h \in H_i$ (see Eq. (11)). Finally, we say that a strategy $s_i \in S_i$ can be *rationally played under common strong belief in rationality* (\mathfrak{F} -RCSBR) whenever $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times T_i^{\mathfrak{F}}))$.

6 Elimination procedure

In this section we introduce a (finite) procedure which, for every player $i \in I$, for every $\theta_i \in \Theta_i$ and every history $h \in H_i$, iteratively eliminates (at each round) strate-

¹² It is well known that the conjunction property implies monotonicity. Therefore, violations of monotonicity—which the strong belief operator exhibits—lead to violations of the conjunction property. We refer to Battigalli and Siniscalchi (2002) for a detailed discussion on this issue.

gies from $S_i(h)$ and first order conditional beliefs from $\Delta(S_{-i}(h))$.¹³ Formally, it is a simultaneous generalization of the iterated conditional dominance procedure (ICDP), originally introduced by Shimoji and Watson (1998), and the backward dominance procedure, originally defined in Perea (2014).¹⁴ Before formally defining our procedure, let us first introduce the notion of a decision problem, which will play a central role throughout this section. In particular, our procedure will be defined as a sequence of decision problems for each $i \in I$, for each $\theta_i \in \Theta_i$ and each history $h \in H_i$.

Decision problem A decision problem for $\theta_i \in \Theta_i$ at a history $h \in H_i$ is a tuple $(B_{\theta_i}(h), D_{\theta_i}(h))$, with $B_{\theta_i}(h) \subseteq S_{-i}(h)$ and $D_{\theta_i}(h) \subseteq S_i(h)$. Intuitively, $B_{\theta_i}(h)$ can be seen as the subset of the opponents' strategies that θ_i could deem possible at h . At this point, we should already make clear that the link between $B_{\theta_i}(h)$ and what θ_i could deem possible at h is only an informal one. The actual relationship between the two will become apparent later on in the paper. Thus, for the time being, $B_{\theta_i}(h)$ and $D_{\theta_i}(h)$ will be merely treated as auxiliary tools, without a concrete meaning.

A strategy $s_i \in D_{\theta_i}(h)$ is said to be rational in the decision problem $(B_{\theta_i}(h), D_{\theta_i}(h))$ whenever there exists a probability measure $\beta_i^h \in \Delta(B_{\theta_i}(h))$ such that $U_i^h(s_i, \beta_i^h) \geq U_i^h(s'_i, \beta_i^h)$ for all $s'_i \in D_{\theta_i}(h)$. Thus, we draw a link between two different notions of rationality, i.e., between rationality of a strategy-type combination in a complete type structure on the one hand, and rationality of a strategy in a decision problem on the other hand.

Now, for an arbitrary focus structure \mathfrak{F} , our procedure will be defined by means of a (weakly) decreasing sequence $(B_{\theta_i}^k(h), D_{\theta_i}^k(h))_{k \geq 0}$ of decision problems for each $i \in I$, each $\theta_i \in \Theta_i$ and each $h \in H_i$. That is, at each step of our procedure, we will simultaneously eliminate strategies from $S_i(h)$ and strategy combinations from $S_{-i}(h)$.

Initial step of the procedure For $k = 0$, we define

$$\begin{aligned} B_{\theta_i}^0(h) &:= S_{-i}(h) \\ D_{\theta_i}^0(h) &:= S_i(h). \end{aligned}$$

Obviously, this initial step does not depend on the choice of \mathfrak{F} .

Inductive step of the procedure Now, fix some $k > 0$ and suppose that for each $i \in I$, each $\theta_i \in \Theta_i$ and each $h \in H_i$ we have undertaken the $(k - 1)$ th step of our procedure, thus having obtained $(B_{\theta_i}^{k-1}(h), D_{\theta_i}^{k-1}(h))$. Then, for an arbitrary $h \in H_i$, define $(B_{\theta_i}^k(h), D_{\theta_i}^k(h))$ by

$$\begin{aligned} B_{\theta_i}^k(h) &:= \begin{cases} C_{\theta_i}^{k-1}(h) & \text{if } C_{\theta_i}^{k-1}(h) \neq \emptyset \\ B_{\theta_i}^{k-1}(h) & \text{if } C_{\theta_i}^{k-1}(h) = \emptyset \end{cases} & (12) \\ D_{\theta_i}^k(h) &:= \{ s_i \in D_{\theta_i}^{k-1}(h) : s_i \text{ is rational in } (B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h)) \}, & (13) \end{aligned}$$

¹³ In fact, as we are going to see later in the paper, it might be the case that different strategies and/or conditional beliefs are eliminated for different Θ_i -types at the same history.

¹⁴ Later in the paper, we discuss the relationship of our procedure with the iterated conditional dominance procedure and with the backward dominance procedure.

where

$$C_{\theta_i}^{k-1}(h) := \bigtimes_{j \neq i} \{ s_j \in S_j(h) : s_j \in D_{\theta_j}^{k-1}(h') \text{ for all } h' \in H_j(s_j) \cap F_{\theta_i}(h) \}, \quad (14)$$

with $(\theta_j)_{j \neq i}$ being the unique element of Θ_{-i} receiving positive probability by θ_i at h in \mathfrak{F} .

The underlying idea behind our procedure is as follows: First, for each $\theta_i \in \Theta_i$ and each $h \in H_i$, we compute $C_{\theta_i}^{k-1}(h)$ which contains all strategy combinations of i 's opponents which (i) are consistent with reaching h , and (ii) have not been eliminated from $D_{\theta_j}^{k-1}(h')$ at any $h' \in F_{\theta_i}(h)$. Notice that in principle $C_{\theta_i}^{k-1}(h)$ might be empty.

Now, having defined $C_{\theta_i}^{k-1}(h)$, we can now proceed to the k th step of our procedure, by first defining $B_{\theta_i}^k(h)$. In particular, a strategy combination $s_{-i} = (s_j)_{j \neq i}$ is eliminated from $B_{\theta_i}^{k-1}(h)$ if and only if (i) there exists some history $h' \in F_{\theta_i}(h) \cap H_j(s_j)$ such that s_j is not in $D_{\theta_j}^{k-1}(h')$ for the Θ_j -type that θ_i deems possible at h , and also (ii) there exists another strategy combination $s'_{-i} = (s'_j)_{j \neq i} \in B_{\theta_i}^{k-1}(h)$ such that for every $j \neq i$ and every $h' \in F_{\theta_i}(h) \cap H_j(s_j)$ it is the case that $s'_j \in D_{\theta_j}^{k-1}(h')$ for the same Θ_j -type θ_j that θ_i deems possible at h , i.e., not all strategy combinations are eliminated from $B_{\theta_i}^{k-1}(h)$.

Then, once we have obtained $B_{\theta_i}^k(h)$, we can define the decision problem $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$, and we eliminate from $D_{\theta_i}^{k-1}(h)$ the strategies that are not rational in this decision problem. Then, it follows from Pearce (1984, Lem. 3) that a strategy is eliminated from $D_{\theta_i}^{k-1}(h)$ if and only if it is strictly dominated by a mixed strategy within this decision problem.

This elimination procedure is called the *iterated conditional dominance procedure* (\mathfrak{F} -ICDP). Obviously, since we consider only finite dynamic games and structures with finitely many Θ_i -types for each player, \mathfrak{F} -ICDP converges after finitely many steps. That is, there exists some $K \geq 0$ such that for each $k \geq K$, for every $i \in I$, every $\theta_i \in \Theta_i$ and every $h \in H_i$, it is the case that $(B_{\theta_i}^k(h), D_{\theta_i}^k(h)) = (B_{\theta_i}^K(h), D_{\theta_i}^K(h))$. Then, we write $(B_{\theta_i}^{\mathfrak{F}}(h), D_{\theta_i}^{\mathfrak{F}}(h)) = (B_{\theta_i}^K(h), D_{\theta_i}^K(h))$. We say that a strategy s_i survives the *iterated conditional dominance procedure* for some θ_i if it is the case that $s_i \in D_{\theta_i}^{\mathfrak{F}}(h)$ for all $h \in H_i(s_i)$.

Example 1 (cont). Let us illustrate the \mathfrak{F} -ICDP using our leading example from Sect. 2. Recall the associated focus structure, described in Example 1. Let us now depict each decision problem $(B_{\theta_i}^k(h), D_{\theta_i}^k(h))$ with a normal form game (see Table 2). The steps of the \mathfrak{F} -ICDP are represented by the lines that cross out the corresponding strategies. Eliminations from $B_{\theta_i}^k(h)$ are represented by dashed lines, whereas eliminations from $D_{\theta_i}^k(h)$ are represented by continuous lines. The corresponding number next to each line refers to the step during which the respective strategy was eliminated.

In particular, at the first step ($k = 1$), RB is eliminated from $D_{\theta_a}^0(h_0)$ and from $D_{\theta_a}^0(h_0)$ because it is strictly dominated by LA at h_0 . Then, at the second step ($k = 2$), θ_b focuses on h_0 while being at h_2 , and therefore the strategy RB is eliminated from

Table 2 \mathfrak{F} -ICDP in the burning-money game

θ_a	h_0	LA	7,3	7,3	2,2	2,2	$(k = 3)$	h_1	LA	7,3	7,3	2,2	2,2	$(k = 3)$	h_2	RA	5,1	0,0	5,1	0,0	$(k = 3)$		
	LB	2,2	2,2	3,7	3,7	$(k = 1)$			LB	2,2	2,2	3,7	3,7			$(k = 3)$	RB	0,0	1,5	0,0		1,5	$(k = 3)$
	RA	5,1	0,0	5,1	0,0				0,0	1,5	0,0	1,5	$(k = 3)$				0,0	1,5	0,0	1,5		$(k = 3)$	
	RB	0,0	1,5	0,0	1,5				0,0	1,5	0,0	1,5					$(k = 3)$	0,0	1,5	0,0			

θ'_a	h_0	LA	7,3	7,3	2,2	2,2	$(k = 3)$	h_1	LA	7,3	7,3	2,2	2,2	$(k = 3)$	h_2	RA	5,1	0,0	5,1	0,0	$(k = 3)$		
	LB	2,2	2,2	3,7	3,7	$(k = 1)$			LB	2,2	2,2	3,7	3,7			$(k = 3)$	RB	0,0	1,5	0,0		1,5	$(k = 3)$
	RA	5,1	0,0	5,1	0,0				0,0	1,5	0,0	1,5	$(k = 3)$				0,0	1,5	0,0	1,5		$(k = 3)$	
	RB	0,0	1,5	0,0	1,5				0,0	1,5	0,0	1,5					$(k = 3)$	0,0	1,5	0,0			

θ_b	h_3	LA	7,3	7,3	2,2	2,2	$(k = 2)$	h_2	RA	5,1	0,0	5,1	0,0	$(k = 2)$
	LB	2,2	2,2	3,7	3,7	0,0		1,5	0,0	1,5	$(k = 2)$			

$B_{\theta_b}^1(h_2)$. Subsequently, CD and DD are eliminated from $D_{\theta_b}^1(h_2)$, as they are both strictly dominated by CC at h_2 . Now, at the third step ($k = 3$), both CD and DD are eliminated from $B_{\theta_a}^2(h_1)$ but not from $B_{\theta'_a}^2(h_1)$, as θ_a focuses on h_2 while being at h_1 whereas θ'_a does not. Likewise, CD and DD are eliminated from both $B_{\theta_a}^2(h_0)$ and $B_{\theta'_a}^2(h_0)$, and consequently LB is eliminated from both $D_{\theta_a}^2(h_0)$ and $D_{\theta'_a}^2(h_0)$.

Notice that the procedure stops after three rounds of elimination. The strategies that survive are those in bold, viz., $\{LA, RA\}$ for θ_a and θ'_a , and $\{CC, DC\}$ for θ_b , as already suggested by our heuristic analysis. \square

Remark 5 Let us stress that at each step of our procedure we perform two types of elimination, viz., for each player $i \in I$, each $\theta_i \in \Theta_i$ and each $h \in H_i$, first we eliminate opponents' strategy combinations from $B_{\theta_i}^{k-1}(h)$, and then we eliminate strategies from $D_{\theta_i}^{k-1}(h)$. Note that these two types of elimination are conceptually very different. Let us for the time being focus on $B_{\theta_i}^{k-1}(h)$.

Eliminating a strategy combination $s_{-i} \in S_{-i}(h)$ from $B_{\theta_i}^{k-1}(h)$ can be thought of as eliminating all of θ_i 's first order conditional beliefs at h that put positive probability to s_{-i} . Consequently, this elimination can be interpreted as a restriction imposed on θ_i 's types, viz., eliminating s_{-i} from $B_{\theta_i}^{k-1}(h)$ essentially means that we are ruling out all types $t_i \in T_{\theta_i}$ with the property that $\text{marg}_{S_{-i}} \lambda_i^h(t_i)(\{s_{-i}\}) > 0$. But then recall that this is exactly what \mathfrak{F} -CSBR does, i.e., it recursively imposes restrictions on θ_i 's

belief types. In the next section we show that there is indeed a very tight relationship between eliminating opponents' strategies from $B_{\theta_i}^{k-1}(h)$ and eliminating types from $T_{\theta_i}^{k-1}(h)$. Thus, it becomes clear why earlier in this section we stated that the strategy combinations in $B_{\theta_i}^k(h)$ can be thought as those that θ_i could deem possible at h after k rounds of reasoning. \square

7 Characterization results

As we have already mentioned in the previous section, there is a very tight relationship between the process of eliminating types from $T_{\theta_i}^{k-1}(h)$ and the process of eliminating opponents' strategy profiles from $B_{\theta_i}^{k-1}(h)$. The following result makes this relationship formal.

Theorem 1 *Fix a structure \mathfrak{F} and consider a complete type structure $\mathcal{T}_{\mathfrak{F}}$. Then, for every player $i \in I$, every $\theta_i \in \Theta_i$, every history $h \in H_i$ and every $k > 0$, the following hold:*

- (i) *If $t_i \in T_{\theta_i}^{k-1}(h)$ then there exists some $\beta_i^h \in \Delta(B_{\theta_i}^k(h))$ with $b_i^h(t_i) = \beta_i^h$.*
- (ii) *If $\beta_i^h \in \Delta(B_{\theta_i}^k(h))$ then there exists some $t_i \in T_{\theta_i}^{k-1}(h)$ with $b_i^h(t_i) = \beta_i^h$.*

Notice that, strictly speaking, (i) and (ii) are not each other's converse. Indeed, by part (i), $t_i \in T_{\theta_i}^{k-1}(h)$ implies $b_i^h(t_i) \in \Delta(B_{\theta_i}^k(h))$. On the other hand, $b_i^h(t_i) \in \Delta(B_{\theta_i}^k(h))$ does not necessarily imply $t_i \in T_{\theta_i}^{k-1}(h)$. For instance, consider a type $t_i \in T_{\theta_i}$ which at $h \in H_i$ puts probability 1 to a strategy-type combination $(s_{-i}, t_{-i}) \in B_{\theta_i}^k(h) \times (T_{-\theta_i}^{k-3}(h') \setminus T_{-\theta_i}^{k-2}(h'))$ where $h' \in F_{\theta_i}(h)$, with such a type always existing (by completeness of the type structure). In this case, we have $b_i^h(t_i) \in \Delta(B_{\theta_i}^k(h))$ and $t_i \notin T_{\theta_i}^{k-1}(h)$. Instead, part (ii) says that, $b_i^h(t_i) \in \Delta(B_{\theta_i}^k(h))$ implies the existence of some $t'_i \in T_{\theta_i}^{k-1}(h)$ —perhaps different than t_i —such that $b_i^h(t'_i) = b_i^h(t_i)$.

With this result at hand, we can then characterize the strategies that can be rationally played under \mathfrak{F} -CSBR, by means of the \mathfrak{F} -ICDP.

Theorem 2 *Fix a structure \mathfrak{F} and consider a complete type structure $\mathcal{T}_{\mathfrak{F}}$. Then, for each player $i \in I$ and each $\theta_i \in \Theta_i$, it is the case that $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times T_{\theta_i}^{\mathfrak{F}}))$ if and only if $s_i \in D_{\theta_i}^{\mathfrak{F}}(h)$ for all $h \in H_i(s_i)$.*

The previous result formally states that a strategy can be rationally played (by a Θ_i -type θ_i) under \mathfrak{F} -CSBR if and only if it survives the \mathfrak{F} -ICDP (for θ_i). For instance, in the context of Example 1, the only strategies that can be rationally played by θ_a under \mathfrak{F} -CSBR are LA and RA , as these are the only ones surviving the \mathfrak{F} -ICDP. Likewise, the strategies that can be rationally played by θ_b under \mathfrak{F} -CSBR are CC and DC , as those are the ones surviving the \mathfrak{F} -ICDP.

8 Special cases revisited

As we have already briefly discussed in Sect. 3.4, it seems natural to associate forward induction (FI) with players who focus on all histories and backward induction (BI) with players who focus only on future histories throughout the game. In this section we study the formal relationship between Battigalli and Siniscalchi’s (2002) common strong belief in rationality and \mathfrak{F} -CSBR with $F_i(h) = H$, as well as the one between Perea’s (2014) common belief in future rationality and \mathfrak{F} -CSBR with $F_i(h) = \text{Fut}(h)$.

8.1 Focusing on all histories: forward induction

The general idea behind forward induction reasoning is that players observe their opponents’ past behavior and use this information in order to form beliefs about their opponents’ future behavior.¹⁵ The most prominent forward induction solution concept is extensive-form rationalizability (EFR), originally introduced by Pearce (1984), subsequently simplified by Battigalli (1997) and later epistemically characterized by Battigalli and Siniscalchi (2002) by means of rationality and common strong belief in rationality (in a complete type structure). The main idea is that players try to rationalize the opponents’ strategies whenever this is possible, thus implicitly postulating that players *focus on the opponents’ rationality at all histories*.

Let us first formally recall the concept of up to k -fold strong belief in rationality, as it was originally defined by Battigalli and Siniscalchi (2002). Consider the following sequences of subsets of T_i :

$$\begin{aligned}
 SB_i^1 &:= SB_i(R_{-i}) \\
 SB_i^2 &:= SB_i^1 \cap SB_i(R_{-i} \cap (S_{-i} \times SB_{-i}^1)) \\
 &\vdots \\
 SB_i^k &:= SB_i^{k-1} \cap SB_i(R_{-i} \cap (S_{-i} \times SB_{-i}^{k-1})) \\
 &\vdots
 \end{aligned}$$

with $SB_{-i}^{k-1} := \times_{j \neq i} SB_j^{k-1}$ for each $k > 1$. Moreover, let

$$CSB_i := \bigcap_{k=1}^{\infty} SB_i^k \tag{15}$$

be the set of types that satisfy *common strong belief in rationality (CSBR)*. Finally, we say that a strategy s_i can be rationally played under CSBR whenever $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times CSB_i))$.

¹⁵ FI is not a solution concept. Rather it is a general principle which is present in different concepts that have appeared in the literature (e.g., Pearce 1984; Battigalli and Siniscalchi 2002; Stalnaker 1998; Battigalli and Friedenberg 2012; Govindan and Wilson 2009; Cho 1987; Cho and Kreps 1987; Reny 1992; McLennan 1985; Hillas 1994).

Let us now consider a structure \mathfrak{F} such that $(F_i)_{i \in I}$ is commonly known, with $F_i(h) = H$ for all $h \in H_i$ and all $i \in I$. Then, we ask whether there is a formal relationship between Battigalli and Siniscalchi's (2002) CSBR on the one hand and our \mathfrak{F} -CSBR on the other. As it turns out the two notions are equivalent, as shown below.

Proposition 1 *Let the structure \mathfrak{F} be such that $(F_i)_{i \in I}$ is commonly known with $F_i(h) = H$ for all $i \in I$ and all $h \in H_i$, and consider an arbitrary type structure $\mathcal{T}_{\mathfrak{F}}$. Then, for every player $i \in I$ and every $k > 0$, it is the case that $SB_i^k = \bigcap_{h \in H_i} T_i^k(h)$.*

Two immediate conclusions follow from the previous result. First, a type satisfies CSBR if and only if it satisfies \mathfrak{F} -CSBR. Then, it naturally follows that a strategy can be rationally played under CSBR if and only if it can be rationally played under \mathfrak{F} -CSBR. This is formally stated in the following corollary. The proof trivially follows from the definition of rationality.

Corollary 1 *Let the structure \mathfrak{F} be such that $(F_i)_{i \in I}$ is commonly known with $F_i(h) = H$ for all $i \in I$ and all $h \in H_i$, and consider an arbitrary type structure $\mathcal{T}_{\mathfrak{F}}$. Then, for every player $i \in I$, it is the case that $\text{Proj}_{S_i}(R_i \cap (S_i \times CSB_i)) = \text{Proj}_{S_i}(R_i \cap (S_i \times T_i^{\mathfrak{F}}))$.*

Another direct consequence of the previous result—combined with Theorem 2 of the previous section and the characterization result of Shimoji and Watson (1998)—is that a strategy survives k steps of our \mathfrak{F} -ICDP if and only if it survives k steps of Shimoji and Watson's (1998) ICDP.

Remark 6 It is known that whenever we restrict attention to complete type structures, Rationality and CSBR epistemically characterize the strategies that are predicted by extensive form rationalizability (EFR) (Battigalli and Siniscalchi 2002). On the other hand, if we allow for an arbitrary type structure, Rationality and CSBR predict the strategy profiles that belong to an extensive form best response set (EFBRS) (Battigalli and Friedenberg 2012). The fact that Proposition 1 holds for an arbitrary type structure—and not just for a complete one—implies that Rationality and \mathfrak{F} -CSBR induce an EFBRS. □

8.2 Focusing on future histories: backward induction

Contrary to forward induction, the general idea behind backward induction is that players maintain the belief that their opponents will continue being rational irrespective of the moves they have observed so far.¹⁶ The two concepts that in our view capture this idea—and nothing more—are the backward dominance procedure (BDP) (Perea 2014) and backward rationalizability (BR) (Penta 2015).¹⁷ Both BDP and BR are epistemically characterized by rationality and common belief in future rationality in

¹⁶ Once again, BI is not a solution concept but rather a general principle embodied in different concepts in the literature (e.g., Selten 1965; Kreps and Wilson 1982; Perea 2014; Baltag et al. 2009; Penta 2015).

¹⁷ Concepts like subgame perfect equilibrium (Selten 1965) or sequential equilibrium (Kreps and Wilson 1982) impose additional equilibrium conditions, whereas the standard backward induction procedure is well-defined only for perfect information extensive-form games without relevant ties.

a complete type structure (Perea 2014), thus implicitly postulating that players *focus on the opponents' rationality at present and future histories only*.¹⁸

First, we define the event that player i believes in the opponents' future rationality by

$$FB_i(R_{-i}) := \bigcap_{h \in H_i} B_i^h(R_{-i}^{\text{Fut}(h)}). \tag{16}$$

Then, we consider the following sequence of subsets of T_i :

$$\begin{aligned} FB_i^1 &:= FB_i(R_{-i}) \\ FB_i^2 &:= FB_i^1 \cap B_i(S_{-i} \times FB_{-i}^1) \\ &\vdots \\ FB_i^k &:= FB_i^{k-1} \cap B_i(S_{-i} \times FB_{-i}^{k-1}) \\ &\vdots \end{aligned}$$

where $FB_{-i}^{k-1} := \times_{j \neq i} FB_j^{k-1}$ for each $k > 1$. We say that FB_i^k contains the types that satisfy *up to k -fold belief in future rationality*.¹⁹ Now, let

$$CFB_i := \bigcap_{k=1}^{\infty} FB_i^k \tag{17}$$

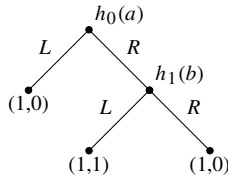
contain the types that satisfy *common belief in future rationality (CBFR)*. We say that a strategy s_i can be rationally played under CBFR whenever $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times CFB_i))$.

Now consider a focus structure \mathfrak{F} such that $(F_i)_{i \in I}$ is commonly known, with $F_i(h) = \text{Fut}(h)$ for all $h \in H_i$ and all $i \in I$. Then, we investigate the formal relationship between Perea's (2014) CBFR and our \mathfrak{F} -CSBR. First, let us point out that whenever $F_i(h) = \text{Fut}(h)$, it is by definition the case that strong belief is directly reduced to standard belief. The reason is that rationality in $F_i(h)$ will never be contradicted by what i has already observed, as $F_i(h)$ does not contain any past history, i.e., formally speaking, it is the case that $(S_{-i}(h) \times T_{-i}) \cap R_{-i}^{F_i(h)} \neq \emptyset$. Thus, we will not be surprised if the two notions are equivalent in terms of the strategy profiles they predict. Still, as the next example illustrates, this is not necessarily the case for the types they induce.

¹⁸ The only difference between the two is that BR postulates Bayesian updating, whereas BDP does not. Throughout the paper, we will mostly focus our discussion on BDP. Nonetheless, our analysis is also valid in the case of BR.

¹⁹ Notice that in the previous definition we have the set $B_i(S_{-i} \times FB_{-i}^{k-1})$ rather than $B_i(R_{-i} \cap (S_{-i} \times FB_{-i}^{k-1}))$. This is in contrast to the respective definition of "up to k -fold strong belief in rationality". This is because—unlike strong belief— B_i is a monotonic operator (Battigalli and Siniscalchi 2002). Hence, it is clear that our definition of FB_i^k contains redundancies, i.e., FB_i^k could have been instead defined to be $B_i(S_i \times FB_{-i}^{k-1})$ instead of $FB_i^{k-1} \cap B_i(S_i \times FB_{-i}^{k-1})$. Nevertheless, such redundancy is helpful in order to illustrate that this is a special case of our general definition.

Example 2 Consider the following dynamic game between Ann and Bob, and let (F_a, F_b) be commonly known, with $F_i(h) = \text{Fut}(h)$ for each $i \in \{a, b\}$ and each $h \in H_i$, i.e., $F_i(h_0) = \{h_0, h_1\}$ and $F_i(h_1) = \{h_1\}$ for both players i .



Now, consider the type structure $\mathcal{T}_{\mathfrak{F}}$ with the type spaces being $T_a = \{t_a, t'_a\}$ and $T_b = \{t_b\}$ and the corresponding conditional beliefs being given by

$$\begin{aligned} \lambda_a^{h_0}(t_a) &= (1 \otimes (R, t_b)) \\ \lambda_a^{h_0}(t'_a) &= (1 \otimes (L, t_b)) \\ \lambda_b^{h_1}(t_b) &= (1 \otimes (R, t_a)) \end{aligned}$$

First notice that the only type of Ann that is consistent with up to onefold belief in future rationality is t'_a , viz., formally, $FR_a^1 = \{t'_a\}$. This is because, Bob's unique rational strategy at h_1 is to choose L . This implies that t_b does not satisfy up to twofold belief in future rationality. Indeed, observe that $\lambda_b^{h_1}(t_b)(S_a \times FR_a^1) = \lambda_b^{h_1}(t_b)(S_a \times \{t'_a\}) = 0$. In fact, it is the case that $FR_b^2 = \emptyset$, i.e., there is no type of Bob satisfying up to twofold belief in future rationality.

Now, switching our attention to \mathfrak{F} -CSBR, observe that again the only type of Ann satisfying 1-fold strong belief in rationality at h_0 is t'_a , viz., $T_a^1(h_0) = \{t'_a\}$. But, then t_b does satisfy up to twofold strong belief in rationality at h_1 . This is because $F_b(h_1) = \{h_1\}$, and therefore $T_b^k(h_1) = T_b$ for all $k > 0$. \square

The reason for the previously illustrated divergence between FB_b^2 and $\bigcap_{h \in H_b} T_b^2(h)$ is that in order for a type t_b to satisfy up to twofold belief in future rationality, it must attach at h_1 probability 1 to $S_a \times FB_a^1$. But then, FB_a^1 contains Ann's types that require Ann to believe at h_0 that Bob will be rational from that point onwards. In other words, t_b must believe at h_1 that Ann believed at the *earlier history* h_0 that Bob would be rational at all histories following h_0 . On the other hand, in order for a type t_b to satisfy up to 2-fold strong belief in rationality, he must believe at h_1 that Ann will believe at all histories *following* h_1 that Bob will be rational at all future histories. However, in the previous example there is no history following h_1 , and hence no requirement is being imposed. This observation is summarized by the following remark.

Remark 7 Our concept of \mathfrak{F} -CSBR with $F_i(h) = \text{Fut}(h)$ is a truly backward induction concept as *it completely disregards the past*. In particular, it postulates that players ignore not only the opponents' past behavior, but also the opponents' reasoning at past histories. \square

Still, even though \mathfrak{F} -CSBR and CBFR differ in the conditional beliefs that they induce, they coincide in the predictions they make. In particular, as we show below, given a complete type structure, a strategy can be rationally played under \mathfrak{F} -CSBR if and only if it can be rationally played under CBFR.

Proposition 2 *Let the structure \mathfrak{F} be such that $(F_i)_{i \in I}$ is commonly known with $F_i(h) = \text{Fut}(h)$ for all $i \in I$ and all $h \in H_i$, and consider a complete type structure $T_{\mathfrak{F}}$. Then, for every player $i \in I$, it is the case that $\text{Proj}_{S_i}(R_i \cap (S_i \times CFB_i)) = \text{Proj}_{S_i}(R_i \cap (S_i \times T_i^{\mathfrak{F}}))$.*

The proof of the result follows almost directly from Lemma B3 in Appendix B, which formally proves that BDP and \mathfrak{F} -ICDP are essentially equivalent.

Remark 8 While \mathfrak{F} -CSBR and CBFR yield the same predicted strategies in a complete type structure, this is not necessarily the case for an arbitrary type structure. To see this recall Example 2. In particular, observe that, given the type structure that we assume, Rationality and CBFR yields an empty set of predictions, whereas Rationality and \mathfrak{F} -CSBR induces a non-empty prediction, viz., $\text{Proj}_{S_b}(R_b \cap (S_b \times T_b^{\mathfrak{F}})) = \{L\}$, while $\text{Proj}_{S_b}(R_b \cap (S_b \times CFB_b)) = \emptyset$. \square

9 Discussion

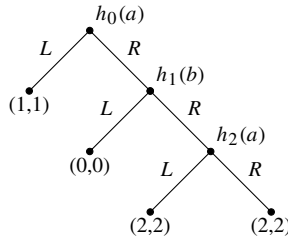
9.1 Relationship to unawareness

As we have already mentioned in the paper, our framework differs from the one concerning dynamic games with unawareness, not only in its technical aspects but also conceptually. In fact, observe that the underlying idea behind all models of unawareness is that players cannot even “see” some parts of the game, such as certain moves or even entire histories (e.g., Feinberg 2012; Heifetz et al. 2013; Halpern and Rêgo 2014). As a consequence, players do not reason about these parts of the game, similarly to what happens in our framework. However, the difference between the two is that in games with unawareness players do not even form beliefs about these parts of the game, whereas in our model the players do form beliefs, but these beliefs are completely arbitrary (see Remark 4).

This distinction is sometimes crucial, not only conceptually, but also for our predictions. To see this consider the following simple game, supposing that at h_0 Ann does not focus on h_2 . Still Ann can figure out that Bob will choose R at h_1 , even without thinking about how Bob (at h_1) will reason about her behavior at h_2 . As a consequence, our concept predicts that she will choose R at h_0 . On the other hand, if she was unaware at h_0 of the existence of R at h_1 , she would choose L at h_0 .

9.2 Infinite focus structures

Throughout the paper we have focused exclusively on finite focus structures. The main reason for doing so is for simplicity. In our view, almost all interesting cases of limited



focus can be already modelled within a finite structure and therefore we do not believe that allowing for arbitrary structures would provide any additional insight. Still, it is natural to ask whether our results also hold in the general case. Our conjecture is that they do, but one should first impose some additional topological/measure-theoretic structure. For instance, if we take some \mathfrak{F} with Θ_i being an arbitrary compact metrizable space, we may need to impose additional restrictions in order to construct the complete type structure $\mathcal{T}_{\mathfrak{F}}$. The reason is that in our analysis different Θ_i -types of player i are often treated as if they were different players. Finally, on a more practical issue, one of the major advantages of our \mathfrak{F} -ICDP is its tractability. This would not be the case anymore if \mathfrak{F} was infinite, and therefore it would become challenging to compute the predictions of the model.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

A. Complete type structures

In this section we construct a canonical type structure $\mathcal{T}_{\mathfrak{F}}$ for any focus structure \mathfrak{F} . This type structure will be complete. For notation simplicity, we prove our claim for two players, but our result can be directly generalized to any finite number of players.

We begin by fixing some $\mathfrak{F} = ((\Theta_i)_{i \in I}, (f_i)_{i \in I}, (g_i)_{i \in I})$, and for each $h \in H_i$ we consider the following sequence of spaces:

$$\begin{aligned} \Omega_i^0(h) &:= \Theta_i \times S_j(h) \\ \Omega_i^1(h) &:= \Omega_i^0(h) \times \left(\prod_{h' \in H_j} \Delta(\Omega_j^0(h')) \right) \\ &\vdots \\ \Omega_i^{k+1}(h) &:= \Omega_i^k(h) \times \left(\prod_{h' \in H_j} \Delta(\Omega_j^k(h')) \right) \\ &\vdots \end{aligned}$$

Obviously, $\Omega_i^k(h)$ is a compact metrizable space for every $k \geq 0$. Then, we define the product spaces $\tilde{T}_i^0(h) := \times_{k \geq 0} \Delta(\Omega_i^k(h))$ and $\tilde{T}_i^0 := \times_{h \in H_i} \tilde{T}_i^0(h)$. For notation simplicity, let us denote the typical element of \tilde{T}_i^0 by t_i . Hence, each $t_i \in \tilde{T}_i^0$ is essentially an abbreviation for $(\mu_{t_i}^1(h), \mu_{t_i}^2(h), \dots)_{h \in H_i}$, with $\mu_{t_i}^k(h) \in \Delta(\Omega_i^{k-1}(h))$ standing for the corresponding coordinate of t_i . As usual, we impose the standard coherency condition, thus restricting attention to collections of conditional beliefs in

$$\tilde{T}_i^1 := \{ t_i \in \tilde{T}_i^0 : \text{marg}_{\Omega_i^k(h)} \mu_{t_i}^{k+2}(h) = \mu_{t_i}^{k+1}(h) \text{ for all } k \geq 0 \text{ and all } h \in H_i \} .$$

Then, it follows from earlier works of Brandenburger and Dekel (1993) and Battigalli and Siniscalchi (1999) that there exists a homeomorphism $\tilde{\pi}_i : \tilde{T}_i^1 \rightarrow \times_{h \in H_i} \Delta(\Theta_i \times S_j(h) \times \tilde{T}_j^0)$, with $\tilde{\pi}_i^h(t_i) := \text{Proj}_{\Delta(\Theta_i \times S_j(h) \times \tilde{T}_j^0)} \tilde{\pi}_i(t_i)$. In fact, this is a direct consequence of Kolmogorov extension theorem. Now, for each $k > 1$, we recursively define

$$\tilde{T}_i^k := \{ t_i \in \tilde{T}_i^{k-1} : \tilde{\pi}_i^h(t_i)(\Theta_i \times S_j(h) \times \tilde{T}_j^{k-1}) = 1 \text{ for all } h \in H_i \}$$

and we let $\tilde{T}_i := \bigcap_{k \geq 0} \tilde{T}_i^k$ be the set of conditional belief hierarchies that satisfy coherency and common certainty in coherency. Then, again from Brandenburger and Dekel (1993) and Battigalli and Siniscalchi (1999) it follows that there exists a homeomorphism $\pi_i : \tilde{T}_i \rightarrow \times_{h \in H_i} \Delta(\Theta_i \times S_j(h) \times \tilde{T}_j)$, once again with $\pi_i^h(t_i) := \text{Proj}_{\Delta(\Theta_i \times S_j(h) \times \tilde{T}_j)} \pi_i(t_i)$. Note that \tilde{T}_i is a compact metrizable space.

Now, for an arbitrary $\theta_i \in \Theta_i$, define

$$\hat{T}_{\theta_i}^1 := \{ t_i \in \tilde{T}_i : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times \tilde{T}_j) = 1 \text{ for all } h \in H_i \}$$

and let $\hat{T}_i^1 := \bigcup_{\theta_i \in \Theta_i} \hat{T}_{\theta_i}^1$. Observe that $\{\theta_i\}$ is closed in Θ_i , and therefore $\hat{T}_{\theta_i}^1$ is closed in \tilde{T}_i (Aliprantis and Border 1994, Cor. 15.6). Thus, \hat{T}_i^1 is also closed in \tilde{T}_i , as it is the finite union of closed subsets. In fact, \hat{T}_i^1 is also open, as the complement of the closed subset $\bigcup_{\theta'_i \in \Theta_i \setminus \{\theta_i\}} \hat{T}_{\theta'_i}^1$. Then, for every $\theta_i \in \Theta_i$ and every $k > 1$, we recursively define

$$\hat{T}_{\theta_i}^k := \{ t_i \in \hat{T}_{\theta_i}^{k-1} : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times \hat{T}_j^{k-1}) = 1 \}$$

and we let $\hat{T}_{\theta_i} := \bigcap_{k \geq 0} \hat{T}_{\theta_i}^k$ and moreover $\hat{T}_i := \bigcup_{\theta_i \in \Theta_i} \hat{T}_{\theta_i}$. Using a similar argument as above, it follows from Aliprantis and Border (1994, Cor. 15.6) that $\hat{T}_{\theta_i}^k$ is closed. Hence, \hat{T}_{θ_i} is also closed as the intersection of closed subsets, and therefore so is \hat{T}_i as the finite union of closed sets.

Now, for an arbitrary $\theta_i \in \Theta_i$, define

$$T_{\theta_i}^1 := \{ t_i \in \hat{T}_{\theta_i} : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times \hat{T}_{\theta_j}) = g_i^h(\theta_i)(\theta_j) \text{ for all } \theta_j \in \Theta_j \text{ and all } h \in H_i \} .$$

Observe that \hat{T}_{θ_j} is clopen, and therefore both $\{t_i \in \hat{T}_{\theta_i} : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times \hat{T}_{\theta_j}) \geq g_i^h(\theta_i)(\theta_j)\}$ and $\{t_i \in \hat{T}_{\theta_i} : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times \hat{T}_{\theta_j}) \leq g_i^h(\theta_i)(\theta_j)\}$ are closed (Aliprantis and Border 1994, Cor. 15.6), thus implying that so is $T_{\theta_i}^1$. Then, for every $\theta_i \in \Theta_i$ and every $k > 1$, we recursively define

$$T_{\theta_i}^k := \{ t_i \in T_{\theta_i}^{k-1} : \pi_i^h(t_i)(\{\theta_i\} \times S_j(h) \times T_j^{k-1}) = 1 \}$$

and let $T_{\theta_i} := \bigcap_{k \geq 0} T_{\theta_i}^k$ and moreover $T_i := \bigcup_{\theta_i \in \Theta_i} T_{\theta_i}$. Following the same steps as above, we show that T_{θ_i} and T_i are clopen and therefore compact metrizable subspaces.

Following Brandenburger and Dekel (1993) and Battigalli and Siniscalchi (1999), we show that there exists a continuous function $\pi_i : T_i \rightarrow \times_{h \in H_i} \Delta(\Theta_i \times S_j(h) \times T_j)$ such that for every $(v_i^h)_{h \in H_i} \in \times_{h \in H_i} \Delta(\Theta_i \times S_j(h) \times T_j)$ with $v_i^h(\{\theta_i\} \times S_j(h) \times T_{\theta_j}) = g_i^h(\theta_i)(\theta_j)$ for some $\theta_i \in \Theta_i$ and for all $h \in H_i$ and all $\theta_j \in \Theta_j$, there exists some $t_i \in T_{\theta_i}$ such that $\pi_i^h(t_i) = v_i^h$ for all $h \in H_i$. Finally, define the type structure $\mathcal{T}_{\mathfrak{F}} = ((T_i)_{i \in I}, (\phi_i)_{i \in I}, (\lambda_i)_{i \in I})$, by letting $\phi_i(t_{\theta_i}) := \theta_i$ and $\lambda_i^h(t_{\theta_i}) := \text{marg}_{S_j(h) \times T_j} \pi_i^h(t_i)$ for every $\theta_i \in \Theta_i$ and every $h \in H_i$. Obviously, $\mathcal{T}_{\mathfrak{F}}$ is complete.

B. Proofs

B.1 Proofs of Sect. 5

We first introduce some additional notation and prove some intermediate results that we will use throughout the proof of our main theorem. Throughout the entire section, without loss of generality we consider a given structure \mathfrak{F} such that for each $i \in I$, each $\theta_i \in \Theta_i$ and each $h \in H_i$ there exists a unique $\theta_{-i} \in \Theta_{-i}$ with $g_i^h(\theta_i)(\theta_{-i}) = 1$.

Lemma B1 (Optimality principle) *Fix a structure \mathfrak{F} , an arbitrary player $i \in I$, an arbitrary $\theta_i \in \Theta_i$, an arbitrary history $h \in H_i$ and some $k > 0$. Then, a strategy $s_i \in D_{\theta_i}^{k-1}(h)$ is rational in $(B_{\theta_i}^k(h), S_i(h))$ if and only if it is rational in $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$.*

Proof Necessity is straightforward, i.e., if s_i is rational in $(B_{\theta_i}^k(h), S_i(h))$, then it is obviously the case that $s_i \in D_{\theta_i}^{k-1}(h)$ and moreover it is rational in the decision problem $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$. Now, let us now prove sufficiency. Take an arbitrary $s_i \in D_{\theta_i}^{k-1}(h)$ and assume that it is rational in $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$. Then, by definition, there exists some $\beta_i^h \in \Delta(B_{\theta_i}^k(h))$ such that

$$U_i^h(s_i, \beta_i^h) \geq U_i^h(s'_i, \beta_i^h) \tag{B.1}$$

for all $s'_i \in D_{\theta_i}^{k-1}(h)$. Now, assume—contrary to what we want to show—that s_i is not rational in $(B_{\theta_i}^k(h), S_i(h))$, and take an arbitrary rational strategy s''_i given β_i^h . Thus, it is the case that

$$U_i^h(s''_i, \beta_i^h) > U_i^h(s_i, \beta_i^h). \tag{B.2}$$

Notice that the last inequality is strict, because otherwise s_i would have been a rational strategy in $(B_{\theta_i}^k(h), S_i(h))$. Moreover, from the previous step it follows that $s_i'' \in D_{\theta_i}^{k-1}(h)$. But then, this contradicts the fact that s_i is rational in $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$, thus implying that s_i must necessarily be rational in $(B_{\theta_i}^k(h), S_i(h))$. \square

Now, let $T_{\theta_i}^k := \bigcap_{h \in H_i} T_{\theta_i}^k(h)$. Then, fix an arbitrary $G \in \mathcal{H} := 2^H \setminus \{\emptyset\}$, and define

$$D_{\theta_i}^k(G) := \{s_i \in S_i : s_i \in D_{\theta_i}^k(h) \text{ for all } h \in H_i(s_i) \cap G\} \tag{B.3}$$

$$\begin{aligned} R_{\theta_i}^k(G) &:= \{s_i \in S_i : \text{there is } t_i \in T_{\theta_i}^k \text{ such that } (s_i, t_i) \in R_i^h \text{ for all } h \in H_i(s_i) \cap G\} \\ &= \text{Proj}_{S_i}(R_i^G \cap (S_i \times T_{\theta_i}^k)). \end{aligned} \tag{B.4}$$

Then, we define the set of θ_i 's strategies that survive \mathfrak{F} -ICDP at all histories in G by

$$D_{\theta_i}(G) := \bigcap_{k=1}^{\infty} D_{\theta_i}^k(G).$$

Likewise, we define the set of θ_i 's strategies that are rational given some type (in T_{θ_i}) that satisfies \mathfrak{F} -CSBR at all histories in G by

$$R_{\theta_i}(G) := \bigcap_{k=1}^{\infty} R_{\theta_i}^k(G).$$

Construction of conditional beliefs Fix an arbitrary $G \in \mathcal{H}$, an arbitrary $\theta_i \in \Theta_i$ and an arbitrary $s_i \in D_{\theta_i}^1(G)$. Then, it follows directly from Pearce (1984, Lem. 3) that for every $h \in H_i(s_i) \cap G$ there exists at least one conditional belief $\beta_{\theta_i, s_i, G}^h \in \Delta(S_{-i}(h))$ such that

$$U_i^h(s_i, \beta_{\theta_i, s_i, G}^h) \geq U_i^h(s_i', \beta_{\theta_i, s_i, G}^h) \tag{B.5}$$

for all $s_i' \in S_i(h)$. Now, consider the following two cases:

- Suppose there exists some $k \in \mathbb{N}$ such that $s_i \in D_{\theta_i}^k(G) \setminus D_{\theta_i}^{k+1}(G)$. Then, it follows by definition that s_i is rational in $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$. Hence, it follows from the optimality principle (Lemma B1) that we can choose some $\beta_{\theta_i, s_i, G}^h \in \Delta(B_{\theta_i}^k(h))$ satisfying Eq. (B.5).
- Suppose that $s_i \in D_{\theta_i}^k(G)$ for all $k \in \mathbb{N}$. Then, it follows by definition that s_i is rational in $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$ for every $k \in \mathbb{N}$. Thus we can choose some $\beta_{\theta_i, s_i, G}^h \in \Delta(B_{\theta_i}^{\mathfrak{F}}(h))$ satisfying Eq. (B.5).

In either of the two cases, complete the collection of conditional beliefs $(\beta_{\theta_i, s_i, G}^h)_{h \in H_i}$ by considering arbitrary conditional beliefs $\beta_{\theta_i, s_i, G}^{h'} \in \Delta(S_{-i}(h'))$ for every $h' \in H_i \setminus (H_i(s_i) \cap G)$.

Construction of types For each player $i \in I$ and each $\theta_i \in \Theta_i$ define the finite set

$$\Psi_{\theta_i} := \{\psi_{\theta_i, s_i, G} \mid (s_i, G) \in S_i \times \mathcal{H}\},$$

and let $\Psi_i := \bigcup_{\theta_i \in \Theta_i} \Psi_{\theta_i}$ and $\Psi_{-i} := \prod_{j \neq i} \Psi_j$. Now, define the function $\phi_i : \Psi_i \rightarrow \Theta_i$ by $\phi_i(\psi_i) = \theta_i$ for each $\psi_i \in \Psi_{\theta_i}$. Moreover, define the mapping $\gamma_i^h : \Psi_i \rightarrow \Delta(S_{-i}(h) \times \Psi_{-i})$ for each $h \in H_i$ as follows: For each $s_i \in D_{\theta_i}^1(G)$, let

$$\begin{aligned} &\gamma_i^h(\psi_{\theta_i, s_i, G})(s_{-i}, \psi_{-i}) \\ &:= \begin{cases} \beta_{\theta_i, s_i, G}^h(s_{-i}) & \text{if } \psi_j = \psi_{\theta_j, s_j, F_{\theta_i}(h)} \text{ for all } j \neq i \text{ and } g_i^h(\theta_i)((\theta_j)_{j \neq i}) = 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

On the other hand, if $s_i \notin D_{\theta_i}^1(G)$, let $\gamma_i^h(\psi_{\theta_i, s_i, G})$ be an arbitrary probability measure over $S_{-i}(h) \times \Psi_{-i}$ such that $(\text{marg}_{\Psi_{-i}} \gamma_i^h(\psi_{\theta_i, s_i, G}))(\prod_{j \neq i} \phi_j^{-1}(\theta_j)) = g_i^h(\theta_i)((\theta_j)_{j \neq i})$ for all $(\theta_j)_{j \neq i} \in \Theta_{-i}$. Now, observe that $((\Psi_i)_{i \in I}, (\phi_i)_{i \in I}, (\gamma_i)_{i \in I})$ is a finite type structure, implying that each $\psi_i \in \Psi_i$ is associated with a hierarchy of conditional beliefs.

Recall that we have assumed $\mathcal{T}_{\mathcal{F}} = ((T_i)_{i \in I}, (\phi_i)_{i \in I}, (\lambda_i)_{i \in I})$ to be a complete type structure. Then, it follows from Appendix A that there exists a function $\xi_i : \Psi_i \rightarrow T_i$ mapping each $\psi_{\theta_i, s_i, G}$ to a (unique) type $t_{\theta_i, s_i, G} := \xi_i(\psi_{\theta_i, s_i, G}) \in T_{\theta_i}$ such that (i) $t_{\theta_i, s_i, G}$ and $\psi_{\theta_i, s_i, G}$ induce the same conditional belief hierarchy, and moreover (ii) it is the case that $\phi_i(\psi_{\theta_i, s_i, G}) = \phi_i(t_{\theta_i, s_i, G})$. Furthermore, notice that by construction it is the case that $\lambda_i^h(t_{\theta_i, s_i, G})(s_{-i}, t_{-i}) = \gamma_i^h(\psi_{\theta_i, s_i, G})(s_{-i}, \xi_i^{-1}(t_{-i}))$ for all $(s_{-i}, t_{-i}) \in S_{-i} \times T_{-i}$. Finally, by construction it is the case that $(s_i, t_{\theta_i, s_i, G}) \in R_{\theta_i}^G$ whenever $s_i \in D_{\theta_i}^1(G)$.

Before moving on, for notation simplicity, let us adopt the convention that $T_{\theta_i}^0(h) := T_{\theta_i}$.

Lemma B2 *For every $i \in I$, every $\theta_i \in \Theta_i$, every $G \in \mathcal{H}$ and every $k > 0$, the following hold:*

- (i) *If $t_i \in T_{\theta_i}^{k-1}(h)$ then $b_i^h(t_i) \in \Delta(B_{\theta_i}^k(h))$.*
- (ii) *If $s_i \in D_{\theta_i}^k(G)$ then $t_{\theta_i, s_i, G} \in T_{\theta_i}^{k-1}(h)$ for all $h \in H_i(s_i) \cap G$.*
- (iii) *$R_{\theta_i}^{k-1}(G) = D_{\theta_i}^k(G)$.*

Proof We prove the result by induction on k .

Initial step First, it is rather trivial to prove the result for $k = 1$. Indeed, observe that by construction it is the case that $B_{\theta_i}^{\mathcal{F}, 1}(h) = B_{\theta_i}^{\mathcal{F}, 0}(h) = S_{-i}(h)$, and therefore $\Delta(B_{\theta_i}^{\mathcal{F}, 1}(h)) = \Delta(S_{-i}(h))$, thus implying that $b_i^h(t_i) \in \Delta(B_{\theta_i}^{\mathcal{F}, 1}(h))$ for all $t_i \in T_{\theta_i}$, which proves (i). Moreover, recall from our convention that $T_{\theta_i}^{\mathcal{F}, 0}(h) = T_{\theta_i}$, thus implying that $t_{\theta_i, s_i, G} \in T_{\theta_i}^{\mathcal{F}, 0}(h)$ for all $h \in H_i(s_i) \cap G$, irrespective of whether $s_i \in D_{\theta_i}^{\mathcal{F}, 1}$ or not, which proves (ii). Finally, notice that

$$R_{\theta_i}^{\mathcal{F}, 0}(G) = \{s_i \in S_i : \text{there is } t_i \in T_{\theta_i}^{\mathcal{F}, 0}(h) \text{ such that } (s_i, t_i) \in R_i^h \text{ for all } h \in H_i(s_i) \cap G\}$$

$$\begin{aligned}
 &= \{s_i \in S_i : \text{there is } t_i \in T_{\theta_i} \text{ such that } (s_i, t_i) \in R_i^G\} \\
 &= D_{\theta_i}^{\mathfrak{S},1}(G)
 \end{aligned}$$

which proves (iii).

Inductive step We assume that the result holds for an arbitrary $k > 0$. We will refer to this as our “induction assumption (IA)”. Then, we are going to prove it for $k + 1$.

Proof of (i) Fix some $h \in H_i$, and assume that $t_i \in T_{\theta_i}^k(h)$. Then, by definition it is the case that

$$t_i \in SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))))).$$

Then, we consider the following two cases:

- (a) Let $R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) \neq \emptyset$.

By the definition of strong belief (at h) it is the case that $\lambda_i^h(t_i)(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h)))) = 1$. Now, recall by Eq. (B.4) that

$$R_{-\theta_i}^{k-1}(F_{\theta_i}(h)) = \text{Proj}_{S_{-i}}(R_{-i}^{F_i(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))))),$$

and therefore it follows that $b_i^h(t_i)(R_{-\theta_i}^{k-1}(F_{\theta_i}(h))) = 1$. Now observe that

$$\begin{aligned}
 R_{-\theta_i}^{k-1}(F_{\theta_i}(h)) &= \bigtimes_{j \neq i} \{s_j \in S_j : s_j \in R_{\theta_j}^{k-1}(F_{\theta_i}(h))\} && \text{(B.6)} \\
 &= \bigtimes_{j \neq i} \{s_j \in S_j : s_j \in D_{\theta_j}^k(F_{\theta_i}(h))\} && \text{(by the IA)} \\
 &= \bigtimes_{j \neq i} \{s_j \in S_j : s_j \in D_{\theta_j}^k(h') \text{ for all } h' \in H_j \cap F_{\theta_i}(h)\},
 \end{aligned}$$

with $(\theta_j)_{j \neq i} \in \Theta_{-i}$ being such that $g_i^h(\theta_i)((\theta_j)_{j \neq i}) = 1$. Thus, it is the case that

$$\begin{aligned}
 C_{\theta_i}^k(h) &= \bigtimes_{j \neq i} \{s_j \in S_j(h) : s_j \in D_{\theta_j}^k(h') \text{ for all } h' \in H_j \cap F_{\theta_i}(h)\} \\
 &= S_{-i}(h) \cap R_{-\theta_i}^{k-1}(F_{\theta_i}(h)).
 \end{aligned} \tag{B.7}$$

Now, there are two possibilities. According to the first possibility we have $C_{\theta_i}^k(h) \neq \emptyset$, in which case we obtain

$$\begin{aligned}
 B_{\theta_i}^{k+1}(h) &= C_{\theta_i}^k(h) \\
 &= S_{-i}(h) \cap R_{-\theta_i}^{k-1}(F_{\theta_i}(h)).
 \end{aligned}$$

Then, by combining $b_i^h(t_i)(R_{-\theta_i}^{k-1}(F_{\theta_i}(h))) = 1$ with $b_i^h(t_i)(S_{-i}(h)) = 1$, it is straightforward to obtain $b_i^h(t_i)(B_{\theta_i}^{k+1}(h)) = 1$. According to the second possibility we have $C_{\theta_i}^k(h) = \emptyset$, in which case we obtain $B_{\theta_i}^{k+1}(h) = B_{\theta_i}^k(h)$. But then,

since $t_i \in T_{\theta_i}^k(h) \subseteq T_{\theta_i}^{k-1}(h)$, it follows from the IA that $b_i^h(t_i)(B_{\theta_i}^{k+1}(h)) = b_i^h(t_i)(B_{\theta_i}^k(h)) = 1$, which completes this part of the proof for the first case.

- (b) Let $R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) = \emptyset$.

Then, it follows by definition that

$$\begin{aligned} R_{-\theta_i}^{k-1}(F_{\theta_i}(h)) \cap S_{-i}(h) &\subseteq R_{-\theta_i}^{k-1}(F_{\theta_i}(h)) \\ &= \text{Proj}_{S_{-i}} \left(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) \right) \\ &= \emptyset \end{aligned} \tag{B.8}$$

Now, using the same reasoning as in Eq. (B.6), combined with Eq. (B.8), we obtain

$$\begin{aligned} R_{-\theta_i}^{k-1}(F_{\theta_i}(h)) \cap S_{-i}(h) &= \bigtimes_{j \neq i} \{ s_j \in S_j(h) : s_j \in D_{\theta_j}^k(h') \text{ for all } h' \in H_j \cap F_{\theta_i}(h) \} \\ &= \emptyset, \end{aligned}$$

again with $(\theta_j)_{j \neq i} \in \Theta_{-i}$ being such that $g_i^h(\theta_j)_{j \neq i} = 1$. Moreover, using the same argument as in Eq. (B.7), we obtain

$$C_{\theta_i}^k(h) = S_{-i}(h) \cap R_{-\theta_i}^{k-1}(F_{\theta_i}(h)).$$

Thus, combining the previous two equations, we conclude that $C_{\theta_i}^k(h) = \emptyset$. Hence, $B_{\theta_i}^{k+1}(h) = B_{\theta_i}^k(h)$. Finally, since $t_i \in T_{\theta_i}^k(h) \subseteq T_{\theta_i}^{k-1}(h)$, it follows from the IA that $b_i^h(t_i)(B_{\theta_i}^{k+1}(h)) = b_i^h(t_i)(B_{\theta_i}^k(h)) = 1$, which completes the proof of part (i).

Proof of (ii): Take an $s_i \in D_{\theta_i}^{k+1}(G)$, and consider some $h \in H_i(s_i) \cap G$. Since $D_{\theta_i}^{k+1}(G) \subseteq D_{\theta_i}^k(G)$, it follows by the IA that $t_{\theta_i, s_i, G} \in T_{\theta_i}^{k-1}(h)$. Hence, it suffices to prove that

$$t_{\theta_i, s_i, G} \in SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h)))). \tag{B.9}$$

The latter amounts to proving that

$$\begin{aligned} &\left[R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) \neq \emptyset \right] \\ &\Rightarrow \left[\lambda_i^h(t_{\theta_i, s_i, G})(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h)))) = 1 \right] \end{aligned} \tag{B.10}$$

First, notice that $t_{\theta_i, s_i, G} \in SB_{\theta_i}^h(R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))))$ is trivially satisfied whenever $R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) = \emptyset$. Hence, we will focus on the case where $R_{-i}^{F_{\theta_i}(h)} \cap (S_{-i} \times T_{-\theta_i}^{k-1}(F_{\theta_i}(h))) \neq \emptyset$. Recall that $(\theta_j)_{j \neq i}$ is the unique element of Θ_{-i} receiving positive probability by $g_i^h(\theta_j)$. Then, for every $j \neq i$, there exists some $(s_j^*, t_j^*) \in S_j(h) \times T_{\theta_j}$ such that (i) $(s_j^*, t_j^*) \in R_j^{h'}$ for all $h' \in H_j(s_j^*) \cap F_{\theta_i}(h)$, and (ii) $t_j^* \in T_{\theta_j}^{k-1}(h')$ for all $h' \in H_j \cap F_{\theta_i}(h)$.

Now, we are going to prove that $s_j^* \in D_{\theta_j}^k(h')$ for every $h' \in H_j(s_j^*) \cap F_{\theta_i}(h)$. To do so, take an arbitrary $t_j^{k-1} \in T_{\theta_j}^{k-1}$, and define the type t_j^{**} by

$$\lambda_j^{h'}(t_j^{**}) := \begin{cases} \lambda_j^{h'}(t_j^*) & \text{for each } h' \in H_j(s_j^*) \cap F_{\theta_i}(h), \\ \lambda_j^{h'}(t_j^{k-1}) & \text{for each } h' \in H_j \setminus (H_j(s_j^*) \cap F_{\theta_i}(h)). \end{cases}$$

Notice that since $T_{\mathcal{F}}$ is a complete type structure, such a type exists. Observe that by construction it is the case that $(s_j^*, t_j^{**}) \in R_j^{F_{\theta_i}(h)}$, and moreover $t_j^{**} \in T_{\theta_j}^{k-1}$. Therefore, we obtain

$$\begin{aligned} s_j^* &\in R_{\theta_j}^{k-1}(F_{\theta_i}(h)) \cap S_j(h) \\ &= D_{\theta_j}^k(F_{\theta_i}(h)) \cap S_j(h) && \text{(by the IA)} \\ &= \{s_j \in S_j(h) : s_j \in D_{\theta_j}^k(h') \text{ for all } h' \in H_j(s_j) \cap F_{\theta_i}(h)\} \\ &\neq \emptyset. \end{aligned}$$

The latter implies directly by definition that $C_{\theta_i}^k(h) \neq \emptyset$. Hence, it is—also by definition—the case that

$$B_{\theta_i}^{k+1}(h) = C_{\theta_i}^k(h). \tag{B.11}$$

Now, notice that by construction $\lambda_i^h(t_{\theta_i, s_i, G})$ put positive probability only to strategy-type pairs (s_j, t_j) such that $t_j = t_{\theta_j, s_j, F_{\theta_i}(h)}$. Moreover, since $s_i \in D_{\theta_i}^{k+1}(G)$ it follows from the construction of the beliefs that $b_i^h(t_{\theta_i, s_i, G}) \in \Delta(B_{\theta_i}^{k+1}(h))$. Therefore, it follows from Eq. (B.11) that $\text{marg}_{S_j \times T_j} \lambda_i^h(t_{\theta_i, s_i, G})$ puts positive probability only to strategy-type pairs $(s_j, t_j) \in S_j(h) \times T_j$ such that $t_j = t_{\theta_j, s_j, F_{\theta_i}(h)}$ and $s_j \in D_{\theta_j}^k(h')$ for all $h' \in H_j(s_j) \cap F_{\theta_i}(h)$. Hence, from the IA it follows that $\text{marg}_{S_j \times T_j} \lambda_i^h(t_{\theta_i, s_i, G})$ assigns probability 1 to

$$R_j^{F_{\theta_i}(h)} \cap \{ (s_j, t_j) \in S_j \times T_j : t_j \in T_{\theta_j}^{k-1}(h') \text{ for all } h' \in H_j \cap F_{\theta_i}(h) \}$$

for every $j \neq i$. Therefore, by definition, $t_{\theta_i, s_i, G} \in T_{\theta_i}^k(h)$, which completes the proof of part (ii).

Proof of (iii): First, we prove that $R_{\theta_i}^{k-1}(G) \subseteq D_{\theta_i}^k(G)$: Take an arbitrary $s_i \in R_{\theta_i}^{k-1}(G)$. By definition there exists a type in $t_i \in T_{\theta_i}^{k-1}$ such that $(s_i, t_i) \in R_i^G$. Now, by part (i) of the result—that we have already proven above—it follows that $b_i^h(t_i)(B_{\theta_i}^k(h)) = 1$ for all $h \in H_i(s_i) \cap G$, implying that at all histories $h \in H_i(s_i) \cap G$, the strategy s_i is rational in the decision problem $(B_{\theta_i}^k(h), D_{\theta_i}^{k-1}(h))$. Thus, we conclude that $s_i \in D_{\theta_i}^k(h)$ for all $h \in H_i(s_i) \cap G$. The latter directly implies that $s_i \in D_{\theta_i}^k(G)$ which completes this part of the proof.

Second, we prove that $D_{\theta_i}^k(G) \subseteq R_{\theta_i}^{k-1}(G)$: Take an arbitrary $s_i \in D_{\theta_i}^k(G)$. Then, by part (ii) that we have already proven above, it follows that $t_{\theta_i, s_i, G} \in T_{\theta_i}^{k-1}(h)$ for

all $h \in G \cap H_i(s_i)$. Now, fix an arbitrary type $t_i^{k-1} \in T_{\theta_i}^{k-1}$, and define the type $t_{\theta_i, s_i, G}^* \in T_{\theta_i}$ by

$$\lambda_i^h(t_{\theta_i, s_i, G}^*) := \begin{cases} \lambda_i^h(t_{\theta_i, s_i, G}) & \text{for each } h \in H_i(s_i) \cap G, \\ \lambda_i^h(t_i^{k-1}) & \text{for each } h \in H_i \setminus (H_i(s_i) \cap G). \end{cases}$$

Notice that since $\mathcal{T}_{\mathfrak{F}}$ is a complete type structure, such a type exists. Then, by construction it is the case that $t_{\theta_i, s_i, G}^* \in T_{\theta_i}^{k-1}$, and we have that $(s_i, t_{\theta_i, s_i, G}^*) \in R_i^h$ for all $h \in G \cap H_i(s_i)$. Hence, we conclude that $s_i \in R_{\theta_i}^{k-1}(G)$, which completes the proof of the lemma. \square

Proof of Theorem 1 Take an arbitrary $i \in I$, an arbitrary $\theta_i \in \Theta_i$ and some $h \in H_i$.

Proof of (i): It follows directly from Lemma B2.i.

Proof of (ii): Fix an arbitrary $\beta_i^h \in \Delta(B_{\theta_i}^{\mathfrak{F}, k}(h))$, and let $s_i^* \in D_{\theta_i}^{\mathfrak{F}, k}(h)$ be such that

$$U_i^h(s_i^*, \beta_i^h) \geq U_i^h(s_i, \beta_i^h) \tag{B.12}$$

for all $s_i \in D_{\theta_i}^{\mathfrak{F}, k-1}(h)$. In fact, notice that Eq. (B.12) holds, not only for every $s_i \in D_{\theta_i}^{\mathfrak{F}, k-1}(h)$, but for every $s_i \in S_i(h)$ (see Lemma B1). Now, we define $\beta_{\theta_i, s_i^*, \{h\}}^h := \beta_i^h$, and construct the type $t_{\theta_i, s_i^*, \{h\}}$ like we did above. Then, by Lemma B2.ii, it is the case that $t_{\theta_i, s_i^*, \{h\}} \in T_{\theta_i}^{\mathfrak{F}, k-1}(h)$, which—together with the fact that $\beta_{\theta_i, s_i^*, \{h\}}^h := \beta_i^h(t_{\theta_i, s_i^*, \{h\}})$ —completes the proof. \square

Proof of Theorem 2 Observe that by construction

$$\begin{aligned} R_{\theta_i}^{\mathfrak{F}}(H) &= \text{Proj}_{S_i}(R_i \cap (S_i \times T_{\theta_i}^{\mathfrak{F}})) \\ D_{\theta_i}^{\mathfrak{F}}(H) &= \{s_i \in S_i : s_i \in D_{\theta_i}^{\mathfrak{F}}(h) \text{ for all } h \in H_i(s_i)\}, \end{aligned}$$

and recall by Lemma B2.iii that $R_{\theta_i}^{\mathfrak{F}}(H) = D_{\theta_i}^{\mathfrak{F}}(H)$, which completes the proof. \square

B.2 Proofs of Sect. 8

In this section, we focus on structures \mathfrak{F} with commonly known $F \in \mathcal{F}$, implying that Θ_i is a singleton for each $i \in I$. Thus, recall that we identify the unique θ_i with i , e.g., we write $F_i(h)$ for $F_{\theta_i}(h)$.

Proof of Proposition 1 We proceed by induction on k . First, note that $SB_i^1 = \bigcap_{h \in H_i} T_i^1(h)$. Then, assume that for every $i \in I$ it is the case that $SB_i^{k-1} = \bigcap_{h \in H_i} T_i^{k-1}(h)$. Now, observe that for every $i \in I$ and $h \in H_i$, it is the case that

$$\begin{aligned}
T_{-i}^{k-1}(F_i(h)) &= \bigtimes_{j \neq i} \{t_j \in T_j : t_j \in T_j^{k-1}(h') \text{ for all } h' \in H_j\} \\
&= \bigtimes_{j \neq i} \left(\bigcap_{h' \in H_j} T_j^{k-1}(h') \right) \\
&= \bigtimes_{j \neq i} SB_j^{k-1} \\
&= SB_{-i}^{k-1}.
\end{aligned}$$

Hence, it is the case that

$$\begin{aligned}
SB_i^k &= SB_i^{k-1} \cap SB_i(R_{-i} \cap (S_{-i} \times SB_{-i}^{k-1})) \\
&= \left(\bigcap_{h \in H_i} T_i^{k-1}(h) \right) \cap \left(\bigcap_{h \in H_i} SB_i^h \left(R_{-i}^{F_i(h)} \cap (S_{-i} \times T_{-i}^{k-1}(F_i(h))) \right) \right) \\
&= \bigcap_{h \in H_i} \left(T_i^{k-1}(h) \cap SB_i^h \left(R_{-i}^{F_i(h)} \cap (S_{-i} \times T_{-i}^{k-1}(F_i(h))) \right) \right) \\
&= \bigcap_{h \in H_i} T_i^k(h)
\end{aligned}$$

which completes the proof. \square

In order to prove Proposition 2, we first recall the formal definition of the backward dominance procedure (BDP), originally introduced by Perea (2014).

Backward dominance procedure For an arbitrary $i \in I$ and an arbitrary $h \in H$, consider the following sequence of subsets of $S_i(h)$:

$$\begin{aligned}
Q_i^1(h) &:= S_i(h) \\
Q_i^2(h) &:= \{s_i \in Q_i^1(h) : s_i \text{ is rational in } (Q_{-i}^1(h'), Q_i^1(h')) \text{ at all } h' \in H_i(s_i) \cap \text{Fut}(h)\} \\
&\vdots \\
Q_i^k(h) &:= \{s_i \in Q_i^{k-1}(h) : s_i \text{ is rational in } (Q_{-i}^{k-1}(h'), Q_i^{k-1}(h')) \text{ at all } h' \in H_i(s_i) \cap \text{Fut}(h)\} \\
&\vdots
\end{aligned}$$

for each $k > 0$, where $Q_{-i}^k(h) = \bigtimes_{j \neq i} Q_j^k(h)$. We say that a strategy s_i survives k steps of the procedure at $h \in H_i$ whenever $s_i \in Q_i^k(h)$. The idea is that a strategy survives k steps of the procedure at some $h \in H_i$ whenever it is not strictly dominated in the corresponding normal form game—that has survived so far—at every history following h where i is active. Then, we define

$$Q_i(h) := \bigcap_{k=1}^{\infty} Q_i^k(h), \quad (\text{B.13})$$

and we say that a strategy survives the procedure whenever it is the case that $s_i \in Q_i(h)$ for all $h \in H_i(s_i)$.

Now, let us prove an intermediate lemma that we will use in the proof of Proposition 2.

Lemma B3 *Let the structure \mathfrak{F} be such that $(F_i)_{i \in I}$ is commonly known with $F_i(h) = \text{Fut}(h)$ for all $i \in I$ and all $h \in H_i$. Then, for every $i \in I$, every $h \in H_i$ and every $k \geq 1$ the following hold:*

- (i) $Q_{-i}^k(h) = B_i^k(h)$.
- (ii) $Q_i^{k+1}(h) = \{s_i \in S_i(h) : s_i \in D_i^k(h') \text{ for all } h' \in \text{Fut}(h) \cap H_i(s_i)\}$.

Proof We proceed to prove the result by induction on k . The result trivially holds for $k = 1$. We assume it holds for $k - 1$ and we will prove it for k . We begin with part (i). Fix an arbitrary $i \in I$ and an arbitrary $h \in H_i$, and observe that

$$\begin{aligned} B_i^k(h) &= C_i^{k-1}(h) && \text{(because } C_i^{k-1}(h) \neq \emptyset) \\ &= \times_{j \neq i} \{s_j \in S_j(h) : s_j \in D_j^{k-1}(h') \text{ for all } h' \in H_j(s_j) \cap \text{Fut}(h)\} \\ &= \times_{j \neq i} Q_j^k(h) && \text{(by the IA)} \\ &= Q_{-i}^k(h), \end{aligned}$$

which completes the inductive step of the proof for part (i).

Now, we move to the inductive step for part (ii). Again, fix an arbitrary $i \in I$ and an arbitrary $h \in H_i$, and take an arbitrary $s_i \in Q_i^{k+1}(h)$. Then, by definition, s_i is rational in $(Q_{-i}^k(h'), Q_i^k(h'))$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$, and by part (i) of the present result, s_i is rational in $(B_i^k(h'), Q_i^k(h'))$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$. Now, notice that for every $s'_i \in S_i(h')$,

$$\begin{aligned} s'_i \text{ is rational in } (B_i^k(h'), Q_i^k(h')) &\Leftrightarrow s'_i \text{ is rational in } (B_i^k(h'), S_i(h')) \\ &\Leftrightarrow s'_i \text{ is rational in } (B_i^k(h'), D_i^{k-1}(h')). \end{aligned}$$

The first equivalence follows from Perea (2012, Lem. 8.14.6), while the second one follows from Lemma B1 above. Hence, s_i is rational in $(B_i^k(h'), D_i^{k-1}(h'))$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$, thus implying that $s_i \in D_i^k(h')$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$. Therefore,

$$Q_i^{k+1}(h) \subseteq \{s_i \in S_i(h) : s_i \in D_i^k(h') \text{ for all } h' \in \text{Fut}(h) \cap H_i(s_i)\}. \tag{B.14}$$

Now, in order to prove the inverse weak inequality, take some s_i with $s_i \in D_i^k(h')$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$. This implies that s_i is rational in $(Q_{-i}^k(h'), D_i^{k-1}(h'))$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$, and by the previous sequence of equivalences, s_i is rational in $(Q_{-i}^k(h'), Q_i^k(h'))$ for every $h' \in \text{Fut}(h) \cap H_i(s_i)$. Then, by definition, $s_i \in Q_i^{k+1}(h)$, thus proving that

$$Q_i^{k+1}(h) \supseteq \{s_i \in S_i(h) : s_i \in D_i^k(h') \text{ for all } h' \in \text{Fut}(h) \cap H_i(s_i)\}. \quad (\text{B.15})$$

Then, inequalities (B.14) and (B.15) complete this part of the proof. \square

Proof of Proposition 2 It follows from Perea (2014, Thm. 5.4) that a strategy can be rationally played under CBF_R (in a complete type structure) if and only if it survives the BDP, i.e., formally, $s_i \in Q_i(h)$ for all $h \in H_i(s_i)$ if and only if $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times CFB_i))$. Moreover, from our Theorem 2, a strategy s_i can be rationally played under \mathfrak{F} -CSBR (in a complete type structure) if and only if it survives the \mathfrak{F} -ICDP, i.e., formally, $s_i \in D_i^{\mathfrak{F}}(h)$ for all $h \in H_i(s_i)$ if and only if $s_i \in \text{Proj}_{S_i}(R_i \cap (S_i \times T_i^{\mathfrak{F}}))$. Thus, it suffice to prove that a strategy survives BDP if and only if it survives \mathfrak{F} -ICDP.

First, consider an arbitrary strategy s_i surviving the BDP. Then, it must be the case that $s_i \in Q_i^k(h)$ for every $k > 0$ and every $h \in H_i(s_i)$. Thus, by Lemma B3, the latter is true if and only if $s_i \in \{s'_i \in S_i(h) : s'_i \in D_i^k(h') \text{ for all } h' \in \text{Fut}(h) \cap H_i(s_i)\}$ for all $k > 0$ and for all $h \in H_i(s_i)$. Obviously, the latter is equivalent to $s_i \in D_i^k(h)$ for every $k > 0$ and every $h \in H_i(s_i)$, which by definition means that s_i survives the \mathfrak{F} -ICDP, thus completing the proof. \square

References

- Aliprantis C, Border K (1994) Infinite dimensional analysis. Springer, Berlin
- Aumann RJ, Sorin S (1989) Cooperation and bounded recall. *Games Econ Behav* 1:5–39
- Baltag A, Smets S, Zvesper JA (2009) Keep hoping for rationality: a solution to the backward induction paradox. *Synthese* 169:705–737
- Barlo M, Carmona M, Sabourian H (2009) Repeated games with one-memory. *J Econ Theory* 312–336
- Battigalli P (1997) On rationalizability in extensive games. *J Econ Theory* 74:40–61
- Battigalli P, Friedenberg A (2012) Forward induction reasoning revisited. *Theor Econ* 57–98
- Battigalli P, Siniscalchi M (1999) Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *J Econ Theory* 88:188–230
- Battigalli P, Siniscalchi M (2002) Strong belief and forward induction reasoning. *J Econ Theory* 106:356–391
- Ben-Porath E, Dekel E (1992) Signaling future actions and potential for sacrifice. *J Econ Theory* 57:36–51
- Bonanno G (2004) Memory and perfect recall in extensive games. *Games Econ Behav* 47:237–256
- Brandenburger A, Dekel E (1993) Hierarchies of beliefs and common knowledge. *J Econ Theory* 59:189–198
- Cho IK (1987) A refinement of sequential equilibrium. *Econometrica* 55:1367–1389
- Cho IK, Kreps DM (1987) Signaling games and stable equilibria. *Q J Econ* 102:179–221
- Cole HL, Kocherlakota NR (2005) Finite memory and imperfect monitoring. *Games Econ Behav* 53:59–72
- Dow J (1991) Search decisions with limited memory. *Rev Econ Stud* 58:1–14
- Feinberg Y (2012) Games with unawareness. In: Stanford Graduate School of Business paper, p 2122
- Friedenberg A (2010) When do type structures contain all hierarchies of beliefs? *Games Econ Behav* 68:108–129
- Govindan S, Wilson R (2009) On forward induction. *Econometrica* 77:1–28
- Grossi D, Turrini P (2012) Short sight in extensive games. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems (AAMAS 2012), pp 805–812
- Halpern JY, Rêgo LC (2014) Extensive games with possibly unaware players. *Math Soc Sci* 70:42–58
- Harsanyi J (1967–68) Games with incomplete information played by Bayesian players, I–III. *Manag Sci* 14:159–182, 320–334, 486–502
- Heifetz A, Meier M, Schipper B (2013) Dynamic unawareness and rationalizable behavior. *Games Econ Behav* 81:50–68
- Hillas J (1994) Sequential equilibria and stable sets of beliefs. *J Econ Theory* 64:78–102
- Jehiel P (2001) Limited foresight may force cooperation. *Rev Econ Stud* 68:369–391

- Johnson EJ, Camerer C, Sen S, Rymon T (2002) Detecting failures of backward induction: monitoring information search in sequential bargaining. *J Econ Theory* 104:16–47
- Kalai E (1990) Bounded rationality and strategic complexity in repeated games. In: Ichiishi T, Neyman A, Tauman Y (eds) *Game theory and applications*. Academic Press, New York, pp 131–157
- Kline JJ (2002) Minimum memory for equivalence between ex ante optimality and time-consistency. *Games Econ Behav* 38:278–305
- Kreps DM, Wilson R (1982) Sequential equilibria. *Econometrica* 50:863–894
- Lehrer E (1988) Repeated games with stationary bounded recall strategies. *J Econ Theory* 46:130–144
- Maenner E (2008) Adaptation and complexity in repeated games. *Games Econ Behav* 63:166–187
- McLennan A (1985) Justifiable beliefs in sequential equilibrium. *Econometrica* 53:889–904
- Mengel F (2014) Learning by (limited) forward looking players. *J Econ Behav Organ* 108:59–77
- Mullainathan S (2002) A memory-based model of bounded memory. *Q J Econ* 117:735–774
- Pearce DG (1984) Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52:1029–1050
- Peeters G, Czapski J (1990) Positive-negative asymmetry in evaluations: the distinction between affective and informational negativity effects. *Eur Rev Soc Psychol* 1:33–60
- Penta A (2015) Robust dynamic implementation. *J Econ Theory* 160:280–316
- Perea A (2012) *Epistemic game theory: reasoning and choice*. Cambridge University Press, Cambridge
- Perea A (2014) Belief in the opponents' future rationality. *Games Econ Behav* 83:231–254
- Piccione M, Rubinstein A (1997) On the interpretation of decision problems with imperfect recall. *Games Econ Behav* 20:3–24
- Reny PJ (1992) Backward induction, normal form perfection and explicable equilibria. *Econometrica* 60:627–649
- Rényi A (1955) On a new axiomatic theory of probability. *Acta Math Acad Sci Hung* 6:285–335
- Roese NJ (1997) Counterfactual thinking. *Psychol Bull* 121:133–148
- Rubinstein A (1991) Comments on the interpretation of game theory. *Econometrica* 59:909–924
- Rubinstein A (1998) *Modeling bounded rationality*. MIT Press, Cambridge
- Selten R (1965) Spieltheoretische behandlung eines oligopolmodells mit nachfragezeit. *Z Gesamte Staatswiss* 121(301–324):667–689
- Shimoji M, Watson J (1998) Conditional dominance, rationalizability, and game forms. *J Econ Theory* 83:161–195
- Soroka SN (2006) Good news and bad news: asymmetric responses to economic information. *J Polit* 68:372–385
- Stalnaker R (1998) Belief revision in games: forward and backward induction. *Math Soc Sci* 36:31–56
- Taylor SE (1991) Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol Bull* 110:67–85
- Turrini P (2015) Computing rational decisions in extensive games with limited foresight. In: *Proceedings of the 13th AAAI conference on artificial intelligence (AAAI 2016)* (forthcoming)
- van Damme E (1989) Stable equilibria and forward induction. *J Econ Theory* 48:476–496