Transparent Restrictions on Beliefs and Forward-Induction Reasoning in Games with Asymmetric Information*

Pierpaolo Battigalli† Andrea Prestipino‡

Final draft, 2013

Abstract

We analyze forward-induction reasoning in games with asymmetric information assuming some commonly understood restrictions on beliefs. Specifically, we assume that some given restrictions Δ on players' initial or conditional first-order beliefs are *transparent*, that is, not only the restrictions Δ hold, but there is common belief in Δ at every node. Most applied models of asymmetric information are covered as special cases whereby Δ pins down the probabilities initially assigned to states of nature. But the abstract analysis also allows for transparent restrictions on beliefs about behavior, e.g. independence restrictions or restrictions induced by the context behind the game. Our contribution is twofold. First, we use dynamic interactive epistemology to formalize assumptions that capture foward-induction reasoning given the transparency of Δ, and show that the behavioral implications of these assumptions are characterized by the Δ-rationalizability solution procedure of Battigalli [5, 1999], [6, 2003]. Second, we study the differences and similarities between this solution concept and a simpler solution procedure put forward by Battigalli and Siniscalchi [12, 2003]. We show that the two procedures are equivalent if Δ is "closed under compositions", a property that holds in all the applications considered by [12, 2003]. We also show that when Δ is not closed under compositions the simpler solution procedure of [12, 2003] may fail to characterize the behavioral implications of forward induction reasoning.

J.E.L. Classification Numbers: C72, C73, D82.

KEYWORDS: Epistemic game theory, Rationalizability, Forward induction, Transparent restrictions on beliefs.

*The authors thank an anonymous referee for his or her thoughtful and constructive review of an earlier draft of this paper. The authors gratefully acknowledge useful comments by Emiliano Catonini, Nicodemo De Vito, Amanda Friedenberg, Elena Manzoni, Sara Negrelli, Jacopo Perego, and Pietro Tebaldi. All errors are their own. Pierpaolo Battigalli thanks NYU-Stern for its hospitality and gratefully acknowledges financial support from ERC grant 324219.

†Università Bocconi and IGIER. E-mail: pierpaolo.battigalli@unibocconi.it

‡New York University. E-mail: andreaprestipino@yahoo.it

1 Introduction

Forward induction reflects the idea that players rationalize their opponents' behavior, whenever possible. In particular, each player forms an assessment about the opponents' private information and future (or yet unobserved) moves, given what he observed about the past play and the presumption that his opponents are strategic. Such strategic reasoning has been studied with the tools of epistemic game theory, the formal analysis of how players' beliefs about each other are formed and – in dynamic games – how they change as the play unfolds (see Battigalli & Siniscalchi [11, 2002] and Battigalli & Friedenberg [9, 2012]). In this paper we consider dynamic games where players may either lack common knowledge of the payoff functions (incomplete information), or have imperfect and asymmetric information about chance moves. Our formalism distinguishes between these two scenarios and can mix elements of both. We provide an epistemic analysis of forward-induction reasoning when some given restrictions Δ on first-order beliefs are transparent to the players, and we show that the behavioral implications of our epistemic assumptions are characterized by an iterated deletion procedure called Δ-rationalizability, first proposed by Battigalli [5, 1999] (see also [6, 2003]). We also discuss a simpler "naïve" Δ-rationalizability algorithm put forward by Battigalli & Siniscalchi [12, 2003], arguing that it is conceptually incorrect for general restrictions Δ, but equivalent to Δ-rationalizability for the type of restrictions most frequently encountered in economic applications (see section 3.3), and in particular for the ones assumed in the examples and results of [12, 2003]. The rest of this introduction provides the background for our contribution and spells it out in more detail.

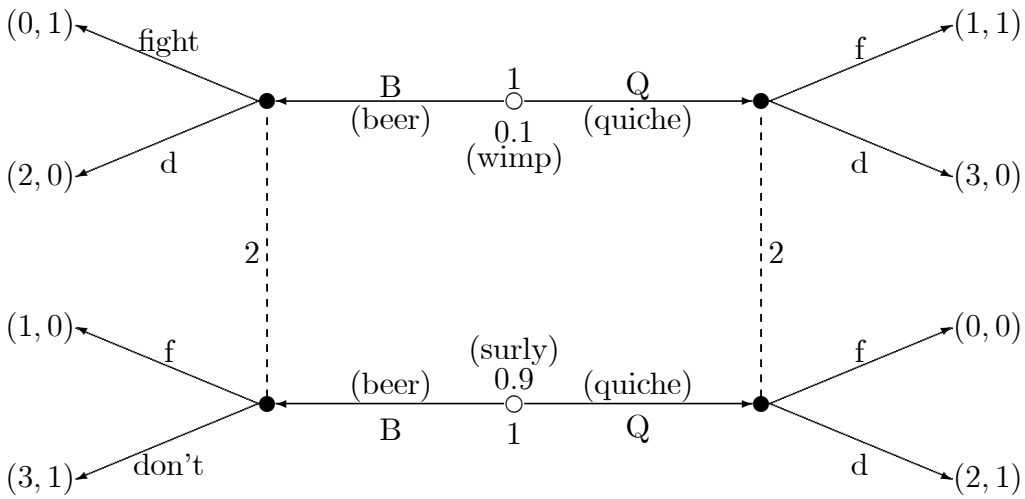


Figure 1

1.1 An illustrative example

To illustrate the kind of belief restrictions and forward-induction reasoning analyzed in this paper we refer to the well known Beer-Quiche game depicted in Figure 1 (see Cho & Kreps [24, 1987]). We will consider two different scenarios that can be formally distinguished

within our framework. (1) In the *complete-information scenario,* which may be easily implemented in the lab, a ball is drawn from an urn with 90 balls labeled *surly* (s) and 10 balls labeled *wimp* (w); Player 1 observes the label, Player 2 does not, and the game is common knowledge. We will make two very natural assumptions regarding the belief restrictions that are transparent in this scenario: Players 1 and 2 *ex ante* assign probability 0.9 to s; furthermore, Player 1 initially believes that the chance move and the strategy of Player 2 are independent (uncorrelated). Although natural, this second restriction has to be explicitly assumed as well, because subjective beliefs about the strategies of different co-players (including chance) may exhibit correlation even though players are aware that there is no causal dependence between such strategies. (2) In the *incomplete-information scenario* there is no *ex ante* stage and there are are no chance moves: Player 1 just knows his true payoff type, s or w, which is unknown to Player 2. It is assumed to be transparent that Player 2 assigns probability 0.9 to the surly type s. But, as we explain below, in this scenario we cannot even state an assumption about first-order beliefs that corresponds to the independence assumption stated for scenario (1).

Why do we distinguish between these two scenarios? Didn't Harsanyi (1967-68) show that they are equivalent? Harsanyi proved that we can find an isomorphism between the game with chance moves (1) and the incomplete-information game (2) so that the Nash equilibria of (1) correspond to the Bayesian Nash equilibria of (2). So, if one is only interested in Nash equilibrium analysis, the differences between (1) and (2) can be overlooked. But in this paper we do not presume that players' choices and beliefs form an equilibrium, and since (1) and (2) are objectively different situations, no equivalence can be taken for granted without first proving an appropriate theorem. For example, the independence restriction assumed within (1) cannot even be stated as such in scenario (2), where there is **no ex ante stage** and hence no prior belief of Player 1 that may have the independence property. The "corresponding" independence assumption that can be expressed within scenario (2) refers to a property of the **second-order** beliefs of **Player 2**: he believes that Player 1's first-order belief about 2's strategy is independent of the payoff state. Thus, these two independence properties are different and they are not necessarily equally plausible. Furthermore, they can be properly compared only by using a formal language that can distinguish between (1) and (2). Such a language is also necessary to provide transparent results of equivalence (according to a well defined criterion) between (1) and (2), although proving such equivalences is not a goal of this paper.[1] That said, in this introduction we provide an informal analysis of forward-induction reasoning in the Beer-Quiche game focusing on the **complete-information** scenario (1). In the Appendix we provide a formal analysis under the two alternative scenarios, assuming different restrictions on first-order beliefs.

Step 1: Player 1 forgoes his preferred meal (which depends on the chance move he observed) only if he thinks that this minimizes the probability of a fight. Independence implies that what he believes about the meal that minimizes the probability of a fight does not depend on the chance move he observes. Hence, for at least one realization of the chance move, the preferred meal is the same as the fight-probability minimizing meal and Player 1 chooses it. In other words, strategy "Q if s, B if w" is irrational under

[1] For a result of this sort see, for example, Battigalli *et al.* [16, 2011, Theorem 3]): absent independent restrictions, rationality and common belief in rationality have equivalent behavioral implications in the two scenarios.

the independence restriction. Player 2 does not want to fight a surly type. Since s is ex ante more likely than w, the strategy "always f" is dominated by the strategy "always d" and hence ruled out. **Step 2:** Suppose Player 2 observes B. By forward induction, he rationalizes this move and, by Step 1, he must believe that Player 1 is playing either "always B" or "Q if and only if w". Hence, his conditional probability of s is between 0.9 (the prior) and 1, which are the two extremes that would result if he were certain of the first, or – respectively – the second strategy. With this, the best reply to B is d. **Step 3:** By Step 2, after s Player 1 chooses B, i.e., the strategy "always Q" is deleted. **Step 4:** By forward induction, **even if he were surprised**, Player 2 would rationalize Q in a way consistent with the previous steps and would conclude that Player 1 is using the strategy "Q if and only if w" and that the state is w. Thus he would fight after Q. **Step 5:** By the previous steps, Player 1 is certain that Player 2 fights if and only if he chooses Q; therefore he plays the best response strategy "always B".[2]

As we show in the Appendix, we can obtain the (B, d) outcome also in the incomplete information scenario (2) under the assumption that, upon observing B (beer), Player 2 thinks that the surly type is more likely, a restriction on first-order conditional beliefs. The same conclusion follows if instead we assume that Player 2's second-order beliefs satisfy the independence assumption stated above. But in this paper we focus on first-order beliefs restrictions.

1.2 Transparent restrictions on beliefs and epistemic game theory

As illustrated by the example, in game theory and economics it is common to assume that certain restrictions on players' beliefs are *transparent*, or more precisely that such restrictions hold and there is common (probability one) belief at every point of the game that they indeed hold.[3] Games with chance moves are an obvious case in point. But an even more prominent example is the use of type structures *à la* Harsanyi [29, 1967-68] to model games with incomplete information as Bayesian games. A type structure in the sense of Harsanyi corresponds to a belief-closed subset of states within the canonical universal structure that contains all conceivable hierarchies of beliefs based on some given set of states of nature Θ;[4] the first-order beliefs in a hierarchy of Player i are elements of $\Delta(\Theta)$, the second-order beliefs are elements of $\Delta\left(\Theta \times (\Delta(\Theta))^{I\setminus\{i\}}\right)$, and so on (see Mertens & Zamir [30, 1985] and Brandenburger & Dekel [19, 1993]). In the analysis of dynamic Bayesian games, it is implicitly and informally assumed to be transparent to the players that initial beliefs about the state of nature and initial beliefs about the exogenous beliefs of others belong to this subset. (In most applied analyses, such transparent restrictions are derived from information partitions and priors.) The reason why such

[2]The strategies and conditional beliefs pinned down by this argument concide with the sequential equilibrium selected by the Intuitive Criterion of Cho and Kreps [24, 1987].

[3]We refrain from saying that such restrictions are "common knowledge". We find the use of the "common knowledge" terminology much too casual in economic theory, as there is often either a terminological or even a conceptual conflation of (common) knowledge and (common, probability one) belief. We find it semantically and conceptually useful to reserve "knowledge" for the justified true belief that comes from observation and logical deduction.

[4]We distinguish between states of *nature*, that parametrize payoff functions, and states of the *world* that describe every relevant aspect of the situation of strategic interaction.

assumption is not only implicit, but also informal, is that the language of Θ-based hierarchies of beliefs is insufficient to formally describe what players think about the state of nature and about each other conditional on some moves in the game. Such conditional beliefs should be derived, whenever possible, from initial beliefs about Θ and about the **behavior** and beliefs of others *via* Bayesian updating. Battigalli & Siniscalchi [10, 1999] show how to construct a so called *canonical* space containing all the conceivable *hierarchies of conditional probability systems* (CPS's) whereby initial first-order beliefs are probability measures on states of nature **and strategies**. Such hierarchies of CPS's provide a language that is sufficiently expressive to formally state assumptions about the transparency of restrictions on beliefs. As an added bonus, this language also allows to express assumptions about how players' mutual beliefs in rationality would evolve for each possible play of the game, and then derive implications about behavior, beliefs about behavior, and so on. The analysis of players' mutual beliefs in rationality by means of hierarchical beliefs about strategies is the goal of *epistemic game theory*, which goes back to the work of Aumann [2, 1987], Brandenburger & Dekel [18, 1987], and Tan & Werlang [36, 1988] for the analysis of static games with complete information. Indeed, the main motivation of [10, 1999] was to extend epistemic game theory to analyze dynamic games, with either complete, or incomplete information. In this paper we focus on assumptions about rationality and hierarchical beliefs whose behavioral implications are captured by versions of the rationalizability solution concept.

1.3 Rationalizability in dynamic games with incomplete, or asymmetric information

It has been argued that analyzing a strategic situation with incomplete information by means of type structures *à la* Harsanyi and looking at the corresponding Bayesian equilibria may be problematic, and that alternative approaches are worth exploring. Battigalli [5, 1999], [6, 2003], and Battigalli & Siniscalchi [12, 2003] propose to use instead a solution concept called Δ-*rationalizability*, a kind of extensive-form-rationalizability[5] deletion procedure for "games with payoff uncertainty". According to this procedure, players' initial and/or conditional beliefs about the true state of nature and their opponents' strategies satisfy some given restrictions Δ, which provide the backdrop for rationalizing observed moves by the opponents. In other words, in non trivial sequential games with asymmetric information this solution concept captures forward-induction reasoning under the assumption that the first-order belief restrictions Δ are in some sense "transparent".[6,7] This approach encompasses the case of games with objective probabilities of

[5] On extensive-form rationalizability see [33, 1984], [3, 1996] and [4, 1997].

[6] [5, 1999], [6, 2003] and [13, 2007] also analyze a less demanding solution concept that only requires **initial** common belief in rationality and the restrictions Δ. To stress the difference between these two solution concepts, these papers call "weak Δ-rationalizability" the one based on assumptions about initial beliefs, and "strong Δ-rationalizability" the one capturing forward-induction reasoning. Like other papers that only analyze the latter solution concept (e.g., [12, 2003] and [9, 2012]), here we simply call it "Δ-rationalizability".

[7] For applications of the Δ-rationalizability approach to economic models see, e.g., [6, 2003], [12, 2003], [7, 2006] and references therein; for applications to robust mechanism design see [35, 2012], [31, 2012] and the survey [17, 2011]; for applications to non-binding agreeements see [23, 2012] and [37, 2011]; for empirical applications see [1, 2008]. Attention is restricted to first-order beliefs for the sake of simplicity,

chance moves, which may be asymmetrically observed, as well as incomplete-information games with hierarchical beliefs implicitly generated by (common or subjective) priors and information partitions on a set of states of nature Θ. But the Δ-rationalizability approach is broader in that it also allows to assume that certain restrictions on beliefs about behavior are made transparent by the context in which the game is played (cf. Battigalli & Friedenberg [8, 2009], [9, 2012]). In an asymmetric-information framework, joint restrictions on beliefs about states of nature, or chance moves, and beliefs about opponents' behavior may also be important. One such restriction – expressible within the complete/asymmetric information framework – is that an individual regards the "strategy" of the chance/nature player as independent of the opponents' strategies (as in the previous illustrative example), which has been shown to be a crucial assumption underlying the traditional rationalizability concept for games with asymmetric information (see Battigalli *et al.* [16, 2011]).

1.4 Epistemic analysis and strong belief

In our view, one can really understand a solution concept $\mathcal{S}$ that is meant to capture some kind of strategic reasoning only if $\mathcal{S}$ can be justified as the result of expressible assumptions about players' rationality and hierarchical beliefs.[8] In other words, solution concept $\mathcal{S}$ should characterize the behavioral consequences of the underlying *epistemic* assumptions.[9] Such results are central to the epistemic game theory program. Here we show that Δ-rationalizability characterizes forward-induction reasoning under transparency of the given first-order restrictions Δ. The rest of this introduction discusses such characterization.

It is routine to extend the results of [18, 1987], [36, 1988] and show that, in **static games**, Δ-rationalizability is justified by the following expressible assumptions:

(a) players are rational and there is correct and common belief in the restrictions Δ (i.e., the restrictions Δ are transparent),

(b) there is common belief of (a).

By the conjunction property of probability-one belief, this in turn is equivalent to the following expressible assumptions:

(a') players are rational and their beliefs satisfy the restrictions Δ,

(b') there is common belief of (a').

These assumptions are represented by events in the canonical type structure containing all the conceivable belief hierarchies generated by the primitive uncertainty about states

as Δ-rationalizability is meant to be a relatively simple reduction procedure whose implementation does not involve type structures and beliefs about beliefs. But restrictions on higher-order beliefs (not implied by the transparency of first-order beliefs restrictions) may well be appropriate in some applications; taking higher-order belief restrictions into account does not change the essential features of the approach. See, for example, the rationalizability analysis of Spence's model in [7, 2006]. Also, Theorem 4 (a straightforward extension of results due to Battigalli & Friedenberg [9, 2012]) provides a kind of equivalence between forward-induction reasoning under transparency of first-order and of higher-order beliefs.

[8] Informally, we call "expressible" an assumption that can be stated using primitive terms and terms derived from primitive terms or other derived terms (see [16, 2011]).

[9] More formally, we say that a collection of expressible assumptions $\mathcal{A}$ *justify* a solution concept $\mathcal{S}$, or equivalently that $\mathcal{S}$ *characterizes* the behavioral implications of $\mathcal{A}$, if for each player i and each piece of private information θ_i, the set of strategies allowed by $\mathcal{A}$ for θ_i coincides with the set of strategies allowed by $\mathcal{S}$ for θ_i.

of nature and actions (see [16, 2011]).

The epistemic justification of Δ-rationalizability in dynamic games is more complex, and also more interesting. As we hinted above, this solution concept captures a forward-induction principle. Loosely stated, Δ-rationalizability rests on the assumption that each player i at each information set h ascribes to his opponents the "highest degree of strategic sophistication" consistent with h given Δ (cf. Battigalli [3, 1996]). To illustrate, in the Beer-Quiche example analyzed above, Player 2 interprets Q as evidence that Player 1 observed w, even though, according to the solution, Q should not be played at all. Building on [10, 1999] and [11, 2002], Battigalli & Siniscalchi [13, 2007] provide a rigorous formalization of this principle based on the notion of "strong belief". Let Ω be the canonical state space constructed in Battigalli & Siniscalchi [10, 1999]. Each $\omega \in \Omega$ specifies a state of nature (and players' private information about it), a strategy profile and a profile of hierarchies of conditional beliefs. Furthermore, no hierarchy satisfying standard coherency conditions is ruled out. Say that player i *strongly believes* an event E if i believes E with probability one *whenever his information set does not contradict* E. It is important to note that a player may strongly believe the conjunction of two events $E \cap F$ even if he does not strongly believe either E or F.[10] For example, suppose $E \cap F$ is the event "Eve and Frank go right", and David strongly believes $E \cap F$. Also assume that one of David's information sets corresponds to "either Eve or Frank did not go right", i.e. $\neg E \cup \neg F = \neg(E \cap F)$. If David observes $\neg E \cup \neg F$, he cannot believe (with probability one) both E and F, thus, he must give up on at least one of them. Suppose he gives up on E, i.e. he does not assign probability one to E conditional on $\neg E \cup \neg F$. Then David does not strongly believe E because he does not assign probability one to E conditional on an observation, $(\neg E \cup \neg F)$, that is consistent with E. By a similar argument, strong belief fails monotonicity: even if $E \subseteq F$, strong belief in E does not imply strong belief in F.[11] Since strong belief does not satisfy the standard conjunction and monotonicity properties, assumptions involving strong belief in multiple events must be stated and analyzed with care.

Say that there is *common strong belief* in event E at state ω if the following assumptions hold at ω:

A_E^1, all players strongly believe E

A_E^2, all players strongly believe $E \cap A_E^1$

...

A_E^{n+1}, all players strongly believe $E \cap A_E^1 \cap ... \cap A_E^n$

etc.

Battigalli & Siniscalchi [13, 2007] proved the following extension of the static-game epistemic justification of Δ-rationalizability given by assumptions (a')-(b'): a strategy profile s is strongly Δ-rationalizable if and only if s is played at some state ω where

(α') players are rational (viz., conditional expected utility maximizers) and the restrictions Δ hold,

(β') there is common strong belief in (α').

[10] On the other hand, the conjunction of strong belief in E and strong belief in F implies strong belief in $E \cap F$.

[11] Indeed, considering the same events and informational setup as above, $E \cap F \subseteq E, F$, but we have just shown that strong belief in $E \cap F$ does not imply strong belief in both E and F. Thus, monotonicity does not hold.

The previous result assumes that the restrictions Δ hold, but it does **not** assume that they are **transparent**. Consider, for example, the first-order belief restrictions Δ for the Beer-Quiche game under the complete information scenario. Since s (surly) is a realization of a chance move given by a commonly known urn, it seems plausible to assume that it is transparent that each player initially assigns probability 0.9 to s. Now, according to epistemic assumption (β'), there must be common belief **at the beginning of the game** that each player initially assigns probability 0.9 to s, but such common belief about initial first-order beliefs does not necessarily hold after a move by Player 1. Furthermore, (β') does not require Player 1 to believe that such common belief would persist. Indeed (β') allows Player 1 to believe that if he made a move that Player 2 cannot "rationalize", then Player 2 would stop believing that 1 believes that 2 initially assigns probability 0.9 to s. Does this matter? Would the behavioral implications change if the restrictions Δ were assumed to be transparent? Answering this question requires careful analysis.

1.5 Our contribution

In this paper we extend the static-game epistemic analysis given by assumptions (a)-(b) above: a strategy profile s is Δ-rationalizable if and only if s is played at some state of the world ω where

(α) players are rational and there is correct common belief in Δ at every node (i.e., Δ is transparent),

(β) there is common strong belief in (α).

We give the following interpretation. Suppose that due to some pre-game history such as public information about the composition of an urn, shared experience, or a social convention, the restrictions Δ are transparent. This provides a backdrop, or *context,* for forward-induction reasoning as expressed by the assumptions (α)-(β) stated above: In other words, while making inferences about the opponents by rationalizing their past moves, players never consider states of the world inconsistent with the transparency of Δ. We show that Δ-rationalizability allows to derive the behavioral consequences of these assumptions when the analyst knows Δ, for example because Δ reflects objective probabilities made transparent in an experimental setting, or Δ is given by some commonly known statistics. Our result also allows to answer the previous question: assumptions (α')-(β'), whereby restrictions Δ are not transparent, have the same behavioral implications of (α)-(β), as both are characterized by Δ-rationalizability.

Besides Δ-rationalizability, we also consider a "naïve" Δ-rationalizability algorithm used in [12, 2003], which is more similar in spirit to Pearce's [33, 1984] original definition of rationalizability for extensive-form games. We show that this algorithm is not conceptually correct for arbitrary restrictions Δ, but it is equivalent to Δ-rationalizability (hence it does capture the epistemic assumptions it was meant to capture) if Δ satisfies a property of "closedness under composition". We argue that this property holds in a wide range of interesting cases. Users of Δ-rationalizability mostly cite [12, 2003] as a reference for this concept. It is therefore important for them to be aware that the "naïve" solution algorithm defined there is sound only when it is equivalent to the Δ-rationalizability solution concept originally defined in [5, 1999] and epistemically justified here.

A minor, but in our view non-negligeble contribution of this paper is that we consider very general extensive-game forms, allowing for several players, imperfectly observed ac-

tions, simultaneous moves by subsets of players, chance moves and lack of common knowledge of the payoff functions, whereas all the papers mentioned above make simplifying assumptions in some of these dimensions. We do not make too much of this, as it is quite clear that in the cited literature such assumptions are made mainly for notational convenience. But we point out that our analysis, besides confirming some claims of possible generalizations made in those papers, allows to represent formally and to compare the complete- *vs* incomplete-information interpretations of well known models and examples of the literature, such as the Beer-Quiche game analyzed above. As argued in the discussion of that example, this helps shed light on the differences between the complete-information scenario with asymmetric information about an initial chance move, and the incomplete information scenario. In particular, our framework and results help to better compare *ex ante* and *interim* rationalizability in dynamic Bayesian games, adding to the static analysis of Battigalli *et al.* [16, 2011].

The rest of the paper is structured as follows: Section 2 gives the preliminary concepts about dynamic games with asymmetric information and interactive conditional beliefs; Section 3 provides the epistemic justification of Δ-rationalizability and compares it with the "naïve" reduction algorithm of [12, 2003]; Section 4 concludes with a discussion of the case when the transparent restrictions on beliefs are not given to the analyst; the Appendix provides a formal analysis of the Beer-Quiche game and collects all the proofs.

2 The framework

In this section we present the building blocks of our analysis: dynamic games with payoff uncertainty (subsection 2.1), systems of conditional probabilities (subsection 2.2), type structures (subsection 2.3), rationality (subsection 2.4), and belief operators, i.e., the language we need to express our assumptions about interactive beliefs (subsection 2.5).

2.1 Dynamic Games with Payoff Uncertainty

We consider dynamic games allowing simultaneous moves, imperfect information about past moves and lack of common knowledge of payoff functions, i.e. incomplete information. Since our formal representation is not standard, we provide a detailed description.

We describe the rules of interaction of the dynamic game under consideration using the following primitive objects:[12]

- A set I of players, plus the chance pseudo Player 0. We let $I_0 = \{0\} \cup I$ denote the extended players' set.
- For each $i \in I_0$, a set of *actions* A_i; for each non-empty subset $J \subseteq I_0$, we let $A_J = \prod_{i \in J} A_i$ denote the set of action profiles for players in J.

[12] Cf. Osborne-Rubinstein [32, 1994], chapters 6, 11.

- A **finite** set of *histories*[13]

$$X \subseteq \left(\bigcup_{\emptyset \neq J \subseteq I_0} A_J \right)^{<\mathbb{N}},$$

that is, sequences of action profiles $x = (a^1, ..., a^\ell)$ with $a^k = (a_i^k)_{i \in J}$ for some non-empty $J \subseteq I_0$.

- A *player correspondence* $\iota : X \to 2^{I_0}$ such that $(a^1, ..., a^{\ell-1}, a^\ell) \in X$ if and only if $\iota(a^1, ..., a^{\ell-1}) \neq \emptyset$ and $a^\ell \in A_{\iota(a^1,...,a^{\ell-1})}$; $\iota(x)$ is the set of *active* players at x.

- For each player $i \in I$, an *information partition* $\bar{H}_i$ of the set $\{x : i \in \iota(x)\}$ of histories at which i is active.

By assumption, every prefix of a history in X (including the *empty history* $\varnothing$) is also in X, thus X endowed with the precedence "prefix-of" relation is a **tree with root** $\varnothing$. The set $\{a_{\iota(x)} : (x, a_{\iota(x)}) \in X\}$ of feasible action profiles at history x is – by assumption – a Cartesian **product** of $|\iota(x)|$ subsets of actions. The information partition $\bar{H}_i$ is such that, for every $h \in \bar{H}_i$ and $x, y \in h$, player i has the same set of feasible actions at x and y; furthermore $\bar{H}_i$ satisfies **perfect recall**.[14] As a matter of notation, we let $\bar{H}_0$ be the finest partition of $\{x \in X : 0 \in \iota(x)\}$. Auxiliary notation and definitions, plus two key additional objects are compiled in the table below; comments will follow. Additional, derived objects (such as "strategies") will be introduced later as needed.

Name	**Notation**	**Definition**
Terminal histories	Z	$\{z \in X : \iota(z) = \emptyset\}$
Actions at h for $i \in \iota(h)$	$A_i(h)$	$\text{proj}_{A_i}\{a_{\iota(x)} : (x, a_{\iota(x)}) \in X\}$ $(x \in h)$
Strict precedence relation	$\prec$	$x \prec y \Leftrightarrow [x \text{ strict prefix of } y]$
Weak precedence relation	$\preceq$	$x \preceq y \Leftrightarrow [x \text{ weak prefix of } y]$
States of nature	Θ	$\Theta = \Theta_0 \times \Theta_1 \times ... \times \Theta_{\lvert I \rvert}$
Payoff function of $i \in I$	u_i	$u_i : \Theta \times Z \to \mathbb{R}$

Note that the information structure defined above **only describes information about past moves**, including chance moves. If each information set h is a singleton, we say that there are *observed actions*. If – on top of this – only one player is active at each history we say that the game has *perfect information*.[15]

In order to model payoff uncertainty, i.e. incomplete information about payoff functions, we introduce a nonempty and **finite** product set of "conceivable" parameter values Θ and the parametrized payoff functions $u_i : \Theta \times Z \to \mathbb{R}$ $(i \in I)$. For every player $i \in I$, each element $\theta_i \in \Theta_i$ represents Player i's private information about the unknown aspects of the game; we call it Player i's *information type*. The set Θ_0 (the "information type of

[13] For any given set Y, $Y^{<\mathbb{N}}$ denotes the set of finite sequences of elements of Y, including the *empty sequence* $\varnothing$, that is, $Y^{<\mathbb{N}} = \bigcup_{n \in \mathbb{N} \cup \{0\}} Y^n$ with $Y^0 = \{\varnothing\}$.

[14] No information set $h \in \bar{H}_i$ contains two ordered histories; furthermore, whenever $y', y'' \in h \in \bar{H}_i$, $x' \in \hat{h} \in \bar{H}_i$ and $(x', a') \preceq y'$, there is a history (x'', a'') such that $x'' \in \hat{h}$, $(x'', a'') \preceq y''$ and $a_i' = a_i''$.

[15] This is not to be confused with "complete information", which means that all the rules of the game and players' preferences over consequences are common knowledge. Indeed we allow for the opposite, if there is payoff uncertainty, there is incomplete information.

Player 0") represents any residual uncertainty about payoffs that remains after pooling players' private information. We often refer to profile $\theta = (\theta_i)_{i \in I_0}$ as the *state of nature*. Appending the states of nature to the game tree specified above we obtain an *arborescence*, that is, a collection of trees $(\Theta \times X, \prec)$ where $(\theta, x) \prec (\theta', x')$ if and only if $\theta = \theta'$ and $x \prec x'$. The elements of the information partition $\bar{H}_i$ are related to, but distinct from the information sets of the graphical representation of the game as an arborescence: when information $h \in \bar{H}_i$ about past moves is combined with the information-type θ_i, we obtain the "traditional" information sets $[\theta_i, h] = \{(\theta', x') : \theta'_i = \theta_i, x' \in h\}$ in the arborescence. Also observe that, with this formulation, the actions available to players at any given history do not depend on their information-type. This restriction could be removed, at the expense of additional notation; our results do not depend upon it in a crucial way.[16]

For instance, in a pure "private-values" setting, Player i's payoff depends solely upon θ_i and the terminal history reached. At the opposite extreme, in a "common-values" environment, players' payoffs at any terminal history depend only upon θ_0; in this case, each player's information type θ_i is interpreted as payoff-irrelevant private information, which may be correlated with θ_0 according to some other player's subjective beliefs. We include "Player 0" and the residual uncertainty θ_0 in our framework (adding some notational complexity) because it is important in economic applications, and because it helps relate our work to the recent literature on rationalizability in games with incomplete information. Also, as we clarify in Example 1, moves by Player 0 are interpreted as chance moves. Our notation allows us to formally distinguish games with imperfect information about an initial chance move (such as poker) from games with incomplete information. We do not specify the probabilities of chance moves in our description of the game because they will be part of the transparent restrictions on beliefs to be introduced later.[17] One can show that under suitable continuity assumptions our analysis can be extended to games with infinite horizon and finite action sets, and to games with compact action sets in the last stage and finite action sets in previous stages.[18]

We call the structure

$$\Gamma = \left\langle I, X, \iota, \Theta_0, \left(\Theta_i, \bar{H}_i, u_i\right)_{i \in I} \right\rangle$$

described so far *game with payoff uncertainty*. It would be perfectly legitimate to call Γ "game with incomplete information",[19] but we refrain from doing so because this expression is mostly used to refer to Bayesian games. The difference between a Bayesian game and a game with payoff uncertainty is that the former specifies (implicitly) players' hierarchies of initial beliefs about the state of nature (or, at least, its payoff-relevant component). We shall introduce hierarchies of beliefs of a much richer kind later on in order to obtain a language that allows us to express assumptions about players' rationality and beliefs, and derive the behavioral implications of these assumptions.

[16] Battigalli [6, 2003] allows for type-dependent actions sets.

[17] Our framework allows for the possibility that players do not have common beliefs about the probabilities of chance moves.

[18] The construction of a canonical type structure *à la* Battigalli & Siniscalchi applies to this more general setting (see [10, 1999] and [6, 2003]). The extension of the main epistemic characterization result of this paper involves, directly or indirectly, a measurable selection argument (see [14, 2012]).

[19] Γ exhibits *complete information* if the payoff map $\theta \longmapsto (u_i(\theta, \cdot))_{i \in I}$ is constant (which is trivially true when Θ is a singleton), otherwise Γ has *incomplete information*.

Example 1 *To illustrate our notation, consider the Beer-Quiche game depicted in Figure 1 of the Introduction. As we said, our formalism allows us to distinguish two different scenarios.*
(1) In the complete-information scenario, *the nodes of Figure 1 correspond to histories: first Player 0 (Chance) chooses* $a_0 \in \{s, w\}$*, Player 1 observes this move, Player 2 does not, and the game is common knowledge. In this case,* $\Theta_i = \{\bar{\theta}_i\}$ *(a singleton) for* $i = 0, 1, 2$*,* $X \backslash Z = \{\varnothing\} \cup \{s, w\} \cup (\{s, w\} \times \{B, Q\})$*,* $Z = \{s, w\} \times \{B, Q\} \times \{f, d\}$*,* $\iota(\varnothing) = \{0\}$*,* $A_0(\varnothing) = \{s, w\}$*,* $\bar{H}_1 = \{\{s\}, \{w\}\}$*,* $\bar{H}_2 = \{\{(s, B), (w, B)\}, \{(s, Q), (w, Q)\}\}$ *and the rest is obvious.*
(2) In the incomplete-information *scenario, the nodes of Figure 1 correspond to state-of-nature/history pairs* (θ, x)*: Player 1 starts the game knowing his true payoff type,* s *or* w*, the pseudo-player 0 can be omitted, and we have:* $\Theta_2 = \{\bar{\theta}_2\}$ *(a singleton),* $\Theta_1 = \{s, w\}$*,* $X \backslash Z = \{\varnothing\} \cup \{B, Q\}$*,* $Z = \{B, Q\} \times \{f, d\}$*,* $\bar{H}_1 = \{\{\varnothing\}\}$*,* $\bar{H}_2 = \{\{B\}, \{Q\}\}$*, and the rest is obvious.*
The probabilities of s *and* w *are not part of the description of the game. They will be described as transparent features of players' systems of beliefs.*

2.2 Conditional Probability Systems

As the game progresses, players update and/or revise their beliefs in light of newly acquired information. In order to account for this process, we represent beliefs by means of *conditional probability systems* (Renyi [34, 1955]).

Fix a player $i \in I$. For a given **compact**[20] **metrizable** topological space Y with Borel sigma-algebra $\mathcal{B}(Y)$, consider a non-empty, **countable** (finite or denumerable) collection $\mathcal{C}_i \subseteq \mathcal{B}(Y)$ of events such that $\emptyset \notin \mathcal{C}_i$. The interpretation is that Player i is uncertain about the "true" element $y \in Y$, and $\mathcal{C}_i$ is a collection of conditioning events (or "relevant hypotheses") concerning a "discrete" component of y observable by Player i.

Definition 1 *A* conditional probability system *(or CPS) on* $(Y, \mathcal{C}_i)$ *is a mapping*

$$\mu(\cdot|\cdot) : \mathcal{B}(Y) \times \mathcal{C}_i \to [0, 1]$$

satisfying the following axioms:[21]

Axiom 1 *For all* $C \in \mathcal{C}_i$*,* $\mu(C|C) = 1$*.*

Axiom 2 *For all* $C \in \mathcal{C}_i$*,* $\mu(\cdot|C)$ *is a probability measure on* Y*.*

Axiom 3 *For all* $E \in \mathcal{B}(Y)$*,* $B, C \in \mathcal{C}_i$*, if* $E \subseteq B \subseteq C$ *then* $\mu(E|B)\mu(B|C) = \mu(E|C)$*.*

The set of probability measures on Y will be denoted by $\Delta(Y)$; we shall endow it with the topology of weak convergence of measures. The set of conditional probability systems on $(Y, \mathcal{C}_i)$ can be regarded as a subset of $[\Delta(Y)]^{\mathcal{C}_i}$ endowed with the product topology, a compact metrizable space.

[20] Compactness of the relevant spaces is assumed for simplicity, it can be relaxed with some additional technical complications.

[21] $(\Omega, \mathcal{A}, \mathcal{C}_i, \mu)$ is called *conditional probability space* in [34, 1955].

Throughout this paper, we shall be interested solely in "relevant hypotheses" corresponding to the event that a certain information set h has occurred, plus the empty information corresponding to the beginning of the game. Each player is uncertain about the combination of information-types and strategies of his opponents, including chance. Beliefs about such primitive uncertainty are called *first-order beliefs*. The "relevant hypotheses" correspond to the sets of type-strategy pairs consistent with information sets. To make this formal, we introduce additional notation and definitions of derived objects, summarized in the following table:

Name	**Notation**	**Definition**
Extended inform. structure of i	H_i	$\overline{H}_i \cup \{\varnothing\}$
Opponents of i (with chance)	$-i$	$I_0 \backslash \{i\}$
Strategies of $i \in I_0$	S_i	$\prod_{h \in \bar{H}_i} A_i(h)$
Strategy profiles (of $-i$)	S (S_{-i})	$\prod_{i \in I_0} S_i$ $(\prod_{j \in I_0 \backslash \{i\}} S_j)$
Outcome function	$\zeta : S \to Z$	$\zeta(s)$ terminal history given by s
Strat. profiles consistent with h	$S(h)$	$\{s \in S : (\exists x \in h)(x \prec \zeta(s))\}$
Strategies of i $(-i)$ *cons. with* h	$S_i(h)$ $(S_{-i}(h))$	$\text{proj}_{S_i} S(h)$ $(\text{proj}_{S_{-i}} S(h))$
Type-strategy pairs	Σ_i	$\Theta_i \times S_i$
Profiles of $(-i)$ *type-strat. pairs*	Σ (Σ_{-i})	$\prod_{i \in I_0} (\Theta_i \times S_i)$ $(\prod_{j \in I_0 \backslash \{i\}} (\Theta_j \times S_j))$
Type-strategy pairs cons. with h	$\Sigma_i(h)$	$\Theta_i \times S_i(h)$
Profiles of $(-i)$ *pairs cons. with* h	$\Sigma(h)$ $(\Sigma_{-i}(h))$	$\prod_{i \in I_0} \Sigma(h)$ $(\prod_{j \in I_0 \backslash \{i\}} \Sigma_j(h))$

Perfect recall implies that the extended information structure H_i (endowed with the obvious precedence relation derived from $\prec$)[22] is a tree with root $\{\varnothing\}$. By assumption, if $i \in \iota(\varnothing)$ then $\{\varnothing\} \in \bar{H}_i = H_i$.[23] The sets $S(h)$ provide a strategic-form representation of the information structure of each player. Perfect recall implies that if $h \in H_i$ then $S(h)$ and $\Sigma(h)$ can be factored as $S(h) = S_i(h) \times S_{-i}(h)$ and $\Sigma(h) = \Sigma_i(h) \times \Sigma_{-i}(h)$. To illustrate, in the incomplete-information version of the Beer-Quiche game in part (2) of Example 1, we have $\Sigma_1 = \{s, w\} \times \{B, Q\}$, $\Sigma_1(\{B\}) = \{s, w\} \times \{B\}$, $\Sigma_1(\{Q\}) = \{s, w\} \times \{Q\}$; these subsets correspond to the conditioning events for Player 2. An information-type/strategy pair in Σ_i is denoted σ_i whenever we do not have to refer specifically to the information component θ_i or the strategy component s_i.

Player i's first-order beliefs about the state of nature and his opponents' behavior may be represented by taking probability measures in $\Delta(\Sigma_{-i})$ and letting $\mathcal{C}_i = \{C \subseteq \Sigma_{-i} : (\exists h \in H_i)(C = \Sigma_{-i}(h))\}$.[24] Since $\mathcal{C}_i$ is indexed by H_i, we denote the collection of CPS's on $(\Sigma_{-i}, \mathcal{C}_i)$ thus defined by $\Delta^{H_i}(\Sigma_{-i})$. Since Σ_{-i} and H_i are finite, $\Delta^{H_i}(\Sigma_{-i})$ is easily seen to be a closed subset of the Euclidean $|H_i| \times |\Sigma_{-i}|$-space.

[22] That is, $(h \prec h') \Leftrightarrow ((\forall x' \in h')(\exists x \in h)(x \prec x'))$.

[23] If the information set of i containing $\varnothing$ also contained another node, then it would contain two nodes on the same path, thus violating perfect recall.

[24] If two information sets $h, h' \in H_i$ differ only because of moves of i, then $\Sigma_{-i}(h) = \Sigma_{-i}(h')$. Thus, the cardinality of H_i may be smaller than the cardinality of $\mathcal{C}_i$: $|H_i| \leq \mathcal{C}_i$. This redundancy is innocuous in our analysis.

2.3 Type Structures

Next we introduce our representation of hierarchical conditional beliefs (see Battigalli & Siniscalchi [10, 1999], [11, 2002]). To represent Player i's *higher-order beliefs*, we consider a compact metrizable set of "possible worlds" $\Omega = \prod_{i \in I_0} \Omega_i$, where $\Omega_i \subseteq \Sigma_i \times T_i$, $\text{proj}_{\Sigma_i} \Omega_i = \Sigma_i$, and $T_0 = \{t_0\}$ is a singleton that we introduce for notational convenience. Elements of the compact metrizable spaces T_i will be interpreted as "epistemic types" of $i \in I$; elements of Ω_i are "states of i" comprising his information type θ_i, strategy s_i and epistemic type t_i. Condition $\text{proj}_{\Sigma_i} \Omega_i = \Sigma_i$ means that every $\sigma_i \in \Sigma_i$ is possible at some state, but we allow for the possibility that some $(\theta_i, s_i, t_i) \in \Sigma_i \times T_i$ does not belong to Ω_i because **the possible beliefs of i about the opponents may depend on his information type**. Consistently with our previous notation, we let $T_{-i} = \prod_{j \in I_0 \setminus \{i\}} T_j$ and $\Omega_{-i} = \prod_{j \in I_0 \setminus \{i\}} \Omega_j$. We let $\Omega_{-i}(h)$ (resp. $\Omega(h)$) denote the event in Ω_{-i} (resp. Ω) that corresponds to information set h, that is, $\Omega_{-i}(h) = \prod_{j \in I_0 \setminus \{i\}} \{(\theta_j, s_j, t_j) \in \Omega_j : s_j \in S_j(h)\}$. Of course, $\Omega_{-i}(\{\varnothing\}) = \Omega_{-i}$ (resp. $\Omega(\{\varnothing\}) = \Omega$). To represent Player i's conditional beliefs about his opponents (including the dummy player 0), we use the collection of observable events $\mathcal{C}_i = \{C \in \mathcal{B}(\Omega_{-i}) \ : (\exists h \in H_i)(C = \Omega_{-i}(h))\}$. The set of CPS's on $(\Omega_{-i}, \mathcal{C}_i)$ is denoted by $\Delta^{H_i}(\Omega_{-i})$. Similarly, to represent Player i's conditional beliefs about the prevailing state of the world (including his own strategy and type), we use the collection of observable events $\mathcal{C}_i = \{C \in \mathcal{B}(\Omega) : (\exists h \in H_i)(C = \Omega(h))\}$. The set of CPS's on $(\Omega, \mathcal{C}_i)$ is denoted $\Delta^{H_i}(\Omega)$. Since $\{\varnothing\} \in H_i$ and $\Sigma(\{\varnothing\}) = \Sigma_i(\{\varnothing\}) \times \Sigma_{-i}(\{\varnothing\}) = \Sigma$, each CPS $\mu \in \Delta^{H_i}(Y)$ (with $Y = \Sigma_{-i}, \Omega_{-i}, \Omega$) contains an "initial" belief $\mu(\cdot|Y)$.

Note that the finite collections of observable events defined above consist of sets that are both **open and closed** in the respective topologies. Battigalli & Siniscalchi [10, 1999] show that, under these conditions, $\Delta^{H_i}(\Omega_{-i})$ and $\Delta^{H_i}(\Omega)$ are closed subsets of the compact metrizable spaces $[\Delta(\Omega_{-i})]^{H_i}$ and, respectively, $[\Delta(\Omega)]^{H_i}$. Hence, they are compact metrizable in the relative topology.

With this, we provide an **implicit** representation of hierarchies beliefs by means of type structures, as is standard in the literature.

Definition 2 *A Γ-based* type structure *is a tuple $\mathcal{T} = (\Omega_0, (\Omega_i, T_i, g_i)_{i \in I})$ such that, $\Omega_0 = \Sigma_0 \times \{t_0\}$ and, for every $i \in I$, T_i is compact metrizable, $\Omega_i \subseteq \Sigma_i \times T_i$ is closed with $proj_{\Sigma_i} \Omega_i = \Sigma_i$, and $g_i = (g_{i,h})_{h \in H_i} : T_i \to \Delta^{H_i}(\Omega_{-i})$ is a continuous mapping given by coordinate mappings $g_{i,h} : T_i \to \Delta(\Sigma_{-i} \times T_{-i})$ such that $g_{i,h}(t_i)(\Omega_{-i}(h)) = 1$ $(h \in H_i)$. The elements of each set T_i are called* epistemic types. *A Γ-based type structure $\mathcal{T} = (T_i, g_i)_{i \in I}$ is* belief-complete *if, for each $i \in I$, $\Omega_i = \Sigma_i \times T_i$ and $g_i : T_i \to \Delta^{H_i}(\Sigma_{-i} \times T_{-i})$ is **onto**.*

Thus, for every possible world $\omega = ((\theta_0, s_0), (\theta_i, s_i, t_i)_{i \in I}) \in \Omega$, we specify a state of nature $(\theta_0, (\theta_i)_{i \in I})$, as well as each player i's *dispositions to act* (his strategy s_i) and (for real players $i \in I$) his *dispositions to believe* (his system of conditional probabilities $g_i(t_i) = (g_{i,h}(t_i))_{h \in H_i}$). These dispositions also include what a player would do and think at histories that are inconsistent with ω.[25] Type structures encode a collection of infinite hierarchies of CPS's, one for each epistemic type of each player. It is natural to ask whether there exists a type structure which encodes *all* "conceivable" hierarchies of conditional beliefs. Battigalli & Siniscalchi [10, 1999] shows that such a type structure can be constructed (for all finite games, and also a large class of infinite games) by taking

[25] History h is inconsistent with (or counterfactual at) $\omega = (\theta, s, t)$ if $s \notin S(h)$.

the sets of epistemic types to be the collection of *all* possible hierarchies of conditional probability systems that satisfy certain intuitive coherency conditions.[26] This is the so called "canonical" type structure, which turns out to be belief-complete. Every type structure may be viewed as a belief-closed substructure of the canonical structure.[27] One possible interpretation of a substructure of the canonical structure is that it encodes a context that makes some belief restrictions (not necessarily first-order belief restrictions) transparent to the players. We find it conceptually appealing to express our results within the canonical structure as this forces the analyst to make it explicit in the formulas that some belief restrictions are transparent to the players. However, as in [11, 2002] and [13, 2007], our results can be stated and proved more generally for any belief-complete type structure, not only the canonical one. This is convenient because it allows us to skip the canonical construction.[28]

Finally, we derive from the type structure extended belief maps so that at each "personal state" (θ_i, s_i, t_i) player i has beliefs about events in $\mathcal{B}(\Omega)$, that is, events about himself as well as the other players. This makes our analysis more easily comparable to the literature on epistemic game theory. Specifically, we assume that for every state of the world $((\theta_i, s_i, t_i), \omega_{-i})$ and every information set $h \in H_i$, Player $i \in I$ would be certain of t_i and θ_i if h occurred, and would also be certain of s_i given h provided that s_i is consistent with h, i.e. $s_i \in S_i(h)$. We also assume that if $s_i \notin S_i(h)$ Player i would still be certain that his continuation strategy agrees with s_i. This means that s_i in state $\omega = ((\theta_i, s_i, t_i), \omega_{-i})$ represents both how Player i would choose at ω conditional any information set h and how he **plans** to play the game and to continue his play after h.

Formally, Player i's conditional beliefs on Ω are given by a continuous mapping

$$g_i^* = (g_{i,h}^*)_{h \in H_i} : \Omega_i \to \Delta^{H_i}(\Omega)$$

derived from g_i by the following formula: for every $(s_i, \theta_i, t_i) \in \Omega_i$, $h \in H_i$, $E \in \mathcal{B}(\Omega)$,

$$g_{i,h}^*(\theta_i, s_i, t_i)(E) = g_{i,h}(t_i)\left(\left\{\omega_{-i} \in \Omega_{-i} : ((\theta_i, s_i^h, t_i), \omega_{-i}) \in E\right\}\right), \tag{1}$$

where s_i^h is the unique strategy in $S_i(h)$ that coincides with s_i at each information set h' that does not strictly precede h (thus, $s_i^h = s_i$ if and only if $s_i \in S_i(h)$). In principle, players are free to choose actions that deviate from their plans. But, our definition of type structure and of the extended belief maps essentially assumes that **it is transparent that players execute their plans**, which is germane to a forward-induction analysis.[29]

2.4 Sequential Rationality

We take the view that a strategy $s_i \in S_i$ for Player i should be optimal, given Player i's beliefs and payoff-type, conditional upon any information set. Two strategies that allow

[26] Battigalli & Siniscalchi [10, 1999] uses a slightly different definition of type structure. But all the arguments in [10, 1999] can be easily adapted to the present framework.

[27] The representation of a type structure as a belief-closed substructure of the canonical one eliminates redundant types, i.e. types that yield the same hierarchy of CPS's. Redundant types do not play any role in our analysis.

[28] A result by Friedenberg [27, 2010] implies that in static games (games where $\bar{H}_i = \{\varnothing\}$ for each i) every compact-continuous complete structure contains all the "conceivable" hierarchies of beliefs and is in a precise sense equivalent to the canonical structure. It can be shown that the same holds more generally for all the dynamic games considered here (De Vito [26, 2012]).

[29] See the discussion section in Battigalli *et al* [15, 2011].

the same collection of information sets and select the same actions at such information sets are indistinguishable by the opponents (or by an external observer). Therefore, in the following definition, we impose the (continuation) optimality condition only at information sets consistent with the given strategy. But we remark that, by a well known dynamic programming argument, this is realization-equivalent to requiring optimality of the selected **action** at **each** information set, given the planned continuation. (Note that in the epistemic framework described in Section 2.3 the planned continuation of Player i at information set h is given by his conditional belief about his own behavior.)

Definition 3 *Fix a player $i \in I$, a CPS $\mu_i \in \Delta^{H_i}(\Sigma_{-i})$ and an information type $\theta_i \in \Theta_i$. A strategy $s_i \in S_i$ is a* sequential best reply to μ_i for θ_i *if and only if, for every $h \in H_i$ with $s_i \in S_i(h)$ and every $s_i' \in S_i(h)$,*

$$\sum_{\theta_{-i}, s_{-i}} [u_i(\theta_i, \theta_{-i}, \zeta(s_i, s_{-i})) - u_i(\theta_i, \theta_{-i}, \zeta(s_i', s_{-i}))]\mu_i(\{(\theta_{-i}, s_{-i})\}|\Sigma_{-i}(h)) \geq 0.$$

For each CPS $\mu_i \in \Delta^{H_i}(S_{-i})$, let $r_{\theta_i}(\mu_i)$ denote the set of sequential best replies to μ_i for θ_i, and let $\rho_i(\mu_i) = \{(\theta_i, s_i) : s_i \in r_{\theta_i}(\mu_i)\}$ denote the set of pairs (θ_i, s_i) such that s_i is a sequential best reply to μ_i for θ_i.

It can be shown by standard arguments that ρ_i is a nonempty-valued and upper-hemicontinuous correspondence (see [6, 2003]). It is convenient to introduce the following additional notation. Fix a Γ-based type structure; for every player $i \in I$, let $f_i = (f_{i,h})_{h \in H_i} : T_i \to [\Delta(\Sigma_{-i})]^{H_i}$ denote his *first-order* belief mapping, that is, for every $t_i \in T_i$ and $h \in H_i$,

$$f_{i,h}(t_i) = \text{marg}_{\Sigma_{-i}} g_{i,h}(t_i).$$

It is easy to see that $f_i(t_i) \in \Delta^{H_i}(\Sigma_{-i})$ for every $t_i \in T_i$; also, f_i is continuous.

Finally, we can introduce our key behavioral assumption. We say that Player i is *rational* at a state $\omega = (\theta, s, t)$ in $\mathcal{T}$ if and only if $(\theta_i, s_i) \in \rho_i(f_i(t_i))$. Then the event

$$R_i = \{\omega = (\theta, s, t) \in \Omega \ : \ (\theta_i, s_i) \in \rho_i(f_i(t_i))\}$$

corresponds to the statement "Player i is rational." (Note that R_i is closed because the correspondence $\rho_i \circ f_i$ is upper hemicontinuous.) We shall also refer to the events $R = \bigcap_{i \in I} R_i$ ("every player is rational") and $R_{-i} = \bigcap_{j \in I \setminus \{i\}} R_j$ ("every opponent of Player i is rational").

A word of caution Events are defined with reference to a specific type structure.

2.5 Belief Operators

The next building block is the epistemic notion of (conditional) *probability-one belief*, or (conditional) *certainty*. Recall that an epistemic type encodes the beliefs a player would hold, should any one of the possible non–terminal histories occur. This allows us to formalize statements such as, "Player i would be certain that Player j is rational, were he to observe h."

Given a Γ-based type structure $\mathcal{T}$, for every $i \in I$, $h \in H_i$, and event E, define the event[30]

$$\mathrm{B}_{i,h}(E) = \{(\sigma, t) \in \Omega \ : \ g^*_{i,h}(\sigma_i, t_i)(E) = 1\}$$

which corresponds to the statement "Player i would be certain of E, were he to observe information h." This definition incorporates the requirement that a player can only be certain of events which are consistent with her own (continuation) strategy and epistemic type (recall how g^*_i was derived from g_i in eq. (1)).

For each player $i \in I$ and history $h \in H_i$, the definition identifies a set–to–set operator $\mathrm{B}_{i,h} : \mathcal{B}(\Omega) \to \mathcal{B}(\Omega)$ which satisfies the usual properties of falsifiable beliefs; in particular, it satisfies

- **Conjunction**: For all events $E, F \in \mathcal{B}(\Omega)$, $\mathrm{B}_{i,h}(E \cap F) = \mathrm{B}_{i,h}(E) \cap \mathrm{B}_{i,h}(F)$;
- **Monotonicity**: For all events $E, F \in \mathcal{B}(\Omega)$: $E \subseteq F$ implies $\mathrm{B}_{i,h}(E) \subseteq \mathrm{B}_{i,h}(F)$.

We say that Player i *strongly believes* that an event $E \neq \emptyset$ is true (i.e., adopts E as a "working hypothesis") if and only if he is certain of E at each information set consistent with E (including the initial information set $\{\varnothing\}$). Formally, for any Γ-based type structure and every $i \in I$ define the belief operator $\mathrm{SB}_i : \mathcal{B}(\Omega) \to \mathcal{B}(\Omega)$ by $\mathrm{SB}_i(\emptyset) = \emptyset$ and

$$\mathrm{SB}_i(E) = \bigcap_{h \in H_i:\ E \cap \Omega(h) \neq \emptyset} \mathrm{B}_{i,h}(E)$$

for every event $E \in \mathcal{B}(\Omega) \backslash \{\emptyset\}$. Note that, as anticipated in the Introduction, in non-trivial games SB_i **fails** Conjunction and Monotonicity (for more on this see [11, 2002]). Sometimes we will say that a CPS μ strongly believes an event $E \neq \emptyset$ if $\mu(E|\Omega(h)) = 1$ whenever $E \cap \Omega(h) \neq \emptyset$.

We also define a *full belief* operator $\mathrm{B}_i : \mathcal{B}(\Omega) \to \mathcal{B}(\Omega)$ as follows

$$\mathrm{B}_i(E) = \bigcap_{h \in H_i} \mathrm{B}_{i,h}(E).$$

Note that belief operator $\mathrm{B}_i(\cdot)$ inherits the conjunction and monotonicity properties of its constituents $\mathrm{B}_{i,h}(\cdot)$ $(h \in H_i)$ operators.

Rationality, strong belief and full belief are the building blocks of our epistemic analysis of forward-induction reasoning with transparent restrictions on beliefs.

3 Forward-induction reasoning with transparent restrictions on beliefs

In this section, we introduce restrictions Δ on first-order beliefs (subsection 3.1). These restrictions are taken as parametrically given in the definition of Δ-rationalizability, which is shown to characterize the behavioral implications of forward-induction reasoning under transparency of Δ (subsection 3.2), and to coincide with a simpler reduction procedure when Δ satisfies a composition property (subsection 3.3).

[30] For any measurable (closed) subset $E \subseteq \Omega$, $\mathrm{B}_{i,h}(E)$ is measurable (closed).

3.1 Restrictions on Beliefs

A player's beliefs may be assumed to satisfy some restrictions that are not implied by mutual, common, or strong belief in rationality. We may distinguish between (i) *restrictions on "exogenous"* (first-order) *beliefs*, that is, beliefs about the state of nature and chance moves, and (ii) *restrictions on "endogenous"* (first-order) *beliefs*, that is, beliefs about behavior; our general theory considers both (i) and (ii). We provided examples in the Introduction; other examples of restrictions on beliefs, as well as applications and additional discussion, can be found in [12, 2003] and [6, 2003].

Formally, for every player $i \in I$ and every information type $\theta_i \in \Theta_i$, we consider a restricted set of beliefs $\Delta_{\theta_i} \subseteq \Delta^{H_i}(\Sigma_{-i})$ and we let $\Delta = (\Delta_{\theta_i})_{i \in I, \theta_i \in \Theta_i}$ denote the restrictions for all information types of all players. Whenever we talk about restrictions Δ on first-order beliefs we take for granted that Δ is a profile of **nonempty and measurable** subsets Δ_{θ_i} ($i \in I$, $\theta_i \in \Theta_i$).

Example 2 *Consider the complete-information version (1) of the Beer-Quiche game of the Introduction (see also Example 1). Since the composition of the urn is common knowledge, it is natural to assume that it is transparent that the prior probability of w is* 0.1*. Another quite natural assumption is that it is transparent that the conjecture of Player 1 about the strategy of Player 2 is* independent *of the observed realization of the chance move. Formally, we have*

$$\begin{aligned} \Delta_1 &= \{\mu_1 \in \Delta^{H_1}(S_{-1}) : \\ (\mu_1(\{w\} \times S_2 | S_0 \times S_2) &= 0.1) \wedge (marg_{S_2}\mu_1(\cdot|w) = marg_{S_2}\mu_1(\cdot|s))\}, \end{aligned}$$

$$\Delta_2 = \{\mu_2 \in \Delta^{H_2}(S_{-2}) : \mu_2(\{w\} \times S_1 | S_0 \times S_1) = 0.1\}.$$

As we made clear, the restrictions Δ only concern first-order beliefs. But in our epistemic analysis we will assume that restrictions Δ are transparent, thus yielding restrictions on higher-order beliefs. Note also that distinct information-types θ_i', θ_i'' of a player may be associated with different belief restrictions. This makes our approach sufficiently flexible to encompass the restrictions on infinite hierarchies of exogenous beliefs implicit in the Bayesian games used in applications (see the discussion of this point in [13, 2007]).

3.2 Common Strong Belief in Rationality and Transparent Restrictions on Beliefs

We adopt a uniform notation for the n-fold composition of operators on $\mathcal{B}(\Omega)$. Formally, fix a map $\mathcal{O} : \mathcal{B}(\Omega) \to \mathcal{B}(\Omega)$; then, for any event $E \in \mathcal{B}(\Omega)$, let $\mathcal{O}^0(E) = E$ and, for $n \geq 1$, let $\mathcal{O}^n(E) = \mathcal{O}(\mathcal{O}^{n-1}(E))$.

To express our epistemic assumptions, we introduce the auxiliary *correct mutual strong belief* operator $\mathrm{CSB}(\cdot)$ defined as follows: for each $E \in \mathcal{B}(\Omega)$

$$\mathrm{CSB}(E) = E \cap \bigcap_{i \in I} \mathrm{SB}_i(\Omega_i \times \mathrm{proj}_{\Omega_{-i}} E).$$

In words, given $E = \prod_{i \in I_0} E_i$, $\mathrm{CSB}(E)$ is the set of states of the world where E holds and each $i \in I$ strongly believes $E_{-i} = \Omega_i \times \prod_{j \in I_0 \setminus \{i\}} E_j$. For our epistemic analysis, we

need not consider events that are not Cartesian products. The reason is that we derive implications about behavior solely from the rationality of each player and assumptions of strong belief of events about the opponents. Our results show that step $(n+1)$ of a solution procedure characterizes the behavioral implications of epistemic assumptions expressed by means of n-th iteration of the auxiliary CSB operator. Observe that, like its constituent operators $\text{SB}_i(\cdot)$ $(i \in I)$, $\text{CSB}(\cdot)$ is not monotone. Also note that, by definition, $\text{CSB}^n(E) = \text{CSB}^{n-1}(E) \cap \text{SB}(\text{CSB}^{n-1}(E)) \subseteq \text{CSB}^{n-1}(E)$. Therefore $\{\text{CSB}^n(E)\}_{n\geq 0}$ is a decreasing sequence of events, and it makes sense to define $\text{CSB}^\infty(E) = \bigcap_{n\geq 0} \text{CSB}^n(E)$.

Similarly, we define the *mutual full belief* operator $\text{B}(\cdot)$ as

$$\text{B}(E) = \bigcap_{i\in I} \text{B}_i(\Omega_i \times \text{proj}_{\Omega_{-i}} E)$$

and we let

$$\text{B}^*(E) = \bigcap_{n\geq 0} \text{B}^n(E)$$

$$= E \cap \text{B}(E) \cap \text{B}(E \cap \text{B}(E)) \cap \text{B}(E \cap \text{B}(E) \cap \text{B}(E \cap \text{B}(E))) \cap ...,$$

where the second equality holds because $\text{B}(\cdot)$, like each $\text{B}_i(\cdot)$, satisfies conjunction. We say that event E is *transparent* at state ω if $\omega \in \text{B}^*(E)$. It follows from the conjunction property of each B_i and the continuity of probability measures that $\text{B}(\text{B}^*(E)) = \bigcap_{n\geq 1} \text{B}^n(E)$. Hence $\text{B}^*(E) = E \cap \text{B}(\text{B}^*(E)) \subseteq \text{B}(\text{B}^*(E))$. This shows that event $F = \text{B}^*(E)$ is *self-evident*, i.e., it satisfies $F \subseteq \text{B}(F)$.

It is easy to show that, whenever $E = \prod_{i\in I_0} E_i$, also $\text{CSB}^n(E)$, $\text{B}^n(E)$, $\text{CSB}^\infty(E)$ and $\text{B}^*(E)$ are Cartesian products.

Denote by $[\Delta]$ the set of states where players' first-order beliefs satisfy the restrictions given by $\Delta = (\Delta_{\theta_i})_{i\in I, \theta_i \in \Theta_i}$; that is,

$$[\Delta_i] = \{(\theta_i, s_i, t_i, \omega_{-i}) \in \Omega : f_i(t_i) \in \Delta_{\theta_i}\}, \; [\Delta] = \bigcap_{i\in I} [\Delta_i].$$

By continuity of the first-order belief function f_i, $[\Delta_i]$ is compact (hence measurable) whenever Δ_{θ_i} is compact for each θ_i. The belief-restrictions Δ are transparent at each state $\omega \in \text{B}^*([\Delta])$. In our epistemic analysis, we will assume for simplicity that $[\Delta]$ is compact.[31]

As explained above, $\text{B}^*([\Delta]) = [\Delta] \cap \text{B}(\text{B}^*([\Delta]))$; a bit more explicitly

$$\text{B}^*([\Delta]) = [\Delta] \cap \left(\bigcap_{i\in I} \bigcap_{h\in H_i} \text{B}_{i,h}(\text{B}^*([\Delta])) \right),$$

that is, transparency of the belief restrictions Δ means that such restrictions hold and are believed to be transparent by every player i conditional on each information set

[31] We can prove our main results without assuming compactness of $[\Delta]$, but we are not able to do it without complicating the analysis. Clearly, compactness of $[\Delta]$ may not hold in interesting applications. In the incomplete-information scenario of the Beer-Quiche example analyzed in the Appendix, Δ_2 is not compact. But exactly the same analysis goes through with any compact subset of Δ_2 sufficiently close to Δ_2 (that is, sufficiently close in the Hausdorff topology to the closure of Δ_2).

$h \in H_i$ even if h contradicts some event that i was previously certain of. Therefore, the transparent restrictions provide a backdrop, or context, for forward-induction reasoning: when player i tries to rationalize the observed moves of the opponents, he does so taking into account that $[\Delta]$ is transparent and he never doubts the transparency of $[\Delta]$, as in the following epistemic assumptions:

(α) players are **rational** and the **restrictions** Δ are **transparent**,
(β.1) players **strongly believe in** (α)
(β.2) players **strongly believe in** (α) **& (and)** (β.1),
...
(β.k+1) players **strongly believe in** (α) & (β.1) & ... & (β.k),
etc.

The main result of the paper is that the behavioral implications of the aforementioned epistemic assumptions are captured by the following solution concept:

Definition 4 *(See [5, 1999], [6, 2003]) Consider the following procedure.*

(Step 0) *For every $i \in \{0\} \cup I$, let $\Sigma^0_{i,\Delta} = \Sigma_i$. Also, let $\Sigma^0_{-i,\Delta} = \prod_{j \neq i} \Sigma^0_{i,\Delta}$ and $\Sigma^0_\Delta = \prod_{i \in I_0} \Sigma^0_{i,\Delta}$.*

(Step $n > 0$) *Let $\Sigma^n_{0,\Delta} = \Sigma_0$; for every $i \in I$, and for every $\sigma_i = (\theta_i, s_i) \in \Sigma_i$, let $\sigma_i \in \Sigma^n_{i,\Delta}$ if and only if there exists a CPS $\mu \in \Delta_{\theta_i}$ such that*

1. *$\sigma_i \in \rho_i(\mu)$;*
2. *for every $m \in \{1, \ldots, n-1\}$ and $h \in H_i$,*

$$\Sigma^m_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset \Rightarrow \mu(\Sigma^m_{-i,\Delta} | \Sigma_{-i}(h)) = 1.$$

Also let $\Sigma^n_{-i,\Delta} = \prod_{j \neq i} \Sigma^n_{i,\Delta}$ and $\Sigma^n_\Delta = \prod_{i \in I_0} \Sigma^n_{i,\Delta}$.

Finally, let $\Sigma^\infty_\Delta = \bigcap_{n \geq 0} \Sigma^n_\Delta$. The profiles in Σ^∞_Δ are called Δ-rationalizable.

Remark 1 *Δ-rationalizability can be more compactly defined as follows: for each $i \in I$ and $\theta_i \in \Theta_i$, let $\Delta^0_{\theta_i} = \Delta_{\theta_i}$; given $\Delta^{n-1} = \left(\Delta^{n-1}_{\theta_j}\right)_{j \in I, \theta_j \in \Theta_j}$, let $\Sigma^n_0 = \Sigma_0$, for each $i \in I$ and $\theta_i \in \Theta_i$, $S^n_{\theta_i} = r_{\theta_i}(\Delta^{n-1}_{\theta_i})$, $\Sigma^n_i = \{(\theta_i, s_i) : s_i \in S^n_{\theta_i}\}$ and*

$$\Delta^n_{\theta_i} = \{\mu_i \in \Delta^{n-1}_{\theta_i} : (\forall h \in H_i)(\Sigma^n_{-i} \cap \Sigma_{-i}(h) \neq \emptyset \Rightarrow \mu_i(\Sigma^n_{-i} | \Sigma_{-i}(h)) = 1)\}.$$

Theorem 1 *Fix a collection $\Delta = (\Delta_{i,\theta_i})_{i \in I, \theta_i \in \Theta_i}$ of compact subsets of first-order CPS's and a belief-complete type structure $\mathcal{T}$. Then, for every $n \geq 0$,*

$$\Sigma^{n+1}_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta])),$$

and

$$\Sigma^\infty_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta])).$$

This characterization result provides an epistemic justification of Δ-rationalizability that makes the transparency of $[\Delta]$ fully explicit in the strong-belief formulas. Next we present a result showing that Δ-rationalizability also characterizes the behavioral implications of assuming **rationality and common strong belief in rationality** within a restricted type structure that exactly captures the transparency of $[\Delta]$. In a sense, the restricted state space of this type structure captures the context for players' strategic reasoning, as if types for whom the restrictions are not transparent were not even conceivable (cf. [8, 2009], [9, 2012]).

Fix an arbitrary type structure $\mathcal{T}$ and (compact) first-order belief restrictions $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$. Since $\mathrm{B}^*([\Delta]) = [\Delta] \cap \mathrm{B}(\mathrm{B}^*([\Delta]))$ and $[\Delta]$ is a Cartesian event, we have

$$\mathrm{B}^*([\Delta]) = \Omega_0 \times \prod_{i\in I} \mathrm{B}_i^*([\Delta])$$

where $\mathrm{B}_i^*([\Delta])$ is the set of $(\theta_i, s_i, t_i) \in \Omega_i$ such that $f_i(t_i) \in \Delta_{i,\theta_i}$ and

$$(\forall h \in H_i)(g_{i,h}(t_i)(\{(\theta_i, s_i^h)\} \times \mathrm{B}_{-i}^*([\Delta])) = 1). \tag{2}$$

Now, for the chance pseudo-player 0 let $\Omega_{0,\Delta} = \Sigma_0 \times T_{0,\Delta} = \Sigma_0 \times \{t_0\} = \Omega_0$ for notational convenience; for each real player $i \in I$, consider the following subsets

$$\begin{aligned} \Omega_{i,\Delta} &= \mathrm{B}_i^*([\Delta]) \subseteq \Omega_i \\ T_{i,\Delta} &= \mathrm{proj}_{T_i}\mathrm{B}_i^*([\Delta]) \subseteq T_i \end{aligned}$$

and define the map $g_{i,\Delta}$ by restriction of g_i: for every (measurable) $E_{-i} \subseteq \Sigma_{-i} \times T_{-i,\Delta} \subseteq \Sigma_{-i} \times T_{-i}$

$$(\forall t_i \in T_{i,\Delta})(\forall h \in H_i)(g_{i,h,\Delta}(t_i)(E_{-i}) = g_{i,h}(t_i)(E_{-i})).$$

By eq. (**??**), $g_{i,h}(t_i)(\Omega_{-i,\Delta}(h)) = 1$ for every $i \in I$, $t_i \in T_{i,\Delta}$, $h \in H_i$. It follows that $g_{i,\Delta}(t_i) \in \Delta^{H_i}(\Sigma_{-i} \times T_{-i,\Delta})$ for every $i \in I$ and $t_i \in T_{i,\Delta}$. Clearly, each function $g_{i,\Delta}$ inherits continuity from g_i. Since the restrictions Δ are compact, the sets $T_{i,\Delta}$ and $\Omega_{i,\Delta}$ are compact metrizable.[32] Therefore

$$\mathcal{T}_\Delta = (\Omega_0, (\Omega_{i,\Delta}, T_{i,\Delta}, g_{i,\Delta})_{i\in I})$$

is a type structure as per Definition 2, and – of course – $\Omega_\Delta = \mathrm{B}^*([\Delta])$. We call $\mathcal{T}_\Delta$ the Δ-*restriction of* $\mathcal{T}$ and we use subscript Δ to denote events and belief operators in $\mathcal{T}_\Delta$. For example, R_Δ is the rationality event in $\mathcal{T}_\Delta$ and $\mathrm{CSB}_\Delta(E_\Delta)$ is correct mutual strong belief of E_Δ in $\mathcal{T}_\Delta$.

Theorem 2 *Fix a collection $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of compact subsets of first-order CPS's and a belief-complete type structure $\mathcal{T}$. Then $\mathcal{T}_\Delta$, the Δ-restriction of $\mathcal{T}$, satisfies the following properties: for every $i \in I$, $\theta_i \in \Theta_i$*

$$\Delta_{i,\theta_i} = \{\mu_i \in \Delta^{H_i}(\Sigma_{-i}) : (\exists t_i \in T_i)(\forall s_i \in S_i)(f_{i,\Delta}(t_i) = \mu_i) \wedge ((\theta_i, s_i, t_i) \in \Omega_{i,\Delta})\}; \tag{3}$$

for every $n \geq 0$,

$$(\mathrm{CSB}_\Delta)^n(R_\Delta) = \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta])); \tag{4}$$

[32] See Lemma 1 in the Appendix.

for every $n \geq 0$,

$$\Sigma_{\Delta}^{n+1} = \text{proj}_{\Sigma}(\text{CSB}_{\Delta})^{n}(R_{\Delta}), \tag{5}$$

and

$$\Sigma_{\Delta}^{\infty} = \text{proj}_{\Sigma}(\text{CSB}_{\Delta})^{\infty}(R_{\Delta}). \tag{6}$$

Theorem 2 says that (a) the Δ-restriction of a belief-complete type structure $\mathcal{T}$ exactly captures transparency of $[\Delta]$ (because $\Omega_{\Delta} = \text{B}^{*}([\Delta])$ and eq.(3) holds), (b) common strong belief of rationality and transparency of $[\Delta]$ in the belief-complete structure $\mathcal{T}$ correspond to common strong belief in rationality in $\mathcal{T}_{\Delta}$ (eq. (4)), and (c) Δ-rationalizability characterizes the behavioral implications of assuming rationality and common strong belief in rationality within type structure $\mathcal{T}_{\Delta}$ (eqs. (5) and (6)).

To illustrate, consider the Beer-Quiche game of Figure 1 under the **complete-information** scenario with the restrictions Δ of Example 2: the prior probability of *wimp* is 0.1 and the conjecture of Player 1 about Player 2 is independent of the observed realization of the chance move. We formally derive the Δ-rationalizability solution in the Appendix. Here it is sufficient to go back to the informal argument provided in the Introduction: **given the transparency of** Δ, if Player 1 is rational, believes that Player 2 is rational, and believes that Player 2 strongly believes that 1 is rational, then he chooses B (Beer) at information set s. Then Player 2 strongly believes that only a wimp could choose Quiche, i.e., a Quiche meal is interpreted by Player 2 as "wimpish". Once this is taken into account, Player 1 chooses B also at information set w in order to prevent a fight. But even though, eventually, Δ-rationalizability prevents Quiche, it still requires that Quiche be interpreted as "wimpish" by Player 2. This is similar to the best-rationalization principle described by [3, 1996] and analyzed epistemically by [11, 2002], but here the rationalization (forward-induction reasoning) is consistent with transparency of the restrictions Δ.

As explained in the Introduction, Battigalli & Siniscalchi [13, 2007] considered different epistemic assumptions whereby the **Δ-restrictions** are **not** assumed to be **transparent**, but rather they are assumed to have the same "epistemic priority" as the rationality assumption: if information set $h \in H_i$ contradicts $R_{-i} \cap [\Delta_{-i}]$ then Player i need not believe that the Δ_{-i} restrictions hold, and similarly for higher orders of mutual beliefs.[33] Formally, the step-1 epistemic assumption of [13, 2007] is

$$R \cap [\Delta] \cap \text{SB}(R \cap [\Delta]) = \text{CSB}(R \cap [\Delta]),$$

and the step-n epistemic assumption is $\text{CSB}^{n}(R \cap [\Delta])$. By definition $\text{B}^{*}([\Delta]) \subseteq [\Delta]$; can we then conclude that $\text{CSB}^{n}(R \cap \text{B}^{*}([\Delta])) \subseteq \text{CSB}^{n}(R \cap [\Delta])$ (for each n) and therefore the epistemic assumptions of [13, 2007] are weaker than those of Theorem 1? The affirmative answer would be obvious if $\text{CSB}(\cdot)$ were monotone, but we know that this is not the case. The following Theorem 3 states that, nonetheless, the answer is Yes, but despite being a weaker epistemic assumption, $\text{CSB}^{n}(R \cap [\Delta])$ has the same behavioral implications as $\text{CSB}^{n}(R \cap \text{B}^{*}([\Delta]))$, for each n. To see why, let us focus on step $n = 1$ for simplicity. First observe that only first-order belief restrictions matter to obtain the behavioral implications

[33] [13, 2007] also puts forward an incorrectly stated conjecture, the correct version of which is Theorem 1 above.

of $R \cap \mathrm{B}^*([\Delta]))$, which therefore must be the same as the behavioral implications of $R \cap [\Delta]$; indeed

$$\mathrm{proj}_\Sigma R \cap \mathrm{B}^*([\Delta])) = \Sigma^1_\Delta = \mathrm{proj}_\Sigma R \cap [\Delta].$$

Therefore, the information sets consistent with $R \cap \mathrm{B}^*([\Delta]))$ and $R \cap [\Delta]$ are the same. In particular, for each player $i \in I$, such information sets are given by the collection

$$H^1_i = \{h \in H_i : \Sigma_{-i}(h) \cap \Sigma^1_{-i,\Delta} \neq \emptyset\}.$$

But then, mutual strong belief in $R \cap \mathrm{B}^*([\Delta])$ implies mutual strong belief in the $R \cap [\Delta]$ despite the non-monotonicity of $\mathrm{SB}(\cdot)$:

$$\mathrm{SB}(R \cap \mathrm{B}^*([\Delta])) = \bigcap_{i \in I, h \in H_i(\Sigma^1_{-i,\Delta})} \mathrm{B}_{i,h}(R_{-i} \cap [\mathrm{B}^*([\Delta])]_{-i})$$

$$\subseteq \bigcap_{i \in I, h \in H_i(\Sigma^1_{-i,\Delta})} \mathrm{B}_{i,h}(R_{-i} \cap [\Delta]_{-i}) = \mathrm{SB}(R \cap [\Delta])$$

(where we let $[E]_{-i} = \Omega_i \times \mathrm{proj}_{\Omega_{-i}} E$),[34] because $[\mathrm{B}^*([\Delta])]_{-i} \subseteq [\Delta]_{-i}$ and each $\mathrm{B}_{i,h}(\cdot)$ is monotone. Thus,

$$\mathrm{CSB}^1(R \cap \mathrm{B}^*([\Delta])) = R \cap \mathrm{B}^*([\Delta]) \cap \mathrm{SB}(R \cap \mathrm{B}^*([\Delta]))$$

$$\subseteq R \cap [\Delta] \cap \mathrm{SB}(R \cap [\Delta]) = \mathrm{CSB}^1(R \cap [\Delta]).$$

The same result holds for each n.

Theorem 3 *Fix a collection $\Delta = (\Delta_{\theta_i})_{i \in I, \theta_i \in \Theta_i}$ of compact subsets of first-order CPS's and a belief-complete type structure $\mathcal{T}$. Then, for every $n \geq 1$,*

$$\begin{aligned} \mathrm{proj}_\Sigma \mathrm{CSB}^{n-1}(R \cap \mathrm{B}^*([\Delta])) &= \Sigma^n_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^{n-1}(R \cap [\Delta]) \\ \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta])) &\subseteq \mathrm{CSB}^n(R \cap [\Delta]), \end{aligned}$$

and

$$\begin{aligned} \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta])) &= \Sigma^\infty_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap [\Delta]), \\ \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta])) &\subseteq \mathrm{CSB}^\infty(R \cap [\Delta]). \end{aligned}$$

3.3 An algorithmic characterization of Δ-rationalizability

The Δ-rationalizability solution concept of Definition 4 is conceptually clear: at each step of the procedure the set of possible first-order beliefs is reduced by adding further strong-belief conditions, and the surviving strategies for each information-type are the sequential best replies to possible beliefs. Theorems 1, 2 and 5 provide transparent epistemic assumptions characterizing (**justifying**) Δ-rationalizability.

However, Δ-rationalizability is a somewhat complex procedure as it requires at step n to keep track of *all* the previous steps $m < n$, not just step $n-1$. Battigalli & Siniscalchi [12, 2003] propose instead a conceptually less transparent, but simpler reduction algorithm, which just requires a "one-step memory" and is closer to the original definition of extensive form rationalizability due to Pearce [33, 1984]. This is the solution concept called "Δ-rationalizability" in Battigalli & Siniscalchi [12, 2003] and Battigalli [7, 2006]. We refer to it as *naïve* Δ-rationalizability.

[34] With this notation, $[\Delta_{-i}] = [\Delta]_{-i}$.

Definition 5 *(See [12, 2003]) Consider the following procedure.*

(Step 0) *For every* $i \in I_0$*, let* $\hat{\Sigma}^0_{i,\Delta} = \Sigma_i$*. Also, let* $\hat{\Sigma}^0_{-i,\Delta} = \prod_{j \neq i} \hat{\Sigma}^0_{i,\Delta}$ *and* $\hat{\Sigma}^0_{\Delta} = \prod_{i \in I_0} \hat{\Sigma}^0_{i,\Delta}$.

(Step $n > 0$**)** *Let* $\hat{\Sigma}^n_{0,\Delta} = \Sigma_0$*. Then, for every* $i \in I$*, and for every* $\sigma_i = (\theta_i, s_i) \in \Sigma_i$*, let* $\sigma_i \in \hat{\Sigma}^n_{i,\Delta}$ *if and only if* $\sigma_i \in \hat{\Sigma}^{n-1}_{i,\Delta}$ *and there exists a CPS* $\mu \in \Delta_{\theta_i}$ *such that*

1. $\sigma_i \in \rho_i(\mu)$*;*
2. *for every* $h \in H_i$*, if* $\hat{\Sigma}^{n-1}_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset$*, then* $\mu(\hat{\Sigma}^{n-1}_{-i,\Delta} | \Sigma_{-i}(h)) = 1$*.*

Also let $\hat{\Sigma}^n_{-i,\Delta} = \prod_{j \neq i} \hat{\Sigma}^n_{i,\Delta}$ *and* $\hat{\Sigma}^n_{\Delta} = \prod_{i \in I_0} \hat{\Sigma}^n_{i,\Delta}$.

Finally, let $\hat{\Sigma}^\infty_{\Delta} = \bigcap_{n \geq 0} \hat{\Sigma}^n_{\Delta}$*. The profiles in* $\hat{\Sigma}^\infty_{\Delta}$ *are called* naïve Δ-rationalizable.

It is known that naïve Δ-rationalizability and Δ-rationalizability are equivalent when no restrictions are imposed on first-order beliefs (essentially, this is Theorem 1 in [4, 1997]). But this is not true in general. By inspection of the definitions, it is obvious that $\Sigma^n_{i,\Delta} = \hat{\Sigma}^n_{i,\Delta}$ for $n = 1, 2$, but the equality need not hold for $n > 2$ if Δ is not "rich enough". Specifically, it is easy to show that $\Sigma^3_{i,\Delta} \subseteq \hat{\Sigma}^3_{i,\Delta}$, but the inclusion may be strict. To see the problem, pick $\sigma_i = (\theta_i, s_i) \in \hat{\Sigma}^3_{i,\Delta}$. Say that a CPS $\mu \in \Delta^{H_i}(\Sigma_{-i})$ *strongly believes* $K_{-i} \subseteq \Sigma_{-i}$ if $\mu(K_{-i} | \Sigma_{-i}(h)) = 1$ whenever $K_{-i} \cap \Sigma_{-i}(h) \neq \emptyset$. With this, $\sigma_i \in \hat{\Sigma}^3_{i,\Delta}$ implies that there are CPS's $\mu^1, \mu^2 \in \Delta_{\theta_i}$ such that μ^n strongly believes $\hat{\Sigma}^n_{-i,\Delta} = \Sigma^n_{-i,\Delta}$ $(n = 1, 2)$ and $\sigma_i \in \rho_i(\mu^1) \cap \rho_i(\mu^2)$. Define the following "composition" $\bar{\mu}$ of μ^1 and μ^2: if $\Sigma^2_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset$ then $\bar{\mu}(\cdot | \Sigma_{-i}(h)) = \mu^2(\cdot | \Sigma_{-i}(h))$, otherwise $\bar{\mu}(\cdot | \Sigma_{-i}(h)) = \mu^1(\cdot | \Sigma_{-i}(h))$. It turns out that $\bar{\mu}$ is a CPS and $\sigma_i \in \rho_i(\bar{\mu})$; if we knew that $\bar{\mu} \in \Delta_{\theta_i}$ we would have shown that $\sigma_i \in \Sigma^3_{i,\Delta}$ because the composition $\bar{\mu}$ strongly believes both Σ^1_{-i} and Σ^2_{-i} as required by Definition 4, part 2. But $\bar{\mu}$ need not be in Δ_{θ_i} (see Example 6). Therefore, absent some assumptions about the structure of Δ, we may have $\Sigma^3_{i,\Delta} \subsetneq \hat{\Sigma}^3_{i,\Delta}$. Given the non-monotonicity of strong belief, the inclusion $\Sigma^3_{\Delta} \subseteq \hat{\Sigma}^3_{\Delta}$ need not be preserved in the following steps, therefore we cannot even conclude that $\Sigma^n_{\Delta} \subseteq \hat{\Sigma}^n_{\Delta}$ for every n. This discussion motivates the following:

Definition 6 *Fix two subsets* $K^1_{-i}, K^2_{-i} \subseteq \Sigma_{-i}$ *and CPS's* $\mu^1, \mu^2 \in \Delta^{H_i}(\Sigma_{-i})$*;* $(K^1_{-i}, \mu^1, K^2_{-i}, \mu^2)$ *is* admissible *if* $K^2_{-i} \subseteq K^1_{-i}$ *and* μ^n *strongly believes* K^n_{-i} $(n = 1, 2)$*. The* (K^1_{-i}, K^2_{-i})-composition *of* μ^1 *and* μ^2 *is the array* $\bar{\mu} = (\mu(\cdot | \Sigma_{-i}(h))_{h \in H_i}$ *such that* $\bar{\mu}(\cdot | \Sigma_{-i}(h)) = \mu^2(\cdot | \Sigma_{-i}(h))$ *whenever* $K^2_{-i} \cap \Sigma_{-i}(h) \neq \emptyset$*, and* $\bar{\mu}(\cdot | \Sigma_{-i}(h)) = \mu^1(\cdot | \Sigma_{-i}(h))$ *otherwise.*

Remark 2 *For every admissible* $(K^1_{-i}, \mu^1, K^2_{-i}, \mu^2)$*, the* (K^1_{-i}, K^2_{-i})*-composition of* μ^1 *and* μ^2 *is a CPS (see the proof of Lemma 4 in the Appendix).*

Definition 7 *Fix* $i \in I$*; a subset* $\Delta_i \subseteq \Delta^{H_i}(\Sigma_{-i})$ *is* closed under compositions *if, for every admissible 4-tuple* $(K^1_{-i}, \mu^1, K^2_{-i}, \mu^2)$ *with* $K^n_{-i} = \Sigma_0 \times \prod_{j \in I \setminus \{i\}} K^n_j$ $(n = 1, 2)$*, whenever* $\mu^1, \mu^2 \in \Delta_i$ *the* (K^1_{-i}, K^2_{-i})-composition *of* μ^1 *and* μ^2 *is also in* Δ_i*.*

Proposition 1 *If each* Δ_{θ_i} *is closed under compositions* $(i \in I,\ \theta_i \in \Theta_i)$*, then naïve* Δ*-rationalizability is equivalent to* Δ*-rationalizability: for every* $i \in I$ *and* n*,* $\hat{\Sigma}^n_{i,\Delta} = \Sigma^n_{i,\Delta}$*.*

Theorems 1, 5 and Proposition 1 implies that naïve Δ-rationalizability characterizes the epistemic assumptions analyzed in subsection 3.2:

Corollary 1 *Fix a collection $\Delta = (\Delta_{\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of compact subsets of CPS's that are closed under compositions and a belief-complete type structure $\mathcal{T}$. Then, for every $n \geq 0$,*

$$\hat{\Sigma}_{\Delta}^{n+1} = \text{proj}_{\Sigma}\text{CSB}^n(R \cap \text{B}^*([\Delta])) = \text{proj}_{\Sigma}\text{CSB}^n(R \cap [\Delta])$$

and

$$\hat{\Sigma}_{\Delta}^{\infty} = \text{proj}_{\Sigma}\text{CSB}^{\infty}(R \cap \text{B}^*([\Delta])) = \text{proj}_{\Sigma}\text{CSB}^{\infty}(R \cap [\Delta]).$$

Of course, closedness under compositions is just a sufficient condition for the above equivalence. Indeed, we are not aware of any application of the theory where Δ-rationalizability differs from naïve Δ-rationalizability. However, we show below a numerical example where the condition fails and the two solution procedures are not equivalent (Example 6).

Next we provide sufficient conditions for closedness under composition that may be useful in applications and are easier to verify. The first one is a strong property used by [13, 2007] to obtain an epistemic justification of naïve Δ-rationalizability (cf. Theorem 5 in the Appendix), the others are interesting special cases.

Definition 8 *(cf. [13, 2007]) Fix $i \in I$ and regard $\Delta^{H_i}(\Sigma_{-i})$ as a subset of $\prod_{h\in H_i}\Delta(\Sigma_{-i}(h))$; a subset $\Delta_i \subseteq \Delta^{H_i}(\Sigma_{-i})$ is* regular *if for each $h \in H_i$ there is a nonempty subset $\Delta_{i,h} \subseteq \Delta(\Sigma_{-i}(h))$ so that $\Delta_i = \Delta^{H_i}(\Sigma_{-i}) \cap \left(\prod_{h\in H_i}\Delta_{i,h}\right)$.*[35]

Remark 3 *If $\Delta_i \subseteq \Delta^{H_i}(\Sigma_{-i})$ is regular then it is closed under composition.*

The following examples show how to apply this remark to the rationalizability analysis of two-person dynamic games with chance moves and two-person dynamic Bayesian games.

Example 3 *Generalizing Example 2 (Beer-Quiche under the complete-information scenario), suppose that Θ is a singleton, but there is an initial chance move ($\iota(\varnothing) = \{0\}$) about which Players 1 and 2 are imperfectly and asymmetrically informed, as in poker; fix $\pi_{i,0} \in \Delta(S_0)$ strictly positive for each $i \in I$ and let*

$$\Delta_i = \{\mu_i \in \Delta^{H_i}(\Sigma_{-i}) : \mu_i(\cdot|S_{-i}) = \pi_{i,0} \times (\text{marg}_{S_j}\mu_i(\cdot|S_{-i}))\}.$$

Since Δ_i is determined by a condition on initial beliefs (the beliefs conditional on information set $\{\varnothing\}$), Δ_i is regular and – by Remark 3 – closed under composition. Formally, the conditions $\text{marg}_{S_0}\mu_i(\cdot|S_{-i}) = \pi_{i,0}$ ($i \in I$) yield a Bayesian game with heterogeneous priors,[36] *and (naïve) Δ-rationalizability corresponds to* ex ante (extensive-form) rationalizability *in this Bayesian game, a concept that is generically equivalent to (maximal)*

[35] We conjecture that if Δ is regular and each $\Delta_{i,h}$ is closed and convex, then the characterization of Δ-rationalizability as iterated Δ-dominance (Cappelletti [22, 2010]) can be extended to the present extensive-form setting.

[36] If the reader is wondering why a complete information game corresponds to a Bayesian game, he should remember that in our terminology (which we claim to be the correct one) "complete information" is a substantive assumption, i.e. common knowledge of the payoff functions. On the other hand, Bayesian games are just mathematical structures that may be used to analyze both games with incomplete information and games with asymmetric, imperfect information about an initial chance move, such as poker. The interpretation of such mathematical structures is immaterial for Harsanyi's equilibrium analysis, but not for rationalizability analysis. The reason is that standard notions of rationalizability for Bayesian games implicitly incorporate independence restrictions, and different restrictions are relevant under different interpretations (see [16, 2011]).

iterated removal of weakly dominated strategies in the ex ante strategic form. (Note that Pearce's [33, 1984] solution concept allows i to have different conjectures about j after different observations about the chance move; hence it violates the independence condition stated above.)

Example 4 *Consider a two-person game with no chance moves (S_0 is a singleton), fix belief maps $(\pi_i : \Theta_i \to \Delta(\Theta_{-i}))_{i=1,2}$ and, for each θ_i, let Δ_{θ_i} be the set of CPS's consistent with $\pi_i(\theta_i)(\cdot)$ that satisfy "conditional independence":*

$$\Delta_{\theta_i} = \left\{ \mu_i \in \Delta^{H_i}(\Sigma_{-i}) : \begin{array}{l} \mathrm{marg}_{\Theta_{-i}}\mu_i(\cdot|\Sigma_{-i}) = \pi_i(\theta_i)(\cdot), \\ \forall \theta_j, \mu_i(\theta_j) > 0 \Rightarrow \frac{\mu_i(\theta_0,\theta_j,s_j)}{\mu_i(\theta_j)} = \mu_i(\theta_0|\theta_j)\mu_i(s_j|\theta_j) \end{array} \right\}$$

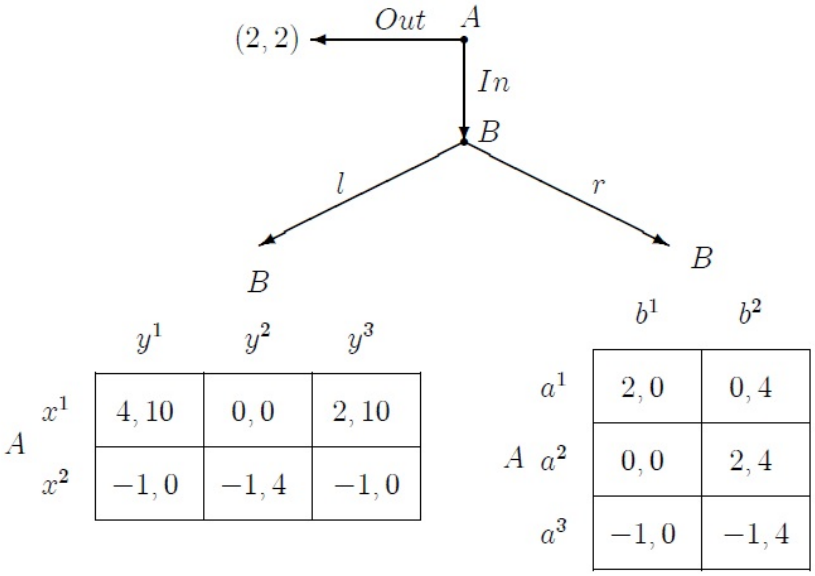
(we use an obvious notation for marginal probabilities and conditional probabilities – whenever well defined). Each set of CPS's Δ_{θ_i} is determined by conditions on initial beliefs, hence it is regular and closed under composition. Formally, the conditions $\mathrm{marg}_{\Theta_{-i}}\mu_i(\cdot|\Sigma_{-i}) = \pi_i(\theta_i)(\cdot)$ $(i \in I)$ yield a Bayesian game, and (naïve) Δ-rationalizability is extensive-form, interim independent rationalizability *for this Bayesian game, which is generically equivalent to the iterated removal of weakly dominated strategies in the interim strategic form. (When players move simultaneously, this is equivalent to the application of Pearce's [33, 1984] solution concept to the extensive form of the Bayesian game. See also* [16, 2011]*).*

Example 5 *Consider a two-person game with no chance moves where Θ_0 is a singleton; fix a distribution $\delta \in \Delta(\Theta \times Z)$ (i.e., a probability density over the terminal nodes of the arborescence representing the game); say that δ is* admissible *if it is obtained as the pushforward of some product measure $\nu_1 \times \nu_2 \in \Delta(\Sigma_1 \times \Sigma_2)$* via *map $(\theta_1, s_1, \theta_2, s_2) \longmapsto (\theta_1, \theta_2, \zeta(s_1, s_2))$. Say that CPS $\mu_i \in \Delta^{H_i}(\Sigma_{-i})$* agrees with *the admissible measure $\delta \in \Delta(\Theta \times Z)$ if there exists some measure $\nu_i \in \Delta(\Sigma_i)$ such that δ is the pushforward of $\nu_i(\cdot) \times \mu_i(\cdot|\Sigma_{-i})$* via *map $(\theta_1, s_1, \theta_2, s_2) \longmapsto (\theta_1, \theta_2, \zeta(s_1, s_2))$. The set $\Delta_i(\delta)$ of CPS's that agree with an admissible $\delta \in \Delta(\Theta \times Z)$ is regular, as it is determined only by a condition on initial beliefs $\mu_i(\cdot|\Sigma_{-i})$. (In particular, this condition specifies the initial belief $marg_{\Theta_{-i}}\delta$ of each player i about the information type of the co-player $-i$, and therefore yields a simple dynamic Bayesian game.)*

Battigalli & Siniscalchi [12, 2003] propose naïve $\Delta(\delta)$-rationalizability as a forward-induction refinement criterion: A candidate equilibrium yields a distribution δ; if $\Sigma^{\infty}_{\Delta(\delta)} \neq \emptyset$ then the candidate equilibrium is "consistent with forward-induction reasoning", otherwise it is not. They show that the Iterated Intuitive Criterion for signaling games is precisely such a refinement, therefore the naïve $\Delta(\delta)$-rationalizability criterion is an extension to more general games of the Iterated Intuitive Criterion. By Proposition 1 this result applies to $\Delta(\delta)$-rationalizability as well. Indeed, all the examples and results of [12, 2003] about naïve Δ-rationalizability feature conditions derived from restrictions on initial beliefs only and therefore they also hold for Δ-rationalizability.

So far we listed examples where Δ-rationalizability and the naïve Δ-algorithm coincide. But the following example shows that, when closedness under compositions fails, the two solution procedures may differ because naïve Δ-rationalizability may fail to capture an inconsistency between the Δ-restrictions and common strong belief in rationality, given

transparency of the Δ-restrictions. This example also shows that naïve Δ-rationalizability may yield a solution set altogether inconsistent with forward-induction reasoning.[37] For notational simplicity, we look at classes of realization-equivalent strategies, which in this game are the strategies of the reduced normal form.



A	y^1	y^2	y^3
x^1	$4, 10$	$0, 0$	$2, 10$
x^2	$-1, 0$	$-1, 4$	$-1, 0$

A	b^1	b^2
a^1	$2, 0$	$0, 4$
a^2	$0, 0$	$2, 4$
a^3	$-1, 0$	$-1, 4$

Figure 2

Example 6 *In the complete-information game depicted in Figure 2, Ann chooses between In and Out; given In, Bob chooses between the matrix game on the left or the matrix game on the right. The given transparent restrictions allow only two possible CPS's for each player:*[38]

$$\begin{aligned} \Delta_a &= \{\mu_a, \nu_a\},\ \textit{with} \\ \mu_a(l.y^1|S_b) &= \mu_a(l.y^2|S_b) = \frac{1}{2},\ \mu_a(r.b^2|\{r.b^1, r.b^2\}) = 1, \\ \nu_a(l.y^3|S_b) &= 1,\ \nu_a(r.b^1|\{r.b^1, r.b^2\}) = 1, \end{aligned}$$

$$\begin{aligned} \Delta_b &= \{\mu_b, \nu_b\}\ \textit{with} \\ \mu_b(O|S_a) &= 1,\ \mu_b(I.x^2.a^3|\{s_a : s_a(\varnothing) = I\}) = 1, \\ \nu_b(O|S_a) &= 1,\ \nu_b(I.x^1.a^2|\{s_a : s_a(\varnothing) = I\}) = 1. \end{aligned}$$

Let us first compute the sequential best replies to each belief system in Δ:

$$\begin{aligned} \rho_a(\mu_a) &= \{O, I.x^1.a^2\},\ \rho_a(\nu_a) = \{O, I.x^1.a^1\}, \\ \rho_b(\mu_b) &= \{l.y^2, r.b^2\},\ \rho_b(\nu_b) = \{l.y^1, l.y^3\}. \end{aligned}$$

[37] Our original example showing the difference between Δ-rationalizability and naive Δ-rationalizability did not have the latter feature. We thank Amanda Friedenberg for providing this example.

[38] We use the following notation: O is the class of realization-equivalent strategies choosing *Out*, otherwise strategies are denoted by lists of action labels in an obvious way.

Note: (1) $I.x^2.a^3$ is strictly dominated by O and $r.b^1$ is strictly dominated by $r.b^2$ given In, but μ_b and ν_a respectively assign positive conditional probability to these dominated strategies, therefore ν_a does not strongly believe $\rho_b(\mu_b) \cup \rho_b(\nu_b)$, and μ_b does not strongly believe $\rho_a(\mu_a) \cup \rho_a(\nu_a)$, (2) μ_a does not strongly believe $\rho_b(\nu_b)$, while ν_a does. These preliminary considerations help explain the following derivation, where S^m_Δ ($\hat{S}^m_\Delta$) denotes Δ-rationalizability (naïve Δ-rationalizability):[39]

Step	*Ann*	*Bob*
1	$S^1_{a,\Delta} = \hat{S}^1_{a,\Delta} = \{O, I.x^1.a^1, I.x^1.a^2\}$	$S^1_{b,\Delta} = \hat{S}^1_{b,\Delta} = \{l.y^2, r.b^2, l.y^1, l.y^3\}$
2	$S^2_{a,\Delta} = \hat{S}^2_{a,\Delta} = \{O, I.x^1.a^2\}$	$S^2_{b,\Delta} = \hat{S}^2_{b,\Delta} = \{l.y^1, l.y^3\}$
3	$S^3_{a,\Delta} = \emptyset$, $\hat{S}^3_{a,\Delta} = \{O\}$	$S^3_{b,\Delta} = \hat{S}^3_{b,\Delta} = \{l.y^1, l.y^3\}$
$n > 3$	$S^n_{a,\Delta} = \emptyset$, $\hat{S}^n_{a,\Delta} = \{O\}$	$S^n_{b,\Delta} = \hat{S}^n_{b,\Delta} = \{l.y^1, l.y^3\}$

As we said above, the two solution procedures may differ only from round 3, as they do here. Both procedures stop at round 3. The naïve Δ-rationalizable set is therefore $\{O\} \times \{l.y^1, l.y^3\}$, whereas we get an empty set with Δ-rationalizability. By Proposition 1, Δ cannot be closed under composition. Let us verify this directly: μ_a strongly believes $S^1_{b,\Delta}$ and ν_a strongly believes $S^2_{b,\Delta} \subseteq S^1_{b,\Delta}$. We can "compose" μ_a and ν_a to obtain a CPS $\bar{\mu}_a$ so that $O \in \rho_a(\bar{\mu}_a)$ and $\bar{\mu}_a$ strongly believes both $S^1_{b,\Delta}$ and $S^2_{b,\Delta}$:

$$\bar{\mu}_a(l.y^3|S_b) = \nu_a(l.y^3|S_b) = 1,\ \bar{\mu}_a(r.b^2|\{r.b^1, r.b^2\}) = \mu_a(r.b^2|\{r.b^1, r.b^2\}) = 1.$$

But $\bar{\mu}_a \neq \mu_a, \nu_a$, therefore $\bar{\mu}_a \notin \Delta_a$.
Finally, we show that there are no restrictions $\bar{\Delta}$ such that $\{O\} \times \{l.y^1, l.y^3\}$ is $\bar{\Delta}$-rationalizable. Suppose, by way of contradiction, that such $\bar{\Delta}$ exists. By finiteness, $\{O\} \times \{l.y^1, l.y^3\} = S^n_{a,\bar{\Delta}} \times S^n_{b,\bar{\Delta}} = S^{n-1}_{a,\bar{\Delta}} \times S^{n-1}_{b,\bar{\Delta}}$ for some n. Define $\bar{\Delta}^n_a$ as in Remark 1. Then we should have $\{O\} = \rho_a(\bar{\Delta}^n_a)$, where $\bar{\mu}_a(\{l.y^1, l.y^3\}) = 1$ for every $\bar{\mu}_a \in \bar{\Delta}_a$. Pick any $\bar{\mu}_a \in \bar{\Delta}^n_a$ with $O \in \rho_a(\bar{\mu}_a)$, then also $I.x^1.a^1 \in \rho_a(\bar{\mu}_a)$ (the ex ante value of $I.x^1.a^1$ is at least 2, and a^1 is the dominant action in the subgame with root (In, l)), therefore $\{O\} \neq \rho_a(\bar{\Delta}^n_a)$.

Indeed, we show in the next section that the set $\{O\} \times \{l.y^1, l.y^3\}$ of Example 6 cannot result from forward-induction reasoning under any transparent restrictions on beliefs, including restrictions on higher-order beliefs.

4 Transparent belief restrictions and extensive-form best response sets

So far we took the perspective of an analyst who knows what belief restrictions are transparent to the players: given the transparent restrictions Δ we obtain the behavioral prediction of forward-induction reasoning, Σ^∞_Δ. This perspective is valid in many, but not all applications of the theory. It is therefore interesting to take the perspective of an analyst who does not know what restrictions are transparent to the players. What can such an analyst say about the behavioral implications of our theory of strategic thinking?

[39] Note that in games with complete information $\Sigma = S$.

If all restrictions are possible, all an analyst can say is that (θ, s) must belong to the union of all solution sets Σ_Δ^∞. In static games, where forward induction does not play any role, this yields an easy answer to our question: just look at the implications of rationality and common belief in rationality without restrictions on beliefs, i.e., take $\Delta_{\theta_i} = \Delta(\Sigma_{-i})$ for each information type of each player. Indeed, let Σ^∞ denote the solution set obtained without restriction; by monotonicity of probability-one belief, $\Sigma_\Delta^\infty \subseteq \Sigma^\infty$ for all restrictions Δ; thus, Σ^∞ is also the union of all the solution sets. But this method does not work in dynamic games, that is, when forward induction matters. The reason is that, as explained earlier, strong belief is not monotone, therefore we may well have $\Sigma_\Delta^\infty \not\subseteq \Sigma^\infty$ for some Δ.[40]

Battigalli & Friedenberg [8, 2009] and [9, 2012] address this issue within a complete-information setting.[41] We report a straightforward adaptation of their concepts and results to the present incomplete-information setting.

Definition 9 *(cf. [9, 2012]) An* extensive-form best response set (EFBRS) *of a game with payoff uncertainty Γ is a set of profiles $Q = \Sigma_0 \times \prod_{i\in I} Q_i \subseteq \Sigma$ such that, for each $i \in I$, $\text{proj}_{\Theta_i} Q_i = \Theta_i$ and, for each $\sigma_i \in Q_i$, there is a CPS $\mu_{\sigma_i} \in \Delta^{H_i}(\Sigma_{-i})$ with*
(i) $\sigma_i \in \rho_i(\mu_{\sigma_i})$,
(ii) μ_{σ_i} strongly believes Q_{-i} ($\forall h \in H_i$, $\Sigma_{-i}(h) \cap Q_{-i} \neq \emptyset \Rightarrow \mu_{\sigma_i}(Q_{-i}|\Sigma_{-i}(h)) = 1$),
(iii) $\rho_i(\mu_{\sigma_i}) \subseteq Q_i$.

Condition (iii) is a kind of maximality property: if we have to use belief μ_{σ_i} to "justify" σ_i as part of the solution set Q, then every other sequential best response to μ_{σ_i} must also be part of Q.

An event about players' beliefs, i.e. an "epistemic" event, is self-evident if and only if it represents the transparency of some restrictions on beliefs (of any order, see Appendix A in [8, 2009]). Battigalli and Friedenberg show that the EFBRS concept characterizes the behavioral implications of common strong belief in rationality when some epistemic event is self-evident. The following definition and theorem make this precise and provide equivalent formal statements of the result. Then an example shows that naïve Δ-rationalizability may fail to yield an EFBRS.

Definition 10 *An event E is called* epistemic *if it only restricts the beliefs of some players for some of their information types θ_i, that is, for each $i \in I$ there is a measurable subset $\mathbf{T}_i \subseteq \Theta_i \times T_i$ with $\text{proj}_{\Theta_i} \mathbf{T}_i = \Theta_i$ so that*

$$E = \Omega_0 \times \prod_{i\in I} \{(\theta_i, s_i, t_i) \in \Omega_i : (\theta_i, t_i) \in \mathbf{T}_i\}.$$

$[\Delta]$ and $\text{B}^*([\Delta])$ are examples of epistemic events. Recall that the latter is also self-evident: $\text{B}^*([\Delta]) \subseteq \text{B}(\text{B}^*([\Delta]))$. Adapting the arguments of [8, 2009] and [9, 2012] one

[40] Battigalli & Friedenberg [9, 2012] provide examples of complete information games where the inclusion also fails for the corresponding sets of paths, that is, $\zeta(S_\Delta^\infty) \not\subseteq \zeta(S^\infty)$ for some game and some restrictions Δ.

[41] Battigalli & Friedenberg [9, 2012] is the abridged published version of [8, 2009]. The latter elaborates more on the context interpretation of incomplete type structures and how they are related to transparent restrictions on beliefs. Battigalli and Friedenberg build on previous work on admissibility by Brandenburger *et al.* [21, 2008] and Brandenburger & Friedenberg [20, 2010].

can show the following:[42]

Theorem 4 *For every game Γ and subset $Q \subseteq \Sigma$, the following are equivalent:*
(1) Q is an EFBRS of Γ;
(2) $Q = \Sigma_\Delta^\infty$ for some first-order belief restrictions Δ;
(3) $Q = \text{proj}_\Sigma \text{CSB}^\infty(R)$ for some type structure $\mathcal{T}$;
(4) $Q = \text{proj}_\Sigma \text{CSB}^\infty(R \cap E)$ for some self-evident epistemic event E in a belief-complete type structure $\mathcal{T}$;
(5) $Q = \text{proj}_\Sigma \text{CSB}^\infty(R \cap \text{B}^*([\Delta]))$ in a belief-complete type structure for some compact first-order belief restrictions Δ.

Example 7 *Consider again the complete-information game of Figure 2 and the belief restrictions Δ of Example 6. The corresponding naïve Δ-rationalizable set is $\hat{S}_\Delta^\infty = Q = \{O\} \times \{l.y^1, l.y^3\}$. It can be directly verified that $\{O\} \times \{l.y^1, l.y^3\}$ is not an EFBRS: if $\mu_O \in \Delta^{H_a}(S_b)$ strongly believes $Q_b = \{l.y^1, l.y^3\}$ and $O \in \rho_a(\mu_O)$, then $I.x^1.a^1 \in \rho_a(\mu_O) \backslash Q_a$; hence at least one of conditions (i)-(iii) of Definition 9 must be violated by Q. Thus, by equivalence (1) $\Leftrightarrow$ (3) of Theorem 4, there is no type structure $\mathcal{T}$ such that $Q = \text{proj}_\Sigma \text{CSB}^\infty(R)$. We can reach the conclusion that Q is not an EFBRS also indirectly, using implication (2) $\Rightarrow$ (1) of Theorem 4: indeed, we have already proved in Example 6 that there are no first-order belief restrictions Δ such that $Q = S_\Delta^\infty$.*

[9, 2012] shows that, in games with complete and perfect information and in other games such as the Finitely Repeated Prisoners' Dilemma, the characterizing properties of EFBRS's can be used to obtain observable implications of forward-induction reasoning, independently of what is transparent to the players. It would be interesting to obtain results of this sort for games with payoff uncertainty.

5 Appendix

Subsection 5.1 provides a complete forward-induction analysis of the Beer-Quiche game under different scenarios and related restrictions on beliefs; 5.2 derives the epistemic justifications of Δ-rationalizability; 5.3 proves the equivalence between Δ-rationalizability and the naïve reduction algorithm when Δ is closed under compositions.

5.1 Two scenarios for Δ-rationalizability in the Beer-Quiche game

Complete information scenario We formalize the informal argument provided in the Introduction for the Beer-Quiche game of Fig. 1 and Example 1 (1). As formally described in Example 2, assume the following restrictions Δ are *transparent*:

- Players ex ante assign 90% (10%) probability to the surly (wimp) state,

[42] The equivalences stated in the Theorem 4 adapt results of Battigalli and Friedenberg as follows: (1) $\Leftrightarrow$ (2) adapts Proposition 1 of [9, 2012]; (1) $\Leftrightarrow$ (3) adapts Theorem 1 of [9, 2012]; (1) $\Leftrightarrow$ (4) adapts Proposition A1 of [8, 2009] applied to the E-restriction $\mathcal{T}_E$ of a belief-complete type structure $\mathcal{T}$; finally (1) $\Rightarrow$ (5) follows from Theorem 2 and (the adaptation of) Proposition 1 in [9, 2012] (the set Δ constructed in the proof is finite, hence compact), (5) $\Rightarrow$ (1) follows from Theorem 1 and (2) $\Rightarrow$ (1).

- According to the ex ante beliefs of Player 1, the chance move and the strategy of Player 2 are independent.

The behavioral implications of (correct) common strong belief of rationality and transparency of Δ can be derived with the the (naïve) Δ-rationalizability algorithm (Theorem 1, Corollary 1). To apply the latter, we first note that we may add wlog the further restriction that, according to the beliefs of Player 2, the strategy of Player 1 and the chance move are independent.[43] With this, Δ-rationalizability is a refinement of rationalizability on the *ex ante strategic form* of the Bayesian game defined by the first restriction.[44] In the table below, we multiply all the expected payoffs by a factor of 10:

$1\backslash 2$	$f\|B,\ f\|Q$	$f\|B,\ d\|Q$	$d\|B,\ f\|Q$	$d\|B,\ d\|Q$
$B\|s,\ B\|w$	9, 1	9, 1	29, 9	29, 9
$B\|s,\ Q\|w$	10, 1	12, 0	28, 10	30, 9
$Q\|s,\ B\|w$	0, 1	18, 10	2, 0	20, 9
$Q\|s,\ Q\|w$	1, 1	21, 9	1, 1	21, 9

Iterated weak dominance on this strategic form deletes $f.f$ and $Q.B$ in step 1, $f.d$ in step 2, $Q.Q$ in step 3, $d.d$ in step 4 and $B.Q$ in step 5.[45] This is equivalent (for this game) to iterated conditional dominance on the strategic form, which gives the Δ-rationalizability solution. We summarize the procedure in the table below. For intuitive explanations based on the epistemic analysis see the Introduction.

Steps	Player 1	Player 2
1	$S^1_{1,\Delta} = \{B.B, B.Q, Q.Q\}$	$S^1_{2,\Delta} = \{f.d, d.d, d.f\}$
2	$S^2_{1,\Delta} = \{B.B, B.Q, Q.Q\}$	$S^2_{2,\Delta} = \{d.d, d.f\}$
3	$S^3_{1,\Delta} = \{B.B, B.Q\}$	$S^3_{2,\Delta} = \{d.d, d.f\}$
4	$S^4_{1,\Delta} = \{B.B, B.Q\}$	$S^4_{2,\Delta} = \{d.f\}$
5	$S^5_{1,\Delta} = \{B.B\}$	$S^5_{2,\Delta} = \{d.f\}$

Incomplete-information scenario In this scenario there is no chance move, the game begins at the "interim" stage, when Player 1 knows the true state; see Example 1 (2). Now there are two conceivable information types of Player 1, s and w, each with strategy space $S_1 = \{B, Q\}$. Thus

$$\Sigma_1 = \{s, w\} \times \{B, Q\},\ \Sigma_2 = S_2 = \{f.d, d.d, d.f, f.f\}.$$

In this scenario, Player 1 has no belief on $\{s, w\}$. Intuitively, the reason is that $\theta \in \{s, w\}$ is just an attribute of Player 1 known to him, not something that he learns. Formally, his primitive uncertainty space is $\Sigma_{-1} = S_2$. With this, the independence assumption of the

[43] The coalition formed by the chance player and Player 1 has perfect recall. Therefore, a correlated strategy of the coalition is realization equivalent to a behavioral strategy of the coalition, which is in turn equivalent to a product measure.

[44] As observed by Battigalli *et al* [16, 2011], the application of solution concepts (such as rationalizability, or iterated dominance) to the *ex ante* strategic form implicitly relies on the independence assumption described above.

[45] We write $X.Y$ for the strategy of Player 1 that selects X in the surly state and Y in the wimp state; similarly we write $x.y$ for the strategy of Player 2 that selects x if B and y if Q.

complete-information scenario cannot be expressed as a property of his first-order beliefs. As we mentioned in the Introduction, there is a property of the *second-order* beliefs of *Player 2* that could replace the independence property expressible in the complete-information scenario: according to the second-order beliefs of Player 2, the information type of Player 1 and his first-order belief about the strategy of Player 2 are independent. We could provide an epistemic analysis of the game based on the transparency of first and second-order restrictions and derive results similar to those obtained for the complete-information scenario. But in this paper we restricted our attention to transparency of first-order restrictions. Therefore, as an illustration of our approach we assume the following first-order restrictions Δ to be *transparent*:

- Player 2 initially assigns 90% probability to $\theta = s$,
- conditional on B, Player 2 believes that s is more likely than w.

Formally,

$$\Delta_{1,s} = \Delta_{1,w} = \Delta(S_2),$$

$$\Delta_2 = \{\mu_2 \in \Delta^{\{B,Q\}}(\{s,w\} \times \{B,Q\}) : (\mu_2(s) = 0.9) \wedge (\mu_2(s|B) > \mu_2(w|B))\},$$

where we used obvious abbreviations for unconditional, conditional, and marginal probabilities. Note, this is just meant to be an example; we do not claim that the assumption in the second bullet is more plausible, or "nicer" than the assumption about the second-order beliefs of Player 2 described above.

The behavioral implications of (correct) common strong belief in rationality and transparency of the restrictions are characterized by a solution procedure that deletes profiles from the set $\Sigma_1 \times S_2$. We summarize the steps in the table below and then comment.

Steps	Player 1	Player 2
1	$\Sigma^1_{1,\Delta} = \{s, w\} \times \{B, Q\}$	$S^1_{2,\Delta} = \{d.d, d.f\}$
2	$\Sigma^2_{1,\Delta} = \{(s,B), (w,B), (w,Q)\}$	$S^2_{2,\Delta} = \{d.d, d.f\}$
3	$\Sigma^3_{1,\Delta} = \{(s,B), (w,B), (w,Q)\}$	$S^3_{2,\Delta} = \{d.f\}$
4	$\Sigma^4_{1,\Delta} = \{(s,B), (w,B)\}$	$S^4_{2,\Delta} = \{d.f\}$

Step 1, *Player 2:* We delete $f.f$ (as in the complete-information scenario) and also $f.d$ (fight if and only if Beer); the reason for the latter is that it is rational to fight if and only if the conditional probability of the surly type is no more that 50%, therefore, given the restrictions on beliefs, if it is rational to fight after B then it is *a fortiori* rational to fight after Q. To sum up, Player 2 does not fight after B.

Step 2, *Player 1:* Given the above, the surly type has his preferred meal, i.e., type/strategy pair (s, Q) is deleted.

Step 3, *Player 2:* By a forward-induction argument similar to Step 4 of the complete-information scenario (see the Introduction), $d.d$ is deleted.

Step 4: Each type of Player 1 is certain of $d.f$, therefore both types choose Beer.

5.2 Proofs of the characterization results

The proof of the main result (Theorem 1) is adapted from [13, 2007]. First, we need some preliminaries. Fix a **belief-complete** type structure and a collection of compact subsets

Δ. $B^*([\Delta])$ is a belief-closed subset of this structure that can be constructed as follows (recall we introduced the singleton $T_0 = \{t_0\}$ for notational convenience): let

$$\begin{aligned} \mathbf{T}^0_{i,\Delta} &= \{(\theta_i, t_i) \in \Theta_i \times T_i : f_i(t_i) \in \Delta_{\theta_i}\} \ (i \in I), \\ \mathbf{T}^0_{0,\Delta} &= \Theta_0 \times T_0 = \Theta_0 \times \{t_0\}, \end{aligned}$$

for each $n > 0$ and $i \in I$

$$\begin{aligned} \mathbf{T}^n_{i,\Delta} &= \left\{ (\theta_i, t_i) \in \mathbf{T}^{n-1}_{i,\Delta} : (\forall h \in H_i) \left(g_{i,h}(t_i)(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^{n-1}_{j,\Delta})) = 1 \right) \right\}, \\ \mathbf{T}^n_{0,\Delta} &= \Theta_0 \times T_0 = \Theta_0 \times \{t_0\}, \end{aligned}$$

$$\mathbf{T}^*_{i,\Delta} = \bigcap_{n \geq 1} \mathbf{T}^n_{i,\Delta}.$$

Clearly, $\cap^n_{m=1} B^m([\Delta]) = \Sigma_0 \times \prod_{i \in I}(S_i \times \mathbf{T}^n_{i,\Delta})$ $(n \geq 0)$,[46] thus $B^*([\Delta]) = \Sigma_0 \times \prod_{i \in I}(S_i \times \mathbf{T}^*_{i,\Delta})$. Note that we had to define $\mathbf{T}^*_{i,\Delta}$ as a subset of $\Theta_i \times T_i$ rather than T_i, because the restrictions on first-order beliefs may depend on the information type θ_i. Elements of $\mathbf{T}^*_{i,\Delta}$ will be simply called "types".

We begin with a few preliminary results. Let $T^n_{\theta_i,\Delta} = \{t_i \in T_i : (\theta_i, t_i) \in \mathbf{T}^n_{i,\Delta}\}$ and $T^*_{\theta_i,\Delta} = \{t_i \in T_i : (\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}\}$ denote the θ_i-section of, respectively, $\mathbf{T}^n_{i,\Delta}$ and $\mathbf{T}^*_{i,\Delta}$. Clearly, $T^*_{\theta_i,\Delta} = \cap_{n \geq 1} T^n_{\theta_i,\Delta}$.

Lemma 1 *For all $i \in I$, $\theta_i \in \Theta_i$ and $n \geq 0$, the sets $T^n_{\theta_i,\Delta}$, $T^*_{\theta_i,\Delta}$, $\mathbf{T}^n_{i,\Delta}$ and $\mathbf{T}^*_{i,\Delta}$ are closed.*

Proof For every $\theta_i \in \Theta_i$, $T^0_{\theta_i,\Delta} = f_i^{-1}(\Delta_{\theta_i})$ is closed because f_i is continuous and Δ_{θ_i} is closed. Therefore, $\mathbf{T}^0_{i,\Delta} = \cup_{\theta_i \in \Theta_i} \{\theta_i\} \times T^0_{\theta_i,\Delta}$ is closed because it is the union of finitely many closed sets.

Now suppose by way of induction that $T^{n-1}_{\theta_i,\Delta}$ is closed for every i and θ_i, so that $\mathbf{T}^{n-1}_{i,\Delta} = \cup_{\theta_i \in \Theta_i} \{\theta_i\} \times T^{n-1}_{\theta_i,\Delta}$ is closed for every i. Then the set of probability measures

$$\Delta \left(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^{n-1}_{j,\Delta}) \right) = \left\{ \mu : \mu \left(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^{n-1}_{j,\Delta}) = 1 \right) \right\}$$

is closed. This implies that the set of conditional beliefs that fully believe $\prod_{j \in I_0 \setminus \{i\}}(S_j \times \mathbf{T}^{n-1}_{j,\Delta})$, that is,

$$\Delta^{H_i}(\Sigma_{-i} \times T_{-i}) \cap \left[\Delta \left(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^{n-1}_{j,\Delta}) \right) \right]^{H_i},$$

is closed as well, because $\Delta^{H_i}(\Sigma_{-i} \times T_{-i})$ is closed. Since g_i is continuous,

$$g_i^{-1} \left(\Delta^{H_i}(\Sigma_{-i} \times T_{-i}) \cap \left[\Delta \left(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^{n-1}_{j,\Delta}) \right) \right]^{H_i} \right),$$

[46] Since $\Sigma_i = \Theta_i \times S_i$ and $\mathbf{T}^n_i \subseteq \Theta_i \times T_i$ we are abusing notation here. This should cause no confusion.

is closed. This implies that, for every θ_i, $T^n_{\theta_i,\Delta}$ is closed, because $T^{n-1}_{\theta_i,\Delta}$ is closed (inductive hypothesis) and

$$T^n_{\theta_i} = T^{n-1}_{\theta_i} \cap g_i^{-1}\left(\Delta^{H_i}(\Sigma_{-i}\times T_{-i}) \cap \left[\Delta\left(\prod_{j\in I_0\setminus\{i\}}(S_j \times \mathbf{T}^{n-1}_{j,\Delta})\right)\right]^{H_i}\right).$$

Hence $T^*_{\theta_i,\Delta} = \cap_{n\geq 0} T^n_{\theta_i,\Delta}$ is closed as well. ■

Lemma 2 *For every $i \in I$, $\text{proj}_{\Theta_i}\mathbf{T}^*_{i,\Delta} = \Theta_i$, therefore there exists a function $\tau^0_i : \Sigma_i \to \mathbf{T}^*_{i,\Delta}$ such that $\text{proj}_{\Theta_i}\tau^0_i(\theta_i, s_i) = \theta_i$.*

Proof We prove below by induction that $T^n_{\theta_i,\Delta} \neq \emptyset$ for every i, θ_i, $n \geq 0$. By Lemma 1, $(T^n_{\theta_i})_{n\geq 0}$ is a decreasing sequence of nonempty closed subsets of a compact space. By the finite intersection property, $T^*_{\theta_i,\Delta} = \cap_{n\geq 0} T^n_{\theta_i,\Delta} \neq \emptyset$. Hence, for every i

$$\text{proj}_{\Theta_i}\mathbf{T}^*_{i,\Delta} = \{\theta_i \in \Theta_i : T^*_{\theta_i,\Delta} \neq \emptyset\} = \Theta_i.$$

Then we can define $\tau^0_i : \Sigma_i \to \mathbf{T}^*_{i,\Delta}$ as follows: for each $(\theta_i, s_i) \in \Sigma_i$ pick $t_{\theta_i} \in T^*_{\theta_i,\Delta}$ and let $\tau^0_i(\theta_i, s_i) = (\theta_i, t_{\theta_i})$.

Now we prove by induction that, for every i, θ_i and n, $T^n_{\theta_i,\Delta} \neq \emptyset$.

Basis step First, for every $i \in I$ and $\theta_i \in \Theta_i$, fix $\bar{t}_i \in T_i$ and $\nu_{\theta_i} \in \Delta_{\theta_i}$ arbitrarily. Also, let $\bar{t}_0 = t_0$. Now define an array of probability measures $\mu^0_{\theta_i} = (\mu^0_{\theta_i}(\cdot|\Sigma_{-i}(h)\times T_i))_{h\in H_i} \in [\Delta(\Sigma_{-i}\times T_{-i})]^{H_i}$ as follows: for every measurable subset $E_{-i} \subseteq \Sigma_{-i}\times T_{-i}$ and $h \in H_i$

$$\mu^0_{\theta_i}(E_{-i}|\Sigma_{-i}(h)\times T_{-i}) = \sum_{(\theta_{-i},s_{-i},):(\theta_{-i},s_{-i},\bar{t}_{-i})\in E_{-i}} \nu_{\theta_i}(\{(\theta_{-i}, s_{-i})\}|\Sigma_{-i}(h)).$$

It can be checked that $\mu^0_{\theta_i}$ is a CPS, that is, $\mu^0_{\theta_i} \in \Delta^{H_i}(\Sigma_{-i}\times T_{-i})$. Since $g_i : T_i \to \Delta^{H_i}(\Sigma_{-i}\times T_{-i})$ is onto (belief-completeness), there is some $t^0_{\theta_i} \in T_i$ such that $g_i(t^0_{\theta_i}) = \mu^0_{\theta_i}$. By construction, $f_i(t^0_{\theta_i}) = \nu_{\theta_i} \in \Delta_{\theta_i}$ thus, $t^0_{\theta_i} \in T^0_{\theta_i,\Delta} \neq \emptyset$.

As a matter of notation, let $t^n_{\theta_0} = t_0$ for every $\theta_0 \in \Theta_0$ and $n \geq 0$.

Induction step Now suppose that, for every i and θ_i, $T^n_{\theta_i,\Delta} \neq \emptyset$ and pick for each $\theta \in \Theta$ a profile $t^n_\theta \in \prod_{i\in I} T^n_{\theta_i,\Delta}$. For every $i \in I$ and θ_i define an array of probability measures $\mu^{n+1}_{\theta_i} = (\mu^{n+1}_{\theta_i}(\cdot|\Sigma_{-i}(h)\times T_i))_{h\in H_i} \in [\Delta(\Sigma_{-i}\times T_{-i})]^{H_i}$ as follows: for every measurable subset $E_{-i} \subseteq \Sigma_{-i}\times T_{-i}$ and $h \in H_i$

$$\mu^{n+1}_{\theta_i}(E_{-i}|\Sigma_{-i}(h)\times T_{-i}) = \sum_{(\theta_{-i},s_{-i},):\left(s_{-i},\theta_{-i},t^n_{\theta_{-i}}\right)\in E_{-i}} \nu_{\theta_i}(\{(\theta_{-i}, s_{-i})\}|\Sigma_{-i}(h)).$$

As in the basis step, it can be checked that $\mu^{n+1}_{\theta_i} \in \Delta^{H_i}(\Sigma_{-i}\times T_{-i})$. Since g_i is onto, there is some $t^{n+1}_{\theta_i} \in T_i$ such that $g_i(t^{n+1}_{\theta_i}) = \mu^{n+1}_{\theta_i}$. By construction, $f_i(t^{n+1}_{\theta_i}) = \nu_{\theta_i} \in \Delta_{\theta_i}$. Furthermore, for every $h \in H_i$ and $m \in \{0, ..., n\}$

$$g_{i,h}(t^{n+1}_{\theta_i})\left(\prod_{j\in I_0\setminus\{i\}}(S_j \times \mathbf{T}^m_{j,\Delta})\right) = \mu^{n+1}_{\theta_i}\left(\prod_{j\in I_0\setminus\{i\}}(S_j \times \mathbf{T}^m_{j,\Delta})|\Sigma_{-i}(h)\times T_{-i}\right) = 1$$

(the equalities hold by construction for $m = n$, then they also hold for $m \in \{0, ..., n\}$ because $\mathbf{T}^n_{j,\Delta} \subseteq \mathbf{T}^m_{j,\Delta}$). Since $T^{n+1}_{\theta_i,\Delta} =$

$$\left\{ t_i : (f_i(t_i) \in \Delta_{\theta_i}) \wedge (\forall m \in \{0, ..., n\}) \left(g_{i,h}(t_i) \left(\prod_{j \in I_0 \setminus \{i\}} (S_j \times \mathbf{T}^m_{j,\Delta}) \right) = 1 \right) \right\},$$

it follows that $t^{n+1}_{\theta_i} \in T^{n+1}_{\theta_i} \neq \emptyset$. ■

Lemma 3 *Fix compact restrictions Δ and maps $\tau_j : \Sigma_j \to \mathbf{T}^*_{j,\Delta}$ $(j \in I_0)$ such that $\text{proj}_{\Theta_j} \tau_j(\theta_j, s_j) = \theta_j$, where $\tau_0(\theta_0, s_0) = (\theta_0, t_0)$. Fix a player $i \in I$ and, for each θ_i a first-order CPS $\mu_{\theta_i} \in \Delta_{\theta_i}$. Then, for each θ_i, there exists an epistemic type $t_i \in T_i$ such that $(\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}$ and, for each $h \in H_i$, $g_{i,h}(t_i)$ has finite support and*

$$\begin{aligned} \forall(\theta_{-i}, s_{-i}) &\in \Sigma_{-i}, \\ g_{i,h}(t_i)\left(\{(s_{-i}, \tau_{-i}(\theta_{-i}, s_{-i}))\}\right) &= \mu_{\theta_i}(\{(\theta_{-i}, s_{-i})\} | \Sigma_{-i}(h)). \end{aligned} \tag{7}$$

Proof. Define a *candidate* CPS ν_i on $\Sigma_{-i} \times T_{-i}$ by setting

$$\nu_i\left(\{(s_{-i}, \tau_{-i}(\theta_{-i}, s_{-i}))\} | \Sigma_{-i}(h) \times T_{-i}\right) = \mu_{\theta_i}(\{(\theta_{-i}, s_{-i})\} | \Sigma_{-i}(h))$$

for every (θ_{-i}, s_{-i}) and $h \in H_i$, and extending the assignments by additivity. Axioms 1 and 2 follow immediately from the observation that the map $(\theta_{-i}, s_{-i}) \mapsto (s_{-i}, \tau_{-i}(\theta_{-i}, s_{-i}))$ yields an embedding of $\bigcup_{h \in H_i} \text{supp}\left[\mu_{\theta_i}(.|\Sigma_{-i}(h))\right]$ (a finite sub setset of Σ_{-i}) in $\Sigma_{-i} \times T_{-i}$, so that, for every $h \in H_i$, $\nu_i(.|\Sigma_{-i}(h) \times T_{-i})$ is indeed a probability measure on $\Sigma_{-i} \times T_{-i}$. By the same argument, ν_i must also satisfy Axiom 3, i.e. it must be a CPS; of course, each $\nu_i(.|\Sigma_{-i}(h) \times T_{-i})$ has finite support by construction. Since g_i is onto, there exists a type $t_i \in T_i$ such that

$$g_{i,h}(t_i)\left(\{(s_{-i}, \tau_{-i}(\theta_{-i}, s_{-i}))\}\right) = \nu_i\left(\{(s_{-i}, \tau_{-i}(\theta_{-i}, s_{-i}))\} | \Sigma_{-i}(h) \times T_{-i}\right)$$

for every $(\theta_{-i}, s_{-i}) \in \Sigma_{-i}$ and $h \in H_i$. To see that $(\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}$ note that by construction $f_i(t_i) = \mu_{\theta_i} \in \Delta_{\theta_i}$ and $g_{i,h}(t_i)(\Sigma_0 \times \prod_{j \in I \setminus \{i\}} (S_j \times \mathbf{T}^*_{j,\Delta})) = 1$ for each h, which implies $g_{i,h}(t_i)(\Sigma_0 \times \prod_{j \in I \setminus \{i\}} (S_j \times \mathbf{T}^n_{j,\Delta})) = 1$ for each h and n. ■

We can now prove our main result.

Proof of Theorem 1: To prove the first part, we rely on Lemma 2 and Lemma 3 to recursively define, for each $n \geq 0$, a profile of functions $\tau^n = (\tau^n_i : \Sigma_i \to \mathbf{T}^*_{i,\Delta})_{i \in I_0}$ such that, for each (θ, s), $\text{proj}_\Theta \tau^n(\theta, s) = \theta$ and $(s, \tau^n(\theta, s)) \in \text{CSB}^{n-1}(R \cap \text{B}^*([\Delta]))$ whenever $(\theta, s) \in \Sigma^n_\Delta$.

For the reader's convenience, we report below the conditions for surviving the $(n+1)$-th step of Δ-rationalizability:

For every $i \in I$ and $n \geq 0$, $(\theta_i, s_i) \in \Sigma^{n+1}_{i,\Delta}$ if and only if there exists a CPS $\mu_{\theta_i} \in \Delta_{\theta_i}$ such that

$$(\theta_i, s_i) \in \rho_i(\mu_{\theta_i}), \tag{8}$$

$$(\forall m \in \{0, \ldots, n\})(\forall h \in H_i)(\Sigma^m_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset \Rightarrow \mu_{\theta_i}(\Sigma^m_{-i,\Delta} | \Sigma_{-i}(h)) = 1). \tag{9}$$

The maps for Player 0 are trivial: $\tau_0^n(\theta_0, s_0) = (\theta_0, t_0)$ for every (θ_0, s_0) and n. As for the real players, let $(\tau_i^0 : \Sigma_i \to \mathbf{T}^*_{i,\Delta})_{i\in I}$ be any profile of functions such that $\text{proj}_{\Theta_i}\tau_i^0(\theta_i, s_i) = \theta_i$; such functions exist by Lemma 2. Next, assuming that $\tau^m = (\tau_i^m)_{i\in I_0}$ has been defined and satisfies $\text{proj}_\Theta \tau^m(\theta, s) = \theta$ for every $m = 0, ..., n$, define τ^{n+1} as follows: for each i and $\sigma_i \in \Sigma_i \backslash \Sigma^{n+1}_{i,\Delta}$ let $\tau_i^{n+1}(\sigma_i) = \tau_i^n(\sigma_i)$; for each $(\theta_i, s_i) \in \Sigma^{n+1}_{i,\Delta}$ there is first-order CPS $\mu_{\theta_i} \in \Delta_{\theta_i}$ such that eq.s (8),(9) hold and, by Lemma 3, there is an epistemic type $t^{n+1}_{\theta_i} \in T_i$ with $(\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}$ such that eq. (7) holds (for each $h \in H_i$) with $\tau_{-i} = (\tau_j^n)_{j\neq i}$, so that – in particular – $f_i(t^{n+1}_{\theta_i}) = \mu_{\theta_i}$; then let $\tau_i^{n+1}(\theta_i, s_i) = (\theta_i, t^{n+1}_{\theta_i})$. Clearly, $\text{proj}_\Theta \tau^{n+1}(\theta, s) = \theta$.

Claim 1 *For every $m \geq 0$ and $(\theta, s) \in \Sigma$,*

$$(\theta, s) \in \Sigma^{m+1}_\Delta \Rightarrow (s, \tau^{m+1}(\theta, s)) \in \text{CSB}^m(R \cap \text{B}^*([\Delta])), \tag{10}$$

$$((\exists t \in T)((\theta, s, t) \in \text{CSB}^m(R \cap \text{B}^*([\Delta]))) \Rightarrow (\theta, s) \in \Sigma^{m+1}_\Delta. \tag{11}$$

Eq. (10) implies

$$\Sigma^{m+1}_\Delta \subseteq \text{proj}_\Sigma \text{CSB}^m(R \cap \text{B}^*([\Delta])).$$

Eq. (11) implies

$$\text{proj}_\Sigma \text{CSB}^m(R \cap \text{B}^*([\Delta])) \subseteq \Sigma^{m+1}_\Delta.$$

Therefore the claim implies that

$$\Sigma^{m+1}_\Delta = \text{proj}_\Sigma \text{CSB}^m(R \cap \text{B}^*([\Delta]))$$

for every m.

Proof of the claim Recall that $\text{CSB}^m(R \cap \text{B}^*([\Delta]))$ is a Cartesian product. To ease notation, for each $i \in I$ and $h \in H_i$, we write

$$\text{CSB}^m_{-i}(R \cap \text{B}^*([\Delta])) = \text{proj}_{\Omega_{-i}} \text{CSB}^m(R \cap \text{B}^*([\Delta])),$$

Basis step Fix (θ, s) arbitrarily. Suppose that $(\theta, s) \in \Sigma^1_\Delta$. By construction of τ^1, for each $i \in I$, $(\theta_i, s_i) \in \rho_i(f_i(\text{proj}_{T_i}\tau_i^1(\theta_i, s_i)))$ and $\tau_i^1(\theta_i, s_i) \in \mathbf{T}^*_{i,\Delta}$. Therefore $(s, \tau(\theta, s)) \in R \cap \text{B}^*([\Delta])$. Since $R \cap \text{B}^*([\Delta]) = \text{CSB}^0(R \cap \text{B}^*([\Delta]))$, this proves (10) for $m = 0$. Next fix any $t \in T$ such that $(\theta, s, t) \in \text{CSB}^0(R \cap \text{B}^*([\Delta]))$. Then, for each $i \in I$, $(\theta_i, s_i) \in \rho_i(f_i(t_i))$ and $f_i(t_i) \in \Delta_{\theta_i}$; therefore $(\theta, s) \in \Sigma^1_\Delta$. This proves (11) for $m = 0$.

Induction step Assume that eq.s (10), (11) hold for each (θ, s) and $m = 0, ..., n-1$. Then, for each $m = 0, ..., n-1$, $i \in I$ and $h \in H_i$,

$$\Sigma_{-i}(h) \cap \Sigma^{m+1}_{-i,\Delta} \neq \emptyset \Leftrightarrow \Omega_{-i}(h) \cap \text{CSB}^m_{-i}(R \cap \text{B}^*([\Delta])) \neq \emptyset. \tag{12}$$

Fix (θ, s) arbitrarily. Suppose that $(\theta, s) \in \Sigma^{n+1}_\Delta$. By construction of τ^{n+1}, for every $i \in I$, $\tau_i^{n+1}(\theta_i, s_i) \in \mathbf{T}^*_{i,\Delta}$ and $(\theta_i, s_i) \in \rho_i(f_i(\text{proj}_{T_i}\tau_i^{n+1}(\theta_i, s_i)))$; therefore $(s, \tau^{n+1}(\theta, s)) \in R \cap \text{B}^*([\Delta])$. By the inductive hypothesis and the construction of τ^n, for every $i \in I$, $m \in \{0, ..., n-1\}$ and $(\theta'_{-i}, s'_{-i}) \in \Sigma_{-i}$

$$\left((\theta'_{-i}, s'_{-i}) \in \Sigma^{m+1}_{-i,\Delta}\right) \Rightarrow ((s'_{-i}, \tau^n_{-i}(\theta'_{-i}, s'_{-i})) \in \text{CSB}^m_{-i}(R \cap \text{B}^*([\Delta]))).$$

Hence, by construction of τ^{n+1} and eq. (12), for every $i \in I$, $m \in \{0, ..., n-1\}$ and $h \in H_i$,

$$\begin{array}{rcl} \left(\Omega_{-i}(h) \cap \mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta])) \neq \emptyset\right) & \Rightarrow & \\ g_{i,h}(\mathrm{proj}_{T_i}\tau_i^{n+1}(\theta_i, s_i))(\mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta])) & = & 1). \end{array}$$

Next note that

$$\mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta])) = R \cap \mathrm{B}^*([\Delta]) \cap \bigcap_{i \in I} \left\{ \bigcap_{m=0}^{n-1} \mathrm{SB}_i(\Omega_i \times [\mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta]))]) \right\}$$

$$= R \cap \mathrm{B}^*([\Delta]) \cap \Omega_0 \times$$

$$\prod_{i \in I} \left\{ \begin{array}{l} (\theta_i, s_i, t_i) \in \Sigma_i \times T_i : (\forall h \in H_i)\,(\forall m \in \{0, ..., n-1\}) \\ \left(\Omega_{-i}(h) \cap \mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta])) \neq \emptyset\right) \Rightarrow \left((g_{i,h}(t_i)\,(\mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta])) = 1)\right) \end{array} \right\}.$$

It follows that $(s, \tau^{n+1}(\theta, s)) \in \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta]))$, showing that eq. (10) holds for $m = n$.

Now fix any $t \in T$ such that $(\theta, s, t) \in \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta]))$. Then, for every $i \in I$, $(\theta_i, s_i) \in \rho_i(f_i(t_i))$, $f_i(t_i) \in \Delta_{\theta_i}$ and (by the inductive hypothesis) for every $m \in \{0, ..., n-1\}$ and $h \in H_i$

$$\left(\Omega_{-i}(h) \cap \mathrm{CSB}^m_{-i}(R \cap \mathrm{B}^*([\Delta])) \neq \emptyset\right) \Rightarrow f_{i,h}(t_i)(\Sigma^{m+1}_{-i,\Delta}) = 1.$$

By eq. (12), the formula above is equivalent to

$$\Sigma_{-i}(h) \cap \Sigma^{m+1}_{-i,\Delta} \neq \emptyset \Rightarrow f_{i,h}(t_i)(\Sigma^{m+1}_{-i,\Delta}) = 1$$

for every $m \in \{0, ..., n-1\}$ and $h \in H_i$. Therefore $(\theta, s) \in \Sigma^{n+1}_\Delta$, showing that eq. (11) holds with $m = n$. □

Next we prove the second part of the thesis: $\Sigma^\infty_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta]))$. Pick any $(\sigma, t) \in \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta]))$. Since, $\Sigma^{n+1}_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^n(R \cap \mathrm{B}^*([\Delta]))$ for every $n \geq 0$, we conclude that $\sigma \in \Sigma^n_\Delta$ for every $n \geq 1$; so $\sigma \in \bigcap_{n \geq 1} \Sigma^n_\Delta = \Sigma^\infty_\Delta$. Hence $\mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta])) \subseteq \Sigma^\infty_\Delta$. Now pick any $\sigma \in \Sigma^\infty_\Delta$ and consider the sequence of sets $K(m, \sigma) = \mathrm{CSB}^m(R \cap \mathrm{B}^*([\Delta])) \cap (\{\sigma\} \times T)$, $m \geq 1$. $\mathrm{B}^*([\Delta])$ is closed and (by standard arguments) R is closed as well. For every closed event E, i and $h \in H_i$, $\mathrm{B}_{i,h}$ is closed, therefore $\mathrm{CSB}^m(E)$ is also closed. $\{\sigma\} \times T$ is obviously closed. Therefore, each $K(m, \sigma)$ is a nonempty (because $\sigma \in \Sigma^{m+1}_\Delta = \mathrm{proj}_\Sigma \mathrm{CSB}^m(R \cap \mathrm{B}^*([\Delta]))$) and closed subset of the compact space $\Omega = \Sigma \times T$; also, the sequence of sets $K(m, \sigma)$ is decreasing, and hence has the finite intersection property. Then $\emptyset \neq \bigcap_{m \geq 1} K(m, \sigma) \subseteq \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta]))$ and $\sigma \in \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta]))$. Therefore $\Sigma^\infty_\Delta \subseteq \mathrm{proj}_\Sigma \mathrm{CSB}^\infty(R \cap \mathrm{B}^*([\Delta]))$. ■

Proof of Theorem 2 First observe that the set of states of player i in $\mathcal{T}_\Delta$ is

$$\Omega_{i,\Delta} = \mathrm{B}^*_i([\Delta]) = \{(\theta_i, s_i, t_i) \in \Sigma_i \times T_i : (\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}\},$$

where $\mathbf{T}^*_{i,\Delta}$ is the set of information/epistemic types defined at the beginning of this subsection.

Proof of eq. (3) Fix i, θ_i. By definition of $\mathcal{T}_\Delta$, the right hand side of eq. (3) is contained in the left hand side. To see that also the converse holds, fix $\mu_i \in \Delta_{i,\theta_i}$; by

Lemma 3 there is a type $t_i \in T_i$ such that $(\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}$ and $f_i(t_i) = \mu_i$. By construction, $f_{i,h}(t_i)$ $=\text{marg}_{\Sigma_{-i}} g_{i,h}(t_i) = \text{marg}_{\Sigma_{-i}} g_{i,h,\Delta}(t_i) = f_{i,h,\Delta}(t_i)$ for each $h \in H_i$ (the second equality holds because $t_i \in \text{proj}_{T_i} \mathbf{T}^*_{i,\Delta} = T_{i,\Delta}$ and $g_{i,\Delta}$ is the restriction of g_i on $T_{i,\Delta}$). Thus $f_{i,\Delta}(t_i) = \mu_i$ for some $t_i \in T_{i,\Delta}$ such that $(\theta_i, s_i, t_i) \in \Omega_{i,\Delta}$ for every s_i. □

Proof of eq. (4) *Basis step.* Each first-order belief map $f_{i,\Delta}$ is the restriction of f_i on $T_{i,\Delta}$ and $\Omega_\Delta = \text{B}^*([\Delta]) = \Omega_0 \times \prod_{i \in I} \{(\theta_i, s_i, t_i) \in \Sigma_i \times T_i : (\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}\}$; therefore

$$\begin{aligned} R_\Delta &= \{(\theta, s, t) \in \Omega_\Delta : (\forall i \in I)(s_i \in r_{\theta_i}(f_{i,\Delta}(t_i))\} \\ &= \{(\theta, s, t) \in \Omega : (\forall i \in I)((s_i \in r_{\theta_i}(f_i(t_i))) \wedge ((\theta_i, t_i) \in \mathbf{T}^*_{i,\Delta}))\} \\ &= R \cap \text{B}^*([\Delta]). \end{aligned}$$

Induction step. Suppose that $(\text{CSB}_\Delta)^n(R_\Delta) = \text{CSB}^n(R \cap \text{B}^*([\Delta]))$. Then

$$(\text{CSB}_\Delta)^{n+1}(R_\Delta) = (\text{CSB}_\Delta)^n(R_\Delta) \cap \text{SB}_\Delta((\text{CSB}_\Delta)^n(R_\Delta)) =$$

$$= \text{CSB}^n(R \cap \text{B}^*([\Delta])) \cap \text{SB}_\Delta(\text{CSB}^n(R \cap \text{B}^*([\Delta]))) =$$

$$= \text{CSB}^n(R \cap \text{B}^*([\Delta])) \cap$$

$$\cap \left\{ (\theta, s, t) \in \Omega_\Delta : \begin{array}{l} (\forall i \in I)(\forall h \in H_i)(\Omega_{-i,\Delta}(h) \cap \text{CSB}^n_{-i}(R \cap \text{B}^*([\Delta])) \neq 0 \\ \Rightarrow g_{i,h,\Delta}(t_i)(\text{CSB}^n_{-i}(R \cap \text{B}^*([\Delta]))) = 1) \end{array} \right\}$$

$$= \text{CSB}^n(R \cap \text{B}^*([\Delta])) \cap$$

$$\cap \left\{ (\theta, s, t) \in \Omega : \begin{array}{l} (\forall i \in I)(\forall h \in H_i)(\Omega_{-i}(h) \cap \text{CSB}^n_{-i}(R \cap \text{B}^*([\Delta])) \neq 0 \\ \Rightarrow g_{i,h}(t_i)(\text{CSB}^n_{-i}(R \cap \text{B}^*([\Delta]))) = 1) \end{array} \right\}$$

$$= \text{CSB}^n(R \cap \text{B}^*([\Delta])) \cap \text{SB}(\text{CSB}^n(R \cap \text{B}^*([\Delta]))) = \text{CSB}^{n+1}(R \cap \text{B}^*([\Delta])),$$

where the second equality follows from the induction hypothesis, the fourth holds because $\text{CSB}^n(R \cap \text{B}^*([\Delta])) \subseteq \Omega_\Delta$ and $g_{i,\Delta}$ is the restriction of g_i on $T_{i,\Delta}$, and the other equalities hold by definition. □

Proof of eq.s (5) and (6). Given eq. (4), (5) and (6) follow from Theorem 1. ■

Proof of Theorem 3. To shorten the proof, we take advantage of the result due to Battigalli & Siniscalchi [13, 2007] mentioned in the Introduction:

Theorem 5 *Fix a collection $\Delta = (\Delta_{i,\theta_i})_{i \in I, \theta_i \in \Theta_i}$ of compact subsets of first-order CPS's and a belief-complete type structure $\mathcal{T}$. Then, for every $n \geq 0$,*

$$\Sigma^{n+1}_\Delta = \text{proj}_\Sigma \text{CSB}^n(R \cap [\Delta])$$

and

$$\Sigma^\infty_\Delta = \text{proj}_\Sigma \text{CSB}^\infty(R \cap [\Delta]).$$

Sketch of proof of Theorem 5. By inspection of the proof of the main characterization result in [13, 2007], it is clear that a separate lemma[47] shows the equivalence of naïve Δ-rationalizability ($\hat{\Sigma}^n_\Delta$) and Δ-rationalizability (Σ^n_Δ) under the assumption that Δ is regular, whereas the rest of the proof shows that Theorem 5 holds. This is done within

[47] Lemma 7 of [13, 2007], which is a special case of our Proposition 1.

the observable-actions framework, but it is clear that the arguments of the rest of the proof are unaffected by considering the more general framework of this paper. □

Theorems 1 and 5 imply

$$\mathrm{proj}_{\Sigma}\mathrm{CSB}^{n-1}(R\cap \mathrm{B}^*([\Delta])) = \Sigma^n_\Delta = \mathrm{proj}_{\Sigma}\mathrm{CSB}^{n-1}(R\cap [\Delta])$$

for each $n \geq 1$, including $n = \infty$. An induction argument shows that all the claimed inclusions hold. Indeed,

$$\mathrm{CSB}^0(R\cap \mathrm{B}^*([\Delta])) = R\cap \mathrm{B}^*([\Delta]) \subseteq R\cap[\Delta] = \mathrm{CSB}^0(R\cap[\Delta]).$$

Suppose that $\mathrm{CSB}^n(R\cap \mathrm{B}^*([\Delta])) \subseteq \mathrm{CSB}^n(R\cap[\Delta])$. Recall from subsection 3.2 that, for each $i \in I$ and $E \subseteq \Omega$, we let $[E]_{-i} = \Omega_i \times \mathrm{proj}_{\Omega_{-i}}E$.

$$\mathrm{CSB}^n_{-i}(R\cap \mathrm{B}^*([\Delta])) := \Omega_i \times \mathrm{proj}_{\Omega_{-i}}\mathrm{CSB}^n_{-i}(R\cap[\Delta]),$$

$\mathrm{CSB}^n_{-i}(R\cap[\Delta])$ is similarly defined. Since $[\mathrm{CSB}^n(R\cap \mathrm{B}^*([\Delta]))]_{-i}$ and $[\mathrm{CSB}^n_{-i}(R\cap[\Delta])]_{-i}$ have the same projection on Σ_{-i}, they are consistent with the same information sets of Player i, even if $[\mathrm{CSB}^n(R\cap \mathrm{B}^*([\Delta]))]_{-i} \subseteq [\mathrm{CSB}^n_{-i}(R\cap[\Delta])]_{-i}$. Let H^n_i denote the collection of such information sets; then, by the inductive hypothesis and monotonicity of the $\mathrm{B}_{i,h}$ operators,

$$\mathrm{SB}_i(\mathrm{CSB}^n_{-i}(R\cap \mathrm{B}^*([\Delta]))) = \bigcap_{h\in H^n_i} \mathrm{B}_{i,h}(\mathrm{CSB}^n_{-i}(R\cap \mathrm{B}^*([\Delta])))$$

$$\subseteq \bigcap_{h\in H^n_i} \mathrm{B}_{i,h}(\mathrm{CSB}^n_{-i}(R\cap[\Delta])) = \mathrm{SB}_i(\mathrm{CSB}^n_{-i}(R\cap[\Delta])).$$

Recall that, for each E, $\mathrm{CSB}^{n+1}(E) = E\cap\bigcap_{i\in I}\mathrm{SB}_i(\Omega_i\times\mathrm{proj}_{\Omega_{-i}}E)$. Therefore, given the above definitions and inclusions,

$$\mathrm{CSB}^{n+1}(R\cap \mathrm{B}^*([\Delta])) = \mathrm{CSB}^n(R\cap \mathrm{B}^*([\Delta]))\cap\bigcap_{i\in I}\mathrm{SB}_i(\mathrm{CSB}^n_{-i}(R\cap \mathrm{B}^*([\Delta])))$$

$$\subseteq \mathrm{CSB}^n(R\cap[\Delta])\cap\bigcap_{i\in I}\mathrm{SB}_i(\mathrm{CSB}^n_{-i}(R\cap[\Delta])) = \mathrm{CSB}^{n+1}(R\cap[\Delta]).$$

■

5.3 Equivalence of Δ-rationalizability and naïve Δ-rationalizability

We begin with a preliminary result about closedness under compositions.

Definition 11 *A finite sequence* $(K^m_{-i},\mu^m)^n_{m=1} \in (2^{\Sigma_{-i}}\times\Delta^{H_i}(\Sigma_{-i}))^n$ *is* admissible *if* $\{K^m_{-i}\}^n_{m=1}$ *is a decreasing sequence of product sets with* $K^m_0 = \Sigma_0$ *(*$K^m_{-i} = \Sigma_0\times\prod_{j\in I\setminus\{i\}}K^m_j$*) and, for every* m, μ^m *strongly believes* K^m_{-i} *(for each* $h\in H_i$, $K^m_{-i}\cap\Sigma_{-i}(h)\neq\emptyset \Rightarrow \mu^m(K^m_{-i}|\Sigma_{-i}(h)) = 1$*).*

Definition 12 *Fix an admissible sequence* $(K^m_{-i},\mu^m)^n_{m=1}$ *and let* $m(h) = \max(\{1\}\cup\{m : K^m_{-i}\cap\Sigma_{-i}(h)\neq\emptyset\})$. *A system of beliefs* $\mu = (\mu(\cdot|\Sigma_{-i}(h))_{h\in H_i} \in [\Delta(\Sigma_{-i})]^{H_i}$ *is the* composition *of* $(K^m_{-i},\mu^m)^n_{m=1}$ *if*

$$\forall h\in H_i,\ \mu(\cdot|\Sigma_{-i}(h)) = \mu^{m(h)}(\cdot|\Sigma_{-i}(h)).$$

Lemma 4 *The composition μ of an admissible sequence $(K^m_{-i}, \mu^m)^n_{m=1}$ is a CPS such that $\mu(K^n_{-i}|\Sigma_{-i}(h)) = 1$ for each $h \in H_i$ with $K^n_{-i} \cap \Sigma_{-i}(h) \neq \emptyset$.*

Proof. We only have to prove that μ satisfies the chain rule, the rest follows immediately by inspection of Definition 12. Fix $g, h \in H_i$ and E_{-i} so that $E_{-i} \subseteq \Sigma_{-i}(h) \subseteq \Sigma_{-i}(g)$ (g is a prefix of h); then $m(g) \geq m(h)$. If $m(g) = m(h)$ then

$$\begin{aligned}
\mu(E_{-i}|\Sigma_{-i}(g)) &= \mu^{m(g)}(E_{-i}|\Sigma_{-i}(g)) \\
&= \mu^{m(g)}(E_{-i}|\Sigma_{-i}(h))\mu^{m(g)}(\Sigma_{-i}(h)|\Sigma_{-i}(g)) \\
&= \mu^{m(h)}(E_{-i}|\Sigma_{-i}(h))\mu^{m(g)}(\Sigma_{-i}(h)|\Sigma_{-i}(g)) \\
&= \mu(E_{-i}|\Sigma_{-i}(h))\mu(\Sigma_{-i}(h)|\Sigma_{-i}(g)).
\end{aligned}$$

If $m(g) > m(h)$ then, by definition of $m(g)$, $K^{m(g)}_{-i} \cap \Sigma_{-i}(h) = \emptyset$ and $K^{m(g)}_{-i} \cap \Sigma_{-i}(g) \neq \emptyset$. By admissibility of the sequence, $\mu^{m(g)}(K^{m(g)}_{-i}|\Sigma_{-i}(g)) = 1$. Therefore $\mu^{m(g)}(E_{-i}|\Sigma_{-i}(g)) = \mu^{m(g)}(\Sigma_{-i}(h)|\Sigma_{-i}(g)) = 0$ and

$$\begin{aligned}
\mu(E_{-i}|\Sigma_{-i}(g)) &= \mu^{m(g)}(E_{-i}|\Sigma_{-i}(g)) = 0 \\
&= \mu^{m(h)}(E_{-i}|\Sigma_{-i}(h)) \times 0 \\
&= \mu^{m(h)}(E_{-i}|\Sigma_{-i}(h))\mu^{m(g)}(\Sigma_{-i}(h)|\Sigma_{-i}(g)) \\
&= \mu(E_{-i}|\Sigma_{-i}(h))\mu(\Sigma_{-i}(h)|\Sigma_{-i}(g)).
\end{aligned}$$

■

Lemma 5 *Let $\Delta_i \subseteq \Delta^{H_i}(\Sigma_{-i})$ be closed under compositions and fix an admissible sequence $(K^m_{-i}, \mu^m)^n_{m=1}$ such that $\mu^m \in \Delta_i$ for each m. Then also the composition of $(K^m_{-i}, \mu^m)^n_{m=1}$ gives a CPS in Δ_i.*

Proof. The statement is true by definition for admissible sequences of length 2. Assume by way of induction that the result holds for admissible sequences of length $n - 1 \geq 2$ and consider an admissible sequence of length n, viz. $(K^m_{-i}, \mu^m)^n_{m=1}$. Let μ be the composition of the (admissible) prefix $(K^m_{-i}, \mu^m)^{n-1}_{m=1}$. By the inductive hypothesis, $\mu \in \Delta_i$. The pair $(K^{n-1}_{-i}, \mu; K^n_{-i}, \mu^n)$ is admissible since $K^n_{-i} \subseteq K^{n-1}_{-i}$ and μ is a composition of $(K^m_{-i}, \mu^m)^{n-1}_{m=1}$ so that, in particular, μ is a CPS such that $K^{n-1}_{-i} \cap \Sigma_{-1}(h) \neq \emptyset$ implies $\mu(K^{n-1}_{-i}|\Sigma_{-i}(h)) = 1$ for each $h \in H_i$ (Lemma 4). Now let μ^* be the composition of $(K^{n-1}_{-i}, \mu; K^n_{-i}, \mu^n)$. Since Δ_i is closed under compositions, $\mu^* \in \Delta_i$. By construction, μ^* is the composition of $(K^m_{-i}, \mu^m)^n_{m=1}$. ■

Proof of Proposition 1. The statement is obvious for $n = 1$. Now pick $n \geq 1$ and assume by way of induction that the statement it is true for each positive integer up to n. We have to show that $\Sigma^{n+1}_{i,\Delta} = \hat{\Sigma}^{n+1}_{i,\Delta}$.

If $\sigma_i = (\theta_i, s_i) \in \Sigma^{n+1}_{i,\Delta}$ then there is some $\mu_{\theta_i} \in \Delta_{\theta_i}$ satisfying (8)-(9); since $\Sigma^{n+1}_{i,\Delta} \subseteq \Sigma^n_{i,\Delta}$, then by the inductive hypothesis $\sigma_i \in \hat{\Sigma}^n_{i,\Delta} = \Sigma^n_{i,\Delta}$; moreover, since $\hat{\Sigma}^n_{-i,\Delta} = \Sigma^n_{-i,\Delta}$ (again by the inductive hypothesis), (9) implies $\hat{\Sigma}^n_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset \Rightarrow \mu_{\theta_i}(\hat{\Sigma}^n_{-i,\Delta}|\Sigma_{-i}(h)) = 1$. We conclude that $\sigma_i \in \hat{\Sigma}^{n+1}_{i,\Delta}$.

In the other direction, suppose $\sigma_i = (\theta_i, s_i) \in \hat{\Sigma}^{n+1}_{i,\Delta}$. Then also $\sigma_i \in \hat{\Sigma}^m_{i,\Delta}$ for $m = 1, \ldots, n$, so we can find CPS's $\mu^m_{\theta_i} \in \Delta_{\theta_i}$, $m = 1, \ldots, n$, such that, for each m, $\sigma_i \in \rho_i(\mu^m_{\theta_i})$

and, $\forall h \in H_i$, $\hat{\Sigma}^m_{-i,\Delta} \cap \Sigma_{-i}(h) \neq \emptyset$ implies $\mu^m_{\theta_i}(\hat{\Sigma}^m_{-i,\Delta}|\Sigma_{-i}(h)) = 1$. Therefore the sequence $(\hat{\Sigma}^m_{-i,\Delta}, \mu^m_{\theta_i})^n_{m=1}$ is admissible. Now let μ_{θ_i} be the composition of this sequence. Clearly $\sigma_i \in \rho_i(\mu_{\theta_i})$ and μ_{θ_i} satisfies (9). Moreover, by Lemma 5 $\mu_{\theta_i} \in \Delta_{\theta_i}$. Therefore $\sigma_i \in \Sigma^{n+1}_{i,\Delta}$. ■

References

[1] Aradillas-Lopez, A. and E. Tamer, "The Identification Power of Equilibrium in Simple Games," *Journal of Business & Economics Statistics* **26** (2008), 261-283.

[2] Aumann R.J., "Correlated Equilibrium as an Expression of Bayesian Rationality", *Econometrica* **55** (1987), 1-18.

[3] Battigalli, P., "Strategic Rationality Orderings and the Best Rationalization Principle," *Games and Economic Behavior* **13** (1996), 178-200.

[4] Battigalli, P., "On Rationalizability in Extensive Form Games," *Journal of Economic Theory* **74** (1997), 40-61.

[5] Battigalli, P., "Rationalizability in Incomplete Information Games," EUI working papers, ECO 99/17 (1999).

[6] Battigalli, P., "Rationalizability in Infinite Dynamic Games with Incomplete Information," *Research in Economics* **57**, (2003), 1-38.

[7] Battigalli, P., "Rationalization in Signaling Games: Theory and Applications," *International Game Theory Review* **8** (2006), 1-27.

[8] Battigalli, P. and A. Friedenberg, "Context-Dependent Forward Induction Reasoning," IGIER w.p. 351 (2009) Bocconi University.

[9] Battigalli, P. and A. Friedenberg, "Forward Induction Reasoning Revisited," *Theoretical Economics* **7** (2012), 57-98.

[10] Battigalli, P. and M. Siniscalchi, "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory* **88** (1999), 188-230.

[11] Battigalli, P. and M. Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory* **106** (2002), 356-391.

[12] Battigalli, P. and M. Siniscalchi, "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, **3** (2003) (1), Article 3 [http://www.bepress.con/bejte/advances/vol3/iss1/art3].

[13] Battigalli, P. and M. Siniscalchi, "Interactive Epistemology in Games with Payoff Uncertainty," *Research in Economics*, **61** (2007), 165-184.

[14] Battigalli, P. and P. Tebaldi, "Interactive Epistemology in Simple Dynamic Games with a Continuum of Strategies," typescript (2012).

[15] Battigalli, P., A. Di Tillio and D. Samet, "Strategies and Interactive Beliefs in Dynamic Games," IGIER w.p. n. 375 (2011), forthcoming in *Advances in Economics and Econometrics* (D. Acemoglu, M. Arellano, E. Dekel Eds.)

[16] Battigalli, P., A. Di Tillio, E. Grillo and A. Penta, "Interactive Epistemology and Solution Concepts for Games with Asymmetric Information," *The B.E. Journal of Theoretical Economics*, **11** (2011): Iss. 1 (Advances), Article 6.

[17] Bergemann, D. and S. Morris, "Robust Mechanism Design: An Introduction," typescript (2011), Princeton and Yale Universities.

[18] Brandenburger, A. and E. Dekel, "Rationalizability and Correlated Equilibria", *Econometrica* **55** (1987)**,** 1391-1402.

[19] Brandenburger, A. and E. Dekel, Hierarchies of Beliefs and Common Knowledge, *Journal of Economic Theory,* **59** (1993), 189-198.

[20] Brandenburger, A. and A. Friedenberg, "Self-Admissible Sets," *Journal of Economic Theory*, **145** (2010), 785-811.

[21] Brandenburger, A., A. Friedenberg, and H.J. Keisler, "Admissibility in Games," *Econometrica* **76** (2008), 307-352.

[22] Cappelletti, G., "A note on rationalizability and restrictions on beliefs," *B.E. Journals in Theoretical Economics*, **10** (2010): Iss. 1 (Topics), Article 40.

[23] Catonini, E., "Non-Binding Agreements and Forward Induction Reasoning," typescript (2012), Bocconi University.

[24] Cho, I.K. and D. Kreps, "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics*, **102** (1987) 179-222.

[25] Dekel, E., D. Fudenberg, and S. Morris, "Interim Correlated Rationalizability," *Theoretical Economics*, **2**, 15-40.

[26] De Vito, N., "Complete Conditional Type Structures," typescript (2012), Bocconi University.

[27] Friedenberg, "When Do Type Structures Contain All Hierarchies of Beliefs?", *Games and Economic Behavior*, **68** (2010), 108-129.

[28] Fudenberg, D. and J. Tirole, *Game Theory.* MIT Press, Cambridge MA, 1991.

[29] Harsanyi, J., Games of incomplete information played by Bayesian players. Parts I, II, III, *Management Science* **14** (1967-68), 159-182, 320-334, 486-502.

[30] Mertens, J.-F. and S. Zamir, Formulation of Bayesian Analysis for Games with Incomplete Information, *International Journal of Game Theory* **14** (1985), 1-29.

[31] Mueller, C., "Robust Virtual Implementation Under Common Strong Belief in Rationality," typescript (2012), Carnegie Mellon University.

[32] Osborne, M. and A. Rubinstein, *A Course in Game Theory*, MIT Press, Cambridge MA, 1994.

[33] Pearce, D., Rationalizable Strategic Behavior and the Problem of Perfection, *Econometrica* **52** (1984), 1029-1050.

[34] Renyi, A., On a New Axiomatic Theory of Probability, *Acta Mathematica Academiae Scientiarum Hungaricae* **6** (1955), 285-335.

[35] Schweinzer, P. and M. Shimoji, "Implementation without Incentive Compatibility: Two Stories with Partially Informed Planners," typescript, University of York, 2012.

[36] Tan, T., and S. Werlang, "The Bayesian Foundation of Solution Concepts of Games," *Journal of Economic Theory* **45** (1988), 370-391.

[37] Tebaldi, P., "Non-Binding Agreements in Dynamic games," MSc Thesis (2011), Bocconi University.