Rationalizability in infinite, dynamic games with incomplete information

Pierpaolo Battigalli

Istituto di Economia Politica, Bocconi University, Via Gobbi 5, 20136 Milan, Italy

Received 2 January 2002; accepted 26 February 2002

Abstract

In this paper, we analyze two nested iterative solution procedures for infinite, dynamic games of incomplete information. These procedures do not rely on the specification of a type space à la Harsanyi. Weak rationalizability is characterized by common certainty of rationality at the beginning of the game. Strong rationalizability also incorporates a notion of forward induction. The solutions may take as given some exogenous restrictions on players’ conditional beliefs. In dynamic games, strong rationalizability is a refinement of weak rationalizability. Existence, regularity properties, and equivalence with the set of iteratively interim undominated strategies are proved under standard assumptions. The analysis mainly focus on two-player games with observable actions, but we show how to extend it to n-player games with imperfectly observable actions. Finally, we briefly survey some applications of the proposed approach.

© 2003 Elsevier Science Ltd. All rights reserved.

JEL classification: C72; D82

Keywords: Incomplete information; Rationalizability; Forward induction

1. Introduction and overview

In a n-person game of incomplete information some of the crucial elements governing strategic interaction—such as individual feasibility constraints, how actions are mapped into consequences and individual preferences over consequences—are represented by a vector of parameters θ which is (partially) unknown to some players. For the sake of simplicity, let us assume that θ determines the shape of each player’s payoff function and that it can be partitioned into subvectors θ₁, …, θₙ whereby each
player $i = 1,\ldots,n$ knows $\theta_i$. We call $\theta$ the state of Nature and $\theta_i$ the private information or payoff-type of player $i$. The form of the parametric payoff functions $u_i(\cdot, \theta)$—or, more generally, the form of the mapping associating each conceivable state of Nature $\theta$ to the ‘true’ (but unknown) game $G(\theta)$—is assumed to be common knowledge. In this paper, we take this mapping $\theta \mapsto G(\theta)$ as the fundamental description of a strategic situation with incomplete information and we put forward and analyze solution concepts associating to any such mapping a set of possible outcomes. Our approach is related to, but different from Harsanyi’s (1967–1968) seminal paper on incomplete information games. Harsanyi’s Bayesian model is now so entrenched in the literature that only a handful of ‘pure’ game theorists still pay attention to its subtleties. In order to motivate and better understand our contribution it is useful to go through Harsanyi’s model in some detail.1

1.1. Harsanyi’s Bayesian model

As Harsanyi noticed, one way to provide a Bayesian analysis of incomplete information games is to endow each player with a hierarchy of beliefs, that is, (i) a subjective probability measure on the set of conceivable states of Nature, or first-order belief, (ii) a subjective probability measure on the set of conceivable first-order beliefs of his opponents, or second-order beliefs, and so on. In principle, a complete description of every relevant attribute of a player should include, not only his payoff-type, but also his epistemic type, that is, an infinite hierarchy of beliefs. Furthermore, (infinitely) many hierarchies of beliefs could be attached to a given payoff-type. This hierarchies-of-beliefs approach is mathematically feasible (see e.g. Mertens and Zamir, 1985), but it does not seem to provide a tractable framework for a direct analysis of incomplete information games. Its usefulness consists mainly in providing a theoretical framework for the analysis of the epistemic foundations of game-theoretic solution concepts.2

Harsanyi’s (1967–1968) contribution was twofold. On the one hand, he put forward a general notion of ‘type space’ which provides an implicit, but relatively parsimonious description of infinite hierarchies of beliefs. On the other hand, he showed how to analyze incomplete information games with the standard tools of game theory. A type space can be defined as follows. For each player $i$ and each payoff-type $\theta_i \in \Theta_i$ ($\Theta_i$ is the set of $i$’s conceivable payoff-types) we add a parameter $e_i$ corresponding to a purely epistemic component of player $i$’s attributes. In general, different values of $e_i$ can be attached to a given payoff-type $\theta_i$. This way we obtain a set $T_i \subseteq \Theta_i \times E_i$ of possible attributes, or Harsanyi-types, of player $i$. A Harsanyi-type encodes the payoff-type and the epistemic type of a player. In fact, the beliefs of any given player $i$ about his opponents’ payoff-types as well as their own beliefs are determined by a function $p_i : T_i \rightarrow \Delta(T_{-i})$, where $T_{-i} = \prod_{j \neq i} T_j$. Note that the array of conditional probabilities $(p_i(t_i))_{t_i \in T_i}$ can always be derived from some ‘prior’, i.e. there is at least one probability measure $P_i \in \Delta(T_{-i} \times T_i)$ such that $p_i(t_i) = P_i(\cdot | t_i)$, but such a ‘prior’ does not represent $i$’s beliefs in a hypothetical ex ante stage, it is only a technical device to express the belief function $p_i(\cdot)$. It is assumed that the vector of

1 For thorough discussion of the Bayesian model see Harsanyi (1995), Gul (1998), and Dekel and Gul (1997).

2 See Dekel and Gul (1997), Battigalli and Bonanno (1999) and the references therein.
functions \((p_1, ..., p_n)\) is common knowledge. Therefore every \(t_i \in T_i\) corresponds to an infinite hierarchy of beliefs: the first-order belief \(p_1^0(t_i)\) is simply the marginal of \(p_i(t_i)\) on \(\Theta_{-i}\), the \((k + 1)\)-order belief implicit in \(t_i\) is derived from \(p_i(t_i)\) and knowledge of the \(n - 1\) functions \(p_j^0(\cdot), j \neq i\), mapping the opponents’ Harsanyi-types into \(k\)-order beliefs. When we add a type space on top of the map \(\theta \mapsto G(\theta)\) we obtain a Bayesian game. A Bayesian equilibrium is a vector of behavioral rules \(b_i : T_i \rightarrow S_i\) (\(i = 1, ..., n\), \(S_i\) is the strategy set for player \(i\)) such that for each player \(i\) and each Harsanyi-type \(t_i = (\theta_i, e_i)\), strategy \(s_i = b_i(\theta_i, e_i)\) maximizes \(i\)'s expected payoff given the payoff-type \(\theta_i\), the subjective belief \(p_i(\theta_i, e_i)\) and the \((n - 1)\) tuple of functions \(b_{-i}\). Note that, for any fixed vector of behavioral rules, a vector of Harsanyi-types \((t_1, ..., t_n)\) provides an implicit, but complete description of every relevant aspect of the world: the state of Nature, each player’s subjective beliefs about the state of Nature and his opponents’ behavior and each player’s subjective beliefs about his opponents beliefs. In other words, once we fix a type space and an equilibrium, we obtain a fully fledged epistemic model of the game, i.e. a model specifying the possible interactive beliefs concerning both payoff-types and players’ choices.

Within this framework, the players’ situation in a game of incomplete information is formally similar to the interim stage of a game with complete, but imperfect and asymmetric information whereby \(t_i\) represents the private information of player \(i\) about the realization of an initial chance move, such as the cards player \(i\) has been dealt in a game of poker. Harsanyi pushed the analogy even further by assuming that all the subjective beliefs \(p_i(t_i)\) (\(i = 1, ..., n\), \(t_i \in T_i\)) can be derived from a common prior \(P \in \Delta(\prod_{j=1}^n T_j)\) so that \(p_i(t_i) = P(\cdot | t_i)\). In this case, Bayesian equilibrium simply corresponds to a Nash equilibrium of a companion game with an imperfect information about a fictitious chance move selecting the vector of attributes according to probability measure \(P\). This is the so-called ‘random vector model’ of the Bayesian game.\(^3\) From the point of view of equilibrium analysis, we can equivalently associate to the given Bayesian game a companion game with complete information whereby for each player/role \(i = 1, ..., n\) there is a population of potential players characterized by the different attributes \(t_i \in T_i\). An actual player is drawn at random from each population \(i\) to play the game. The joint distribution of attributes in the \(n\) populations is given by the common prior \(P\). This is the ‘prior lottery model’ of the Bayesian game.

1.2. Drawbacks of standard Bayes–Nash equilibrium analysis

Harsanyi’s analysis of incomplete information games has offered invaluable insights to economic theorists and applied economists, but its success should not make us overlook some potential drawbacks of this approach and of its standard applications to economic models. These potential drawbacks are all related to the following facts: (a) a Bayesian game provides only an implicit and (in general) non-exhaustive—or
non-universal—representation of the conceivable epistemic types; (b) representing a Bayesian game with the ‘random vector model’ or the ‘prior lottery model’ blurs the fundamental distinction between games with genuine incomplete information and games with imperfect, asymmetric information: in the former there is no ex ante stage at which the players analyze the situation before receiving some piece of information selected at random.

(a) Non-transparent assumptions about beliefs. We mentioned that for every Harsanyi-type in a Bayesian game we can derive a corresponding infinite hierarchy of beliefs. The derivation makes sense if it is assumed that the Bayesian game is common knowledge. Mertens and Zamir (1985) shows that this informal assumption is without loss of generality because (i) the space of n-tuples of (consistent) infinite hierarchies of beliefs is a well-defined type space in the sense of Harsanyi and (ii) every type space is essentially a belief-closed subspace of the space of infinite hierarchies of beliefs, which is therefore a universal type space. This means that the class of all Bayesian models is sufficiently rich, but whenever we consider a particular (non-universal) model, or a subclass of models, we rule out some epistemic types. If, on top of this, we add the equilibrium hypothesis that players’ conjectures about their opponents’ behavioral rules are correct, we end up making assumptions about players’ interactive beliefs, which are often questionable and—due to the implicit representation of epistemic types—non-transparent.

For example, ‘agreement’ and ‘no-trade’ results hold for Bayesian models satisfying the common prior assumption, but the meaning of this assumptions as a restriction on players’ hierarchies of beliefs is not obvious. For the sake of tractability, applied economists often restrict their attention to an even smaller class of Bayesian models by assuming that there is a one-to-one correspondence between payoff-types and Harsanyi-types. These strong and yet only implicit assumptions about players’ hierarchies of beliefs may affect the set of equilibrium outcomes in an important way. But we have a hard time reducing these assumptions to more primitive and transparent axioms.

(b1) No ex ante stage and plausibility of assumptions about beliefs. The formal similarity between Bayesian games and games with asymmetric information may be misleading. We are quite ready to accept that in the ‘random vector model’ players assign the same prior probabilities to chance moves. Similarly, assuming a common probability measure over players’ attributes is meaningful and plausible, if not compelling, in the ‘prior lottery model.’ For example, it can be justified by assuming that the statistical

---

4 If we regard the Bayesian game itself as a subjective model of a given player, then we have to assume that this player is certain that everybody shares the same model (cf. Harsanyi, 1967–68).

5 See also Brandenburger and Dekel (1993) and references therein. Battigalli and Siniscalchi (1999a) provides analogous results for infinite hierarchies of systems of conditional beliefs in dynamic games of incomplete information.

6 As clarified in Battigalli and Siniscalchi (2001b), it is the interaction between restricted type spaces and the equilibrium assumption that yields restrictions on behavior beyond those implied by common certainty of rationality.

7 For more on this see, for example, Gul (1998) and Dekel and Gul (1997). Bonanno and Nehring (1999) ‘makes sense’ of the common prior assumption in incomplete information games, characterizing it as a very strong ‘agreement’ property.

8 For a discussion of the common prior assumption in situations with asymmetric, but complete information see Morris (1995).
distribution of characteristics in the population of potential players is commonly known. But in games with genuine incomplete information there is no ex ante stage and prior probabilities are only a convenient, but unnecessary notational device to specify players’ infinite hierarchies of beliefs. Thus, the common prior assumption and the conflation of payoff-types and Harsanyi-types are much harder to accept.

(b2) No ex ante stage and learning. The lack of an ex ante stage also makes the equilibrium concept more problematic. A Nash equilibrium of a given ‘objective’ game $G$ may by interpreted as a stationary state of a learning process as the players repeatedly play $G$. Furthermore, it is possible to provide sufficient conditions such that learning eventually induces a Nash equilibrium outcome. We cannot provide a similar justification for equilibria of Bayesian games representing genuine incomplete information. Let $\theta$ be the actual state of Nature in a game of incomplete information $\Gamma$ and recall that $G(\theta)$ denotes the ‘true objective game’ corresponding to $\theta$. Let us assume that the players interact repeatedly. By the very nature of the problem we are considering, we have to assume that the state of Nature $\theta$ is fixed once and for all at the beginning of time rather than being drawn at random according to some i.i.d. process. By repeatedly playing $G(\theta)$ the players can learn (at most) to play a Nash equilibrium of $G(\theta)$, not a Bayesian equilibrium of (some Bayesian game based on) $\Gamma$.10,11

1.3. Rationalizable outcomes of incomplete information games

To summarize what we said so far, in order to analyze an economic model with incomplete information $\Gamma$ using Harsanyi’s approach we have to specify a type space based on $\Gamma$ and then look for the Bayesian equilibria of the resulting Bayesian game. The specification of the type space is hardly related to the fundamentals of the economic problem and yet may crucially affect the set of equilibrium outcomes. This raises several related theoretical questions. Can we analyze incomplete information games without specifying a type space? Can we provide an independent justification for the Bayesian equilibrium concept? Which results of the Bayesian analysis are independent of the exact specification of the type space? Is it possible to provide a relatively simple characterization of the set of all Bayesian equilibrium outcomes?

9 In general, convergence is not guaranteed and, even if the play converges, the limit outcome is a self-confirming (or conjectural) equilibrium, which need not be equivalent to a Nash equilibrium. See Fudenberg and Levine (1998) and references therein.

10 More generally, their pattern of behavior may converge to what Battigalli and Guaitoli (1997) call ‘a conjectural equilibrium at $\theta$’ which need not correspond to a Nash equilibrium of $G(\theta)$.11

11 Dekel et al. (2001) shows that the Bayes–Nash equilibrium concept is very hard to justify in terms of learning even for games with asymmetric information where the ex ante stage is real, but players have subjective heterogeneous priors on the state of Nature, which is drawn at random in each repetition according to and i.i.d. process. The reason is rather obvious: even if in equilibrium conjectures about opponents’ behavioral rules are correct, the subjective probabilities assigned to opponent’s actions may be incorrect. If there is enough ex post monitoring, the players will eventually find it out and revise their beliefs about Nature moves. If there is little ex post monitoring, there is no reason why players should come to have correct conjectures about their opponents behavioural rules.

12 By ‘fundamentals’ we mean the conceivable configurations of technologies and tastes, corresponding to the states of Nature.
The answer to these questions can be found in the literature on rationalizability. Let us consider complete information games first, i.e. games with only one conceivable state of Nature. The set of rationalizable strategies in a static game with complete information is obtained by an iterative deletion procedure which (in two-person games) coincides with iterated strict dominance (Pearce, 1984). Rationalizability exactly characterizes the strategies consistent with common certainty of rationality (Tan and Werlang, 1988) and also the set of subjective correlated equilibrium outcomes (Brandenburger and Dekel, 1987). Note that, according to the terminology used so far, a subjective correlated equilibrium is simply a Bayesian equilibrium of a model with a unique state of Nature and hence with payoff-irrelevant Harsanyi-types.

This paper puts forward and analyzes some notions of rationalizability for games with genuine incomplete information, but the proposed solutions are also relevant for games with asymmetric information where the statistical distribution of attributes in the population of potential players is not known. We focus mainly on the analysis of dynamic games, where players can signal their types and strategic intent. But the basic idea is more easily understood if we consider static games first. Consider the following procedure: (Basis Step) For every player \( i \), payoff-type \( u_i \) and strategy \( s_i \) in \( G \), we check whether \( s_i \) can be justified as a feasible best response for \( u_i \) to some probabilistic beliefs about the opponents’ payoff-types and behavior. If the pair \( (u_i, s_i) \) does not pass this test it is ‘removed.’ (Inductive Step) For every \( i \), \( u_i \) and \( s_i \) we check whether \( s_i \) is a feasible best response for \( u_i \) to some probabilistic beliefs about the opponents assigning probability zero to the (vectors of) pairs \( (\theta_{-i}, s_{-i}) \) removed so far. Note that (epistemic) type spaces are not mentioned. The procedure depends only on the ‘fundamentals’ of the economic model. Not surprisingly, this solution is equivalent to an iterative ‘interim’ dominance procedure. Furthermore, it turns out that it exactly characterizes the set of all possible equilibrium outcomes of the Bayesian games based on \( G \) (Battigalli and Siniscalchi, 2001b). It is also easy to provide an epistemic characterization à la Tan and Werlang (1988) of the rationalizable outcomes as those consistent with common certainty of rationality (see Battigalli and Siniscalchi (1999a) in the context of dynamic games).

Let us see how the solution procedure works in a textbook example. Consider a Cournot duopoly with one-sided incomplete information. The inverse demand schedule \( P(Q) \) is linear and firms have constant marginal costs. The marginal cost firm 1, \( c_1 \), is common knowledge, but \( c_2 \) the marginal cost of firm 2, is unknown to firm 1. The range of conceivable values of \( c_2 \) is a closed interval strictly contained in \([0, P(0)]\) and containing \( c_1 \) in its interior. Both firms are expected profit maximizers. Fig. 1 shows the reaction functions for firm 1 \( (r_1(q_2)) \), for the most efficient type of firm 2 \( (r_2(\theta, q_1)) \), and for the least efficient type of firm 2 \( (r_2(\bar{\theta}, q_1)) \). In this model, there is no loss of generality in considering only best responses to deterministic beliefs.\(^{13}\) The first step of the rationalizability procedure eliminates, for each type of each firm, all the outputs above the monopolistic choice (e.g. \( r_2(\theta_2, 0) \) for type \( \theta_2 \) of firm 2), which is the best response to the most optimistic conjecture about the opponent (assuming that the opponent might also be irrational). In fact, all the eliminated outputs are strictly dominated for type \( \theta_i \) by the monopolistic choice of \( \theta_i \), while the remaining outputs are best responses to some conjecture. In

\(^{13}\) This is true in a large class of games. See Proposition 3.9.
the second step of the procedure we eliminate, for each type of each firm, all the outputs below the best response to the most pessimistic conjecture consistent with rationality of the opponent. For example, for firm 1 we eliminate all the outputs below \( r_1(r_2(\hat{\theta}, 0)) \). In the third step, we eliminate, e.g. for firm 1, all the outputs above \( r_1(r_2(\hat{\theta}, r_1(0))) \), which is the best response to the most optimistic conjecture consistent with the opponent being rational and certain that everybody is rational. In the limit we obtain a set of rationalizable outcomes represented by the rectangle ABCD in Fig. 1.

Let us compare rationalizable outcomes and standard Bayesian equilibrium outcomes. The standard Bayesian model specifies the belief of player 1 about \( u_2 \), say \( \pi \in \Delta(Q_2) \). It is assumed that it is common knowledge that \( \pi \) indeed represents the belief of player 1. The Bayesian equilibrium strategy for player 1 is given by the intersection between the graph of \( r_1(\cdot) \) and the graph of \( r_2(E(\hat{\theta}; \pi), \cdot) \), where \( E(\hat{\theta}; \pi) \) denotes the expected value of \( \hat{\theta} \) given \( \pi \). The set of Bayesian equilibrium outcomes for all possible \( \pi \in \Delta(Q_2) \) is the parallelogram \( A'B'C'D \) in Fig. 1. But if we consider all the possible specifications of a type space à la Harsanyi, the set of Bayesian equilibrium outcomes coincides with the set of rationalizable outcomes.14

The procedure described above is relevant if we do not want to rule out any conceivable epistemic type. However, it may be plausible to assume that players’ beliefs satisfy some qualitative restrictions. The iterative solution concept can be easily modified to accommodate restrictions on first-order beliefs (informally) assumed to be commonly known. In the general definition of the solution procedure these exogenous restrictions on players’ beliefs are parametrically given.

---

14 Battigalli and Siniscalchi (2001b) shows how to construct a type space such that, in the resulting Bayesian game, each rationalizable outcome is a Bayesian equilibrium outcome. Here we provide a simpler example. Assume that there are two epistemic types for each payoff-type. Thus \( T_1 = \{t_1^1, t_1^2\} \) and \( T_2 = \Theta_2 \times \{e_2^1, e_2^2\} \). Assume \( p_1(t_1^i) \) is degenerate on \( (\theta, e_2^j) \), \( p_1(t_1^i) \) is degenerate on \( (\theta, e_2^j) \), and \( p_2(\theta, e_2^j) \) assigns probability one to \( t_1^i \) for all \( \theta \) and \( j \). (These belief functions are consistent with a ‘correlated’ common prior.) In the Bayesian equilibrium where type \( t_1^i(t_2^j) \) chooses the lowest (highest) rationalizable output for firm 1, all the points in the vertical segments AD and BC are equilibrium outcomes.
The analysis of incomplete information games is particularly interesting when they have a dynamic structure, because in this case a player can make inferences about the types and/or strategic intents of his opponents by observing their behavior in previous stages of the game. As in the complete information case, there are several possible definitions of the rationalizability solution concept for dynamic games, corresponding to different assumptions about how players would update their beliefs if they observed unexpected behavior. Here we consider two nested solution concepts for (possibly infinite) multi-stage games with incomplete information, called weak rationalizability and strong rationalizability. Rigorous axiomatizations of these solution concepts involve the definition of extensive-form epistemic models and are given elsewhere (Ben Porath, 1997; Battigalli and Siniscalchi, 1999a,b, 2001a). Intuitively, weak rationalizability simply assumes that players choose sequential best responses to their systems of conditional beliefs, updating via Bayes rule whenever possible, and this is common certainty at the beginning of the game. On top of this, strong rationalizability also assumes that each player keeps believing that his opponents are rational even when they behave in an unexpected way, provided that their behavior can somehow ‘rationalized’ (a more detailed account is provided in Section 3). Thus, unlike weak rationalizability, strong rationalizability incorporates a forward induction criterion.

1.4. Related literature

The solution concepts developed in this paper extend notions of rationalizability for extensive-form games with complete information put forward and analyzed by Pearce (1984), Battigalli (1996, 1997) and Ben Porath (1997). The idea of using some notion of rationalizability to analyze games of incomplete information is a quite natural development of Bernheim (1984) and Pearce’s (1984) work on complete information games and it appears in some papers in the literature (although several papers take for granted the common prior assumption and/or identify payoff-types and Harsanyi-types). Battigalli and Guaitoli (1997) analyzes the extensive-form rationalizable paths of a simple macroeconomic game with incomplete information and no common prior. This paper also puts forward a notion of conjectural (or self-confirming) equilibrium at a given state of Nature of an incomplete information game. Battigalli and Siniscalchi (2001b) relate rationalizability in incomplete information games to Bayesian equilibria and to the iterated intuitive criterion. Cho (1994) and Watson (1998) use a notion of subform rationalizability to analyze dynamic bargaining with incomplete information. Watson (1993, 1996) obtains reputation and/or cooperation results for perturbed repeated games under mild restrictions on players’ beliefs. Perry and Reny (1999) consider some specific social choice problems with incomplete information and propose extensive-form mechanisms to introduce desirable outcomes in iteratively undominated strategies. Rabin (1994) proposes to combine rationalizability and exogenous restrictions on players’ beliefs to introduce behavioral assumptions in game-theoretic analysis. A different approach to incomplete information games is proposed in Sákovics (2001). He considers Bayesian models with finite hierarchies of beliefs and puts forward a novel solution concept, called ‘mirage equilibrium.’
In the last section, we will comment more specifically on a number of applications of
our approach.

1.5. Structure of the paper

The rest of the paper is organized as follows. Section 2 contains the game-theoretic set
up. Weak and strong rationalizability are defined and analyzed in Section 3 focusing on
two-person games with observable actions. Existence and regularity properties are proved
for a class of ‘simple’, but possibly infinite games. We also extend to the present
framework some known results relating rationalizability and iterative dominance. Section 4
shows how the analysis can be extended to \( n \)-person games with imperfectly observable
actions. Section 5 briefly reports on a number of applications of the proposed
methodology. Section 6 concludes. The appendix contains some details about infinite
dynamic games of incomplete information and all the proofs.

2. Game-theoretic framework

2.1. Games of incomplete information with observable actions

A game of incomplete information with observable actions is a structure
\[ \Gamma = \left( N, (\Theta_i)_{i \in N}, (A_i)_{i \in N}, \mathcal{H}^*(\cdot), (u_i)_{i \in N} \right) \]
given by the following elements:\footnote{The following model generalizes Fudenberg and Tirole (1991, pp 331–332) and Osborne and Rubinstein (1994, pp 231–232). The Appendix A provides further details.}

- \( N \) is a non-empty, finite set of players.
- For each \( i \in N \), \( \Theta_i \subseteq \mathbb{R}^{m_i} \) is a non-empty set of possible types for player \( i \) and \( A_i \subseteq \mathbb{R}^{n_i} \) is a non-empty set of possible actions for player \( i \) (\( \mathbb{R}^k \) is the \( k \)-dimensional Euclidean space).
- Let \( \Theta = \prod_{i \in N} \Theta_i \) and \( A = \prod_{i \in N} A_i \). Then
  \[ A^* = \{ \phi \} \cup \left( \bigcup_{t=1}^{\infty} A^t \right), \]
  that is, \( A^* \) is the set of finite and countably infinite sequences of action profiles, including the empty sequence \( \phi \), and

\[ \mathcal{H}^*(\cdot) : \Theta \rightarrow 2^{A^*} \]

\( (2^{A^*}) \) is the power set of \( A^* \) is a non-empty valued correspondence assigning to each
profile of types \( \theta \) the set \( \mathcal{H}^*(\theta) \) of feasible histories given \( \theta \). For every history \( h \in \mathcal{H}^*(\theta) \) one can derive the set \( A(\theta, h) = \prod_{i \in N} A_i(\theta_i, h) \) of feasible action profiles. A
history \( h \in H^*(\theta) \) is terminal at \( \theta \) if \( A(h, \theta) = \emptyset \) (every infinite feasible history is
terminal). We let

\[ \mathcal{H}(\theta) = \{ h \in A^* : A(\theta, h) \neq \emptyset \}, \]

\[ \mathcal{H}(\theta_i) = \{ h \in A^* : \exists \theta_{-i} \in \Theta_{-i}, A((\theta_i, \theta_{-i}), h) \neq \emptyset \}, \]

\[ \mathcal{H} = \bigcup_{\theta \in \Theta} \mathcal{H}(\theta), \]

respectively, denote the set of feasible non-terminal histories at \( \theta \), or for \( \theta_i \), and the set of a priori feasible non-terminal histories.

- Define the set \( \mathcal{Z} \) of outcomes as follows:\(^{16}\)

\[ \mathcal{Z} = \{ (\theta, h) : h \in \mathcal{H}^*(\theta), A(\theta, h) = \emptyset \}. \]

For all \( i \in N \),

\[ u_i : \mathcal{Z} \rightarrow \mathbb{R} \]

is the payoff function for player \( i \) (\( \mathbb{R} \) denotes the set of real numbers).

Parameter \( \theta \) represents player \( i \)’s private information about the feasibility constraints and payoffs. For brevity, we call \( \theta_i \) the ‘payoff-type’ of player \( i \). It is assumed that \( \Gamma \) is common knowledge. The array \( \theta = (\theta_i)_{i \in N} \) is interpreted as a state of Nature; it completely specifies the unknown parameters of the game and the players’ interactive knowledge about them. Player \( i \) at \( (\theta, h) \) knows \( (\theta_i, h) \) and whatever can be inferred from history \( h \) given that \( \Gamma \) (hence \( \mathcal{H}^*(\cdot) \)) is common knowledge. Chance moves and residual uncertainty about the environment can be modeled by having a pseudo-player \( c \in N \) with a constant payoff function. The ‘type’ \( \theta_c \) of this pseudo-player represents the residual uncertainty about the state of Nature which would remain after pooling the private information of the real players. Players’ common or heterogeneous beliefs about chance moves can be modeled as exogenous restrictions on beliefs (see below).

Game \( \Gamma \) is static if for all \( \theta \in \Theta \) and \( a \in A(\theta, \phi) \), \( (a) \) is a terminal history at \( \theta \). Game \( \Gamma \) has private values if, for all \( i \in N \), \( u_i(\theta_i, \theta_{-i}, \cdot) \) is independent of \( \theta_{-i} \). A player of type \( \theta_i \) is active at history \( h \) if \( A_i(\theta_i, h) \) contains at least two elements. \( \Gamma \) has no simultaneous moves if for every state of Nature \( \theta \) and every history \( h \in \mathcal{H}(\theta) \) there is only one active player. In this case, \( \Gamma \) can be represented by an extensive form with decision nodes \( (\theta, h) \), \( \theta \in \Theta \), \( h \in \mathcal{H}(\theta) \) (pairs \( (\theta, \phi) \) are the initial nodes of the arborescence) and information sets for player \( i \) of the following form:

\[ I(\theta_i, h) = \{ (\theta_i, \theta_{-i}, h) : h \in \mathcal{H}(\theta_i, \theta_{-i}) \}, \]

where \( \theta_i \) is active at \( h \). Game \( \Gamma \) has (incomplete but) perfect information if it has no simultaneous moves and \( \mathcal{H}^*(\theta) \) is independent of \( \theta \).\(^ {17}\)

Note that the basic model \( \Gamma \) does not specify players’ beliefs about the state of Nature \( \theta \). This is what makes \( \Gamma \) different from the standard notion of a Bayesian game. As mentioned in Section 1, if we want to provide a general (albeit implicit) representation of players’

\(^{16}\) The feasibility correspondence is such that, if \( ((\theta_i, \theta_{-i}), h) \in Z \), then \( ((\theta_i, \theta_{-i}), h) \in Z \), for all \( \theta_{-i} \).

\(^{17}\) In this case, \( \Gamma \) can also be represented by a game tree (with decision nodes \( h \in \mathcal{H} \)) featuring perfect information and payoff functions \( \nu_i : \Theta \times Z \rightarrow \mathbb{R} \), where \( Z \) is the set of terminal nodes.
beliefs about the state of Nature and of their hierarchies of beliefs, we have to embed each set $\Theta$ in a possibly richer set $T_i$ of ‘Harsanyi-types’ and specify belief functions $p_i : T_i \to \Delta(T_{-i})$.

Turning to the topological properties of $T$, we endow $A^*$ and $Z$ with the standard ‘discounting’ metrics (see Appendix A) and throughout the paper we rely on the following assumption:

**Assumption 0.** $A$ and $\Theta$ are closed, $H^*(\cdot)$ is a continuous correspondence and, for all $i \in N$, $u_i$ is a continuous function.

### 2.2. Strategic forms

A feasible strategy for type $\theta_i$ is a function $s_i : \mathcal{H} \to A_i$ such that $s_i(h) \in A_i(\theta_i, h)$ for all $h \in \mathcal{H}(\theta_i)$.\(^{18}\) The set of feasible strategies for type $\theta_i$ is denoted $S_i(\theta_i)$ and

$$S_i = \bigcup_{\theta \in \Theta} S_i(\theta)$$

denotes the set of a priori feasible strategies. (By definition of $\mathcal{H}$, for all $h \in \mathcal{H}$, $A_i(\theta_i, h)$ is nonempty. Therefore $S_i(\theta_i)$ is also nonempty.)

The basic elements of our analysis are feasible type-strategy pairs: $(\theta_i, s_i)$ is a feasible pair if $s_i \in S_i(\theta_i)$. A generic feasible pair for player $i$ is denoted $s_i$ and the set of such feasible pairs for player $i$ is the graph of the correspondence $S_i(\cdot) : \Theta_i \to 2^\mathcal{H}$, i.e.

$$\Sigma_i := \{(\theta_i, s_i) : \theta_i \in \Theta_i \times A_i : s_i \in S_i(\theta_i)\}$$

The sets of profiles of feasible pairs for all players and for the opponents of a player $i$ are, respectively, $\Sigma = \prod_{i \in N} \Sigma_i$ and $\Sigma_{-i} = \prod_{i \neq i} \Sigma_i$. Each profile $\sigma = [(\theta_i, s_i)]_{i \in N}$ induces a terminal history $\zeta(\sigma) \in \mathcal{H}(\theta)$ and hence an outcome $\zeta^* (\sigma) = (\theta_i, \zeta(\theta)) \in \mathcal{Z}$. Therefore, for each player $i$, we can derive the following strategic form payoff function:

$$U_i = u_i + \zeta^* : \Sigma \to R.$$  

Furthermore, for each a priori feasible history $h \in \mathcal{H}$ we can define the set of profiles of feasible pairs consistent with $h$:

$$\Sigma(h) = \{ \sigma \in \Sigma : h \text{ is a prefix of } \zeta(\sigma) \}.$$  

Clearly, $\Sigma(\phi) = \Sigma$. We let $\Sigma_i(h)$ denote the projection of $\Sigma(h)$ on $\Sigma_i$, that is, the set of $(\theta_i, s_i)$ such that strategy $s_i$ is feasible for type $\theta_i$ and does not prevent history $h$. It can be easily checked that, for all $h \in \mathcal{H}$,

$$\Sigma(h) = \prod_{i \in N} \Sigma_i(h) \neq \emptyset.$$  

The information of player $i$ about his opponents at history $h$ is represented in strategic form by $\Sigma_{-i}(h)$, the projection of $\Sigma(h)$ on $\Sigma_{-i}$.

We endow the sets $\Sigma_i$ ($i \in N$) with the standard metrics derived from the metric on $\mathcal{Z}$ (see Appendix A).

\(^{18}\) We let the domain of $s_i$ be $\mathcal{H}$ (instead of $H(\theta_i)$) only for notational simplicity.
Lemma 2.1. For all \( h \in \mathcal{H} \), \( \Sigma_i(h) \) is closed.

2.3. Conditional beliefs

Players’ beliefs in dynamic games can be represented as systems of conditional probabilities. Let \( \Sigma \) be a metric space with Borel sigma-algebra \( \mathcal{S} \). Fix a nonempty collection of subsets \( \mathcal{B} \subseteq \mathcal{S}\setminus\{\emptyset\} \), to be interpreted as ‘relevant hypotheses.’

Definition 2.2. (cf. Rényi, 1955) A conditional probability system (or CPS) on \( (\Sigma, \mathcal{S}, \mathcal{B}) \) is a mapping

\[
\mu(\cdot|\cdot) : \mathcal{S} \times \mathcal{B} \to [0, 1]
\]
satisfying the following axioms:

Axiom 1. For all \( B \in \mathcal{B} \), \( \mu(B|B) = 1 \).

Axiom 2. For all \( B \in \mathcal{B} \), \( \mu(\cdot|B) \) is a probability measure on \( (\Sigma, \mathcal{S}) \).

Axiom 3. For all \( A \in \mathcal{S}, B, C \in \mathcal{B}, A \subseteq B \subseteq C \Rightarrow \mu(A|B)\mu(B|C) = \mu(A|C) \).

The set of probability measures on \( (\Sigma, \mathcal{S}) \) is denoted by \( \Delta(\Sigma) \); the set of conditional probability systems on \( (\Sigma, \mathcal{S}, \mathcal{B}) \) can be regarded as a subset of \( [\Delta(\Sigma)]^{\mathcal{B}} \) (the set of mappings from \( \mathcal{B} \) to \( \Delta(\Sigma) \)) and it is denoted by \( \Delta^{\mathcal{B}}(\Sigma) \). The topology on \( \Sigma \) and \( \mathcal{S} \) (the smallest sigma-algebra containing this topology) are understood and need not be explicit in our notation. It is also understood that \( \Delta(\Sigma) \) is endowed with the topology of weak convergence of measures and \( [\Delta(\Sigma)]^{\mathcal{B}} \) is endowed with the product topology.

A relatively simple way to represent the beliefs of a player \( i \) in a dynamic game with incomplete information is to consider the set \( \Delta^{\mathcal{B}_i}(\Sigma_{-i}) \) of conditional probability systems on \( (\Sigma_{-i}, \mathcal{S}_{-i}, \mathcal{B}_i) \), where \( \Sigma_{-i} \) is the set of type-strategy profiles for his opponents, \( \mathcal{S}_{-i} \) is the Borel sigma algebra of \( \Sigma_{-i} \), and

\[
\mathcal{B}_i = \{ B \subseteq \Sigma_{-i} : \exists h \in \mathcal{H}, B = \Sigma_{-i}(h) \}
\]
is the family of ‘strategic-form information sets’ for player \( i \). By Lemma 2.1, \( \mathcal{B}_i \) is a collection of closed subsets and thus \( \Delta^{\mathcal{B}_i}(\Sigma_{-i}) \) is indeed a well-defined space of conditional probability systems.

Two points are worth discussing. (1) In a situation of incomplete information, when player \( i \) forms his beliefs he already knows his private information \( \theta_i \). Therefore it would be more germane to the analysis of incomplete information games to consider the set \( \Delta^{\mathcal{B}_i}(\Sigma_{-i}) \) of conditional beliefs for type \( \theta_i \), where

\[
\mathcal{B}_i(\theta_i) = \{ B \subseteq \Sigma_{-i} : \exists h \in \mathcal{H}(\theta_i), B = \Sigma_{-i}(h) \}.
\]

(2) A player also has beliefs about himself and they may be relevant when we discuss the epistemic foundations of a solution concept. Once again, we do not explicitly consider such beliefs for notational simplicity. This does not alter the analysis in any essential way. Our representation of a player’s beliefs and our game theoretic analysis are consistent with the following epistemic assumption: at a state of the world where player \( i \) is type \( \theta_i \) and \( i \)’s plan is \( s_i \in \mathcal{S}(\theta_i) \), player \( i \) would be certain of \( \theta_i \) at each history \( h \in \mathcal{H}(\theta_i) \) and would be certain to follow plan \( s_i \) at each history \( h \) consistent with \( s_i \).
An element of $\Delta^i(\Sigma_{-i})$ only describes the first-order conditional beliefs of player $i$. Only such beliefs are explicit in the game-theoretic analysis of this paper, but the motivations and epistemic foundations of the solution concepts to be proposed below at least implicitly consider higher order beliefs. Battigalli and Siniscalchi (1999a) shows how to construct infinite hierarchies of conditional beliefs which represent the epistemic type of a player, that is, the beliefs that this player would have, conditional on each history, about the state of Nature, his opponents’ strategies and his opponents’ epistemic types. This construction allows one to define formal notions of conditional common certainty and strong belief which are informally used in this paper to motivate and clarify the proposed solution concepts. Formal epistemic characterizations of solution concepts in terms of infinite hierarchies of conditional beliefs can be found in Battigalli and Siniscalchi (1999a, b, 2001a).

2.4. Sequential rationality

A strategy $\hat{s}_i$ is sequentially rational for a player of type $\hat{\theta}_i$ with conditional beliefs $\mu^i$ if it maximizes the conditional expected utility of $\hat{\theta}_i$ at every history $h$ consistent with $\hat{s}_i$. Note that this is a notion of rationality for plans of actions\(^{20}\) rather than strategies (see for example, Reny, 1992). Let

$$\mathcal{H}(\theta_i, s_i) = \{ h \in \mathcal{H}(\theta_i) : (\theta_i, s_i) \in \Sigma_i(h) \}$$

and

$$S_i(\theta_i, h) = \{ s_i \in S_i(\theta_i) : (\theta_i, s_i) \in \Sigma_i(h) \}$$

respectively denote the set of histories consistent with $(\theta_i, s_i)$ and the set of strategies consistent with $(\theta_i, h)$. Given a CPS $\mu^i \in \Delta^i(\Sigma_{-i})$ and a history $h \in \mathcal{H}(\theta_i, s_i)$, let

$$U_i(\theta_i, s_i, \mu^i(\cdot|\Sigma_{-i}(h))) = \int_{\Sigma_i(h)} U(\theta_i, s_i, \sigma_{-i}) \mu^i(d\sigma_{-i}|\Sigma_{-i}(h))$$

denote the expected payoff for type $\theta_i$ from playing $s_i$ given $h$, provided that the integral on the right hand side is well-defined.\(^{21}\)

**Definition 2.3.** A strategy $\hat{s}_i$ ($i = 1, 2, \ldots$) is sequentially rational for type $\hat{\theta}_i$ with respect to beliefs $\mu^i \in \Delta^i(\Sigma_{-i})$, written $(\hat{\theta}_i, \hat{s}_i) \in \rho_i(\mu^i)$ or equivalently $\hat{s} \in r_i(\hat{\theta}_i, \mu^i)$, if for all $h \in \mathcal{H}(\hat{\theta}_i, \hat{s}_i)$ where player $i$ is active and all $s_i \in S_i(\hat{\theta}_i, h)$ the following inequality is well-defined and satisfied:

$$U_i(\hat{\theta}_i, \hat{s}_i, \mu^i(\cdot|\Sigma_{-i}(h))) \geq U_i(\hat{\theta}_i, s_i, \mu^i(\cdot|\Sigma_{-i}(h))).$$

**Lemma 2.4.** If $S_i(\hat{\theta}_i)$ is compact and $U_i(\hat{\theta}_i, s_i, \sigma_{-i})$ is upper-semicontinuous in $s_i$, bounded and measurable in $\sigma_{-i}$, then $r_i(\hat{\theta}_i, \mu^i) \neq \emptyset$.

\(^{20}\)Formally, a plan of action is a maximal set of strategies consistent with the same histories and prescribing the same actions at such histories.

\(^{21}\)Even in well-behaved games (e.g. the Ultimatum Game with a continuum of offers), for some choices of $\mu^i$ and/or $s_i$, the strategic form payoff function $U_i$ is not integrable.
2.5. Exogenous restrictions on beliefs

A player’s beliefs may be assumed to satisfy some restrictions that are not implied by mutual or common belief in rationality. We call such restrictions *exogenous*, although they may be related to some structural properties of the model. We may distinguish between (i) restrictions on beliefs about the state of Nature and chance moves and (ii) restrictions on beliefs about behavior. Our general theory and the applications mentioned in Section 5 consider both (i) and (ii). Some examples of restrictions of the first kind are the following:

- Some ‘objective probabilities’ of chance moves might be known or satisfy some known restrictions such as positivity or independence across nodes.\(^22\)
- It may be common belief that all the opponents’ payoff-types are considered possible a priori by each player (cf. Dekel and Wolinsky, 2001; Siniscalchi, 1998, Ch. 5). Or it may be common belief that the prior probability of a ‘crazy type’ \(\theta^*_i\) committed to play a strategy \(s^*_i\) (either because \(s^*_i\) is dominant for \(\theta^*_i\) or because \(S_i(\theta^*_i) = \{\theta^*_i\}\)) is either positive or bounded below by a given positive number \(e_i(\theta^*_i)\). This kind of restriction is considered in the analysis of reputation by Battigalli and Watson (1997) and Battigalli (2001).

The following are examples of restrictions of the second kind:

- Specific structural properties of the game such as stationarity or monotonicity may be somehow reflected in players’ beliefs. Stationarity restrictions are considered in Cho’s (1994) analysis of the Coase’s conjecture. Restrictions related to monotonicity play a role in the analysis of signaling (Battigalli, 2000) and the analysis of rationalizable bidding in first price auctions with interdependent values (Battigalli and Siniscalchi, 2001c).
- In a first price auction, it may be common belief that every bid strictly above the reservation price yields a positive probability of winning the object, this implies that a rational player whose valuation is above the reservation price would never bid (weakly) above his valuation or (weakly) below the reservation price (cf. Battigalli and Siniscalchi, 2001c).
- It may be common belief that each player’s beliefs about the types and strategies of different opponents satisfy stochastic independence (Battigalli and Siniscalchi, 1999b).
- It may be common belief that each player’s conditional beliefs have countable support (Watson, 1996; Battigalli, 2000).
- It may be common belief that each player’s first-order beliefs agree with a given distribution over the set of outcomes \(\mathcal{Z}\) (Battigalli and Siniscalchi, 2001b).

In general, we assume that, for each state of Nature \(\theta\), the conditional probability system of each player \(i\) belongs to a given, nonempty subset \(\Delta^i\). In order to make sense of

\(^{22}\) Börgers (1991) considers perturbed games with ‘small trembles’ whereby the true trembling probabilities are unknown, but it is common belief that the actual choice is very likely to coincide with the intended choice. He stresses the difference between correlated and uncorrelated trembles.
the solution concepts discussed in the next section it is sufficient (but not necessary) to assume that the restrictions \((\mathcal{D}_i)_{i \in \mathbb{N}}\) are ‘common knowledge’ in the following sense: for every sequence of players and histories \((i_1, h_1, \ldots, i_l, h_l, i_{l+1})\) player \(i_1\) would be certain at \(h_1\) that \(\ldots\) player \(i_l\) would be certain at \(h_l\) that the first-order CPS of player \(i_{l+1}\) belongs to \(\Delta^{l+1}\). Weaker sufficient epistemic assumptions are discussed in the next section.

3. Weak and strong \(\Delta\)-rationalizability

In this section we define and analyze two nested extensions of the rationalizability solution concept to dynamic games of incomplete information, which take as given some exogenous restrictions on players’ beliefs represented by sets of CPSs \(\Delta' \subseteq \Delta^i_{\mathcal{B}_i}(\Sigma^{-i}), i \in \mathbb{N}\). Weak rationalizability is an extension of a solution concept put forward and analyzed by Ben Porath (1997) for games of perfect and complete information. Strong rationalizability is a generalization of the notion of extensive-form rationalizability proposed by Pearce (1984) and further analyzed by Battigalli (1996, 1997) (see also Reny (1992)). We focus mainly on two-person games (i.e. \(N = \{1,2\}\)) to avoid discussing the issue of correlated vs independent beliefs, which would distract the readers’ attention from more important points. The analysis is extended to \(n\)-person games in Section 4. The two solution concepts are defined by procedures which iteratively eliminate feasible type-strategy pairs. These procedures coincide on the class of static games. Epistemic assumptions are crucial for the motivation of these solution concepts, but a formal epistemic analysis is beyond the scope of this paper and is provided elsewhere. Nevertheless, we will be explicit and clear about the epistemic assumptions underlying each solution concept.

A given state of the world describes the state of Nature (hence each player’s private information) and the players’ dispositions to act and to believe conditional on each history, that is, their strategies and their infinite hierarchies of conditional beliefs. Let \(\Delta = (\Delta')_{i \in \mathbb{N}}\). Each \(\Delta\)-rationalizability solution concept characterizes the feasible type-strategy realized at states where (a) every player \(i \in \mathbb{N}\) is sequentially rational and has first-order beliefs in \(\Delta^i\); and (b) the players’ higher order conditional beliefs satisfy conditions concerning mutual certainty of (a) and/or robustness of beliefs about (a).

3.1. Weak \(\Delta\)-rationalizability

Weak \(\Delta\)-rationalizability characterizes the set of feasible type-strategy pairs realized at states of the world where all the following events are true:\(^{25}\)

\(^{23}\) See also Dekel and Fundenberg (1990), Brandenburger (1992), Börgers (1994) and Gul (1996).

\(^{24}\) Ben Porath (1997) analyzes weak rationalizability using finite, non-universal, extensive form type spaces. Battigalli and Siniscalchi (1999a) analyzes universal and non-universal type spaces for dynamic games of incomplete information and provides epistemic characterizations of solution concepts. Battigalli and Siniscalchi (1999b, 2001a) uses an extensive-form, universal (or belief-complete) type space to provide an epistemic characterization of strong \(\Delta\)-rationalizability with correlated and independent beliefs.

\(^{25}\) The conditions are indexed by the assumed order of mutual certainty of rationality.
(0) every player $i$ has first-order conditional beliefs in $\Delta_i^t$ and is sequentially rational, (W1) every player $i$ is certain of (0) at the beginning of the game (i.e. conditional on $\phi$), (W2) every player $i$ is certain of (W1) at the beginning of the game, 

(Wk) every player $i$ is certain of (W$k - 1$)) at the beginning of the game, 

**Definition 3.1.** Let $W_i(0, \Delta) = \Sigma_i, i = 1, 2$. Assume that the subsets $W_i(k, \Delta), i = 1, 2$, have been defined, $k = 0, 1, \ldots$. Then for each $i = 1, 2$, $W_i(k + 1, \Delta)$ is the set of feasible $(\theta, s_i)$ such that $s_i$ is sequentially rational for $\theta_k$ with respect to some CPS $\mu^t \in \Delta_i^t$ such that $\mu^t(W_{-i}(k, \Delta) \mid \Sigma_{-i}) = 1$.\(^{26}\) A feasible pair $(\theta, s_i) \in W_i(k, \Delta)$ is called weakly $(k, \Delta)$-rationalizable. A feasible pair is weakly $\Delta$-rationalizable if it is weakly $(k, \Delta)$-rationalizable for all $k = 1, 2, \ldots$. The set of weakly $\Delta$-rationalizable pairs for player $i$ is denoted by $W_i(\infty, \Delta)$.

There is a convenient way to reformulate Definition 3.1. For any subset $B_{-i} \subseteq \Sigma_{-i}$, let

$$
\Lambda_i^t(B_{-i}) = \{ \mu^i \subseteq \Delta_i^t : \mu^i(B_{-i} \mid \Sigma_{-i}) = 1 \}.
$$

Note that (a) $\Lambda_i^t(B_{-i}) = \emptyset$ whenever $B_{-i}$ is not measurable, (b) operator $\Lambda_i^t$ is monotone\(^{27}\) on the Borel sigma-algebra of $\Sigma_{-i}$ and is also monotone with respect to $\Delta_i^t$, and (c) $W_i(k + 1, \Delta) = \rho_i(\Lambda_i^t(W_{-i}(k, \Delta)))$.

$W_i(k, \Delta) \times W_2(k, \Delta)$ is the set of profiles consistent with assumptions (0)–(k – 1) above. Note that these assumptions are silent about how the players would change their beliefs if they observed a history $h$ which they believed impossible at the beginning of the game, even if $h$ is consistent with rationality or mutual certainty of rationality of any order. Therefore weak rationalizability satisfies only a very weak form of backward induction (e.g. in two-stage games with perfect information) and cannot capture any kind of forward induction reasoning. This is what makes weak rationalizability different from strong rationalizability.

**3.2. Strong $\Delta$-rationalizability**

According to strong rationalizability each player believes that his opponent is rational as long as this is consistent with his observed behavior. More generally, each player bestows on his opponent the highest degree of ‘strategic sophistication’ consistent with his observed behavior (see Remark 1 below). This ‘best rationalization principle’ is a form of forward induction reasoning; it also induces the backward induction path in games of perfect and complete information (cf. Battigalli (1996, 1997)). To make the epistemic assumptions underlying strong rationalizability more transparent recall that a state of the world describes the players’ *dispositions* to believe, that is, it describes not only how the players’ actual beliefs evolve along the actual path, but also the beliefs the players would

---

\(^{26}\) It goes without saying that whenever we write a condition like $\mu^t(E \mid \Sigma_{-i}(h)) \equiv \alpha$ and $E$ is not measurable, the condition is not satisfied.

\(^{27}\) A set to set operator $\Lambda$ is monotone if $E \subseteq F$ implies $\Lambda(E) \subseteq \Lambda(F)$. 

have at histories off the actual path. We say that player $i$ strongly believes an event $E$ if $i$ will or would be certain of $E$ at each history $h$ consistent with $E$ (see Battigalli and Siniscalchi (2001a) and references therein). Strong $\Delta$-rationalizability characterizes the feasible type-strategy pairs realized at states of the world where all the following events are true:

(0) every player $i$ has first-order conditional beliefs in $\Delta^i$ and is sequentially rational, 
(S1) every player $i$ strongly believes (0), 
(S2) every player $i$ strongly believes (0) & (S1), 

... 

(Sk) every player $i$ strongly believes (0) & (S2) &…& (S(k − 1)), 

....

**Definition 3.2.** Let $\Sigma_i(0, \Delta) = \Sigma_i$ and $\Phi_i(0, \Delta) = \Delta^i, i = 1, 2$. Suppose that $\Sigma_i(k, \Delta)$ and $\Phi_i(k, \Delta)$ have been defined for each $i = 1, 2$. Then for each $i = 1, 2$,

$$
\Phi_i(k + 1, \Delta) = \{ \mu^i \in \Phi_i(k, \Delta) : \forall h \in \mathcal{H}, \Sigma_i(h) \cap \Sigma_i(k, \Delta) \neq \emptyset \Rightarrow \mu^i(\Sigma_i(k, \Delta)|\Sigma_i(h)) = 1 \}, \quad \Sigma_i(k + 1, \Delta) = \rho_i(\Phi_i(k, \Delta)).
$$

A feasible pair $(\theta, s_i) \in \Sigma_i(k, \Delta)$ is called strongly $(k, \Delta)$-rationalizable. A feasible pair is strongly $\Delta$-rationalizable if it is strongly $(k, \Delta)$-rationalizable for all $k = 1,2,\ldots$. The set of strongly $\Delta$-rationalizable pairs for player $i$ is denoted by $\Sigma_i(\infty, \Delta)$.

Note that $W_i(1, \Delta) = \rho_i(\Delta^i) = \Sigma_i(1, \Delta)$. We show below that under appropriate regularity conditions, as the terminology suggests, the set of strongly $(k, \Delta)$-rationalizable profiles is contained in the set of weakly $(k, \Delta)$-rationalizable profiles and that the two sets coincide in static games (in general, it is sufficient that all the sets $W_i(k, \Delta)$ and $\Sigma_i(k, \Delta)$ ($i \in N$, $k = 1,2,\ldots$) are nonempty and measurable).

**Remark 1.** (‘Best rationalization’) The set $\Phi_i(n + 1, \Delta)$ can be characterized as follows: let $\kappa(-i, h, n)$ denote the highest index $k \leq n$ such that strongly $(k, \Delta)$-rationalizable behavior by $-i$ is consistent with $h \in \mathcal{H}$,28 then

$$
\Phi_i(n + 1, \Delta) = \{ \mu^i \in \Delta^i : \forall h \in \mathcal{H}, \mu^i(\Sigma_i(h)|\Sigma_{-i}(\kappa(-i, h, n), \Delta)|\Sigma_{-i}(h)) = 1 \}
$$

$$
= \bigcap_{k=0}^{n} \{ \mu^i \in \Delta^i : \forall h \in \mathcal{H}, \Sigma_i(h) \cap \Sigma_{-i}(k, \Delta) \neq \emptyset \Rightarrow \mu^i(\Sigma_i(k, \Delta)|\Sigma_{-i}(h)) = 1 \}.
$$

3.3. Examples

In Section 1 we analyzed a duopoly à la Cournot with one-sided incomplete information to illustrate the rationalizability procedure for static games without exogenous restrictions on beliefs. In such a model, the set of rationalizable outcomes is quite large.

28 That is, $\kappa(-i, h, n) = \max\{ k \in \{0,\ldots,n\} : \Sigma_i(h) \cap \Sigma_{-i}(k, \Delta) \neq \emptyset \}$. 
Now we consider two examples where each state of Nature corresponds to a unique rationalizable outcome.

**An exchange game.** Two artists meet at a fair in the morning and have to decide whether to exchange the works of art they are going to produce in the afternoon. The value of a work of art by individual $i$ is equal to his or her ability $u_i \in \{0, 1\}$, which is private information. Each individual has to choose whether to propose to exchange or not. Proposals are simultaneous and become binding if and only if both individuals propose. In order to propose an exchange, an individual has to pay a small transaction cost $e \in (0, 1)$.

The payoffs are given by Table 1.

We do not assume any exogenous restriction on beliefs.

The rationalizable solution is that, independently of his ability, no individual proposes to exchange. To see this, first note that a rational individual $i$ whose ability is $u_i > 1 - e$ will not propose to exchange, because this action is strictly dominated given his type (on the other hand, $p$ is a best response to belief $\mu'$ for $\theta_j \leq 1 - e$ if $\mu'(\{(1, p)\}) = 1$). It follows that a rational individual $i$ whose ability is $\theta_i > 1 - 2e$ and who is certain of the rationality of individual $j$ will not propose to exchange, because he is certain that $j$ will propose to exchange only if $\theta_j \leq 1 - e$. More generally, it can be easily shown by induction that

$$W_i(k) = [0, 1] \times \{n\} \cup \{(\theta_i, p) : 0 \leq \theta_i \leq 1 - k e\}.$$ 

Therefore

$$W_i(\infty) = [0, 1] \times \{n\} = W(k), \quad \forall k \geq \frac{1}{e}.$$ 

A similar result obtains in a variant of this game where the set of types is finite and there is no transaction cost, provided that we assume the following restrictions on beliefs: each player $i$ assigns positive probability to the pair $(\theta, p) \in \Sigma_j$ where $\theta$ is the lowest ability. Since $p$ is a weakly dominant action for the lowest type, this seems a very weak restriction. (For an analysis of rationalizable trade in exchange games see Morris and Skiadas (2000).)

**A game of disclosure.** Consider the following signaling game: the Sender’s type can be either high ($\theta^H$) or low ($\theta^L$). The Sender can either credibly reveal his type or not. This means that the message space for type $\theta^H$ is $A_1(\theta^H) = \{H, N\}$ and the message space for

---

Table 1

<table>
<thead>
<tr>
<th>An exchange game</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1, \theta_1) \times (2, \theta_2)$</td>
<td>$p$ (propose)</td>
<td>$n$ (not)</td>
</tr>
<tr>
<td>$p$ (propose)</td>
<td>$\theta_1 - \epsilon, \theta_1 - \epsilon$</td>
<td>$\theta_1 - \epsilon, \theta_2$</td>
</tr>
<tr>
<td>$n$ (not)</td>
<td>$\theta_1, \theta_2 - \epsilon$</td>
<td>$\theta_1, \theta_2$</td>
</tr>
</tbody>
</table>

---

---

---
type $\theta^L$ is $A_1(\theta^L) = \{L, N\}$, where $N$ is the neutral message meaning ‘no information’ and $H$ and $L$ have the obvious meaning.

The Receiver’s maximizes his expected payoff by estimating the probability that the Sender’s type is high: $A_2(a_1) = [-k, 1 + k]$ ($k > 0$), $u_2(\theta, a_1, a_2) = -(I_{\theta^H}(\theta) - a_2)^2$, for all messages $a_1 \in \{H, L, N\}$, where $I_{\theta^H}$ is the indicator function for the high type (i.e. $I_{\theta^H}(\theta^H) = 1$ and $I_{\theta^H}(\theta^L) = 0$). The Sender’s utility is increasing in $a_2$, e.g. $u_1(\theta, a_1, a_2) = v(\theta) + a_2$; thus the Sender has an incentive to convince the Receiver that his type is high.

We assume that the Receiver is ‘mildly skeptical’ in the sense that if he gets the neutral message $N$ he assigns a positive conditional probability to the low type: $\Delta^2 = \{\mu \in \Delta_{\theta^H}(\Sigma_1) : \mu(\theta^L|N) > 0\}$, whereas $\Delta^1 = \Delta_{\theta^H}(\Sigma_2)$.

The essentially unique strongly $\Delta$-rationalizable outcome of this game is that the Sender discloses if his type is high and the Receiver infers from the neutral message that the Sender’s type must be low. To see this, note that mild skepticism and sequential rationality imply that the Receiver’s response to $a_1 = N$ is $a_2 = s_2(N) < 1$.

Obviously, sequential rationality also implies $s_2(H) = 1$ and $s_2(L) = 0$. Since the Sender is rational and certain that the Receiver is rational and mildly skeptical, he strictly prefers to disclose if his type is high. According to strong rationalizability, we assume that the Receiver is initially certain of this and maintains this belief conditional on message $N$ even if he initially assigned probability zero to $N$. Therefore $s_2(N) = 0$.

Formally, we obtain:

\[
\Sigma(1, \Delta) = \Sigma_1 \times \{s_2 : 0 \leq s_2(N) < 1, s_2(H) = 1, s_2(L) = 0\},
\]

\[
\Sigma(2, \Delta) = \{(\theta, a_1) : \theta = \theta^H \Rightarrow a_1 = H\} \times \Sigma_2(1, \Delta),
\]

\[
\Sigma(3, \Delta) = \Sigma_1(2, \Delta) \times \{s_2^*\} \text{ where } s_2^*(N) = s_2^*(L) = 0, s_2^*(H) = 1.
\]

(On strong rationalizability and disclosure see Battigalli (2000).)

### 3.4. Existence and regularity

It is well-known that even for well-behaved dynamic games with a continuum of actions the strategic-form payoff functions need not be continuous or measurable and hence the sequential best response correspondences $r_i(\theta_i, \mu^I)$ ($i \in N$) need not be well-behaved. We first provide simple conditions on the ‘fundamentals’ implying that the correspondences $r_i(\cdot, \cdot)$ are nonempty-valued and upper-hemicontinuous. Then we show that, if the latter properties are satisfied and $\Delta$ (exogenous restrictions on beliefs) is regular, weak and strong $\Delta$-rationalizability are well-behaved.

**Definition 3.3.** A game

\[
\Gamma = \langle N, (\Theta_i)_{i \in N}, (A_i)_{i \in N}, H^*(\cdot), (u_i)_{i \in N} \rangle
\]

is ‘simple’ if $\Theta$ is compact and either (a) $A$ is finite or (b) $A$ is compact and for some integer $T$, (b1) $\Gamma$ has $T$ stages (that is, every terminal history $h$ has length $\ell(h) = T$), (b2) for every $\theta \in \Theta$ and $h \in H(\theta)$, if $\ell(h) < T - 1$ then $A(\theta, h)$ is finite.
Clearly, finite games and infinitely repeated games with a finite stage game are simple. Signaling games with a finite message space are simple if \( A_2 \) and \( \Theta \) are compact. Signaling games with a continuum of messages are not simple.

**Lemma 3.4.** For every simple game \( \Sigma \) is compact and, for each player \( i \), \( r_i(\cdot, \cdot) \) is an upper-hemicontinuous, nonempty-valued correspondence.

Even in simple games, the set of (weakly or strongly) \( \Delta \)-rationalizable profiles may be empty because the exogenous restrictions on beliefs represented by \( \Delta \) may conflict with common certainty of rationality (or mutual strong belief in rationality). But we can obtain a simple existence result and other regularity properties for the case where \( \Delta \) only represents restrictions on the (marginal) initial beliefs about the opponent’s type (existence results with a more general set of restrictions on beliefs can be obtained for specific models; see Section 5). For any subset \( C \) of a product set \( X \times Y \) and for any probability measure \( \mu \) on \( C \) let \( \text{proj}_X \) and \( \text{marg}_X \mu \) respectively denote the projection of \( C \) on \( X \) and the marginal of \( \mu \) on \( X \), that is,

\[
\text{proj}_X C = \{ x \in X : \exists y \in Y, (x, y) \in C \}
\]

\[
(\text{marg}_X \mu)(E) = \mu(\{ (x, y) \in C : x \in E \}), E \subseteq X(\text{measurable}).
\]

\( \Delta \) is regular if, for each player \( i \), \( \Delta^i \) is nonempty and closed, and there is a set \( \prod_{-i} \subseteq \Delta(\Theta_{-i}) \) such that

\[
\Delta^i = \{ \mu^i \in \Delta^\Theta(\Sigma_{-i}) : \text{marg}_\Theta \mu^i(\cdot|\Sigma_{-i}) \in \prod_{-i} \}.
\]

The following propositions are jointly proved in Appendix A (see Proof of Proposition 8.1):

**Proposition 3.5.** Suppose that \( \Delta \) and \( \Delta^i \) are regular, \( \Sigma \) is compact, \( r_i(\cdot, \cdot) \) is nonempty-valued and upper-hemicontinuous and \( \Delta^i \subseteq (\Delta^i)' \) for every player \( i \). Then for every player \( i \) and all \( k = 0, 1, \ldots, \infty \),

(a) the sets \( W_i(k, \Delta) \) and \( \Lambda^1_\Delta(W_i(k, \Delta)) \) of weakly \((k, \Delta)\)-rationalizable pairs and beliefs are nonempty and compact, \( \text{proj}_i W_i(k, \Delta) = \Theta_i \);

(b) weak \((k, \Delta)\)-rationalizability implies weak \((k, \Delta^i)\)-rationalizability: \( W_i(k, \Delta) \subseteq W_i(k, \Delta^i) \);

(c) \( W_1(\infty, \Delta) \times W_2(\infty, \Delta) \) is the largest measurable subset \( F_1 \times F_2 \subseteq \Sigma \) such that

\[
F_1 \times F_2 \subseteq \rho_1(\Lambda^1_\Delta(F_2)) \times \rho_2(\Lambda^2_\Delta(F_1)).
\]

Furthermore,

\[
W_1(\infty, \Delta) \times W_2(\infty, \Delta) = \rho_1(\Lambda^1_\Delta(W_2(\infty, \Delta))) \times \rho_2(\Lambda^2_\Delta(W_1(\infty, \Delta))).
\]
Proposition 3.6. Suppose that $\Delta$ is regular, $\Sigma$ is compact and $r_i(\cdot, \cdot)$ is nonempty-valued and upper-hemicontinuous for every player $i$. Then for every player $i$ and all $k = 0, 1, \ldots, \infty$,

1. the sets $\Sigma_i(k, \Delta)$ and $\Psi(k, \Delta)$ of strongly $(k, \Delta)$-rationalizable pair and beliefs are nonempty and compact, $\text{proj}_Q \Sigma_i(k, \Delta) = \Theta_i$;
2. strong $(k, \Delta)$-rationalizability implies weak $(k, \Delta)$-rationalizability: $\Sigma_i(k, \Delta) \subseteq W_i(k, \Delta)$ (the inclusion holds as an equality if the game is static).

Proposition 3.5 (a) (3.6 (1)) says that there is a weakly (strongly) rationalizable strategy for each payoff-type. (b) says that weak rationalizability is monotone with respect to exogenous restrictions on beliefs. This does not hold for strong rationalizability. In fact, if stronger restrictions on beliefs make fewer histories consistent with strongly $k$-rationalizable strategies, the $k$-forward induction criterion applies only to this smaller set of histories and the set of $(k + 1)$-rationalizable profiles need not be smaller. (c) says the set of weakly rationalizable profiles is the largest set with the ‘best response property.’

As an immediate consequence of Lemma 3.4 and Propositions 3.5 and 3.6 we obtain the following:

Corollary 3.7. In every simple game, if $\Delta$ is regular then (a), (c) of Proposition 3.5 and (1), (2) of Proposition 3.6 hold.

3.5. Rationalizability and iterated interim dominance

The set of weakly and strongly rationalizable pairs can be further characterized for generic finite games in terms of dominance relations. We say that a game has no relevant tie if the following holds: for each player $i$ and all pairs of outcomes $(\theta, z^\prime)$, $(\theta, z^\prime) \in \mathcal{X}$, if there are $h \in \mathcal{H}(\theta)$, $a', a'' \in A(\theta, h)$ such that $a_i' \neq a_i''$, $z'$ follows $(h, a')$ and $z''$ follows $(h, a'')$, then $u_i(\theta, z^\prime) \neq u_i(\theta, z^\prime)$. This means that if player $i$, immediately after history $h$, has deterministic beliefs about the true parameter $\theta$ and the continuation of the game, then he cannot be indifferent between any two feasible actions.

A strategy $s_i \in S_i(\theta_i)$ is weakly dominated\(^\text{30}\) by mixed strategy $m_i \in \Delta(S_i(\theta_i))$ for type $\theta$, on $B_{-i} \subseteq \Sigma_{-i}$ if

$$\forall \sigma_{-i} \in B_{-i}, U_i(\theta_i, s_i, \sigma_{-i}) \leq \sum_{s'_i} m_i(s'_i) U_i(\theta_i, s'_i, \sigma_{-i})$$

and

$$\exists \sigma_{-i} \in B_{-i}, U_i(\theta_i, s_i, \sigma_{-i}) \leq \sum_{s'_i} m_i(s'_i) U_i(\theta_i, s'_i, \sigma_{-i}).$$

\(^\text{30}\) This is also called ex post dominance, because the dominance relation between $s_i$ and $m_i$ would hold even if the state of Nature were revealed to player $i$. 

The definition of strict dominance is analogous (all weak inequalities are replaced by strict inequalities). For any given rectangular subset \( B \subseteq \Sigma \) let \( \mathcal{W}(B) \) denote the set of \((\theta_i, s_i)_{i \in N} \subseteq \Sigma\) such that, for each \( i \), \( s_i \) is not weakly (strictly) dominated for \( \theta_i \) on \( B_{-i} \) and let \( \mathcal{W}^n(B) = \mathcal{W}(B) \cap \mathcal{W}^{n-1}(B) \). The iterated operator \( \mathcal{W}^n \) is defined in the usual way: \( \mathcal{W}^n(B) = \mathcal{W}(\mathcal{W}^{n-1}(B)) \), where \( \mathcal{W}^0(B) = B \). A subscript \( p \) denotes that we only consider weak domination by pure strategies. Thus \( \mathcal{W}_p(B) \) is the set of profiles \((\theta_i, s_i)_{i \in N}\) such that \( s_i \) is not weakly dominated for \( \theta_i \) by another pure strategy on \( B_{-i} \), and \( \mathcal{W}_p(B) = \mathcal{W}(\mathcal{W}_p(B)) \). Note that \( \mathcal{W} \) is a monotone operator. Therefore, also \( \mathcal{W}_p \) and \( \mathcal{W}_k \) are monotone operators. \( \mathcal{W}(k) \) and \( \Sigma(k) \) denote the subsets of weakly and strongly \( k \)-rationalizable profiles, without exogenous restrictions on beliefs. The following proposition extends to games with incomplete information results proved by Pearce (1984) and Ben Porath (1997).

**Proposition 3.8.** (a) (Cf. Pearce, 1984) In every finite and static game,

\[
\Sigma(k) = W(k) = \mathcal{W}^k(\Sigma), \quad k = 1, 2, \ldots
\]

(b) In every finite game with no relevant ties,

\[
\Sigma(k) \subseteq W(k) \subseteq \mathcal{W}_p^k(\Sigma), \quad k = 1, 2, \ldots
\]

(c) (Cf. Ben Porath, 1997) In every finite game with no relevant ties, perfect information and private values,

\[
\Sigma(k) \subseteq W(k) = \mathcal{W}_p^k(\Sigma), \quad k = 1, 2, \ldots
\]

An exact characterization of strong rationalizability can be obtained using a notion of iterated conditional dominance for each payoff-type. The characterization result can be easily adapted from Shimoji and Watson (1998). These characterizations of rationalizability through iterative dominance procedures can be used to compute the set of rationalizable strategies solving a sequence of linear programming problems (cf. Shimoji and Watson (1998), Section 4). The computation algorithm can also incorporate exogenous restrictions on conditional beliefs (Siniscalchi (1997)).

Finally, we can easily extend known results about rationalizability, best replies to deterministic beliefs and dominance in infinite static games. These results provide sufficient conditions implying that the set of strictly dominated actions (for a given type and domain) coincides with the set of actions that are not a best reply to any deterministic belief. This implies that rationalizability coincides with iterated strict dominance and can be computed easily, as in the duopoly example of the introduction. A well-known set of such conditions goes under the general heading of ‘supermodularity’ (Milgrom and Roberts (1990)). Here we generalize a perhaps less well-known result due to Moulin (1984).

In the following we write \( \rho_i(\sigma_{-i}) \) for the set of type-strategy pairs \((\theta_i, s_i)\) such that \( s_i \) is a best reply for \( \theta_i \) to the deterministic belief assigning probability one to \( \sigma_{-i} \). Similarly, \( \rho_i(B_{-i}) = \bigcup_{\sigma_{-i} \in B_{-i}} \rho_i(\sigma_{-i}) \).
Proposition 3.9. (Cf. Moulin (1984)) Consider a static game with incomplete information. Suppose that, for each \( i \in N \), \( \Theta_i \subseteq \mathbb{R}^n \) is a connected compact set, \( A_i = [\tilde{q}_i, \tilde{a}_i] \subseteq \mathbb{R} \) and \( u_i(\theta, a_i, a_{-i}) \) is (continuous in all its arguments and) strictly quasi-concave in \( a_i \). Then, for every connected product subset \( B = \prod_{i \in N} B_i \subseteq \Sigma \)

\[ \mathcal{S}(B) = \prod_{i \in N} \rho_i(B_{-i}). \]

and \( \mathcal{S}(B) \) is connected. It follows that

\[ \prod_{i \in N} \rho_i(\Sigma_{-i}(k-1)) = W(k) = \Sigma(k) = \mathcal{S}^k(\Sigma), \quad k = 1, 2, \ldots. \]

4. Generalizations

The solution concepts defined in Section 3 for two-person games with observable actions can be extended to general \( n \)-person games with imperfect information about past actions. While the introduction of imperfect information is conceptually straightforward, considering more than two players forces a modeling choice between correlated and independent belief and poses the problem of providing a satisfactory definition of independence for conditional probability systems and an appropriate formalization of the forward induction principle for players with multiple opponents. In this section we briefly describe how to deal with these problems.

4.1. Imperfectly observed actions

In a game with observed actions the set of partial histories \( \mathcal{H} \) can be regarded as a common collection of information sets for all the players. In games with imperfectly and asymmetrically observed actions each player \( i \) has his own collection of information sets \( \mathcal{H}_i \), whereby a typical element \( h \in \mathcal{H}_i \) now represents a (maximal) set of partial histories that player \( i \) cannot distinguish. Of course, \( \mathcal{H}_i \) need only contain the information sets where player \( i \) is active. In order to adapt the analysis of the previous section to this situation it is sufficient to redefine \( \Sigma(h) \) as the set of feasible profiles consistent with at least one history contained in \( h \). Perfect recall implies that \( \Sigma(h) = \Sigma_i(h) \times \Sigma_{-i}(h) \) for each \( h \in \mathcal{H}_i \). The collection \( \mathcal{B}_i \) of ‘relevant hypotheses’ for player \( i \) is then defined as

\[ \mathcal{B}_i = \{ B \subseteq \Sigma_{-i} : \exists h \in \mathcal{H}_i, B = \Sigma_{-i}(h) \} \]

and this determines the space of conditional probability systems \( \Delta^\mathcal{B}_i(\Sigma_{-i}) \). Given these modifications, the other formal definitions are virtually unchanged.

4.2. \( n \)-Person games and independent beliefs

Extending the previous analysis to \( n \)-person games is quite straightforward if it is assumed that each player’s beliefs concerning the type and strategy of different opponents may exhibit correlation. Therefore we consider here only the case of independent beliefs.
Recall that in games with observable actions the set $\Sigma(h)$ of feasible profiles consistent with a given history/information set $h$ has a Cartesian structure: $\Sigma(h) = \prod_{i \in H} \Sigma_i(h)$. The same is true whenever $h$ is an information set of a game with observable deviators. For the sake of simplicity, we limit our analysis to this class of games. For any two players $i$ and $j$ let

$$R_{ij} = \{B_j \subseteq \Sigma_j : \exists h \in H, B_j = \Sigma_j(h)\}$$

be the collection of ‘strategic form’ pieces of information about player $j$ that player $i$ might obtain and let $\Delta^{\delta_i}(\Sigma_j)$ be the associated set of $i$’s marginal CPS’s about player $j$. A CPS $\mu_i \in \Delta^{\delta_i}(\Sigma_{-i})$ is independent if there exists a vector of marginal CPS’s $(\mu_j)_{j \neq i} \in \prod_{j \neq i} \Delta^{\delta_j}(\Sigma_j)$ such that, for all $h \in H_i$, $\mu_i(\cdot | \Sigma_j(h))$ is the product measure obtained from the vector of marginal probability measures $(\mu_j(\cdot | \Sigma_j(h)))_{j \neq i}$ (cf. Rényi (1955), p 303).

Assuming that the players are rational and have independent conditional beliefs and that this is common certainty at the beginning of the game, we obtain a notion of weak rationalizability with independent beliefs. The formal definition is essentially the same as in Section 2 except that now it has to be assumed that, for each player $i$, the restricted set of beliefs $\mathcal{A}_i$ contains only independent CPS’s.32

Let us now turn to strong rationalizability. Since we assume that players’ conditional beliefs are independent, we also incorporate in the definition of strong rationalizability a principle of independent best rationalization: each player $i$ ascribes to every opponent $j$ the ‘highest degree of strategic sophistication’ consistent with $j$’s observed behavior independently of any information about other players.33 The formal, inductive definition of strong rationalizability (without exogenous restrictions on beliefs beyond independence) can be given as follows. Let $\mu^i_j$ denote the marginal on $\Sigma_j$ of a given independent CPS $\mu^i$.

(0) For all $i \in N$, $\Sigma^0_i = \Sigma_i$ and $\Phi^i(0) = \{\mu^i \in \Delta^{\delta_i}(\Sigma_{-i}) : \mu^i$ is independent$\}$. 

(k + 1) For all $i \in N$, $\Sigma^{k+1}_i = \rho_i(\Phi^i(k))$ and $\Phi^i(k+1) = \{\mu^i \in \Phi^i(k) : \forall h \in H_i, \forall j \neq i, \Sigma_j(h) \cap \Sigma^k_j \neq \emptyset \Rightarrow \mu^i_j(\Sigma^k_j | \Sigma_j(h)) = 1\}$.

5. Applications

The methodology proposed in this paper has been applied to a number of economic models concerning reputation, disclosure, market signaling and auctions. Here we briefly report on the results.

32 This notion of rationalizability is used in Battigalli and Watson (1997) and in Siniscalchi’s (1998) analysis of ‘Japanese’ auctions. See the next section.

33 Battigalli and Siniscalchi (1999b) provides a rigorous epistemic axiomatization of the independent best rationalization principle.
5.1. Reputation in repeated games

Building on Watson (1993), Battigalli and Watson (1997) and Battigalli (2001) analyze reputation in repeated games, relying on minimal assumptions about beliefs and mutual certainty of rationality.

Battigalli and Watson (1997) consider a long-run player facing a sequence of short-run opponents. It is assumed that the beliefs of the short-run players may be heterogeneous, but are not too different from each other, and furthermore that they satisfy stochastic independence across opponents (see Section 4.2) and assign at least $\epsilon$ probability to the long-run player being a commitment type who always plays his stage-game ‘Stackelberg action.’ These are the exogenous restrictions on beliefs $\Delta$. It is shown that if (a) the short-run opponents are rational and (b) have beliefs in $\Delta$, then the long-run player can make them choose the best response to the Stackelberg action, simply by playing such action for a long enough time. If the long-run player is rational, believes in (a) and (b) and is patient, then his long-run average expected payoff is approximately bounded below by the Stackelberg payoff. This result can be obtained with two steps of the weak $\Delta$-rationalizability procedure.\footnote{As noted in Battigalli and Watson (1997), a similar, but simpler result holds when the sequence of short-run players is replaced by a long-run, relatively impatient player, provided that the payoffs of the constituent matrix game satisfy a property called ‘conflicting interests’.}

Battigalli (2001) proves a similar result for the case of a long-run patient player with a long-run (impatient) opponent. According to the assumed restrictions $\Delta$, the impatient player assigns at least $\epsilon$ probability to the patient player being a commitment type that plays a history-dependent strategy ‘teaching’ to choose the best response to the Stackelberg action.

Battigalli (2000) applies strong $\Delta$-rationalizability to signaling games. The first application is a model of disclosure generalizing the second example of Section 3 (in particular, the number of types is finite, but arbitrary). The assumed exogenous restriction on beliefs is that the Receiver is mildly skeptical, i.e. assigns a positive probability to the worst type consistent with any given message. Then strong $\Delta$-rationalizability implies that the Receiver interprets any given message $m$ as being sent by the worst type consistent with $m$ and this belief is indeed correct. In other words, the weak restriction of mild skepticism, when combined with the forward induction logic of strong rationalizability yields extreme skepticism.

The second application is a version of Spence’s job market signaling model where productivity depends on ability and education. The assumed exogenous restriction on beliefs is that the conditional expectation of ability is weakly increasing with observed education. If high and low types are sufficiently different, strong $\Delta$-rationalizability yields the most efficient separating equilibrium outcome. Otherwise, strong $\Delta$-rationalizability only yields bounds on the level of education that the high and low types would choose.

5.2. Signaling games

Battigalli (2000) applies strong $\Delta$-rationalizability to signaling games. The first application is a model of disclosure generalizing the second example of Section 3 (in particular, the number of types is finite, but arbitrary). The assumed exogenous restriction on beliefs is that the Receiver is mildly skeptical, i.e. assigns a positive probability to the worst type consistent with any given message. Then strong $\Delta$-rationalizability implies that the Receiver interprets any given message $m$ as being sent by the worst type consistent with $m$ and this belief is indeed correct. In other words, the weak restriction of mild skepticism, when combined with the forward induction logic of strong rationalizability yields extreme skepticism.

The second application is a version of Spence’s job market signaling model where productivity depends on ability and education. The assumed exogenous restriction on beliefs is that the conditional expectation of ability is weakly increasing with observed education. If high and low types are sufficiently different, strong $\Delta$-rationalizability yields the most efficient separating equilibrium outcome. Otherwise, strong $\Delta$-rationalizability only yields bounds on the level of education that the high and low types would choose.

\footnote{As noted in Battigalli and Watson (1997), a similar, but simpler result holds when the sequence of short-run players is replaced by a long-run, relatively impatient player, provided that the payoffs of the constituent matrix game satisfy a property called ‘conflicting interests’.}
Battigalli and Siniscalchi (2001b) provide a more abstract result about strong $\Delta$-rationalizability in signaling games. Let $\Delta(\zeta)$ represent the restriction that the players beliefs ‘agree’ with a given outcome distribution $\zeta \in \Delta(\Theta \times A_1 \times A_2)$. It is shown that $\zeta$ is a self-confirming equilibrium distribution satisfying the iterated intuitive criterion (IIC)\textsuperscript{35} if and only if the set of $\Delta(\zeta)$-rationalizable outcomes is not empty. This in turn yields an epistemic characterization of the IIC via the results of Battigalli and Siniscalchi (2001a).

5.3. Auctions

Dekel and Wolinsky (2000) analyze first-price, IPV auctions with a discrete set of possible bids and types. On top of stochastic independence of opponents’ types, they assume that each player assigns probability at least $d$ to each possible valuation of each competitor. They prove a limit result: if the number of participants is large enough $\Delta$-rationalizability implies that each player submits the highest bid below his private value.

On the other hand, Battigalli and Siniscalchi (2001c) show that if the set of possible valuations and bids is continuous (an interval), then the set of rationalizable bids is quite large even if there are many participants. They assume that players know the ‘true’ distribution of types and that they believe that a strictly positive bid yields a strictly positive probability of winning (this rules out weakly dominated bids).\textsuperscript{36} They show how to compute the upper bound on the set of $(\Delta,k)$-rationalizable as a function of a player’s valuation. The upper bound is increasing, concave and strictly above the (symmetric) equilibrium bidding function. Every bid between zero and this upper bound is $(\Delta,k)$-rationalizable. These results seem to be consistent with the experimental evidence.

Siniscalchi (1998, Ch. 5) analyzes ascending bid (‘Japanese’) auctions with a common discrete set of valuations and bids. It is well known that the canonical solution of such auctions is that each participant plays the weakly undominated strategy of ‘staying in’ as long as the price called by the auctioneer is below his valuation. However, there are sequential equilibria in weakly dominated strategies where bidders stay in even at prices above their valuation. Without exogenous restrictions on beliefs, these sequential equilibria are not ruled out by strong rationalizability. But Siniscalchi shows that, introducing rather weak exogenous restrictions on beliefs, $\Delta$-rationalizability yields the standard solution. In particular, this result holds when $\Delta$ represents the following assumptions: (i) players’ beliefs about opponents types and strategies satisfy stochastic independence, (ii) every valuation of every competitor is assigned positive prior probability, (iii) if, when price $v^n$ is called, player $i$ assigns positive conditional probability to $j$’s valuation being $v^n$, then $i$ assigns positive probability to player $j$ quitting before a higher price is called.

\textsuperscript{35} The original definition and informal motivation of the IIC can be found in Cho and Kreps (1987).

\textsuperscript{36} Their analysis is extended to auctions with interdependent (affiliated) valuations. The basic insights also apply to the case of unknown distributions of valuations.
6. Conclusion

We proposed a methodology to analyze models of strategic interaction where some of the ‘fundamental’ parameters (preferences and technology) are not common knowledge. We called ‘payoff-types’ the possible pieces of private information about such fundamentals.

Our methodology is different from Harsanyi’s (1967–68) analysis of incomplete information games, but it is consistent with it. In order to apply Harsanyi’s approach, it is necessary to append to a given model with unknown fundamentals a type space, i.e. a compact, implicit specification of the set of possible infinite hierarchies of beliefs (beliefs about the opponents’ payoff-types, beliefs about such beliefs, and so on). We call such hierarchies of beliefs ‘epistemic types’ and the extended model ‘Bayesian game.’ In an equilibrium of a Bayesian game, each player best responds to a correct conjecture about how his opponents would choose, given their payoff + epistemic types. According to Harsanyi’s approach we should study the equilibria of the ‘appropriate’ Bayesian extension of a given model with incomplete information.

But how can we choose the ‘appropriate’ type space? As we show elsewhere,37 if the set of possible hierarchies of beliefs about payoff-types is unrestricted (i.e. ‘universal’), then the equilibrium assumption (correct conjectures about opponents’ choice functions) has no more bite than just assuming common certainty of rationality. On the other hand, analyzing the equilibria of Bayesian games with restricted, but still large type spaces may be quite complex. Thus, to obtain sharp results and for tractability reasons, most economic applications consider very small ‘Micky Mouse’ type spaces. But the epistemic assumptions implicit in this ‘small-type-space-plus-equilibrium' approach are often implausible and/or non-transparent. Therefore we propose to explore the consequences of alternative assumptions.

We believe that, in principle, the analysis of any model of interactive decisions should be based on explicit assumptions about beliefs (including assumptions about how beliefs change) and rationality (how choices are related to beliefs). Behavioral implications should be derived from such assumptions. The epistemic analysis of games shows that some interesting constellations of assumptions about rationality and beliefs exactly characterize corresponding solutions concepts, which can be used as ‘shortcuts’ to obtain results about specific models. In this paper we take advantage of this work on the epistemic foundations of game theory and we focus on solution concepts.

Rather than explicitly enrich the given economic model with a type space and compute Bayesian equilibria, we propose to apply a sort of iterated interim dominance deletion procedure called ‘Δ-rationalizability.’ This procedure is parametrized by some exogenous restrictions on first-order beliefs (beliefs about the payoff-types and/or strategies of the opponents), represented by some belief set Δ. In static games, the procedure corresponds to considering higher and higher degrees of mutual certainty that (a) players are rational and (b) their beliefs satisfy the restrictions Δ. The above mentioned result about Bayesian equilibrium and common certainty of rationality implies that, without exogenous restrictions, rationalizability characterizes the set of

37 Battigalli and Siniscalchi (2001b).
all Bayesian equilibrium outcomes. This shows that our approach is fully consistent with Harsanyi’s one.

Rationalizability can be extended from static to dynamic games in several ways. These extensions have in common the assumption that players carry out sequential best replies to their system of conditional beliefs. Assuming initial common certainty of sequential rationality and of the exogenous restrictions on beliefs we obtain a solution concept called ‘weak $\Delta$-rationalizability.’ On top of this, we may also want to formalize elements of strategic reasoning related to the general principle that observed actions are interpreted as signals about private information and/or strategic intent. Hence we also define a notion of strong $\Delta$-rationalizability featuring this forward induction principle.

We provided a unified analysis of weak and strong $\Delta$-rationalizability in incomplete information games where the set of payoff-types and actions may be uncountably infinite and the time horizon may be infinite as well. We obtained existence and regularity conditions for these solution concepts and we analyzed how they are related with iterated dominance procedures. For the sake of simplicity, most of the analysis focused on two-person games with observable actions, but we also show how to extend the solution concepts when there are several players and actions are not perfectly observed. We hope that the technical results and examples contained in the paper, and the brief survey of applications mentioned in Section 5 will convince the reader that we proposed a viable and interesting methodology for the analysis of incomplete information games.

Acknowledgements

This paper is a revision of the more theoretical part of ‘Rationalizability in Incomplete Information Games’. Helpful comments from Patrick Bolton, Giacomo Bonanno, Tilman Bürgers, Françoise Forges, Faruk Gul, Marciano Siniscalchi, Juuso Välimäki, Joel Watson and seminar participants at the University of Valencia, Northwestern University, Caltech, McGill University, SITE (Stanford University), Université de Cergy Pontoise, University of North Carolina and European University Institute are gratefully acknowledged.

Appendix A

A.1. Incomplete information games: feasibility correspondence and topological structure

The sets of feasible actions for a given state of Nature $\theta$ and (feasible) history $h$ are derived from the feasibility correspondence $\mathcal{M}^*(\cdot) : \Theta \rightarrow 2^A$ as follows:

$A(\theta, h) = \{a \in A : (h, a) \in \mathcal{M}^*(\theta)\},$

$A_i(\theta_i, h) = \{a_i \in A_i : \exists a_{-i} \in A_{-i}, \exists \theta_{-i} \in \Theta_{-i}, (a_i, a_{-i}) \in A((\theta_i, \theta_{-i}), h)\}.$

The feasibility correspondence satisfies the following properties (recall that $A^*$ is the set of finite or countable infinite sequences of action profiles):

1. $\mathcal{M}^*$ is a correspondence from $\Theta$ to $2^A$.
2. $A(\theta, h)$ is a subset of $A^*$.
3. $A_i(\theta_i, h)$ is a subset of $A_i^*$. 

These properties ensure the consistency and applicability of the solution concepts in dynamic games.
1. for every \( h \in A^* \) and \( \theta \in \Theta \), if \( h \in \mathcal{H}^*(\theta) \), every initial subsequence (prefix) of \( h \) belongs to \( \mathcal{H}^*(\theta) \), in particular, \( \phi \in \mathcal{H}^*(\theta) \) for all \( \theta \in \Theta \).

2. for every infinite sequence \( h^* \in A^\omega \) and every \( \theta \in \Theta \), if for every finite initial subsequence \( h \) of \( h^* \), \( h \in \mathcal{H}^*(\theta) \), then \( h^* \in \mathcal{H}^*(\theta) \).

3. for every \( \theta = (\theta_i)_{i \in \mathbb{N}} \in \Theta \), \( h \in A^\omega \)

\[
A(\theta, h) = \prod_{i \in \mathbb{N}} A_i(\theta_i, h),
\]

\[
A(\theta, h) = \emptyset \text{ if and only if for all } i \in \mathbb{N}, \ A_i(\theta_i, h) = \emptyset.
\]

We endow \( A^* \) and the set of outcomes \( \mathcal{Y} \subset \Theta \times A^* \) with the following metrics \( d_{A^*} \) and \( d_\mathcal{Y} \): Recall that \( A_i \) and \( A_i \) are subsets of \( \mathbb{R}^{m_i} \) and \( \mathbb{R}^{n_i} \), respectively \((i \in \mathbb{N})\). Let \( d_m \) be the Euclidean metric in \( \mathbb{R}^k \) and \( m = \sum_{i \in \mathbb{N}} m_i \), \( n = \sum_{i \in \mathbb{N}} n_i \). Denote by \( l(h) \) the length of a history \((l(h) = \infty \text{ if } h \text{ is a finite history})\) and let \( \alpha(h) \) be the action profile at position \( t \) in history \( h \) \((t \leq l(h))\). If \( l(h) \leq l(h') \), then

\[
d_{A^*}(h, h') = \sum_{i=1}^{l(h)} (1/2)^i d_m(\alpha^i(h), \alpha^i(h')) + \sum_{i=l(h)+1}^{l(h')} (1/2)^i
\]

(the second summation is zero if \( l(h) = l(h') \))

\[
d_\mathcal{Y}((\theta, h), (\theta', h')) = d_m(\theta, \theta') + d_{A^*}(h, h').
\]

\( d_{A^*} \) is the natural metric for games with discounting. It can be checked that \((A^*, d_{A^*})\) and \((\mathcal{Y}, d_\mathcal{Y})\) are complete, separable, metric spaces.

The sets of strategies and strategy-type pairs are endowed with the ‘discounted’ sup-metrics \( d_{\Sigma_i}, d_{\Sigma_i} \) and \( d_{\Sigma_j} \) \((i \in \mathbb{N}, \emptyset \neq J \subset \mathbb{N}, \Sigma_j = \prod_{i \in J} \Sigma_i)\):

\[
d_{\Sigma_i}(s_i, s'_i) = \sum_{i=t}^{\infty} (1/2)^i \left( \sup_{h : l(h) = t} d_m(s_i(h), s'_i(h)) \right),
\]

\[
d_{\Sigma_j}(\sigma_j, \sigma'_j) = \sum_{i \in J} d_{\Sigma_i}(\sigma_i, \sigma'_i).
\]

A.2. Proofs

**Proof of Lemma 2.1.** Let \( S_i(h) \) be the set of strategies consistent with history \( h \). Clearly \( S_i(h) \) is closed. Since

\[
\Sigma_i(h) = \{ (\theta_i, s_i) : s_i \in S_i(\theta_i) \cap S_i(h) \}
\]

we only have to show that \( S_i(h) \) is upper-hemicontinuous in \( \theta_i \). Suppose that \((\theta_i^k, s_i^k) \rightarrow (\theta_i, s_i)\) and \( s_i^k \in S_i(\theta_i^k)\) for all \( k \). Then for all \( h' \in \mathcal{H} \), \( s_i^k(h') \rightarrow s_i(h') \) and \( s_i^k(h) \in A_i(\theta_i^k, h') \) for all \( k \). Since \( \mathcal{H}^*(\cdot) \) is continuous, each \( A_i(\cdot, h') \) \((h' \in \mathcal{H})\) is also continuous. Therefore for all \( h' \in \mathcal{H} \), \( s_i(h') \in A_i(\theta_i, h') \) and \( s_i \in S_i(\theta_i) \). \( \square \)
Proof of Lemma 2.4. For each history $h^t$ of length $t$, $h^t \in \mathcal{H}(\theta)$, define

$$r_i(\theta_i, \mu^i, h^t) = \arg \sup_{s_i \in S_i(\theta_i, h^t)} U_i(\theta_i, s_i, \mu^i h^t (\Sigma_{-i}(h^t))).$$

It follows from the assumptions on $U_i$ that the expectation $E_{\mu^i}[U_i(\theta_i, s_i, \sigma_{-i}(h^t))]$ is well defined and upper-semicontinuous in $s_i$. Thus, by compactness of $S_i(\theta_i)$, $r_i(\theta_i, \mu^i, h^t)$ is a non-empty and compact set. Construct the following decreasing sequence of compact subsets of $S_i(\theta_i)$:

- $R^0 = r_i(\theta_i, \mu^i, \emptyset) \neq \emptyset$.
- Pick $s^0_i \in R^0$ and let

$$R^1 = R^0 \cap \left( \bigcap_{h^1 \in \mathcal{H}(\theta_i, s^0_i)} r_i(\theta_i, \mu^i, h^1) \right).$$

Clearly $R^1$ is a compact subset of $R^0$. By dynamic consistency of expected utility maximization $R^1$ is non-empty.

- Assume that $(R^0, ..., R^{t-1})$ has been defined and is a decreasing (nested) sequence of non-empty compact subsets. Pick $s^{t-1}_i \in R^{t-1}$ and let

$$R^t = \left( \bigcap_{k=0}^{t-1} R^k \right) \cap \left( \bigcap_{h^t \in \mathcal{H}(\theta_i, s^{t-1}_i)} r_i(\theta_i, \mu^i, h^t) \right).$$

Then again $R^t$ is a compact subset of $R^{t-1}$ and by dynamic consistency of expected utility maximization $R^t$ is non-empty.

Therefore we can construct a decreasing sequence $(R^t)_{t=0}^{\infty}$ of non-empty and compact subsets. By the finite intersection property, the infinite intersection is non-empty; by construction, it is a subset of the set of sequentially rational strategies for type $\theta_i$ given CPS $\mu^i$:

$$\emptyset \neq \bigcap_{t=0}^{\infty} R^t \subseteq r_i(\theta_i, \mu^i).$$

\[\square\]

Proof of Lemma 3.4. In a simple game $\Theta$ and $A$ are compact and either $A$ is finite (case (a)) or $\mathcal{H}$ is finite (case (b)). If $A$ is finite, $S$ is a totally bounded, complete metric space. Therefore $S$ is compact. If $\mathcal{H}$ is finite, $S$ is topologically equivalent to a compact subset of a Euclidean space. In both cases $\Sigma \subseteq \Theta \times S$ is compact. By Lemma 2.1 each $\Sigma(h)$ is closed, hence compact.

We consider the rest of the proof for case (b) ($A$ compact, finite horizon, finite sets of feasible actions through the second-to-last stage). The proof for case (a) is similar. Since $\Sigma(h)$ is the graph of the correspondence $S_i(\cdot, h)$, this correspondence is
non-empty-compact-valued and upper-hemicontinuous. Now we show that it is also lower-hemicontinuous. Fix \( h \in \mathcal{H} \) and suppose that \( \Theta^h_i \rightarrow \theta_i \) and \( s_i \in S_i(\theta_i, h) \). By Assumption 0, each \( A_i(\cdot, h') \), \( (h' \in \mathcal{H}) \) is continuous, hence lower-hemicontinuous. Therefore we can find a sequence of actions \( (a^k_{i,h'})_{k=1}^{\infty} \) such that \( a^k_{i,h'} \rightarrow s_i(h') \) and \( a^k_{i,h'} \in A_i(\Theta^h_i, h') \). Let \( s^k_i(h') = a^k_{i,h'} \) for all \( h' \in \mathcal{H} \). By construction \( s^k_i \in S_i(\theta_i) \) and \( (s^k_i)_{k=1}^{\infty} \) converges pointwise to \( s_i \). Since \( \mathcal{H} \) is finite \( s_i^k \rightarrow s_i \). If \( h' \neq h \) is a prefix of \( h \), then by assumption all \( A_i(\Theta^h_i, h') \) and \( A_i(\theta_i, h') \) are finite. Thus, by continuity of \( A_i(\cdot, h') \), \( A_i(\Theta^h_i, h') = A_i(\theta_i, h') \) and \( s^k_i(h') = s_i(h') \) for \( k \) large. This implies that \( s^k_i \in S_i(\Theta^h_i, h) \). Therefore \( S_i(\cdot, h) \) is lower-hemicontinuous.

The outcome function \( \zeta^\ast : \Sigma \rightarrow \mathcal{Y} \) is continuous: suppose that \( (\Theta^h_i, s^k_i)_{k\in\mathbb{N}} \) converges to \( (\theta_i, s_i)_{k\in\mathbb{N}} \), then for \( k \) large \( s^k_i \) and \( s_i \) induce the same action profile through the second-to-last stage and in the last stage the action profile induced by \( s^k_i \) converges to the action profile induced by \( s_i \). Therefore the strategic payoff functions \( U_i = u^\ast_i \zeta^\ast \) are also continuous and (by compactness of \( \Sigma \)) bounded.

Since \( S_i(\cdot, h) \) is non-empty-compact-valued and continuous and \( U_i \) is continuous and bounded, the conditional expected payoff \( U_i(\theta_i, s_i, \mu^i(\cdot|\Sigma_{-i}(h))) \) is always well-defined and continuous in \( (\theta_i, s_i, \mu^i) \) and the correspondence

\[
\quad r_i(\theta_i, \mu^i, h) = \arg \max_{s_i \in S_i(\theta_i, h)} U_i(\theta_i, s_i, \mu^i(\cdot|\Sigma_{-i}(h)))
\]

is nonempty-valued (for \( h \in \mathcal{H}(\theta_i) \)) and upper-hemicontinuous in \( (\theta_i, \mu^i) \). We have shown above that \( r_i(\theta_i, \mu^i) \) is non-empty (Lemma 2.4). We show that \( r_i(\cdot, \cdot) \) is upper-hemicontinuous. Suppose that \( (\Theta^h_i, \mu^{i,k}, s^k_i) \rightarrow (\theta_i, \mu^i, s_i) \) and, for all \( k \), \( s^k_i \in r_i(\theta_i^h, \mu^{i,k}) \).

Since the game is simple, for \( k \) large \( s^k_i \) and \( s_i \) prescribe the same action through the second-to-last stage, which implies that \( \mathcal{H}(\Theta^h_i, s^k_i) = \mathcal{H}(\theta_i, s_i) \). This and upper-hemicontinuity of each correspondence \( r_i(\cdot, \cdot, h)(h \in \mathcal{H}) \) imply that, for each history \( h \in \mathcal{H}(\theta_i, s_i) \), \( s_i \in r_i(\theta_i, \mu^i, h) \). Therefore \( s_i \in r_i(\theta_i, \mu^i) \). \( \square \)

The following result summarizes Propositions 3.5 and 3.6.

**Proposition 8.1.** Suppose that \( \Delta \) and \( \Delta' \) are regular, \( \Sigma \) is compact, \( r_i(\cdot, \cdot) \) is nonempty-compact and upper-hemicontinuous and \( \Delta' \subseteq (\Delta')^i \) for every player \( i \). Then for every player \( i \) and all \( k = 0, 1, \ldots, \infty \)

(a) the sets \( W_i(k, \Delta) \) and \( \Sigma_i(k, \Delta) \) of weakly and strongly \( (k, \Delta) \)-rationalizable profiles are nonempty and compact with \( \text{proj}_{\Theta_i} W_i(k, \Delta) = \text{proj}_{\Theta_i} \Sigma_i(k, \Delta) = \Theta_i \), the sets \( \Lambda^1_i(W_i(k, \Delta)) \) and \( \Phi^i_k(W_i(k, \Delta)) \) are non-empty and compact as well;

(b) \( \Sigma_i(k, \Delta) \subseteq W_i(k, \Delta) \),

(c) \( W_i(k, \Delta) \subseteq W_i(k, \Delta') \),

(d) \( W_i(\infty, \Delta) \times W_i(\infty, \Delta) \) is the largest measurable subset \( F_1 \times F_2 \subseteq \Sigma \) such that

\[
F_1 \times F_2 \subseteq \rho_1(\Lambda^1_2(F_2)) \times \rho_2(\Lambda^2_2(F_2)).
\]
Furthermore,
\[ W_1(\infty, \Delta) \times W_2(\infty, \Delta) = \rho_1(A_D^1(W_2(\infty, \Delta))) \times \rho_2(A_D^2(W_1(\infty, \Delta))). \]

**Proof.** First note that compactness of \( \Sigma \) and regularity of \( \Delta \) imply that each set \( \Delta^i \) is compact as well. Then observe that for every measurable subset \( \emptyset \neq E_{-i} \subseteq \Sigma_{-i} \) such that \( \text{proj}_{\Theta_{-i}} E_{-i} = \Theta_{-i} \) the following holds:
\[
\emptyset \neq \{ \mu^i \in \Delta^\Theta_i(\Sigma_{-i}) : \forall h \in \mathcal{H}, E_{-i} \cap \Sigma_{-i}(h) \neq \emptyset \Rightarrow \mu^i(E_{-i} | \Sigma_{-i}(h)) = 1 \} \cap \Delta^i
\]
non-emptiness follows from measurability and the fact that, since \( \text{proj}_{\Theta_{-i}} E_{-i} = \Theta_{-i} \) and \( \Delta \) is regular, we are taking the intersection of non-empty sets characterized by logically independent properties. The inclusion holds because \( \Sigma_{-i}(\phi) = \Sigma_{-i} \) and \( E_{-i} \cap \Sigma_{-i}(\phi) \neq \emptyset \). The last equality is true by definition. Finally note that (a), (b) and (c) are true by definition for \( k = 0 \). Assume that (a), (b) and (c) hold for all \( k = 0, \ldots, n \).

(a, \( n + 1 \)) By the inductive hypothesis, the argument above implies the sets of weakly and strongly \((n, \Delta)\)-rationalizable beliefs \( A_D^i(W_{-i}(n, \Delta)) \) and
\[
\Phi^i(n, \Delta) = \bigcap_{k=0}^{n} \{ \mu^i \in \Delta^i : \forall h \in \mathcal{H}, \Sigma_{-i}(h) \cap \Sigma_{-i}(k, \Delta) \neq \emptyset \Rightarrow \mu^i(\Sigma_{-i}(k, \Delta) | \Sigma_{-i}(h)) = 1 \}
\]
are non-empty and compact. Since \( r_i(\cdot, \cdot) \) is a non-empty-valued, upper-hemicontinuous and \( \Theta_{-i} \) is closed, each set \( \rho_i(\mu^i) = \bigcup_{\theta_i \in \Theta_i} \{ \theta_i \} \times r_i(\theta_i, \mu^i) \) is non-empty and closed and correspondence \( \rho_i(\cdot) \) is upper-hemicontinuous. Therefore the sets of weakly and strongly \((n + 1, \Delta)\)-rationalizable pairs
\[
\rho_i(A_D^i(W_{-i}(n, \Delta)))
\]
and
\[
\Sigma_i(n + 1, \Delta) = \rho_i(\Phi^i(n, \Delta))
\]
are non-empty and compact. Furthermore, non-emptiness of \( r_i(\cdot, \cdot) \) implies that their projections on \( \Theta_i \) coincide with \( \Theta_i \). This proves that (a) holds for all non-negative integers \( k \). Clearly, compactness and the projection property hold also for \( k = \infty \). Since the sequences of weakly and strongly \((k, \Delta)\)-rationalizable sets are nested, non-emptiness of
\[
W_i(\infty, \Delta) = \bigcap_{k \geq 0} W_i(k, \Delta)
\]
and
\[
\Sigma_i(\infty, \Delta) = \bigcap_{k \geq 0} \Sigma_i(k, \Delta)
\]
follows from the finite intersection property of compact sets.
Clearly the inclusion holds in the limit as \( k \to \infty \). We first show that (d) (The following argument is a simple generalization of the proof of Proposition 3.1 in Bernheim (1984).) We now suppose that
\[
W_i(n + 1, \Delta) = p_i(A_i'(W_i(n, \Delta)) \subseteq p_i(A_i'(W_i(n, \Delta))) \subseteq p_i(A_i'(W_i(n, \Delta)))
\]
Thus we obtain
\[
\Sigma_i(n + 1, \Delta) = \rho_i(\Phi_i(n, \Delta)) \subseteq \rho_i(\Phi_i(n, \Delta)) \subseteq \rho_i(\Phi_i(n, \Delta))
\]
Clearly the inclusion holds in the limit as \( k \to \infty \).

(c, \( n + 1 \)) By the inductive hypothesis and part (a) \( W_\Delta(n, \Delta) \subseteq W_\Delta(n, \Delta') \) and both sets are measurable. By monotonicity of operator \( \rho_i A_i' \) we obtain
\[
W_i(n + 1, \Delta) = p_i(A_i'(W_i(n, \Delta))) \subseteq p_i(A_i'(W_i(n, \Delta))) \subseteq p_i(A_i'(W_i(n, \Delta)))
\]
Thus monotonicity of the operator \( \rho_i A_i' \) on the Borel sigma algebra of \( \Sigma_i \) (\( i = 1, 2 \)) implies
\[
F_1 \times F_2 \subseteq p_i(1(F_2)) \times p_i(2(F_1)) \subseteq p_i(1(W_i(k, \Delta))) \times p_i(2(W_i(k, \Delta)))
\]
Clearly the inclusion holds in the limit as \( k \to \infty \).

Now we show that \( W_i(\Delta) \times W_i(\Delta) \) has the 'best-reply property' and is a 'fixed set.' By part (a) each set \( W_i(k, \Delta) \) is measurable. Thus, monotonicity of the operator \( \rho_i A_i' \) on the Borel sigma algebra of \( \Sigma_i \) (\( i = 1, 2 \)) implies
\[
W_i(k + 1, \Delta) = W_i(k + 1, \Delta) \times W_i(k + 1, \Delta).
\]
Now suppose that \( \sigma_i \subseteq W_i(\Delta) \). Then there exists a sequence of CPSs \( \mu^{k} \) such that for all \( k, \mu^{k} \subseteq \Delta' \mu^{k}(W_i(k, \Delta) \Sigma_i) = 1 \) and \( \sigma_i \subseteq \mu^{k} \). Since \( \Delta^{k}(\Sigma_i) \) is compact, we may assume w.l.o.g. that \( \mu^{k} \to \mu \). Since \( \Delta' \) is closed, \( \mu' \subseteq \Delta' \). Furthermore, it must be
the case that $\mu^i(W_{-i}(k, \Delta) \mid \Sigma_{-i}) = 1$ for all $k$ (otherwise, $\mu^{ik}$ could not converge to $\mu^i$) and thus (by continuity of the measure $\mu^i(\cdot \mid \Sigma_{-i})$) $\mu^i(W_{-i}(\infty, \Delta) \mid \Sigma_{-i}) = 1$. Since $\rho_i$ is upper-hemicontinuous, $\sigma_i \in \rho_i(\mu^i)$. This shows that $W_1(\infty, \Delta) \times W_2(\infty, \Delta)$ has the ‘best reply property’

$$W_1(\infty, \Delta) \times W_2(\infty, \Delta) \subseteq \rho_1(A_1^1(W_2(\infty, \Delta))) \times \rho_2(A_2^1(W_1(\infty, \Delta))).$$

Hence $W_1(\infty, \Delta) \times W_2(\infty, \Delta)$ is a ‘fixed set.’ □

**Remark 2.** The proof of part (b) uses only the fact that the sets of weakly and strongly $(k, \Delta)$-rationalizable profiles are measurable and nonempty. The proof of part (c) relies only on measurability of the sets of weakly $(k, \Delta)$-rationalizable profiles.

**Proof of Proposition 3.8.** By Proposition 8.1 (b) we only have to consider the relationship between weak rationalizability and dominance. Take an arbitrary finite game. If $(\theta_i, s_i) \in \rho_i(\mu^i)$, then $s_i$ is a best reply to the (prior) belief $\mu^i(\cdot \mid \Sigma_{-i})$ for type $\theta_i$. This implies that $s_i$ cannot be strictly dominated for type $\theta_i$. Thus for every rectangular subset $B \subseteq \Sigma$

$$\rho_1(A_1^1(B_2)) \times \rho_2(A_2^1(B_1)) \subseteq \mathcal{F}(B).$$

(a) If the game is static, then it is also true that

$$\mathcal{F}(B) \subseteq \rho_1(A_1^1(B_2)) \times \rho_2(A_2^1(B_1))$$

(the proof can be easily adapted from Pearce, 1984, Lemma 3) and a standard inductive argument proves (a).

(b) If we assume that the game has no relevant ties, then $W(1) \subseteq \mathcal{H}_p(\Sigma)$ (the proof can be adapted from Battigalli, 1997, Lemma 3). Thus $W(1) \subseteq \mathcal{F}(\Sigma) \cap \mathcal{H}_p(\Sigma) = \mathcal{F}\mathcal{H}_p(\Sigma)$. Suppose that

$$W(n) \subseteq \mathcal{F}\mathcal{H}_p^n(\Sigma).$$

Then

$$W(n + 1) = \rho_1(A_1^1(W_2(n))) \times \rho_2(A_2^1(W_1(n))) \subseteq \mathcal{F}(\mathcal{H}_p^n(\Sigma)) \cap \mathcal{H}_p(\Sigma) = \mathcal{F}\mathcal{H}_p^{n+1}(\Sigma).$$

This proves statement (b).

(c) In every perfect information game with private values, $\mathcal{H}_p(\Sigma) = \mathcal{H}(\Sigma)$ (Battigalli, 1997, Lemma 4, shows this result for games with perfect and complete information, the proof can be easily adapted to cover the present more general case). Thus, if the game has no relevant tie, part (b) implies $W(k) \subseteq \mathcal{F}\mathcal{H}_p^k(\Sigma)$ for all $k$.\(^\text{38}\)

\(^{38}\) Ben Porath (1997, Lemma 2.1) independently proved that, in generic games with perfect (and complete) information, $W(1) \subseteq W(\Sigma)$.
for type \( \theta_i \) (Pearce (1984, Lemmas 3 and 4)). Construct \( \mu_i^t \in [\Delta(\Sigma)]^{\theta_i} \) as follows: for all \( h \in \mathcal{H}, B_{-i} \subseteq \Sigma_{-i}(h), \)

\[
\mu_i^t(B_{-i}\mid \Sigma_{-i}(h)) = \frac{\nu(B_{-i})}{\nu(\Sigma_{-i}(h))},
\]

where \( \nu = \nu' \), if \( \nu'(\Sigma_i(h)) > 0 \), and \( \nu = \nu'' \) otherwise. It can be checked that \( \mu_i^t \) is indeed a CPS \( (\mu_i^t \in \Delta^{\theta_i}(\Sigma_{-i})), \mu_i^t(W_{-i}(n)\mid \Sigma_{-i}) = 1 \) and \( (\theta_i, s_i) \in \rho_i(\mu_i^t) \). Thus \( (\theta_i, s_i) \in W_i(n) \). \( \square \)

**Proof of Proposition 3.9.** Fix a connected product subset \( B = \prod_{i \in N} B_i \subseteq \Sigma \) and a player \( i \).

First note that under the stated assumptions \( r_i(\cdot, \cdot) \) is a continuous function. Therefore, for each \( \theta_i, \) the image of \( B_{-i} \) through \( r_i(\theta_i, \cdot) \) is a closed connected subset of \( A_i \subseteq \mathbb{R} \); hence it is a compact interval, say \( r_i(\theta_i, B_{-i}) = [\tilde{b}_i, \tilde{b}_i] \). Clearly, no action in \( [\tilde{b}_i, \tilde{b}_i] \) is strictly dominated for type \( \theta_i \) on \( B_{-i} \). Now we show that every action outside \( [\tilde{b}_i, \tilde{b}_i] \) is strictly dominated on \( B_{-i} \) for type \( \theta_i \).

Suppose that \( a_i > \tilde{b}_i \) and fix \( (\theta_{-i}, a_{-i}) \in B_{-i} \) arbitrarily. By assumption

\[
u_i(\theta_i, \theta_{-i}, r_i(\theta_i, \theta_{-i}, a_{-i}), a_{-i}) > u_i(\theta_i, \theta_{-i}, a_i, a_{-i}).
\]

Since \( u_i \) is strictly quasi-concave and \( \tilde{b}_i = \alpha_i r_i(\theta_i, \theta_{-i}, a_{-i}) + (1 - \alpha_i)a_i \) for some \( \alpha_i \in (0, 1], \) then

\[
u_i(\theta_i, \theta_{-i}, \tilde{b}_i, a_{-i}) > u_i(\theta_i, \theta_{-i}, a_i, a_{-i}).
\]

Therefore \( \tilde{b}_i \) strictly dominates \( a_i \) for type \( \theta_i \) on \( B_{-i} \). A similar argument shows that every action \( a_i < \tilde{b}_i \) is strictly dominated on \( B_{-i} \) by \( \tilde{b}_i \) (for type \( \theta_i \)). This proves that

\[
\mathcal{F}(B) = \prod_{i \in N} \rho_i(B_{-i}).
\]

Furthermore, each subset \( \rho_i(B_{-i}) \) is the graph of the correspondence \( \theta_i \mapsto r_i(\theta_i, B_{-i}), \) which—by continuity of \( r_i \)—is upper-hemicontinuous with connected values. Since \( \Theta_i \) is connected, also the graph \( \rho_i(B_{-i}) \) must be connected. Hence \( \mathcal{F}(B) \) is connected.

The second claim of the proposition follows by induction. First recall that \( \Sigma(k) = W(k) = \prod_{i \in N} \rho_i(\Delta(\Sigma_{-i}(k - 1))).^{39} \) Then note that by definition, for every product subset \( B \subseteq \Sigma, \prod_{i \in N} \rho_i(B_{-i}) \subseteq \prod_i [\rho_i(\Delta(\Sigma_{-i})) \subseteq \mathcal{F}(B). \) By assumption \( \Sigma = \Sigma(0) = W(0) \) is a connected product set. Therefore the previous result implies

\[
\prod_{i \in N} \rho_i(\Sigma_{-i}(0)) = \Sigma(1) = W(l) = \mathcal{F}(\Sigma).
\]

and \( \mathcal{F}(\Sigma) \) is connected. Assume by way of induction that

\[
\prod_{i \in N} \rho_i(\Sigma_{-i}(k)) = \Sigma(k) = W(k + 1) = \mathcal{F}(\Sigma)
\]

\( ^{39} \Sigma(k) = W(k) \) because the game is static and hence weak and strong rationalizability coincide.
and $S^k(\Sigma)$ is connected. Then the previous result implies that

$$\prod_{i \in N} \rho_i(\Sigma_{-i}(k + 1)) = \Sigma(k + 1) = W(k + 1) = S^k(\Sigma) = S^{k+1}(\Sigma)$$

and $S^{k+1}(\Sigma)$ is connected. □

References


Stanford University.