

Rationalizability and Epistemic Priority Orderings*

Emiliano Catonini[†]

October 2017

At the beginning of a dynamic game, players may have exogenous theories about how the opponents are going to play. Suppose that these theories are commonly known. Then, players will refine their first-order beliefs and challenge their own theories through strategic reasoning. I develop and characterize epistemically a new solution concept, Selective Rationalizability, which accomplishes this task under the following assumption: when the observed behavior is not compatible with the beliefs in rationality and in the theories of all orders, players keep the orders of belief in rationality that are per se compatible with the observed behavior, and drop the incompatible orders of belief in the theories. Thus, Selective Rationalizability captures Common Strong Belief in Rationality (Battigalli and Siniscalchi, 2002) and refines Extensive-Form Rationalizability (Pearce, 1984; BS, 2002), whereas Strong- Δ -Rationalizability (Battigalli, 2003; Battigalli and Siniscalchi, 2003) captures the opposite epistemic priority choice. Selective Rationalizability is extended to encompass richer epistemic priority orderings among different theories of opponents' behavior. This allows to establish a surprising connection with strategic stability (Kohlberg and Mertens, 1986).

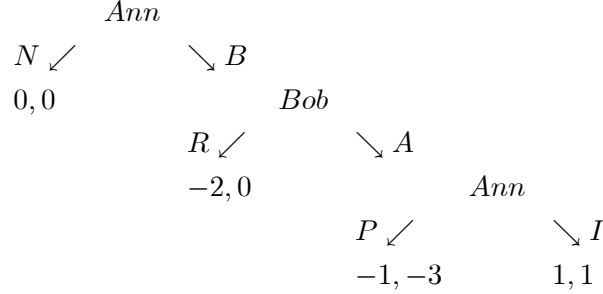
Keywords: Forward induction, Strong Belief, Strong Rationalizability, Strong- Δ -Rationalizability, Strategic Stability.

*A special thanks goes to Pierpaolo Battigalli. Thank you also to two anonymous referees and to all the attendants of the presentations at LOFT 2016, TARK 2017, and the reading group on epistemic game theory at Bocconi University, in particular to Andres Perea, Burkhard Schipper, Roberto Corrao and Nicodemo De Vito for their comments.

[†]Higher School of Economics, ICEF, emiliano.catonini@gmail.com

1 Introduction

Consider the following dynamic game with perfect information.



Ann can try to *Bribe* Bob, a public officer, or *Not*. If she does, Bob can *Accept* or *Report* her, so that Ann loses two utils. If Bob accepts, Ann can *Implement* her plan, achieving the Pareto dominating outcome, or repent (*P*) and speak with a prosecutor, harming both Bob and herself.

Suppose that Ann is rational¹ and, at the beginning of the game, believes with probability 1 that Bob would play *R* after *B*. I call this belief "(first-order belief) restriction". Then, she plays *N*. Suppose that Bob believes that Ann is rational and that the restriction holds. Then, he expects Ann to play *N*. So, what would he believe after observing *B*? He cannot believe at the same time that Ann is rational and that the restriction holds: the two things are at odds given *B*. Which of the two beliefs will Bob keep? This is the epistemic priority issue. Suppose that he keeps the belief that the restriction holds. So, he drops the belief that Ann is rational. Then, he can also expect Ann to play *P* after (*B*, *A*) and thus play *R*. If Ann believes that Bob reasons in this way, she can keep her restriction and play *N*.

These lines of strategic reasoning are captured by Strong- Δ -Rationalizability (Battigalli, [4]; Battigalli and Siniscalchi, [10]). In this reasoning process, the faith in the restrictions is so strong that Bob is ready to deem Ann irrational after the bribing attempt. This could be the case if, for instance, the belief that Bob would report Ann is induced by a commonly known social convention that always holds in context of the game (see Battigalli and Friedenberg [5]). Suppose instead that, in the context of the game, public officers are not commonly believed to be incorruptible. However, Bob declares that he would play *R* after *B*. If Bob observes that Ann plays *B* anyway, he might think that Ann has not taken his words seriously, rather than thinking that Ann is irrational. Then,

¹i.e. subjective expected utility maximizer given her beliefs at every information set.

he would expect Ann to play I after A , hence he would play A instead of R . If Ann believes that Bob is rational and keeps believing that she is rational after B , she must believe that Bob will play A , differently than what the restriction suggests. Hence, under this reasoning scheme, such restriction to first-order beliefs cannot hold.

Note that opposite conclusions were reached without any uncertainty about payoffs: the two situations do not represent different types of Bob, but only different strengths of the belief that he would report Ann.

In Section 3, I construct an elimination procedure, Selective Rationalizability, that captures these instances of forward induction reasoning in all dynamic games with perfect recall and countably many conditioning events,² although for notational simplicity the formal analysis focuses on finite games with complete information. Selective Rationalizability refines a notion of Extensive-Form Rationalizability (Pearce [29], Battigalli [2], Battigalli and Siniscalchi [9]), which I will call "Rationalizability" for brevity. Thus, Selective Rationalizability represents a natural way for players to refine their beliefs through (*possibly partial*) coordination and consequent forward induction considerations when lone strategic reasoning about rationality does not pin down a unique plan of actions. As above, Selective Rationalizability delivers an empty set when the "tentative" first-order belief restrictions of a player are at odds with strategic reasoning. Strong- Δ -Rationalizability, instead, does not refine Rationalizability: in the example, N is not a rationalizable outcome.³ It is worth noting that Selective Rationalizability can also be seen as an instance of Strong- Δ -Rationalizability, where the restrictions are the conjunction of the exogenous theories and the rationalizable first-order beliefs. However, keeping the two separate has both conceptual and technical advantages. The separation allows to investigate the epistemic priority issue between the two different sources of beliefs, and to compare Strong- Δ -Rationalizability and Selective Rationalizability for the *same* restrictions. It turns out that Strong- Δ -Rationalizability and Selective Rationalizability are outcome-equivalent when the restrictions correspond to the belief in a specific path of play.⁴ In general, one could expect Selective Rationalizability to always yield a (possibly empty) subset of the strongly- Δ -rationalizable outcomes. A counterexample in the Discussion Section shows that, opposite to the example above, Selective Rationalizability can

²For instance, infinitely repeated games with a finite stage game, or games with uncountably many available actions only at preterminal histories.

³The game has no simultaneous moves and no relevant ties. Therefore, as shown by Battigalli [3] first and Chen and Micali [16], Heifetz and Perea [22], and Perea [30] later, extensive-form rationalizability (in all its variants, including the one of this paper) delivers the unique backward induction outcome.

⁴The proof of this result is rather sophisticated and it is presented in [15], where the focus is on the algorithms and not on their epistemic foundations. The proof cannot be performed if Selective Rationalizability is formalized as a special case of Strong- Δ -Rationalizability.

yield non-empty predictions when Strong- Δ -Rationalizability rejects the first-order belief restrictions; a counterexample in the Appendix shows that Selective Rationalizability and Strong- Δ -Rationalizability can even yield non-empty disjoint predictions.

In Section 4, I clarify with an epistemic characterization the strategic reasoning hypotheses that motivate Selective Rationalizability. Selective Rationalizability captures the behavior of rational players who restrict their beliefs about opponents' behavior for some exogenous reason. Moreover, at the beginning of the game, players believe that opponents are rational and have their own restrictions; that opponents believe that everyone else is rational and has the own restrictions; and so on. These beliefs are tentative because at some information set of a player, the observed behavior of one opponent may be incompatible, say, with the opponent being rational and, at the same time, having beliefs in her restricted set. In this case, our player will drop the belief that the opponent has such restrictions, rather than dropping the belief that the opponent is rational. More generally, players always keep all orders of belief in rationality that are per se compatible with the observed behavior, and drop all orders of belief in the restrictions that are at odds with them. I call this choice *epistemic priority to rationality*. Strong- Δ -Rationalizability predicts instead the behavior of players who assign epistemic priority to the beliefs in the restrictions, and drop the incompatible beliefs in rationality. Thus, Selective Rationalizability captures a version of Common Strong Belief in Rationality (Battigalli and Siniscalchi, [9]), whereas Strong- Δ -Rationalizability does not. However, both solution concepts capture all orders of belief in rationality and in the restrictions along the induced paths, if non-empty. Since the epistemic priority issue materializes only off-path, it is hard to grasp why Strong- Δ -Rationalizability and Selective Rationalizability can yield radically different predictions. A deeper look into their epistemic characterizations and the Discussion Section will clarify how the epistemic priority affects predictions.

In Section 5, I extend the analysis to finer epistemic priority orderings. Each player can have multiple theories, say two, about opponents' behavior: a weaker theory and a stronger theory (in the sense of more restrictive). Players reason according to everyone's weaker theory like under Selective Rationalizability. On top of this, as long as compatible with strategic reasoning about the weaker theories, players reason according to the stronger theories. So, when a player displays behavior which is not compatible with strategic reasoning about both theories, the opponents keep believing that the player is reasoning according to the weaker theories, and drop the belief that the opponent is reasoning according to the stronger ones.⁵ When the two theories correspond to an equilibrium

⁵By non-monotonicity of strong belief, strategic reasoning about the stronger theories can potentially

path and an equilibrium strategy profile, a surprising connection with strategic stability (Kohlberg and Mertens [23]) can be established. In Section 5 I provide an example and I lay the foundations of this bridge with an extended version of Selective Rationalizability, which encompasses nested restrictions in an epistemic priority order.

Since players' theories of opponents' behavior are assumed to be commonly known, the most natural application of Selective Rationalizability is probably explicit pre-play coordination among players. A non-binding agreement is purely cheap talk; hence, if a player displays behavior which is not compatible with rationality and belief in the agreement, the opponents are, in my view, more likely to abandon the belief that the player believes in the agreement, rather than the belief that the player is rational. Or, as in the example, the source of belief restrictions can be a public announcement.⁶ Thus, Selective Rationalizability seems to be an appropriate tool to combine strategic reasoning and equilibrium play, especially when the motivation for equilibrium is explicit coordination. The application of Selective Rationalizability to agreements and its relationship with equilibrium are deeply investigated in [14]. In particular, the outcomes that Selective Rationalizability uniquely pins down for some restrictions do not include and are not included in the set of subgame perfect equilibrium outcomes. It is worth noting that the flexibility of Selective Rationalizability, which allows to model incomplete coordination instead of coordination on full strategy profiles, can be crucial to induce an outcome of the game (see [14] for details).

The Appendix contains the proofs of the results and the formal analysis of the counterexample mentioned above.

2 Preliminaries

Description of the game. Consider a finite dynamic game with complete information and perfect recall $\Gamma = \langle I, X, (\bar{A}_i, H_i, u_i)_{i \in I} \rangle$ where:⁷

- I is the finite set of *players*, and for any *profile* $(X_i)_{i \in I}$ and any $\emptyset \neq J \subseteq I$, I write $X_J := \times_{j \in J} X_j$, $X := X_I$, $X_{-i} := X_{I \setminus \{i\}}$, $X_{-i,j} := X_{I \setminus \{i,j\}}$;

lead to behavior that cannot be rationalized under the weaker theories. For this reason, the epistemic priority issue arises.

⁶Or, extending Selective Rationalizability to games with incomplete information, the restrictions can model public news about a state of nature. For instance, in a financial market, players can tentatively believe that everyone is behaving according to the same public information about the value of an asset. Yet, if a player does not behave accordingly, the opponents may believe that the player has different information rather than deeming the player irrational.

⁷The notation for the game is mainly taken from Osborne and Rubinstein [28].

- \bar{A}_i is the finite set of *actions* of player i ;
- $X \subseteq \bigcup_{t \in \{0, \dots, T\}} \left(\bigcup_{\emptyset \neq J \subseteq I} \bar{A}_J \right)^t$ is the finite set of *histories*, where T is the finite horizon, and:
 1. $\bar{A}_J^0 := \emptyset =: h^0 \in X$, i.e. X contains the initial empty history;
 2. for every $(\tilde{a}^1, \dots, \tilde{a}^l) \in X$ and every $t < l$, $(\tilde{a}^1, \dots, \tilde{a}^t) \in X$, and I write $(\tilde{a}^1, \dots, \tilde{a}^t) \prec (\tilde{a}^1, \dots, \tilde{a}^l)$;
 3. there exist a correspondence $J(x) : X \rightrightarrows I$ and, for every $i \in I$, a non-empty-valued correspondence $\tilde{A}_i : \{x \in X : i \in J(x)\} \rightrightarrows \bar{A}_i$ such that for every $x \in X$, $(x, a) \in X$ if and only if $a \in \times_{j \in J(x)} \tilde{A}_j(x)$;
 4. $Z := \{x \in X : J(x) = \emptyset\}$ is the set of *terminal histories*;
- $H_i \subset 2^X$ is the set of *information sets* of player i where:
 1. it partitions $\{x \in X : i \in J(x)\}$;
 2. for every $h \in H_i$ and $x, x' \in h$, $\tilde{A}_i(x) = \tilde{A}_i(x') =: A_i(h)$;
 3. (*perfect recall*) for every $h \in H_i$ and $x, x' \in h$, $x \not\prec x'$; moreover, for every $(\tilde{x}, \tilde{a}) \preceq x$ with $\tilde{x} \in \tilde{h}$ for some $\tilde{h} \in H_i$, there exists $(\tilde{x}', \tilde{a}') \preceq x'$ such that $\tilde{x}' \in \tilde{h}$ and $\text{Proj}_{\bar{A}_i} \tilde{a}' = \text{Proj}_{\bar{A}_i} \tilde{a}$;⁸
- $u_i : Z \rightarrow \mathbb{R}$ is the *payoff function* of player i .

Perfect recall implies that each H_i inherits the partial order \prec from X .

A *strategy* is a function $s_i : H_i \rightarrow \bar{A}_i$ such that for every $h \in H_i$, $s_i(h) \in A_i(h)$. The set of all strategies is denoted by S_i . A strategy profile clearly induces one and only one terminal history; let $\zeta : S \rightarrow Z$ denote the map that associates each strategy profile with the induced terminal history. Fix $h \in H_i$. The set of strategy profiles compatible with h is

$$S(h) := \{s \in S : \exists x \in h, x \prec \zeta(s)\}.$$

Then, for each $J \subset I$, the set of J 's strategy profiles that are compatible with h is $S_J(h) := \text{Proj}_{S_J} S(h)$. Perfect recall implies that $S(h) = S_i(h) \times S_{-i}(h)$; $S_{-i}(h)$ represents the partial observation by player i of opponents' moves up to h . To keep the strategic

⁸The first statement means that players cannot end up in the same information set twice, because they remember having moved from it the first time. The second statement means that players always distinguish two histories if they were able to distinguish two predecessors or if they follow two different own moves.

reasoning hypotheses simple, I further assume the "observable deviators" property (Fudenberg and Levine, [20]): for each $i \in I$ and $h \in H_i$,

$$S(h) = \times_{j \in I} S_j(h).$$

Observable deviators is always satisfied by games with observable actions, where information sets are singletons. I will clarify later the simplifying role of observable deviators and how it can be removed.

For any profile of strategy sets $\bar{S}_J \subset S_J$, let $\bar{S}_J(h) := S_J(h) \cap \bar{S}_J$. The set of information sets of i compatible with \bar{S}_J is

$$H_i(\bar{S}_J) := \{h \in H_i : \bar{S}_J(h) \neq \emptyset\}.$$

Beliefs. Players update their beliefs about opponents' strategies and beliefs as the game unfolds. A Conditional Probability System (Renyi, [31]; henceforth CPS) assigns to each information set a belief, conditional on the observed opponents' behavior. Here I define CPS's over the opponents' state space $\Omega_{-i} := \times_{j \neq i} (S_j \times T_j)$, where epistemic type spaces $(T_j)_{j \in I}$ will be defined in Section 4.

Definition 1 *A CPS on $(\Omega_{-i}, (T_{-i} \times S_{-i}(h))_{h \in H_i})$, with Borel sigma algebra $\mathbf{B}(\Omega_{-i})$, is an array of probability measures $(\mu_i(\cdot|h))_{h \in H_i}$ on $(\Omega_{-i}, \mathbf{B}(\Omega_{-i}))$ such that for each $h \in H_i$, $\mu_i(T_{-i} \times S_{-i}(h)|h) = 1$, and (chain rule) for every $E \in \mathbf{B}(\Omega_{-i})$ and $C, D \in (T_{-i} \times S_{-i}(h))_{h \in H_i}$, if $E \subseteq D \subseteq C$ then $\mu_i(E|D)\mu_i(D|C) = \mu_i(E|C)$.*

The set of all CPS's of player i is denoted by $\Delta^{H_i}(\Omega_{-i})$.⁹ CPS's on strategies are defined by replacing Ω_{-i} with S_{-i} and $(T_{-i} \times S_{-i}(h))_{h \in H_i}$ with $(S_{-i}(h))_{h \in H_i}$.

For any $J \subseteq I \setminus \{i\}$ and $\bar{S}_J \subseteq S_J$, I say that $\mu_i \in \Delta^{H_i}(S_{-i})$ *strongly believes* (Battigalli and Siniscalchi, [9])¹⁰ \bar{S}_J if $\mu_i(\bar{S}_J \times S_{I \setminus (J \cup \{i\})}|h) = 1$ for all $h \in H_i(\bar{S}_J)$. Thanks to observable deviators, there always exists a CPS μ_i that strongly believes \bar{S}_J and at the same time any given $\bar{S}_K \subseteq S_K$, $K \subseteq I \setminus (J \cup \{i\})$. This is because, under observable deviators, if $h \in H_i(\bar{S}_J) \cap H_i(\bar{S}_K)$, then $h \in H_i(\bar{S}_J \times \bar{S}_K)$.

Rationality. I consider players who reply rationally to their beliefs. By rationality I mean that players, at every information set, choose an action that maximizes expected

⁹If each Ω_j is compact metrizable, endowing the set $\Delta(\Omega_{-i})$ of Borel probability measures on Ω_{-i} with the topology of weak convergence and $(\Delta(\Omega_{-i}))^{H_i}$ with the product topology, Battigalli and Siniscalchi [7] proved that $\Delta^{H_i}(\Omega_{-i})$ is a compact metrizable subset of $(\Delta(\Omega_{-i}))^{H_i}$.

¹⁰Battigalli and Siniscalchi make a stricter use of the term strong belief, by referring only to Borel subsets of Ω_{-i} or S_{-i} .

utility given their belief about how the opponents will play and the expectation to choose rationally again in the continuation of the game. By standard dynamic programming arguments, this is equivalent to playing a *sequential best reply* to the CPS.

Definition 2 Fix $\mu_i \in \Delta^{H_i}(S_{-i})$. A strategy $s_i \in S_i$ is a sequential best reply to μ_i if for each $h \in H_i(s_i)$, s_i is a continuation best reply to $\mu_i(\cdot|h)$, i.e. for all $\tilde{s}_i \in S_i(h)$,

$$\sum_{s_{-i} \in S_{-i}(h)} u_i(\zeta(s_i, s_{-i})) \mu_i(s_{-i}|h) \geq \sum_{s_{-i} \in S_{-i}(h)} u_i(\zeta(\tilde{s}_i, s_{-i})) \mu_i(s_{-i}|h).$$

The set of sequential best replies to μ_i is denoted by $\rho(\mu_i)$.

3 Selective Rationalizability

In dynamic games, forward induction reasoning about rationality has already been studied under different assumptions. Pearce [29] defines Extensive-Form Rationalizability under the hypothesis that conditional beliefs satisfy *structural consistency* (Kreps and Wilson [24]), that is, that they can be generated by a prior product distribution on S_{-i} . Battigalli [2] assumes *strategic independence*, which (roughly speaking) requires players to maintain the belief about each opponent as long as her individual behavior does not contradict it. Battigalli and Siniscalchi [9] remove any assumption of independence and require players to maintain each order of belief in rationality only until *none* of the opponents contradict it. Then, they give to the resulting elimination procedure, *Strong Rationalizability*, an epistemic characterization based on the notion of *strong belief*. For this reason, I adopt Strong Rationalizability as a starting point, but I amend it by introducing *independent rationalization*: players maintain an order of belief in rationality of an opponent as long as her individual behavior does not contradict it. The motivation for this choice is two-fold. First, it is coherent with the emphasis on the persistence of beliefs in rationality. Second, it will allow to better compare Selective Rationalizability with equilibrium refinements, as discussed later. As far as Strong Rationalizability is concerned, it is easy to observe that independent rationalization is immaterial for the predicted outcomes, since it kicks in at an information set only when it is not reached anymore by some player. However, the whole analysis can be read without independent rationalization by simply substituting j , which will indicate one opponent of player i , with $-i$, i.e. all opponents of i jointly considered. Instead, I do not adopt strategic independence. This is not in contradiction with independent rationalization: there can be correlations¹¹ also among the choices of

¹¹For instance, a player can believe that two opponents get the same signal of her own intentions, regardless of their strategic sophistication. See also Aumann [1] and Brandenburger and Friedenberg [13]

players with different levels of sophistication. However, assuming strategic independence would complicate the notation but not alter the results.

For brevity and to distinguish it from the original notion of Strong Rationalizability, I will call this version simply "Rationalizability".

Definition 3 (Rationalizability) *Consider the following procedure.*

(Step 0) *For each $i \in I$, let $S_i^0 = S_i$.*

(Step $n > 0$) *For each $i \in I$ and $s_i \in S_i$, let $s_i \in S_i^n$ if and only if there is $\mu_i \in \Delta^{H_i}(S_{-i})$ such that:*

R1 $s_i \in \rho(\mu_i)$;

R2 μ_i strongly believes S_j^q for all $j \neq i$ and $q < n$.

Finally let $S_i^\infty = \cap_{n \geq 0} S_i^n$. The profiles in S^∞ are called rationalizable.

Note that R2 can always be satisfied thanks to observable deviators. Therefore, Rationalizability always yields a non-empty output. In absence of observable deviators, in place of strong belief in S_j^q for all $j \neq i$, a player can instead strongly believe that (i) all opponents play a strategy in S_j^q , (ii) all opponents but one play a strategy in S_j^q , and so on. These hypotheses correspond to nested sets of opponents' strategy profiles, which can be strongly believed at the same time regardless of the structure of information sets. In another paper, Battigalli and Siniscalchi [8] adopt instead a weaker but more complicated notion of independent rationalization ("independent best rationalization").

Selective Rationalizability refines Rationalizability in the following way. Each player has an exogenous theory of opponents' behavior and refines the rationalizable first-order beliefs according to this theory. The theory of player i is represented by a set of CPS's $\Delta_i \subseteq \Delta^{H_i}(S_{-i})$ over opponents' strategies. Players are aware of the theories of everyone else. Therefore, they can also expect *each* opponent to refine her first-order beliefs according to the own theory. This belief towards an opponent is maintained as long as the opponent herself is not observed making a move that contradicts it. Moreover, players expect *each* opponent to reason about everyone else in the same way. Also this belief is maintained as long as the opponent herself does not make a move that contradicts it. And so on. Thus, Selective Rationalizability is defined under independent rationalization. This allows better comparability with the equilibrium literature. Without independent rationalization, if a player deviates from the agreed-upon path, each opponent is free to believe that any

for motivation of spurious correlations among players' strategies.

other opponent is not going to implement her threat. In this way, no coordination of threats would be required. These issues are widely discussed in [14]. Note however that independent rationalization is immaterial for the message of this paper and for the analysis of all the examples: players are only two in all games except for the game of Section 5, where independent rationalization plays no role anyway.

Definition 4 (Selective Rationalizability) *Fix a profile $(\Delta_i)_{i \in I}$ of compact subsets of CPS's. Let $((S_i^m)_{i \in I})_{m=0}^\infty$ denote the Rationalizability procedure. Consider the following procedure.*

(Step 0) *For each $i \in I$, let $S_{i,R\Delta}^0 = S_i^\infty$.*

(Step $n > 0$) *For each $i \in I$ and $s_i \in S_i$, let $s_i \in S_{i,R\Delta}^n$ if and only if there is $\mu_i \in \Delta_i$ such that:*

S1 $s_i \in \rho(\mu_i)$;

S2 μ_i strongly believes $S_{j,R\Delta}^q$ for all $j \neq i$ and $q < n$;

S3 μ_i strongly believes S_j^q for all $j \neq i$ and $q \in \mathbb{N}$.

Finally, let $S_{i,R\Delta}^\infty = \cap_{n \geq 0} S_{i,R\Delta}^n$. The profiles in $S_{R\Delta}^\infty$ are called selectively-rationalizable.

Step 0 initializes Selective Rationalizability with the rationalizable strategy profiles. This is only to stress that Selective Rationalizability refines Rationalizability: S3 already implies that players strongly believe in the rationalizable strategies of each opponent, and that the strategies surviving step 1 are rationalizable. Indeed, Selective Rationalizability can also be seen as an "extension" of Rationalizability, in a unique elimination procedure where the restrictions kick in once no more strategies can be eliminated otherwise.

In absence of observable deviators, S2 and S3 can be modified in the same fashion of R2.

Selective Rationalizability can be simplified in different ways according to the structure of the restrictions. S3 can be eliminated by requiring strategies to be rationalizable when first-order beliefs are not restricted at the non-rationalizable information sets. Let $((\hat{S}_{j,R\Delta}^n)_{j \in I})_{n=0}^\infty$ denote Selective Rationalizability redefined with $s_i \in S_i^\infty$ in place of S3.

Definition 5 *I say that $\Delta_i \subseteq \Delta^{H_i}(S_{-i})$ is maximal if for every $\mu_i \in \Delta_i$ and μ_i^* with $\mu_i^*(S_{-i}(z)|h) = \mu_i(S_{-i}(z)|h)$ for all $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$, $\mu_i^* \in \Delta_i$.*

Proposition 1 *Suppose that for every $i \in I$, Δ_i is maximal. Then, $\zeta(\hat{S}_{R\Delta}^\infty) = \zeta(S_{R\Delta}^\infty)$.*

Proposition 2 *Fix compact $\Delta = (\Delta_i)_{i \in I}$ with $S_{R\Delta}^\infty \neq \emptyset$. There exists a profile $(\Delta_i^*)_{i \in I}$ of compact maximal subsets of CPS's such that $\zeta(S_{R\Delta^*}^\infty) = \zeta(S_{R\Delta}^\infty)$.*

Thus, the class of maximal restrictions suffices to yield all the possible behavioral implications of Selective Rationalizability.¹²

Selective Rationalizability is an elimination procedure. So, a classical question is whether it can be defined as a *reduction* procedure, i.e. a procedure where step n can be computed based on step $n - 1$ only. Battigalli and Prestipino [6] identify a class of restrictions, "closed under composition", under which Strong- Δ -Rationalizability can be defined as a reduction procedure. For instance, think of restrictions Δ_i where, for some map $\eta : h \in H_i \mapsto \eta(h) \subseteq S_{-i}(h)$, $\mu_i \in \Delta_i$ if and only if $\mu_i(\eta(h)|h) = 1$ for all $h \in H_i$.¹³ Also for Selective Rationalizability, if the restrictions are maximal and closed under composition, S3 and S2 can be substituted by $s_i \in S_{i,R\Delta}^{n-1}$ and strong belief in just $(S_{j,R\Delta}^{n-1})_{j \neq i}$, in two-players games (and R2 by $s_i \in S_i^{n-1}$ and strong belief in $(S_j^{n-1})_{j \neq i}$ for Rationalizability). The same would hold in games with more than two players in absence of independent rationalization, but not under independent rationalization. The reason is the following. Call j, k two opponents of player i , and fix $h \in H_i$ with $S_{j,R\Delta}^{n-1}(h) \neq \emptyset = S_{k,R\Delta}^{n-1}(h)$ and $S_{k,R\Delta}^{n-2}(h) \neq \emptyset$. A strategy $s_i \in S_{i,R\Delta}^{n-1}(h)$ may be a sequential best reply to some μ_i that strongly believes $S_{j,R\Delta}^{n-1}$ and $S_{k,R\Delta}^{n-1}$, but not to any μ'_i that strongly believes $S_{j,R\Delta}^{n-1}$ and $S_{k,R\Delta}^{n-2}$, because it may not be a continuation best reply to any belief over $S_{j,R\Delta}^{n-1}(h) \times S_{k,R\Delta}^{n-2}(h)$.

It will be useful to compare Selective Rationalizability with Strong- Δ -Rationalizability, both in terms of epistemics (in Section 4) and in terms of predictions (in the Discussion Section). Thus, I provide here the formal definition of Strong- Δ -Rationalizability.

Definition 6 (Strong- Δ -Rationalizability, Battigalli and Prestipino [6]) *Fix a profile $\Delta = (\Delta_i)_{i \in I}$ of compact subsets of CPS's. Consider the following procedure.*

(Step 0) *For each $i \in I$, let $S_{i,\Delta}^0 = S_i$;*

(Step $n > 0$) *For each $i \in I$ and $s_i \in S_i$, let $s_i \in S_{i,\Delta}^n$ if and only if there is $\mu_i \in \Delta_i$ such that:*

D1 $s_i \in \rho(\mu_i)$;

D2 μ_i strongly believes $S_{-i,\Delta}^q$ for all $q < n$.

¹²In [14] I identify a class of agreements that suffices to yield all the possible behavioral implications of agreements. This class of agreements gives rise to restrictions that are equivalent to the corresponding maximal restrictions. The same applies to the agreements that correspond to a *Self-Enforcing Set* ([14]).

¹³The restrictions generated by agreements in [14] fall in this class.

Finally let $S_{i,\Delta}^\infty = \cap_{n \geq 0} S_{i,\Delta}^n$. The profiles in S_Δ^∞ are called *strongly- Δ -rationalizable*.

Note that, if D2 is modified like S2 to assume independent rationalization (or vice versa), Selective Rationalizability can be seen as a special case of Strong- Δ -Rationalizability, since S3 is a constant restriction on CPS's which could be incorporated in a compact Δ_i . However, to compare the two under the same restrictions and to better analyze the epistemic priority issues, Selective Rationalizability will be kept as a stand-alone solution concept, isolating the exogenous restrictions.

Selective Rationalizability and Strong- Δ -Rationalizability, differently than Rationalizability, can yield the empty set. This happens when at some step there is no $\mu_i \in \Delta_i$ that satisfies S2 and S3, or D2. This means that the restrictions are not compatible with strategic reasoning about rationality and the restrictions themselves.

To see all three procedures formally at work and yield non-empty predictions, consult the example in the Appendix.

4 Epistemic analysis

I adopt the epistemic framework of Battigalli and Prestipino [6], dropping the incompleteness of information dimension.¹⁴ Players' beliefs over strategies of all orders are given an implicit representation through a compact, complete, and continuous *type structure* $(\Omega_i, T_i, g_i)_{i \in I}$,¹⁵ where for every $i \in I$, $\Omega_i = S_i \times T_i$, T_i is a compact metrizable space of *epistemic types*, and $g_i : t_i \in T_i \mapsto (g_{i,h}(t_i))_{h \in H_i} \in \Delta^{H_i}(\Omega_{-i})$ is a continuous and onto *belief map*. I will call "events" the elements of the Borel sigma-algebras on each Ω_i , and of the product sigma algebras on the Cartesian spaces $\Omega_J := \times_{i \in J \subseteq I} \Omega_i$.

The first-order belief map of player i , $f_i : t_i \in T_i \mapsto (f_{i,h}(t_i))_{h \in H_i} \in \Delta^{H_i}(S_{-i})$, is defined as $f_{i,h}(t_i) = \text{Marg}_{S_{-i}} g_{i,h}(t_i)$ for all $i \in I$ and $h \in H_i$, so it inherits continuity from g_i . The event in Ω_i where the restrictions of player i hold is

$$[\Delta_i] := \{(s_i, t_i) \in \Omega_i : f_i(t_i) \in \Delta_i\};$$

¹⁴Note that, within the same framework, I define events in a slightly different way: a player's beliefs are not extended over the own strategy and type, and events that restrict only her strategies and types are defined in her own strategy-type space. This makes it easier to deal with independent rationalization. However, given the absence of major conceptual differences, the reader is invited to consult [6] for interesting and detailed explanations about this framework.

¹⁵Friedenberg [18] proves that in static games, such a type structure represents all hierarchies of beliefs about strategies. Although this result has not been formally extended to dynamic games, to the best of my knowledge, no counterexample has been found. However, the canonical type structure for CPS's of Battigalli and Siniscalchi [7] is compact, complete, and continuous.

$[\Delta_i]$ is compact because Δ_i is compact and f_i is continuous. The cartesian set where the restrictions of all players hold is $[\Delta] := \times_{i \in I} [\Delta_i]$.

From now on, fix a Cartesian (across players) event $E = \times_{j \in I} E_j \subseteq \Omega$. The closed¹⁶ event where player i believes in E_{-i} at an information set $h \in H_i$ is defined as

$$B_{i,h}(E_{-i}) := \{(s_i, t_i) \in \Omega_i : g_{i,h}(t_i)(E_{-i}) = 1\}$$

The closedness of $B_{i,h}(E_{-i})$ implies the closedness of all the following belief events. If $E = \times_{j \in I} (\hat{T}_j \times S_j)$ for some $(\hat{T}_j)_{j \in I}$, E is an *epistemic event* and can be believed at every information set:

$$\begin{aligned} B_i(E_{-i}) &: = \cap_{h \in H_i} B_{i,h}(E_{-i}); \\ B(E) &: = \times_{j \in I} B_j(E_{-j}). \end{aligned}$$

An epistemic event E is *transparent* when it holds and is commonly believed at every information set:

$$\begin{aligned} B^0(E) &: = E, \\ B^{n+1}(E) &: = B(B^n(E)), \\ B^*(E) &: = \cap_{n \geq 0} B^n(E). \end{aligned}$$

If E is not an epistemic event, it could be impossible for player i to believe in E_{-i} at some information set $h \in H_i$, because $\text{Proj}_{S_{-i}} E_{-i} \cap S_{-i}(h) = \emptyset$. However, player i may want to believe in E_{-i} as long as not contradicted by observation. The event where this persistency of the belief holds is:

$$\overline{SB}_i(E_{-i}) := \bigcap_{h \in H_i : \text{Proj}_{S_{-i}} E_{-i} \cap S_{-i}(h) \neq \emptyset} B_{i,h}(E_{-i}).$$

The "strong belief" operator \overline{SB}_i is non-monotonic: if $E_{-i} \subset F_{-i}$, it needs not be the case that $\overline{SB}_i(E_{-i}) \subset \overline{SB}_i(F_{-i})$. This will explain why Strong- Δ -Rationalizability is not a refinement of Strong Rationalizability, and Selective Rationalizability, for given restrictions, is not a refinement of Strong- Δ -Rationalizability.

Suppose now that, for each opponent j , player i believes that the true pair (s_j, t_j) is in E_j , as long as not contradicted by observation. Then I say that i strongly believes in

¹⁶Battigalli and Prestipino [6] provide an argument for the closedness of $B_{i,h}(E)$ based on the Portmanteau theorem. A direct proof based on the Prokhorov metric is available upon request.

E_j for all $j \neq i$. Formally, I define the operator

$$SB_i(E_{-i}) := \cap_{j \neq i} \overline{SB}_i(E_j \times \Omega_{-j,i}),$$

and given the independent rationalization hypothesis of the paper, from now on I will refer to this operator and not to \overline{SB}_i as the "*strong belief*" operator. Note that $(s_i, t_i) \in SB_i(E_{-i})$ if and only if, for each $j \neq i$, $g_i(t_i)$ strongly believes in E_j , i.e. $g_{i,h}(t_i)(E_j \times \Omega_{-i,j}) = 1$ for all $h \in H_i$ with $\text{Proj}_{S_j} E_j \cap S_j(h) \neq \emptyset$.

Recalling that for a profile $(X_j)_{j \in I}$, $X = \times_{j \in I} X_j$, define inductively:

$$\begin{aligned} CSB_i(E) & : = E_i \cap SB_i(E_{-i}), \\ CSB_i^0(E) & : = E_i, \\ CSB_i^{n+1}(E) & : = CSB_i(CSB^n(E)), \\ CSB_i^\infty(E) & : = \cap_{n \in \mathbb{N}} CSB_i^n(E). \end{aligned}$$

The event $CSB^\infty(E)$ is "correct and common strong belief in E ."

First-order and higher-order beliefs are epistemic events, so they have no bite in terms of behavior and predictions about opponents' behavior without rationality and beliefs in rationality. The "rationality of player i " event is denoted by

$$R_i := \{(s_i, t) \in \Omega_i : s_i \in \rho(f_i(t))\},$$

and it is closed whenever $\rho \circ f_i$, as in finite games, is upper-hemicontinuous. The *rationality* event is $R := \times_{i \in I} R_i$.

Here I consider rational players who keep, as the game unfolds, the highest order of belief in rationality of each opponent that is consistent with her observed behavior. Players further refine their first-order beliefs through the own theories. All this is captured by the event $[\Delta] \cap CSB^\infty(R)$. The event "rationality and common strong belief in rationality", $CSB^\infty(R)$, characterizes Rationalizability.¹⁷ Furthermore, players believe, as long as not contradicted by observation, that each opponent: (1) reasons in the same way; (2) believes, as long as not contradicted by observation, that everyone else reasons in the same way; and so on. The n -th order of this belief is captured by the event $CSB^n([\Delta] \cap CSB^\infty(R))$, and it characterizes the $n+1$ -th step of Selective Rationalizability. The event $CSB^\infty([\Delta] \cap CSB^\infty(R))$ captures all the steps of reasoning at once.

¹⁷ Battigalli and Siniscalchi [9] characterize Strong Rationalizability with rationality and common strong belief in rationality, where strong belief is meant without independent rationalization.

Theorem 1 Fix a profile $\Delta = (\Delta_i)_{i \in I}$ of compact subsets of CPS's. Then, for every $n \geq 0$,

$$S_{R\Delta}^{n+1} = \text{Proj}_S \text{CSB}^n([\Delta] \cap \text{CSB}^\infty(R)),$$

and

$$S_{R\Delta}^\infty = \text{Proj}_S \text{CSB}^\infty([\Delta] \cap \text{CSB}^\infty(R)).$$

Therefore, Selective Rationalizability delivers the behavioral implications of rationality, common strong belief in rationality, first-order belief restrictions, and common strong belief in their conjunction. That is, step by step:

1. each player is rational and her beliefs are compatible with common strong belief in rationality and with the first-order belief restrictions;
2. 1 holds and each player believes that 1 holds for each opponent as long as not contradicted by observation;
3. 1 and 2 hold and each player believes that 1 and 2 hold for each opponent as long as not contradicted by observation;
4. ...

A deeper understanding of Selective Rationalizability and epistemic priority requires a closer look at the event that characterizes Selective Rationalizability and a comparison with the characterization of Strong- Δ -Rationalizability proposed by Battigalli and Prestipino [6]. Since independent rationalization is immaterial for this analysis, to simplify exposition I assume that there are only two players.

A simple preliminary observation: In the event $\text{CSB}^\infty(R \cap B^*([\Delta])) \subset B^*([\Delta])$ that characterizes Strong- Δ -Rationalizability, players keep at every information set every order of belief in the restrictions; in the event $\text{CSB}^\infty([\Delta] \cap \text{CSB}^\infty(R)) \subset \text{CSB}^\infty(R)$ that characterizes Selective Rationalizability players keep at every information set the highest order of strong belief in rationality which is per se compatible with the observed behavior.¹⁸ But what about the beliefs in the restrictions under $\text{CSB}^\infty([\Delta] \cap \text{CSB}^\infty(R))$ and the beliefs in rationality under $\text{CSB}^\infty(R \cap B^*([\Delta]))$? Let $E = \times_{i \in I} E_i := [\Delta] \cap \text{CSB}^\infty(R)$.

Fix $i \in I$ and $h \in H_i(\text{Proj}_S \text{CSB}^\infty(E))$. Consider the belief of player i at h in the restrictions of $-i$. That is, consider the event $B_{i,h}([\Delta_{-i}])$. Does it hold under

¹⁸Note that in both events, the own restrictions are never dropped: $\text{CSB}^\infty(R \cap B^*([\Delta])) \subset [\Delta] \supset \text{CSB}([\Delta] \cap \text{CSB}^\infty(R))$. If they are at odds with the behavioral implications of opponents' strategic reasoning, the events are empty. That is, the theories of opponents' behavior represented by the restrictions are rejected by strategic reasoning.

$CSB_i^\infty(E)$? That is, does it hold that $CSB_i^\infty(E) \subset B_{i,h}([\Delta_{-i}])$? Fix $(s_i, t_i) \in SB_i(E_{-i}) \supset CSB_i^\infty(E)$. Since $h \in H_i(\text{Proj}_{S_{-i}}E)$, then $1 = g_{i,h}(t_i)(E_{-i}) \leq g_{i,h}(t_i)([\Delta_{-i}])$. Thus, $(s_i, t_i) \in B_{i,h}([\Delta_{-i}])$. The first-order belief in the opponent's restrictions holds at all selectively rationalizable information sets.

Now, let $\tilde{B}_{-i}([\Delta_i]) := \cap_{h' \in H_{-i}(\text{Proj}_S CSB^\infty(E))} B_{-i,h'}([\Delta_i])$. Consider the belief of player i at h in the belief of $-i$ at all $h' \in H_{-i}(\text{Proj}_S CSB^\infty(E))$ in i 's restrictions. That is, consider the event $B_{i,h}(\tilde{B}_{-i}([\Delta_i]))$. Does it hold that $CSB_i^\infty(E) \subset B_{i,h}(\tilde{B}_{-i}([\Delta_i]))$? Fix $(s_i, t_i) \in SB_i(CSB_{-i}(E)) \supset CSB_i^\infty(E)$. As shown for i , $SB_{-i}(E_i) \subseteq \tilde{B}_{-i}([\Delta_i])$. Since $h \in H_i(\text{Proj}_{S_{-i}} CSB_{-i}(E))$, then $1 = g_{i,h}(t_i)(CSB_{-i}(E)) \leq g_{i,h}(t_i)(\tilde{B}_{-i}([\Delta_i]))$. Thus, $(s_i, t_i) \in B_{i,h}(\tilde{B}_{-i}([\Delta_i]))$.

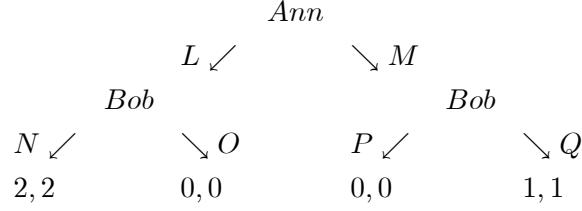
Proceeding by induction, one shows that at every selectively rationalizable information sets, there is common belief that the restrictions hold at all selectively rationalizable information sets themselves. A similar argument shows that at every information set compatible with $CSB^\infty(R \cap B^*([\Delta]))$ (i.e. with Strong- Δ -Rationalizability), there is common belief in rationality at all the strongly- Δ -rationalizable information sets themselves.

Put down in this way, it seems that the epistemic priority issue does not actually arise at the relevant information sets, so that Selective Rationalizability and Strong- Δ -Rationalizability, when both non-empty, should predict the same outcomes. This is false. Common belief in rationality and in the restrictions holds along different sets of paths, depending on which off-path beliefs sustain them: common strong belief in rationality or transparency of the restrictions. This phenomenon is illustrated concretely in the Discussion Section.

Since an order of belief in the restrictions is immaterial without the corresponding order of belief in rationality, Strong- Δ -Rationalizability can also be characterized without requiring transparency of the first-order belief restrictions at all information sets. That is, Strong- Δ -Rationalizability is also characterized by the event $CSB^\infty(R \cap [\Delta])$ (see Battigalli and Prestipino [6]),¹⁹ which puts rationality and the restrictions on the same epistemic priority level. To complete the picture, one may wonder whether the event $CSB^\infty(R) \cap B^*([\Delta])$, which does not assign epistemic priority to rationality or the restrictions either, also characterizes one of the two procedures. The answer is negative.

¹⁹In games with more than 2 players, since Strong- Δ -Rationalizability is defined without independent rationalization, CSB_i has to be redefined with \overline{SB}_i in place of SB_i .

Consider the following game.



Under the restriction that Ann believes that Bob plays N , both Strong- Δ -Rationalizability and Selective Rationalizability yield L for Ann and $N.Q$ for Bob. Yet, $CSB^\infty(R) \cap B^*([\Delta])$ is empty, because it requires Bob to believe at (M) both that Ann is rational and that her restriction holds, and the two things are clearly at odds. This is because restrictions and rationality are not under the same strong belief operator, so Bob is not allowed to drop the belief in their conjunction.²⁰

To conclude this section, it is worth to stress which assumptions on the game and on the type structure are crucial for the characterization result. Completeness, compactness and continuity of the type structure play a crucial role in the proof of Theorem 1. Finiteness of the game, instead, is only instrumental for the existence of such type structure and the upper-hemicontinuity of the best response correspondences, which guarantees closedness of the rationality event. A complete, compact and continuous type structure exists not only for finite games, but also for the class of "simple dynamic games" introduced by Battigalli [4], i.e. all games where the sets of available actions are finite at all histories (such as infinitely repeated games with a finite stage game), except possibly for preterminal histories where they can be any compact metric space. Indeed, the canonical type structure for CPS's constructed by Battigalli and Siniscalchi [7] exists in all such games. Under continuity of the payoff functions, Battigalli and Tebaldi [11] extend the epistemic characterization of Strong- Δ -Rationalizability to simple dynamic games. The same could be done here for Selective Rationalizability. The proof of Theorem 1 (and 2) can be easily adapted to simple dynamic games by using Lemma 3 of Battigalli and Tebaldi [11] to claim the existence of CPS's over strategies and types with the desired marginal CPS over strategies. On the other hand, finiteness allows to provide a self-contained proof of the main results, so it is maintained.

²⁰Friedenberg [19] obtains predictions for a two-players bargaining game by intersecting common strong belief in rationality with the event "on path strategic certainty" and common strong belief in it. On path strategic certainty selects the states where players have correct beliefs about the path of play induced by their strategies. Thus, it does not fix beliefs on a particular path and it is not an epistemic event. So, at (M) , on path strategic certainty does not force Bob to keep the belief that Ann believed in N and no contradiction with the belief in the rationality of Ann arises.

5 Finer epistemic priority orderings

Consider the following game, where after I Cleo chooses the matrix.

$Cleo$ $\downarrow I$	$\longrightarrow O \longrightarrow$	$A \backslash B$	W	E
		N	2 2 3.6	0 0 0
		S	0 0 0	2 2 4

$M1$	L	R	$M2$	L	R
U	1 1 3.3	0 0 3.3	U	0 0 0	1 1 8.1
D	0 0 3.3	1 1 3.9	D	1 1 8.1	0 0 0

All strategies are rationalizable. Suppose that players agree on the subgame perfect equilibrium $(S.U, E.L, O.M1)$. Consider the corresponding first-order-belief restrictions for all players. Then, Selective Rationalizability yields the desired outcome $(O, (S, E))$. Upon observing I , Ann and Bob drop the belief that Cleo believes in (S, E) and (U, L) . In particular, they can believe that Cleo did not believe in (S, E) and could rationally play $M1$ after I . In this case, they have the incentive to play U against L and vice versa.

Suppose now instead that Ann and Bob have an alternative theory to rationalize Cleo's move I . They believe that Cleo believed that they would have complied with the agreement on path (i.e. that they would have played (S, E) after O), but does not believe that they will implement the threat off-path (i.e. that they will play (U, L) after I). If Ann and Bob rationalize I under this light, they expect Cleo to pick $M2$, because $(I, M1)$ is not rational given the belief in (S, E) . Under $M2$, Ann and Bob cannot coordinate on (U, L) .

Suppose now that players agree on the subgame perfect equilibrium $(N.U, W.L, O.M1)$, and that upon observing I , they believe that Cleo believed in (N, W) , but does not believe in (U, L) . This time, this does not exclude that Cleo would play $M1$, hoping for (D, R) . Thus, Ann may play U when she believes that Bob will play L , and vice versa. So, the restrictions are compatible with this kind of strategic reasoning and yield the desired outcome $(O, (N, W))$ as unique prediction.²¹

Two important questions arise now. First: Does the exclusion of $(O, (S, E))$ and not of $(O, (N, W))$ correspond to some existing equilibrium refinement? Note that both outcomes are induced by a subgame perfect equilibrium in (extensive-form/strongly) rationalizable strategies. Second, and most importantly: Can this kind of strategic reasoning be modeled

²¹Note a seemingly paradoxical but quite customary consequence of forward induction reasoning: to convince Cleo to play O , Ann and Bob must promise to play (N, W) , which gives Cleo a payoff of 3.6, instead of (S, E) , which yields Cleo a payoff of 4.

as an epistemic priority order between different theories of opponents' behavior, and be captured by a solution concept analogous to Selective Rationalizability?

The answer to the first question is yes: strategic stability à la Kohlberg and Mertens [23].²²

Definition 7 (Kohlberg and Mertens [23]) *For each $i \in I$, let Σ_i be the set of mixed strategies of i , i.e. the set of probability distributions over S_i . A closed set of mixed equilibria $\widehat{\Sigma} \subseteq \Sigma$ is stable if it is minimal with respect to the following property: for any $\varepsilon > 0$, there exists $\delta_0 > 0$ such that for any completely mixed $(\sigma_i)_{i \in I} \in \Sigma$ and $(\delta_i)_{i \in I}$ with $0 < \delta_i < \delta_0$ for all $i \in I$, the perturbed game where for every $i \in I$, every $s_i \in S_i$ is substituted by $(1 - \delta_i)s_i + \delta_i\sigma_i$ has a mixed equilibrium ε -close to $\widehat{\Sigma}$.*

Consider first a set of two mixed equilibria $\widehat{\Sigma} = \{(\sigma_i)_{i \in I}, (\sigma'_i)_{i \in I}\}$ inducing outcome $(O, (N, W))$:

$$\begin{aligned} \sigma_C(O) &= 1, \sigma_A(N.D) = \sigma_B(W.R) = \frac{1}{\sqrt{2}}, \sigma_A(N.U) = \sigma_B(W.L) = 1 - \frac{1}{\sqrt{2}}; \\ \sigma'_C(O) &= 1, \sigma'_A(N.D) = \sigma'_B(W.R) = \frac{2}{3}, \sigma'_A(N.U) = \sigma'_B(W.L) = \frac{1}{3}. \end{aligned}$$

Under σ , Cleo is actually indifferent between O and $I.M1$, while under σ' , she is indifferent between O and $I.M2$. I show that $\widehat{\Sigma}$ is stable. Fix any completely mixed $(\tilde{\sigma}_i)_{i \in I} \in \Sigma$, an arbitrarily small δ_0 , and $(\delta_i)_{i \in I}$ with $0 < \delta_i < \delta_0$ for all $i \in I$. Consider the game perturbed as in Definition 7 and indicate with tilde the perturbed strategies. If $\tilde{\sigma}_A(I.M2) > \tilde{\sigma}_A(I.M1)$ (resp., $\tilde{\sigma}_A(I.M2) < \tilde{\sigma}_A(I.M1)$), assign small probability to $\widetilde{I.M1}$ (resp., $\widetilde{I.M2}$) and the complementary probability to \widetilde{O} in such a way that, overall, $I.M1$ and $I.M2$ are played with equal probability. Then, after I , Ann and Bob are indifferent between their actions regardless of the belief about the action of the other. Thus, since all strategies are perturbed in the same way, Ann and Bob are indifferent between $\widetilde{N.U}$ and $\widetilde{N.D}$, and between $\widetilde{W.L}$ and $\widetilde{W.R}$. Assign probability to these strategies in such a way that Cleo is indifferent between \widetilde{O} and $\widetilde{I.M1}$ (resp., $\widetilde{I.M2}$).²³ For any $\varepsilon > 0$, by picking a small enough δ_0 , we have an equilibrium in the perturbed game where the induced probabilities

²²Strategic stability has been chosen over Forward Induction equilibria of Govindan and Wilson [21] or Man [26] because the latter do not refine extensive-form rationalizability, hence do not capture all orders of strong belief in rationality. Strategic stability, instead, refines iterated admissibility, which in generic games corresponds to extensive-form rationalizability (Shimoji, [32]).

²³Since the perturbed strategies assign positive probability to S and E , the expected payoff of Cleo after O is lower than 3.6. The payoff of Cleo after $I.M1$ (resp., $I.M2$) can be lowered by the same amount by assigning probabilities to $\widetilde{N.U}$ and $\widetilde{N.D}$ and $\widetilde{W.L}$ and $\widetilde{W.R}$ in such a way that $N.D$ and $W.R$ have probability lower than $1/\sqrt{2}$ (resp., higher than $2/3$). This can be done with probabilities of $\widetilde{N.D}$ and $\widetilde{W.R}$ close to $1/\sqrt{2}$ (resp., $2/3$).

over the original strategies are ε -close to those assigned by σ (resp., σ').²⁴

Instead, there is no stable set of equilibria inducing $(O, (S, E))$: any perturbation of O that gives negligible probability to $I.M1$ with respect to $I.M2$ cannot be compensated by giving positive probability to $\widetilde{I.M1}$, because $\widetilde{I.M1}$ cannot be optimal under belief in (S, E) (albeit perturbed). Thus, Ann and Bob must play close to an equilibrium of matrix $M2$, which cannot discourage a deviation to $\widetilde{I.M2}$.

This is not the first time that a connection between rationalizability and equilibrium refinements à la strategic stability is established. In signaling games, Battigalli and Siniscalchi [10] show that when an equilibrium outcome satisfies the Iterated Intuitive Criterion (Cho and Kreps [17]), Strong- Δ -Rationalizability yields a non-empty set for the corresponding restrictions (i.e. the belief that opponents play compatibly with the path). In [14] I prove that Selective Rationalizability yields the empty set for a class of non strategically stable equilibrium paths: those that *can be upset by a convincing deviation* (Osborne [27]). So, one could think that strategic stability simply requires non-emptiness of Selective Rationalizability/Strong- Δ -Rationalizability²⁵ under the belief in the equilibrium path. This is false. In the example above, both procedures yield a non-empty set under the belief in $(O, (S, E))$.²⁶ Thus, there is no incompatibility between the belief in the path and the rationalization of deviations based on it (unlike for equilibrium paths that can be upset by a convincing deviation). The problem is the incompatibility between the rationalization of deviations based on the belief in the path and the threats that sustain the path in equilibrium. This calls for a rationalizability procedure that takes both into account. The remainder of this section is dedicated to construct and characterize epistemically such rationalizability procedure. The scope is expanded to an arbitrary number of theories of opponents' behavior, of an arbitrary nature (i.e. not just path versus off-path behavior). Without the ambition to perfectly characterize strategic stability, the application of this rationalizability procedure to equilibrium path and profile captures in a general and transparent way the spirit of the strategic reasoning stories in the background of strategic stability and related refinements.

When players have competing theories of opponents' behavior, the first issue to solve is the *epistemic priority ordering* between them. Suppose, for instance, that each player has two overlapping theories. Thus, some predictions may be consistent with both theories, and both theories can be used at the same time to refine beliefs. At the beginning of the

²⁴The set is minimal because, depending on $\tilde{\sigma}_A(I.M2) \leq \tilde{\sigma}_A(I.M1)$, only one of the two equilibria can be approximated.

²⁵The two conditions are equivalent: see Proposition 3.

²⁶ $S_{R\Delta}^1 = S_{\Delta}^1 = \{S.U, S.D\} \times \{E.L, E.R\} \times \{O, I.M2\} = S_{R\Delta}^\infty = S_{\Delta}^\infty$.

game, each player believes that opponents refine beliefs according to their theories. Yet, as the game unfolds, some player may display behavior which cannot be optimal under both her theories at the same time. Then, for the opponents, the epistemic priority issue arises. Which theory is the player following? Suppose that everyone solves this dilemma in favour of the same theory. Then, this theory receives epistemic priority, and strategic reasoning about the other theory alone kicks in only at information sets that are not compatible with strategic reasoning about the first theory. Thus, when the theories are compatible with strategic reasoning, strategic reasoning about the second theory alone is immaterial for the induced outcomes. Hence, the problem is simplified by taking as second theory the overlap with the first.

Definition 8 *A chain of restrictions is a finite sequence $((\Delta_i^1)_{i \in I}, \dots, (\Delta_i^k)_{i \in I})$ such that for each $i \in I$, $\Delta_i^k \subset \dots \subset \Delta_i^1 \subset \Delta^{H_i}(S_{-i})$, and for each $l \leq k$, Δ_i^l is compact.*

Note that, in the equilibrium path – equilibrium profile motivating case, under this formalization the equilibrium path is the primary theory of opponents' behavior. The equilibrium profile yields a more restrictive, secondary theory. In the intuitive narration, players "resort" to the path theory when a deviator displays disbelief in the whole equilibrium profile. But since believing in the equilibrium profile actually implies believing in its path, the belief that the deviator believes in the path holds all along, and receives higher epistemic priority. Attributing the highest epistemic priority to the beliefs in rationality, the theories are then considered in their epistemic priority order according to the following extension of Selective Rationalizability.

Definition 9 *Fix a chain of restrictions $((\Delta_i^1)_{i \in I}, \dots, (\Delta_i^k)_{i \in I})$. Let $((S_{i,0}^q)_{i \in I})_{q=0}^\infty$ denote Rationalizability. Fix $1 \leq l \leq k$ and for each $p = 0, \dots, l-1$, suppose that $((S_{i,p}^q)_{i \in I})_{q=0}^\infty$ has already been defined. Consider now the following procedure.*

(Step 0) *For each $i \in I$, let $S_{i,l}^0 = S_{i,l-1}^\infty$.*

(Step n) *For each $i \in I$ and $s_i \in S_i$, let $s_i \in S_{i,l}^n$ if and only if there exists $\mu_i \in \Delta_i$ such that:*

E1(l) $s_i \in \rho(\mu_i)$;

E2(l) μ_i strongly believes $S_{j,l}^q$ for all $j \neq i$ and $q < n$;

E3(l) μ_i strongly believes $S_{j,p}^q$ for all $p < l$, $j \neq i$, and $q \in \mathbb{N}$.

For every $i \in I$, let $S_{i,l}^\infty := \bigcap_{q \in \mathbb{N}} S_{i,l}^q$.

Similarly to Selective Rationalizability, the procedure starts with Rationalizability. Then, the first-order belief restrictions are gradually introduced, following the descending epistemic priority order, when strategic reasoning about the weaker theories does not refine the strategy sets anymore. E3(l) guarantees that strategic reasoning according to the weaker theories is maintained. This is reflected in the following epistemic characterization.

Theorem 2 *Fix a chain of restrictions $((\Delta_i^1)_{i \in I}, \dots, (\Delta_i^k)_{i \in I})$. For each $n \in \mathbb{N} \cup \{\infty\}$, it holds*

$$S_k^n = \text{Proj}_S \text{CSB}^{n-1}([\Delta^k] \cap \text{CSB}^\infty([\Delta^{k-1}] \cap \dots \text{CSB}^\infty([\Delta^1] \cap \text{CSB}^\infty(R)))).$$

Let $\text{CSB}_0 := \text{CSB}^\infty(R)$ and, for each $l = 1, \dots, k$, let $\text{CSB}_l := \text{CSB}^\infty([\Delta^l] \cap \text{CSB}_{l-1})$. Note that $\text{CSB}_k \subseteq \text{CSB}^\infty(R)$: the highest epistemic priority is still assigned to rationality. As long as compatible with the beliefs in rationality and with the observed behavior, players believe in the first-order belief restrictions Δ^1 at every order. As long as compatible with this and with the observed behavior, players believe in the restrictions Δ^2 at every order. And so on. The own restrictions can never be dropped: if for some $i \in I$, $l \leq k$ and $n \geq 0$ the restrictions Δ_i^l are not compatible with the behavioral implications of $\text{CSB}_{-i}^{n-1}([\Delta^l] \cap \text{CSB}_{l-1})$, the event $\text{CSB}_i^n([\Delta^l] \cap \text{CSB}_{l-1})$ is empty.

Back to the equilibrium path – equilibrium profile case, if instead of considering the belief in the equilibrium path one considers the belief in the path *but not* in the equilibrium threats, an *alternative* theory with respect to the belief in the whole equilibrium profile is obtained. Then, the belief in the equilibrium profile by a player is here considered by the opponents *infinitely more likely* (Blume et al. [12], Lo [25]) than the belief in the path but not in the threats,²⁷ because the former is believed to hold at the beginning of the game, and the latter only after a deviation that contradicts the former. Counterintuitively, the infinitely more likely order seems inverted with respect to the epistemic priority one. Note, though, that the belief in the path *but not* in the threats is not the original primary theory Δ_i^1 that represents belief in the path, but the difference $\Delta_i^1 \setminus \Delta_i^2$, where Δ_i^2 is the secondary theory that represents belief in the whole equilibrium profile. However, there seems to be a tight connection between the notion of epistemic priority and the notion of infinitely more likely. Exploring this connection is an avenue for future research.

²⁷The notion of Infinitely More Likely applies to Lexicographic Probability Systems (Blume et al. [12]), but a CPS can be transformed into a Lexicographic Probability System. Siniscalchi [33] uses this connection, but first introduces a notion of "at least as plausible as" between theories of opponents' behavior that applies directly to CPS's.

6 Discussion

To discuss some issues related to epistemic priority, consider the following game.

$$\begin{array}{c}
 \text{Ann} \\
 A \swarrow \quad \searrow B \\
 2, 2 \qquad \qquad \text{Bob} \\
 \qquad \qquad I \swarrow \quad \searrow O \\
 \qquad \qquad \qquad 3, 3
 \end{array}$$

<i>Ann</i> \Bob	<i>L</i>	<i>C</i>	<i>R</i>
<i>U</i>	0, 1	0, 0	1, 0
<i>M</i>	1, 0	0, 4	0, 0
<i>D</i>	0, 0	1, 0	1, 4

One could expect that if the restrictions are not compatible with strategic reasoning when the epistemic priority is on them, a fortiori they will not when the epistemic priority is on rationality. This is false. Selective rationalizability can yield a non-empty set when Strong- Δ -Rationalizability does not. Suppose that Bob promises to play *O* and Ann threatens not to play *D* otherwise. Fix the corresponding restrictions.

At step 1 of Strong- Δ -Rationalizability, Ann eliminates *A* and Bob eliminates *I.L* and *I.R*. At step 2, Ann eliminates *B.U* and *B.M*. At step 3, Bob obtains the empty set: he cannot believe that Ann will not play *D* after *I*.

At step 1 of Rationalizability, Bob eliminates *I.L*. At step 2, Ann eliminates *B.M*. At step 3, Bob eliminates *I.C*. All other strategies are rationalizable.²⁸ At step 1 of Selective Rationalizability, Ann further eliminates *A* and Bob further eliminates *I.R*. The remaining strategies, *B.U* and *B.D* for Ann and *O* for Bob, are selectively rationalizable.

The algorithmic reason why Selective Rationalizability yields a non-empty set while Strong- Δ -Rationalizability does not is that S3 and the restrictions prevent Bob from reaching (B, I) already at the first step of Selective Rationalizability, while Bob still reaches (B, I) after the first step of Strong- Δ -Rationalizability. In this way, at the second step and at (B, I) , D2 becomes stricter than S2 (which is vacuous), and forces Ann to believe that Bob will play *C* after *I*. The epistemic reason is that, at (B, I) , Ann is forced to believe that Bob is rational and has the restriction under Strong- Δ -Rationalizability (so that he would play *C*), but not under Selective Rationalizability.²⁹ Keeping the highest possible

²⁸The tie between *D* and *U* against *R* is only to keep the game small; it can be eliminated by introducing another action of Bob in the subgame.

²⁹In abstract terms, this is an effect of the non-monotonicity of strong belief: strong belief in the event "Bob is rational, has the restriction, and strongly believes that Ann is rational and strongly believes that he is rational" is less restrictive for Ann's beliefs at (B, I) than strong belief in the larger event "Bob is rational and has the restriction".

order of belief in rationality may require to drop an order of belief in the restrictions, even if it is compatible with the same order of belief in rationality at the information set. At (B, I) , if Ann believes in rationality up to the third order, she cannot believe that Bob has the restriction, although this is compatible with just believing in Bob's rationality. Under Selective Rationalizability, S3 imposes from step 1 all orders of belief in rationality that are per se compatible with the observed behavior. In this way, through S2, any order of belief in the restrictions is maintained only as long as compatible with them (see Section 4). Under Strong- Δ -Rationalizability, the n -th order belief in rationality and in the restrictions is always maintained at an information set that can be reached if this belief is correct. But then, as it happens to Ann at (B, I) , beliefs may fall in a subset of n -th order belief in rationality where some higher order of belief in rationality never holds.

However, as shown in Section 4, when Selective Rationalizability or Strong- Δ -Rationalizability yields a non-empty set, it captures all orders of belief in rationality and in the restrictions along the induced paths. Then, one could expect Selective Rationalizability and Strong- Δ -Rationalizability to yield the same paths when they both yield a non-empty set. This is, again, false. The game after (B, I) and the restrictions are modified in the Appendix, and the two procedures are formally shown to yield non-empty, disjoint predictions. At (B) , when Bob believes that Ann would interpret his move I by giving epistemic priority to the beliefs in the restrictions, he prefers to play I instead of O . But then, anticipating this, Ann plays A . When instead Bob believes that Ann would interpret I by giving epistemic priority to rationality, he prefers O over I , and then Ann plays B . So, the selectively rationalizable information sets are the root and (B) , and all orders of belief in rationality and in the restrictions at the root and (B) hold. Although common belief in rationality and in the restriction can hold along the (B, O) path, the path is not sustained by off-the-path beliefs under epistemic priority to the restrictions.

In games with observable actions, there are very interesting restrictions under which Selective Rationalizability and Strong- Δ -Rationalizability predict the same outcomes (or both deliver the empty set). Such restrictions correspond to the belief in a path of play. That is, players strongly believe that each opponent plans to remain on-path.

Proposition 3 (Catonini [15]) *Fix a game with observable actions and $z \in Z$. For each $i \in I$, let Δ_i be the set of all μ_i that strongly believe $(S_j(z))_{j \neq i}$. Then $\zeta(S_{R\Delta}^\infty) = \zeta(S_\Delta^\infty)$.*

The proof of this result is rather sophisticated, and it is provided in [15].³⁰

³⁰The result is formally proved without independent rationalization, but it is possible to prove that independent rationalization is immaterial under path restrictions.

7 Appendix

7.1 A game

In this game, Selective Rationalizability and Strong- Δ -Rationalizability yield non-empty, yet disjoint predictions for the same first-order belief restrictions.

					<i>Ann</i>	
					$A \swarrow$	$\searrow B$
					(4, 6)	<i>Bob</i>
					$I \swarrow$	$\searrow O$
<i>Ann</i> \ <i>Bob</i>	<i>W</i>	<i>N</i>	<i>S</i>	<i>E</i>	(5, 5)	
<i>T</i>	(3, 6)	(0, 7)	(0, 0)	(0, 0)		
<i>U</i>	(0, 7)	(3, 6)	(0, 0)	(0, 0)		
<i>M</i>	(2, 1)	(2, 1)	(0, 0)	(0, 0)		
<i>D</i>	(0, 0)	(0, 6)	(1, 0)	(0, 0)		
<i>B</i>	(0, 0)	(0, 0)	(0, 6)	(1, 0)		

First-order-beliefs restrictions:

$$\Delta_A := \Delta^{H_A}(S_B); \Delta_B := \{\mu_B \in \Delta^{H_B}(S_A) : \forall h \in H_B, \mu_B(B.T|h) = \mu_B(B.U|h) = 0\}.$$

Rationalizability:

$$\begin{aligned} S_A^1 &= S_A, S_B^1 = \{I.W, I.N, I.S, O\}; \\ S_A^2 &= \{A, B.T, B.U, B.M, B.D\}, S_B^2 = S_B^1; \\ S_A^3 &= S_A^2, S_B^3 = \{I.W, I.N, O\}; \\ S_A^4 &= \{A, B.T, B.U, B.M\}, S_B^4 = S_B^3; \\ S_A^5 &= S_A^4 = S_A^\infty, S_B^5 = S_B^4 = S_B^\infty. \end{aligned}$$

Selective Rationalizability:

$$\begin{aligned} S_{A,R\Delta}^1 &= \{A, B.T, B.U, B.M\}, S_{B,R\Delta}^1 = \{O\}; \\ S_{A,R\Delta}^2 &= \{B.T, B.U, B.M\}, S_{B,R\Delta}^2 = S_{B,R\Delta}^1; \\ S_{A,R\Delta}^3 &= S_{A,R\Delta}^2 = S_{A,R\Delta}^\infty, S_{B,R\Delta}^3 = S_{B,R\Delta}^2 = S_{B,R\Delta}^\infty. \end{aligned}$$

Strong- Δ -Rationalizability:

$$\begin{aligned} S_{A,\Delta}^1 &= S_A, S_{B,\Delta}^1 = \{I.N, I.S, O\}; \\ S_{A,\Delta}^2 &= \{A, B.U, B.D\}, S_{B,\Delta}^2 = S_{B,\Delta}^1; \\ S_{A,\Delta}^3 &= S_{A,\Delta}^2, S_{B,\Delta}^3 = \{I.N\}; \\ S_{A,\Delta}^4 &= \{A\}, S_{B,\Delta}^4 = S_{B,\Delta}^3; \\ S_{A,\Delta}^5 &= S_{A,\Delta}^4 = S_{A,\Delta}^\infty, S_{B,\Delta}^5 = S_{B,\Delta}^4 = S_{B,\Delta}^\infty. \end{aligned}$$

7.2 Proofs

Proof of Proposition 1. I prove by induction a stronger statement.

Induction hypothesis (n): For every $m \leq n$, $i \in I$, and $s_i \in S_{i,R\Delta}^m$ (resp., $\hat{s}_i \in \hat{S}_{i,R\Delta}^m$), there exists $\hat{s}_i \in \hat{S}_{i,R\Delta}^m$ (resp., $s_i \in S_{i,R\Delta}^m$) such that $s_i(h) = \hat{s}_i(h)$ for all $h \in H_i(S^\infty)$.

Basis step ($n = 0$): It follows from $\hat{S}_{R\Delta}^0 = S_{R\Delta}^0 = S^\infty$.

Inductive step ($n + 1$): Fix $i \in I$, $s_i \in S_{i,R\Delta}^{n+1}$ (resp., $\hat{s}_i \in \hat{S}_{i,R\Delta}^{n+1}$), and $\mu_i \in \Delta_i$ that strongly believes $((S_{j,R\Delta}^m)_{j \neq i})_{m=0}^n$ (resp., $((\hat{S}_{j,R\Delta}^m)_{j \neq i})_{m=0}^n$) such that $s_i \in \rho(\mu_i)$. By the induction hypothesis, for each $j \neq i$, I can construct a map $\eta_j : S_j^\infty \rightarrow S_j^\infty$ such that for each $m \leq n$ and $s_j \in S_{j,R\Delta}^m \subseteq S_j^\infty$ (resp., $s_j \in \hat{S}_{j,R\Delta}^m \subseteq S_j^\infty$), $\eta_j(s_j) \in \hat{S}_{j,R\Delta}^m$ (resp., $\eta_j(s_j) \in S_{j,R\Delta}^m$) and $\eta_j(s_j)(h) = s_j(h)$ for all $h \in H_j(S^\infty)$. For each $h \in H_i(S^\infty)$ and $s_{-i} \in S_{-i}^\infty$, note that $s_{-i} \in S_{-i}^\infty(h)$ if and only if $(\times_{j \neq i} \eta_j)(s_{-i}) \in S_{-i}^\infty(h)$.³¹ Moreover, by the induction hypothesis, for each $j \neq i$ and $m = 1, \dots, n$, $H_i(S_{j,R\Delta}^m) \cap H_i(S^\infty) = H_i(\hat{S}_{j,R\Delta}^m) \cap H_i(S^\infty)$. Then, there exists $\bar{\mu}_i$ that strongly believes $((\hat{S}_{j,R\Delta}^m)_{j \neq i})_{m=0}^n$ (resp., $((S_{j,R\Delta}^m)_{j \neq i})_{m=0}^n$) and $((S_j^q)_{j \neq i})_{q=0}^\infty$ such that $\bar{\mu}_i(s_{-i}|h) = \mu_i((\times_{j \neq i} \eta_j)^{-1}(s_{-i})|h)$ for all $h \in H_i(S^\infty)$ and $s_{-i} \in (\times_{j \neq i} \eta_j)(S_{-i}^\infty(h))$. So, since $\mu_i(S_{-i}^\infty|h) = 1$ (by strong belief in $S_{-i,R\Delta}^0 = \hat{S}_{-i,R\Delta}^0 = S_{-i}^\infty$), $\bar{\mu}_i(S_{-i}(z)|h) = \mu_i(S_{-i}(z)|h)$ for all $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$. Hence: by maximality of Δ_i , $\bar{\mu}_i \in \Delta_i$; by $\zeta(\rho(\bar{\mu}_i) \times S_{-i}^\infty) \subseteq \zeta(S^\infty)$ for any $\bar{\mu}_i$ that strongly believes S_{-i}^∞ ,³² there is $\bar{s}_i \in \rho(\bar{\mu}_i) \subseteq \hat{S}_{i,R\Delta}^{n+1}$ (resp., $\bar{s}_i \in \rho(\bar{\mu}_i) \subseteq S_{i,R\Delta}^{n+1}$) such that $\bar{s}_i(h) = s_i(h)$ for all $h \in H_i(S^\infty)$. ■

Proof of Proposition 2. For each $i \in I$, let $\bar{\Delta}_i$ be the (compact³³) set of all $\mu_i \in \Delta_i$ that satisfy S3 and S2 under $(\Delta_j)_{j \in I}$ for all $n \in \mathbb{N}$. By finiteness,³⁴ $S_{R\Delta}^\infty = \times_{i \in I} \rho(\bar{\Delta}_i) = S_{R\Delta}^1$, and then each $\mu_i \in \bar{\Delta}_i$ strongly believes $(S_{j,R\Delta}^1)_{j \neq i}$, thus $S_{R\Delta}^1 = S_{R\Delta}^\infty$. Let $\mu_i \in \Delta_i^*$ if and only if there exists $\bar{\mu}_i \in \bar{\Delta}_i$ such that $\mu_i(S_{-i}(z)|h) = \bar{\mu}_i(S_{-i}(z)|h)$ for all $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$. It is easy to observe that for each $i \in I$, Δ_i^* is compact³⁵ and maximal, and $\Delta_i^* \supseteq \bar{\Delta}_i$. Now I show by induction that $\zeta(S_{R\Delta}^1) = \zeta(S_{R\Delta}^\infty)$.

³¹For each $s_{-i} \in S_{-i}^\infty(h)$, by perfect recall $S_i^\infty(h) \times \{s_{-i}\} \subseteq S^\infty(h)$, hence there is a history $x \in h$ such that $x \prec z$ for some $z \in \zeta(S_i^\infty(h) \times \{s_{-i}\}) \subseteq \zeta(S^\infty)$, and then for all $x' \prec x$ and $j \neq i$ with $x' \in h'$ for some $h' \in H_j$, it holds $h' \in H_j(S^\infty)$.

³²If this was not the case, then there would be $\tilde{\mu}_i$ that strongly believes also $((S_j^q)_{j \neq i})_{q=0}^\infty$, and $\tilde{s}_i \in \rho(\tilde{\mu}_i) \subseteq S_{-i}^\infty$ such that $\zeta(\{\tilde{s}_i\} \times S_{-i}^\infty) \not\subseteq \zeta(S^\infty)$, a contradiction.

³³Compactness can be indirectly argued from the epistemic characterization and Lemma 1: $\bar{\Delta}_i = f_i(\text{Proj}_{T_i} \text{CSB}_i^\infty([\Delta] \cap \text{CSB}^\infty(R)))$.

³⁴Or milder conditions which guarantee that every $s_i \in S_{i,R\Delta}^\infty$ is a sequential best reply to some μ_i that strongly believes $((S_{j,R\Delta}^q)_{j \in I})_{q=0}^\infty$ and $((S_j^q)_{j \in I})_{q=0}^\infty$.

³⁵For each sequence of CPS's in Δ_i^* and any corresponding sequence of CPS's in $\bar{\Delta}_i$, the equalities for each $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$ are preserved in the limit. By compactness of $\bar{\Delta}_i$, the sequence in $\bar{\Delta}_i$ converges to a CPS in $\bar{\Delta}_i$, so the limit of the sequence in Δ_i^* is a CPS (by compactness of $\Delta^{H_i}(S_{-i})$) that satisfies the conditions to be in Δ_i^* .

Induction hypothesis (n): For every $m \leq n$, $i \in I$, and $s_i \in S_{i,R\bar{\Delta}}^1$ (resp., $s_i^* \in S_{i,R\Delta^*}^m$), there exists $s_i^* \in S_{i,R\Delta^*}^m$ (resp., $s_i \in S_{i,R\bar{\Delta}}^1$) such that $s_i(h) = s_i^*(h)$ for all $h \in H_i(S^\infty)$.

Basis step ($n = 1$). Fix $i \in I$. By $\Delta_i^* \supseteq \bar{\Delta}_i$, $S_{i,R\Delta^*}^1 \supseteq S_{i,R\bar{\Delta}}^1$. Fix $s_i^* \in S_{i,R\Delta^*}^1$ and $\mu_i \in \Delta_i^*$ that strongly believes S_{-i}^∞ such that $s_i^* \in \rho(\mu_i)$. By definition of Δ_i^* , there exists $\bar{\mu}_i \in \bar{\Delta}_i$ such that $\mu_i(S_{-i}(z)|h) = \bar{\mu}_i(S_{-i}(z)|h)$ for all $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$. Since $\zeta(\rho(\bar{\mu}_i) \times S_{-i}^\infty) \subseteq \zeta(S^\infty)$ for any $\bar{\mu}_i$ that strongly believes S_{-i}^∞ , there exists $s_i \in \rho(\bar{\mu}_i) \subseteq S_{i,R\bar{\Delta}}^1$ such that $s_i(h) = s_i^*(h)$ for all $h \in H_i(S^\infty)$.

Inductive step ($n + 1$). Fix $i \in I$, $s_i \in S_{i,R\bar{\Delta}}^1$, and $\mu_i \in \bar{\Delta}_i$ such that $s_i \in \rho(\mu_i)$. By the induction hypothesis, for each $j \neq i$, I can construct a map $\eta_j : S_j^\infty \rightarrow S_j^\infty$ such that for each $s_j \in S_j^\infty \setminus S_{j,R\bar{\Delta}}^1$, $\eta_j(s_j) = s_j$, and for each $s_j \in S_{j,R\bar{\Delta}}^1 \subseteq S_j^\infty$, $\eta_j(s_j) \in S_{j,R\Delta^*}^n$ and $\eta_j(s_j)(h) = s_j(h)$ for all $h \in H_j(S^\infty)$. For each $h \in H_i(S^\infty)$ and $s_{-i} \in S_{-i}^\infty$, note that $s_{-i} \in S_{-i}^\infty(h)$ if and only if $(\times_{j \neq i} \eta_j)(s_{-i}) \in S_{-i}^\infty(h)$.³¹ Moreover, by the induction hypothesis, for each $j \neq i$ and $m = 1, \dots, n$,

$$H_i(S_{j,R\Delta^*}^m) \cap H_i(S^\infty) = H_i(S_{j,R\bar{\Delta}}^1) \cap H_i(S^\infty) = H_i(S_{j,R\Delta^*}^n) \cap H_i(S^\infty).$$

Then, recalling that μ_i strongly believes $(S_{j,R\bar{\Delta}}^1)_{j \neq i}$, there exists μ_i^* that strongly believes $((S_{j,R\Delta^*}^m)_{j \neq i})_{m=0}^n$ and $((S_j^q)_{j \neq i})_{q=0}^\infty$ such that $\mu_i^*(s_{-i}|h) = \mu_i((\times_{j \neq i} \eta_j)^{-1}(s_{-i})|h)$ for all $h \in H_i(S^\infty)$ and $s_{-i} \in (\times_{j \neq i} \eta_j)(S_{-i}^\infty(h))$. So, since $\mu_i(S_{-i}^\infty|h) = 1$ (by strong belief in S_{-i}^∞), $\mu_i^*(S_{-i}(z)|h) = \mu_i(S_{-i}(z)|h)$ for all $h \in H_i(S^\infty)$ and $z \in \zeta(S^\infty)$. Hence: by definition of Δ_i^* and $\mu_i \in \bar{\Delta}_i$, $\mu_i^* \in \Delta_i^*$; by $\zeta(\rho(\bar{\mu}_i) \times S_{-i}^\infty) \subseteq \zeta(S^\infty)$ for any $\bar{\mu}_i$ that strongly believes S_{-i}^∞ ,³² there is $s_i^* \in \rho(\mu_i^*) \subseteq S_{i,R\Delta^*}^{n+1}$ such that $s_i^*(h) = s_i(h)$ for all $h \in H_i(S^\infty)$. The other direction is identical to the basis step. ■

PROOFS OF THE THEOREMS.

First, I prove a generalized version of Theorem 1. Applying this result to Rationalizability yields the conditions to apply it to Selective Rationalizability and prove Theorem 1, and with further iterations, Theorem 2.

Consider this generalized rationalizability procedure (without a step 0).

Definition 10 Fix two profiles of subsets of CPS's, $(\Delta_i)_{i \in I}$ and $(\Delta_i^G)_{i \in I}$. Fix $n \geq 1$ and, if $n > 1$, suppose that $((S_{i,G}^q)_{i \in I})_{q=1}^{n-1}$ has already been defined. For every $i \in I$ and $s_i \in S_i$, let $s_i \in S_{i,G}^n$ if and only if there exists $\mu_i \in \Delta_i$ such that:

$$G1 \quad s_i \in \rho(\mu_i);$$

$G2$ μ_i strongly believes $S_{j,G}^q$ for all $j \neq i$ and $1 \leq q < n$,³⁶

$G3$ $\mu_i \in \Delta_i^G$.

Call $\Delta_i^{n,G}$ the set of all $\mu_i \in \Delta_i$ that satisfy $G2$ and $G3$.

Finally, let $S_{i,G}^\infty = \cap_{n \geq 1} S_{i,G}^n$ and $\Delta_i^{\infty,G} = \cap_{n \geq 1} \Delta_i^{n,G}$.

Consider now the following property for a Cartesian event $E = \times_{i \in I} E_i \subseteq \Omega$.

Definition 11 A Cartesian event $E = \times_{i \in I} E_i$ satisfies the "completeness property" if for every $i \in I$, $t_i \in \text{Proj}_{T_i} E_i$, $s_i \in \rho(f_i(t_i))$, and maps³⁷ $(\tau_j)_{j \neq i}$ with $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} E_j \mapsto (\bar{s}_j, t_j) \in E_j$ for all $j \neq i$, there exists $t'_i \in T_i$ such that $(s_i, t'_i) \in E_i$, $f_i(t'_i) = f_i(t_i)$, and $g_{i,h}(t'_i) [\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i) [\{s_j\} \times S_{-i,j}]$ for all $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} E_j$.

In the proof of the following, generalized characterization result, the completeness property (assumed here and shown to hold for rationality later) allows to retrieve the desired types from the induction hypothesis, instead of constructing them from scratch in the inductive step (differently than, for instance, in [6]).

Lemma 1 Fix a closed, Cartesian event $E = \times_{i \in I} E_i \subseteq R$ with the completeness property³⁸ such that for each $i \in I$, $f_i(\text{Proj}_{T_i} E_i) = \Delta_i \cap \Delta_i^G$ ⁽³⁹⁾ (which implies $S_{i,G}^1 = \text{Proj}_{S_i} E_i$).⁴⁰

Then, for every $n \in \mathbb{N}$, $CSB^{n-1}(E)$ has the completeness property and for each $i \in I$, $f_i(\text{Proj}_{T_i} CSB_i^{n-1}(E)) = \Delta_i^{n,G}$ (which implies $S_{i,G}^n = \text{Proj}_{S_i} CSB_i^{n-1}(E)$).

Moreover, $CSB^\infty(E)$ has the completeness property and for each $i \in I$, $f_i(\text{Proj}_{T_i} CSB_i^\infty(E)) = \Delta_i^{\infty,G}$ and $S_{i,G}^\infty = \text{Proj}_{S_i} CSB_i^\infty(E)$.

Proof. For finite n , the proof is by induction.

Induction Hypothesis (n=1,...,m): the Lemma holds for $n = 1, \dots, m$.

Basis step (n=1): the Lemma holds for $n = 1$ by assumption.

Inductive step (n=m+1): For each $i \in I$, let $F_i = CSB_i^{m-1}(E)$ and $G_i = CSB_i^m(E)$

³⁶ Vacuous for $n = 1$.

³⁷ Note that the maps are injective.

³⁸ The event E can be empty, just like $CSB^\infty(R) \cap [\Delta]$ in Theorem 1.

³⁹ Since f_i is continuous and $\text{Proj}_{T_i} E_i$ is compact (because E_i is closed and T_i is compact), compactness of $\Delta_i \cap \Delta_i^G$ is implied.

⁴⁰ \subseteq is guaranteed by the completeness property of E ; \supseteq is guaranteed by the fact that $E \subseteq R$.

Fix $i \in I$ and $\mu_i \in \Delta_i^{m+1,G}$. Since $\mu_i \in \Delta_i^{m,G}$, by the Induction Hypothesis there exists $t_i \in \text{Proj}_{T_i} F_i$ such that $f_i(t_i) = \mu_i$. Fix maps $(\tau_j)_{j \neq i}$ with $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} F_j \mapsto (\bar{s}_j, t_j) \in F_j$ for all $j \neq i$. By the Induction Hypothesis, F has the completeness property. So, there exists $(s'_i, t'_i) \in F_i$ such that $f_i(t'_i) = f_i(t_i)$, and for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} F_j$, $g_{i,h}(t'_i)[\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i)[\{s_j\} \times S_{-i,j}]$. Then, since $f_i(t_i)$ strongly believes $S_{j,G}^m = \text{Proj}_{S_j} F_j$ (by the Induction Hypothesis), $g_i(t'_i)$ strongly believes F_j . So, $(s'_i, t'_i) \in SB_i(F_{-i}) \cap F_i = G_i$.

Fix $i \in I$ and $t_i \in \text{Proj}_{T_i} G_i$. Since $t_i \in \text{Proj}_{T_i} F_i$, by the Induction Hypothesis $f_i(t_i) \in \Delta_i^{m,G}$. Since $t_i \in \text{Proj}_{T_i} SB_i(F_{-i})$, $g_i(t_i)$ strongly believes F_j for all $j \neq i$, hence $f_i(t_i)$ strongly believes $\text{Proj}_{S_j} F_j$. By the Induction Hypothesis $\text{Proj}_{S_j} F_j = S_j^m$. So $f_i(t_i) \in \Delta_i^{m+1,G}$.

Now I show that G has the completeness property. Fix $i \in I$, $t_i \in \text{Proj}_{T_i} G_i \subseteq \text{Proj}_{T_i} F_i$, $s_i \in \rho(f_i(t_i))$, and maps $(\tau_j)_{j \neq i}$ with $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} G_j \mapsto (\bar{s}_j, t_j) \in G_j \subseteq F_j$ for all $j \neq i$. Extend each τ_j to $\tau'_j : \bar{s}_j \in \text{Proj}_{S_j} F_j \mapsto (\bar{s}_j, t_j) \in F_j$ in such a way that for every $s_j \in \text{Proj}_{S_j} G_j$, $\tau'_j(s_j) = \tau_j(s_j)$. By the Induction Hypothesis, F has the completeness property. So, there exists $t'_i \in T_i$ such that $(s_i, t'_i) \in F_i$, $f_i(t'_i) = f_i(t_i)$, and for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} F_j$, $g_{i,h}(t'_i)[\{\tau'_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i)[\{s_j\} \times S_{-i,j}]$. Since $t_i \in \text{Proj}_{T_i} SB_i(F_{-i})$, $f_i(t_i)$ strongly believes $\text{Proj}_{S_j} F_j$ for all $j \neq i$. Then, $g_i(t'_i)$ strongly believes F_j . So $(s_i, t'_i) \in SB_i(F_{-i}) \cap F_i = G_i$. \square

Now I prove that the lemma holds for $n = \infty$. By finiteness, there is $M \in \mathbb{N}$ such that $S_G^\infty = S_G^M$. For each $i \in I$, let $F_i := CSB_i^M(E)$ and $G_i := CSB_i^\infty(E)$.

For each $j \in I$, $s_j \in S_{j,G}^\infty$, and $q \in \mathbb{N}$, as shown above $(\{s_j\} \times T_j) \cap CSB_j^{q-1}(E)$ is non-empty, and also closed (see Section 4). Thus, $((\{s_j\} \times T_j) \cap CSB_j^{q-1}(E))_{q \in \mathbb{N}}$ is a sequence of nested, non-empty closed sets, so it has the finite intersection property. Then, since Ω_j is compact,

$$\bigcap_{q \in \mathbb{N}} ((\{s_j\} \times T_j) \cap CSB_j^{q-1}(E)) = (\{s_j\} \times T_j) \cap G_j \neq \emptyset.$$

So, $S_{j,G}^\infty \subseteq \text{Proj}_{S_j} G_j$, and as shown above $S_{j,G}^\infty = \text{Proj}_{S_j} F_j \supseteq \text{Proj}_{S_j} G_j$. Hence, $S_{j,G}^\infty = \text{Proj}_{S_j} G_j$ and there exists a map $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} F_j \mapsto (\bar{s}_j, t_j) \in G_j \subset F_j$.

Fix $i \in I$ and $\mu_i \in \Delta_i^{\infty,G}$. Since $\mu_i \in \Delta_i^{M+1,G}$, as shown above there exists $t_i \in \text{Proj}_{T_i} F_i$ such that $f_i(t_i) = \mu_i$, and F has the completeness property. So, there exists $(s'_i, t'_i) \in F_i$ such that for every $i \in I$, $f_i(t'_i) = f_i(t_i)$, and for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} F_j$, $g_{i,h}(t'_i)[\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i)[\{s_j\} \times S_{-i,j}]$. Then, since $f_i(t_i)$ strongly believes $S_{j,G}^M = S_{j,G}^\infty = \text{Proj}_{S_j} G_j$, $g_i(t'_i)$ strongly believes G_j . Hence, for each $q \geq M$, since

$\text{Proj}_{S_j} \text{CSB}_j^q(E) = S_{j,G}^\infty$ and $\text{CSB}_j^q(E) \supset G_j$, $g_i(t'_i)$ strongly believes $\text{CSB}_j^q(E)$. So, $(s'_i, t'_i) \in \text{SB}_i(\text{CSB}_{-i}^q(E))$. Then, inductively, $(s'_i, t'_i) \in \text{CSB}_i^{q+1}(E)$ for all $q \geq M$. Thus $(s'_i, t'_i) \in G_i$.

Fix $i \in I$ and $t_i \in \text{Proj}_{T_i} G_i$. For every $q \geq 1$, $t_i \in \text{Proj}_{T_i} \text{CSB}_i^{q-1}(E)$, thus, as shown above, $f_i(t_i) \in \Delta_i^{q,G}$. Then, $f_i(t_i) \in \Delta_i^{\infty,G}$.

Finally, I show that G has the completeness property. Fix $i \in I$, $t_i \in \text{Proj}_{T_i} G_i \subseteq \text{Proj}_{T_i} F_i$, $s_i \in \rho(f_i(t_i))$, and maps $(\tau_j)_{j \neq i}$ with $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} G_j \mapsto (\bar{s}_j, t_j) \in G_j \subseteq F_j$ for all $j \neq i$. As shown above, $\text{Proj}_{S_j} G_j = \text{Proj}_{S_j} F_j$, and F has the completeness property. So, there exists $t'_i \in T_i$ such that $(s_i, t'_i) \in F_i$, $f_i(t'_i) = f_i(t_i)$, and for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} F_j$, $g_{i,h}(t'_i)[\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i)[\{s_j\} \times S_{-i,j}]$. Since $t_i \in \text{Proj}_{T_i} \text{SB}_i(F_{-i})$, $f_i(t_i)$ strongly believes $\text{Proj}_{S_j} F_j = \text{Proj}_{S_j} G_j$ for all $j \neq i$. Then, $g_i(t'_i)$ strongly believes G_j . Hence, for each $q \geq M$, since $\text{Proj}_{S_j} \text{CSB}_j^q(E) = \text{Proj}_{S_j} G_j$ and $\text{CSB}_j^q(E) \supset G_j$, $g_i(t'_i)$ strongly believes $\text{CSB}_j^q(E)$. So, $(s_i, t'_i) \in \text{SB}_i(\text{CSB}_{-i}^q(E))$. Then, inductively, $(s_i, t'_i) \in \text{CSB}_i^{q+1}(E)$ for all $q \geq M$. Thus $(s_i, t'_i) \in G_i$. ■

Proof of Theorems 1 and 2.

Let $(\Delta_i^0)_{i \in I} := (\Delta^{H_i}(S_{-i}))_{i \in I}$. Fix a chain of restrictions $((\Delta_i^1)_{i \in I}, \dots, (\Delta_i^k)_{i \in I})$. (For Theorem 1, let $(\Delta_i^1)_{i \in I} = (\Delta_i)_{i \in I}$.) Let $\text{CSB}_{-1}^\infty := R$ and, for each $l = 0, \dots, k$ and $n \in \mathbb{N}_0 \cup \{\infty\}$, let $\text{CSB}_l^n := \text{CSB}^n([\Delta^l] \cap \text{CSB}_{l-1}^\infty)$ (so, $\text{CSB}_0^n = \text{CSB}^n(R)$). For each $i \in I$, let $\Delta_{i,0}^G := \Delta^{H_i}(S_{-i})$. For each $l = 1, \dots, k$, let $\Delta_{i,l}^G$ be the set of CPS's that satisfy E3(l) (so, $\Delta_{i,1}^G$ is the set of CPS's that satisfy S3). Then, Theorem 1 is given by Lemma 1 with $E := [\Delta^1] \cap \text{CSB}_0^\infty = [\Delta] \cap \text{CSB}^\infty(R)$ and $(\Delta_i^G)_{i \in I} := (\Delta_{i,1}^G)_{i \in I}$.⁴¹ Theorem 2 is given by Lemma 1 with $E := [\Delta^k] \cap \text{CSB}_{k-1}^\infty$, $(\Delta_i)_{i \in I} = (\Delta_i^k)_{i \in I}$, and $(\Delta_i^G)_{i \in I} := (\Delta_{i,k}^G)_{i \in I}$. I am going to show inductively that each $E = [\Delta^l] \cap \text{CSB}_{l-1}^\infty$ satisfies the conditions of Lemma 1.

Induction Hypothesis ($l = 0, \dots, k$): Lemma 1 holds for $E := [\Delta^l] \cap \text{CSB}_{l-1}^\infty$, $(\Delta_i)_{i \in I} = (\Delta_i^l)_{i \in I}$, and $(\Delta_i^G)_{i \in I} = (\Delta_{i,l}^G)_{i \in I}$.

Basis step ($l = 0$):

The event $E = [\Delta^0] \cap \text{CSB}_{-1}^\infty = R$ is closed (see Section 4). Now I show that it has the completeness property. Fix $i \in I$, $t_i \in \text{Proj}_{T_i} R_i$, $s_i \in \rho(f_i(t_i))$, and, for each $j \neq i$, $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} R_j \mapsto (\bar{s}_j, t_j) \in R_j$. Extend each τ_j to $\tau'_j : \bar{s}_j \in S_j \mapsto (\bar{s}_j, t_j) \in \Omega_j$ in such a way that for every $s_j \in \text{Proj}_{S_j} R_j$, $\tau'_j(s_j) = \tau_j(s_j)$. Define $\nu_i \in (\Delta(S_{-i} \times T_{-i}))^{H_i}$ as

⁴¹Differently than Selective Rationalizability, the generalized procedure used here does not have a step 0. However, as already observed in Section 3, step 0 is immaterial in presence of S3.

$\nu_i((\tau'_j(s_j))_{j \neq i} | h) = f_{i,h}(t_i)[s_{-i}]$ for all $h \in H_i$ and $s_{-i} = (s_j)_{j \neq i} \in S_{-i}$ (it is well defined because each τ'_j is injective). It is easy to verify that ν_i is a CPS given that $f_i(t_i)$ is a CPS. By ontteness of g_i , there exists $t'_i \in T_i$ such that $g_i(t'_i) = \nu_i$. Clearly, $f_i(t'_i) = f_i(t_i)$, which implies $(s_i, t'_i) \in R_i$, and $g_{i,h}(t'_i) [\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i) [\{s_j\} \times S_{-i,j}]$ for all $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} R_j$.

Moreover, $f_i(\text{Proj}_{T_i} R_i) = \Delta^{H_i}(S_{-i}) = \Delta_i^0 \cap \Delta_{i,0}^G$. So, the conditions of Lemma 1 are satisfied.

Inductive step (l):

By assumption, $[\Delta^l]$ is closed, and CSB_{l-1}^∞ is closed too (see Section 4). Thus, $E = [\Delta^l] \cap \text{CSB}_{l-1}^\infty$ is closed. Now I show that it has the completeness property. Fix $i \in I$, $t_i \in \text{Proj}_{T_i} E$, $s_i \in \rho(f_i(t_i))$, and, for each $j \neq i$, $\tau_j : \bar{s}_j \in \text{Proj}_{S_j} E \mapsto (\bar{s}_j, t_j) \in \text{Proj}_{\Omega_j} E$. Extend each τ_j to $\tau'_j : \bar{s}_j \in \text{Proj}_{S_j} \text{CSB}_{l-1}^\infty \mapsto (\bar{s}_j, t_j) \in \text{Proj}_{\Omega_j} \text{CSB}_{l-1}^\infty$ in such a way that for every $s_j \in \text{Proj}_{S_j} E$, $\tau'_j(s_j) = \tau_j(s_j)$. By the Induction Hypothesis, CSB_{l-1}^∞ has the completeness property. So, there exists $t'_i \in T_i$ such that $(s_i, t'_i) \in \text{Proj}_{\Omega_i} \text{CSB}_{l-1}^\infty$, $f_i(t'_i) = f_i(t_i)$ and for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} \text{CSB}_{l-1}^\infty$, $g_{i,h}(t'_i) [\{\tau'_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i) [\{s_j\} \times S_{-i,j}]$. Thus, for every $h \in H_i$, $j \neq i$, and $s_j \in \text{Proj}_{S_j} E$, $g_{i,h}(t'_i) [\{\tau_j(s_j)\} \times \Omega_{-i,j}] = f_{i,h}(t_i) [\{s_j\} \times S_{-i,j}]$. Since $f_i(t'_i) = f_i(t_i)$, $f(t'_i) \in \Delta_i^{l+1} \subset \Delta^l$. Thus, $(s_i, t'_i) \in \text{Proj}_{\Omega_i} E$.

By the Induction Hypothesis, $f_i(\text{Proj}_{T_i} \text{CSB}_{l-1}^\infty) = \Delta_i^{l-1} \cap \Delta_{i,l}^G \supseteq \Delta_i^l \cap \Delta_{i,l}^G$ for all $i \in I$. Thus, $f_i(\text{Proj}_{T_i} E) = \Delta_i^l \cap \Delta_{i,l}^G$.

So, the conditions of Lemma 1 are satisfied. ■

References

- [1] Aumann, R. “Correlated Equilibrium as an Expression of Bayesian Rationality”, *Econometrica*, **55** (1987), 1–18.
- [2] Battigalli, P., “Strategic Rationality Orderings and the Best Rationalization Principle,” *Games and Economic Behavior*, **13** (1996), 178-200.
- [3] Battigalli, P., “On rationalizability in extensive games”, *Journal of Economic Theory*, **74** (1997), 40-61.
- [4] Battigalli, P., “Rationalizability in Infinite, Dynamic Games of Incomplete Information,” *Research in Economics*, **57** (2003), 1-38.
- [5] Battigalli, P. and A. Friedenberg, “Forward induction reasoning revisited”, *Theoretical Economics*, **7** (2012), 57-98.

- [6] Battigalli, P. and A. Prestipino, “Transparent Restrictions on Beliefs and Forward Induction Reasoning in Games with Asymmetric Information”, *The B.E. Journal of Theoretical Economics* (Contributions), **13** (2013), Issue 1.
- [7] Battigalli, P. and M. Siniscalchi, “Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games”, *Journal of Economic Theory*, **88** (1999), 188-230.
- [8] Battigalli, P. and M. Siniscalchi, “Interactive Beliefs, Epistemic Independence and Strong Rationalizability”, *Research in Economics*, **53** (1999), 247-273.
- [9] Battigalli, P. and M. Siniscalchi, “Strong Belief and Forward Induction Reasoning”, *Journal of Economic Theory*, **106** (2002), 356-391.
- [10] Battigalli P. and M. Siniscalchi, “Rationalization and Incomplete Information”, *The B.E. Journal of Theoretical Economics*, **3(1)** (2003), 1-46.
- [11] Battigalli, P. and P. Tebaldi, “Interactive Epistemology in Simple Dynamic Games with a Continuum of Strategies”, 2017, IGIER working paper 602.
- [12] Blume, L., A. Brandenburger, and E. Dekel, “Lexicographic Probabilities and Choice Under Uncertainty”, *Econometrica*, **59** (1991), 61-79.
- [13] Brandenburger, A., and A. Friedenberg, “Intrinsic correlation in games”, *Journal of Economic Theory*, **141** (2008), 28-67.
- [14] Catonini, E., “Self-enforcing agreements and forward induction reasoning”, 2017, working paper.
- [15] Catonini, E., “On non-monotonic strategic reasoning”, 2017, working paper.
- [16] Chen, J., and S. Micali, “The order independence of iterated dominance in extensive games”, *Theoretical Economics*, **8** (2013), 125-163.
- [17] Cho, I.K. and D. Kreps, “Signaling Games and Stable Equilibria”, *Quarterly Journal of Economics*, **102** (1987), 179-222.
- [18] Friedenberg, A., “When Do Type Structures Contain All Hierarchies of Beliefs?”, *Games and Economic Behavior*, **68** (2010), 108-129.
- [19] Friedenberg, A., “Bargaining under strategic uncertainty”, working paper, 2014.
- [20] Fudenberg, D., and D. Levine, “Self-confirming equilibrium”, *Econometrica*, **61** (1993), 523–546.

- [21] Govindan, S., and R. Wilson, “On Forward Induction,” *Econometrica*, **77** (2009), 1-28.
- [22] Heifetz, A. and A. Perea, “On the Outcome Equivalence of Backward Induction and Extensive Form Rationalizability”, *International Journal of Game Theory*, **44**(1) (2015), 37-59.
- [23] Kohlberg, E. and J.F. Mertens, “On the Strategic Stability of Equilibria”, *Econometrica*, **54** (1986), 1003-1038.
- [24] Kreps, D., and R. Wilson, “Sequential Equilibria”, *Econometrica*, **50** (1982), 863-894.
- [25] Lo, K.C., “Nash Equilibrium without Mutual Knowledge of Rationality”, *Economic Theory*, **14** (1999), 621-633.
- [26] Man, P. “Forward Induction Equilibrium”, *Games and Economic Behavior*, **75**(1) (2012), 265-276.
- [27] Osborne, M., “Signaling, Forward Induction, and Stability in Finitely Repeated Games”, *Journal of Economic Theory*, **50** (1990), 22-36.
- [28] Osborne, M. J. and A. Rubinstein, “A Course in Game Theory”, 1994, Cambridge, Mass.: MIT Press.
- [29] Pearce, D., “Rational Strategic Behavior and the Problem of Perfection”, *Econometrica*, **52** (1984), 1029-1050.
- [30] Perea, A., “Why Forward Induction Leads to the Backward Induction Outcome: A New Proof for Battigalli’s Theorem”, 2017, working paper.
- [31] Renyi, A., “On a New Axiomatic Theory of Probability”, *Acta Mathematica Academiae Scientiarum Hungaricae*, **6** (1955), 285-335.
- [32] Shimoji, M., “On the Equivalence of Weak Dominance and Sequential Best Response”, *Games and Economic Behavior*, **48** (2004), 385-402.
- [33] Siniscalchi, M., “Structural Rationality in Dynamic Games”, 2016, working paper.