Strong Rationalizability, Learning, and Equilibrium in Repeated Games with Imperfect Feedback

Pierpaolo Battigalli and Davide Bordoli

February 21, 2022

Abstract

We analyze the infinite repetition with imperfect feedback of a simultaneous or sequential game, assuming that players are strategically sophisticated (but possibly impatient) expected-utility maximizers. Players correctly frame the repetition of the one-period game as a grand repeated game. Sophisticated strategic reasoning in the repeated game is combined with belief updating, or revision, to provide a foundation for a refinement of self-confirming equilibrium. In particular, we provide an epistemic analysis of *rationality and common strong belief in rationality*, extending some existing results to our context. Then, we combine beliefs updating and sophisticated reasoning to provide sufficient conditions for a kind of learning—that is, the ability, in the limit, to exactly forecast the sequence of future observations—thus showing that impatient agents end up playing a sequence of *self-confirming equilibria with strongly rationalizable beliefs* of the one-period game. Moreover, any such sequence is the possible outcome of a repeated interaction among sophisticated, impatient players that satisfy the sufficient conditions for learning. Irrespective of whether individuals value the future, if they are able to learn then they will play in the limit a self-confirming equilibrium with strongly rationalizable beliefs of the continuation game.

1. Introduction

In this paper we analyze the limits of learning dynamics in the infinite repetition with imperfect monitoring of a one-period game played by strategically sophisticated agents. The one-period game may be sequential or with simultaneous moves. Focusing on the case of impatient agents who maximize their subjective expected one-period payoff, we relate such limits to solutions of the one-period game, that is, self-confirming equilibrium and rationalizability.

In a **self-confirming equilibrium** (**SCE**), players best reply to confirmed beliefs about co-players' behavior, where "confirmed" means that each player, given her beliefs, correctly predicts her observations about play. The SCE concept characterizes the limits of learning dynamics in games played recurrently given the possibly imperfect feedback about play obtained by each player at the end of each period (e.g., Fudenberg & Kreps 1995 and Gilli 1999). Note that the SCE term was coined by Fudenberg & Levine (1993), but the concept was also previously or simultaneously called "conjectural equilibrium" (Battigalli 1987, Battigalli & Guaitoli 1988, Rubinstein & Wolinsky 1994) and "subjective equilibrium" (Kalai & Lehrer 1993, 1995). Here we stick to the more explicative SCE terminology (see the discussion in Battigalli et al 2015). In an SCE, beliefs about others may be incompatible with strategic reasoning based on what is commonly known about the game. Indeed, in an environment with possibly incomplete information and private values,[1] the SCE set is independent of players' interactive knowledge of the profile of payoff functions. It is then natural to ask how one can characterize the limits of learning dynamics when beliefs are shaped by sophisticated strategic reasoning, which we take to mean some form of *common belief in rationality.*

The literature offers two kinds of answers that directly focus on refinements of SCE, neglecting an explicit analysis of learning dynamics. The simplest one can be found in the works that first put forward a version of the SCE concept (Battigalli 1987, and Battigalli & Guaitoli 1988): SCE should be refined by requiring that players' beliefs about co-players' behavior assign probability 1 to co-players' rationalizable strategies, a condition that follows from common belief in rationality. Yet, such **SCE in rationalizable beliefs** allows for the possibility that confirmation of beliefs is not commonly believed, which may be thought to jeopardize the stability of the equilibrium. Intuitively, if confirmed beliefs is a pre-requisite to play again the same strategies, why should a sophisticated player who is unsure whether her co-players' beliefs are confirmed expect that they behave in the future as in the current period? And if they don't, why should she? Motivated by such informal considerations, Rubinstein & Wolinsky (1994) proposed an even more refined notion of SCE: while an SCE in rationalizable beliefs obtains if there is common belief in rationality and beliefs are confirmed, in a **rationalizable SCE**[2] players beliefs about behavior are compatible with common belief of *both* rationality and confirmation of beliefs. Rationalizable SCE is elegant and intuitive, but—unlike the mere SCE concept, to the best of our knowledge—there is no formal result relating it to learning in recurrent interaction. Instead, here we obtain a kind of learning foundation for SCE in rationalizable beliefs.

To formally represent rationality and strategic sophistication, we adopt the approach of dynamic epistemic game theory[3] extended to infinitely repeated games as in Battigalli & Tebaldi (2019). To ease notation, we assume *complete information*: the rules of the game and players' expected-utility

[1]Knowledge of one's own payoff function.
[2]In their words, "rationalizable conjectural equilibrium."
[3]See the survey of Dekel & Siniscalchi (2015).

preferences over streams of stochastic outcomes are commonly known. Since the one-period game being repeated may have a sequential (multistage) structure, we need to distinguish between strategies of the one-period game and strategies of the repeated game; we call the latter **superstrategies**. Players are endowed with **conditional probability systems (CPSs)**, which specify subjective beliefs about the behavior and beliefs of co-players *in the infinitely repeated game* conditional on every personal history (roughly, information set) so as to satisfy the *chain rule*. We assume that players are **rational**, that is, they carry out strategies that maximize their subjective expected utility conditional on every personal history, including those that they did not expect to observe according to earlier beliefs specified by their CPSs. Of course, assumptions about intertemporal preferences are crucial. We mostly focus on the extreme case of **impatient** players who do not value future payoffs, as in much of the literature on learning in games, but we also consider the more general case of a positive discount factor. To model strategic sophistication, we assume *common strong belief in rationality* (Battigalli & Siniscalchi 2002): each player **strongly believes** in the co-players' rationality, i.e., she assigns probability 1 to it conditional on every personal history that does not contradict it; furthermore, she strongly believes that, on top of being rational, her co-players also strongly believe in the rationality of others; analogous assumptions hold for higher and higher levels of beliefs about beliefs. With this, in every period impatient agents play (strongly) rationalizable strategies, and *assign probability 1 to the (strongly) rationalizable strategies of others* even if they are surprised.[4] The reason is that, on a rationalizable path, unexpected observations cannot be due to deviations from rationalizability; therefore, common strong belief in rationality implies that even surprised players keep believing in rationalizability. To obtain convergence to SCE play, we assume that the profile of superstrategies and CPSs satisfy an **"observational grain of truth"** condition (cf. Kalai & Lehrer 1993, 1995): after some date T, each player assigns positive probability to what she is actually going to observe in the continuation (infinitely repeated) game.[5] This implies that in the long-run limit players assign probability 1 to what they observe, i.e., their *beliefs are confirmed*. Since impatient players maximize their one-period subjective expected utility, there must be *convergence to playing an SCE in rationalizable beliefs in each period*. However, *the SCE played in the limit may change from period to period*, because convergence of beliefs about co-players' superstrategies does not imply convergence of marginal one-period beliefs about co-players' strategies. We also show a converse: for every sequence of one-period SCEs in rationalizable beliefs there is a profile of superstrategies and CPS satisfying the aforementioned conditions that yields such sequence in the limit. We also extend our results to allow for a positive discount factor: under rationality, common strong belief in rationality and observational grain of truth, players' *behavior and beliefs converge to an SCE in rationalizable beliefs of the repeated game*, that is, in the long-run limit players best respond to confirmed beliefs assigning probability 1 to co-players' rationalizable superstrategies, and the appropriate version of the converse mentioned above also holds.

The rest of the paper is organized as follows. Section 2 contains some mathematical preliminaries. Section 3 describes one-period multistage games with imperfectly observable actions and their infinite repetition. Section 4 analyzes rationality for the one-period game and its repetition, and characterizes

[4]strong rationalizability is akin to the notion of rationalizability for sequential games put forward by Pearce (1984); thus, it coincides with the usual rationalizability concept in games with simultaneous moves. Since it is the only version of the rationalizability idea considered here, we sometimes simplify our language and omit the adjective "strong."

[5]Absent randomization, which we exclude because expected utility maximizers have no need to randomize, our assumption is a generalization of the "grain of truth condition of Kalai & Lehrer (1993).

the behavioral and first-order-beliefs implications of rationality and common strong belief in rationality. Section 5 analyzes convergence of beliefs. Section 6 contains the main result of the paper, i.e., convergence to SCE with strongly rationalizable beliefs. Section 7 discusses in detail the related literature and some possible extensions of our work.

2. Preliminaries

2.1 Mathematical notation

We denote with $[n] = \{1, \ldots, n\}$ the set of the first n natural numbers. Given a finite set X, we denote by $X^{[t]}$ the set of functions from $[t]$ to X, i.e., the sequences of length t of elements of X, by $X^{\mathbb{N}}$ the set of infinite sequences of elements of X, by $X^{[0]} = \{\varnothing\}$ the singleton containing the empty sequence $\varnothing$, by $X^{<\mathbb{N}_0} = \cup_{t\in\mathbb{N}_0} X^{[t]}$ the set of finite sequences of elements of X, and by $X^{\leq\mathbb{N}_0} = X^{<\mathbb{N}_0} \cup X^{\mathbb{N}}$ the set of finite and infinite sequences of elements of X.[6] To make the elements of the sequence explicit, we write $x^{[t]} = (x^k)_{k=1}^t$ for any $t \in \mathbb{N} \cup \{\infty\}$.

We endow every finite set X with the discrete topology and the corresponding Borel σ-algebra $\mathcal{B}(X)$, which coincides with its power set 2^X. We endow any Cartesian product of sets with the product topology and the corresponding Borel σ-algebra. Then, given a countable sequence of finite sets $(X_t)_{t\in\mathbb{N}}$, the σ-algebra $\mathcal{B}(X)$ on their product $X = \prod_{t\in\mathbb{N}} X_t$ is the one generated by all the cylinders of the form $\{x_1\} \times \ldots \times \{x_t\} \times X_{t+1} \times \ldots$, with $t \in \mathbb{N}$.[7] Given any topological space Y endowed with its Borel σ-algebra $\mathcal{B}(Y)$, we denote by $\Delta(Y)$ the space of probability measures defined over $(Y, \mathcal{B}(Y))$, which we endow with the topology of weak convergence.

On the space of sequences we can define a natural partial order, the "prefix of" relation, denoted by $\preceq$, in the following way: for all $\ell \in \mathbb{N}$, $t \in \mathbb{N} \cup \{\infty\}$, $x^{[\ell]}, y^{[t]} \in X^{\leq\mathbb{N}_0}$,

$$x^{[\ell]} \preceq y^{[t]} \Leftrightarrow (t \geq \ell) \wedge \left(\exists z^{[t-\ell]} \in X^{\leq\mathbb{N}_0},\ y^{[t]} = (x^{[\ell]}, z^{[t-\ell]})\right)$$

where $(x^{[\ell]}, z^{[t-l]})$ is the concatenation of $x^{[\ell]}$ and $z^{[t-l]}$ (for $t = l$, $z^{[0]} = \varnothing$ is the empty sequence and $(x^{[\ell]}, z^{[0]}) = x^{[l]}$). Given a sequence $x^{[t]} \in X^t$, we define its length as $\ell(x^{[t]}) = t$.

2.2 beliefs representation and properties

In this subsection we introduce conditional probability systems, representing players beliefs and hierarchies of beliefs, and give the definition of strong belief.

Definition 1: *Let Y be a Polish space and $\mathcal{C} \subseteq \mathcal{B}(Y)$ be a countable collection of Borel subsets of Y. A **conditional probability system (CPS)** on $(Y, \mathcal{C})$ is an array of probability measures $\mu = (\mu(\cdot|C))_{C\in\mathcal{C}} \in [\Delta(Y)]^{\mathcal{C}}$ such that:*

(i) *for all $C \in \mathcal{C}$, $\mu(C|C) = 1$;*

[6] That is, we regard such sequences as functions with domain $[n]$ or $\mathbb{N}$ and codomain X.
[7] See, for example, Aliprantis & Border, *Infinite Dimensional Analysis*

(ii) *for all* $E \in \mathcal{B}(Y)$ *and* $C, D \in \mathcal{C}$ *such that* $E \subseteq D \subseteq C$,

$$\mu(E|C) = \mu(E|D)\mu(D|C).$$

We denote with $\Delta^{\mathcal{C}}(Y)$ the set of all CPSs on $(Y, \mathcal{C})$. CPSs will be used to represent the beliefs of a player about opponents' behavior and beliefs, compactly modeling the way in which, upon observing some sequence of messages (from which a conditioning event can be inferred), the player updates her beliefs. The interpretations are: (i) upon observing a certain sequence of messages, a player is certain about it; (ii) beliefs comply with the chain rule of conditional probabilities, which connects two beliefs conditioned on two different events, when one event is contained in the other (i.e. one sequence of messages is a possible continuation of the other) and belief conditioned on the larger event gives positive probability to the smaller event.

Definition 2: (Battigalli and Siniscalchi (2002)) *Fix an event* $E \in \mathcal{B}(Y)$ *and a CPS* $\mu \in \Delta^{\mathcal{C}}(Y)$*. We say that* μ ***strongly believes*** E *if, for every* $C \in \mathcal{C}$,

$$E \cap C \neq \emptyset \Rightarrow \mu(E|C) = 1.$$

Take a decreasing sequence of events $\mathcal{E}$*, then* μ *strongly believes* $\mathcal{E}$ *if it strongly believes every element of* $\mathcal{E}$*.*

3. Set up

3.1 One-period game

A finite multistage game with feedback is a game that may last for more than one stage, where at each stage every player chooses an action and then observes a message about the play. We represent the information accruing to agents as the play unfolds with a formalism that is similar to the one used to represent information (monitoring) in repeated games.[8] Stages are indexed by natural numbers: stage k starts after the end of stage $k-1$ and ends with the realization of the action and message profiles $(a_i^k)_{i \in I}$ and $(m_i^k)_{i \in I}$ (of messages). At a stage in which a player i is inactive, we adopt the convention that her set of available actions is the singleton $\{w\}$, where w is interpreted as the action "wait". A finite multistage game has necessarily a finite horizon, that is, a maximum number of stages $L \in \mathbb{N}$ after which the game ends. In order to simplify the formal representation of the infinite repetition of the game, we adopt the convention that, each time the one-period interaction is played, the play lasts L stages. If at some history shorter than L the game ends, then players are assumed to play the action "wait" for all the following stages, until the *L-th*. The rules of a multistage game are represented by the primitive elements

$$\langle I, (A_i, M_i, \mathcal{A}_i, F_i)_{i \in I} \rangle,$$

where:

[8] See Battigalli & Generoso (2021).

- I is the finite set of players;
- A_i is the finite set of all actions player i may ever take at any point in the game;
- M_i is the finite set of all messages player i may ever observe at any point in the game, including the "message" observed at the beginning of the game;
- $\mathcal{A}_i = (\mathcal{A}_i^k : M_i \rightrightarrows A_i)_{k=1}^L$ is a *sequence of constraint correspondences*: for every k and for every possible message m_i^{k-1} observed at the end of stage $k-1$, $\mathcal{A}_i^k\left(m_i^{k-1}\right)$ specifies the set of i's feasible actions at stage k. The set of actions feasible for a player at a given stage must depend solely on what the player has just observed. With this representational assumption, any informative effect of available actions is internalized in the messages;
- $F_i = (f_i^k : A^{[k]} \to M_i)_{k=0}^L$ is the *incremental feedback function*, representing the informational flow structure of the game. For every k and every conceivable sequence of action profiles $a^{[k]}$, $f_i^k\left(a^{[k]}\right)$ is message observed by player i at the end of stage k after $a^{[k]}$. Then, the sequence of such feedbacks up to the current stage, which is the incremental feedback function, determines the information potentially available to a player, besides the sequence of her own actions, that are automatically observed as soon as they are irreversibly chosen. The initial message $f_i^0(\varnothing)$ informs i of her feasible actions in the first stage.[9]

The primitives allow the derivation of the sets of one-period histories and personal histories, that is, the *objective* and the *subjective trees* (see Appendix B) generated by the one-period game form. Recalling the notation for sequences, we label the sequence of messages received until the end of stage k by player i, when the sequence of action profiles $a^{[k]}$ has been played, as $f_i^{[k]}(a^{[k]})$. The corresponding sequence of actions played by i is instead written as $a_i^{[k]}(a^{[k]})$. Then, we can define the following sets:

- $\overline{H}$ is the set of **histories,** that is, the feasible sequences of action profiles including the empty sequence $a^{[0]} = \varnothing$ (root):

$$\overline{H} = \{\varnothing\} \cup \left\{ (a^k)_{k=1}^l : l \leq L, \forall k \in [l]\,, a^k \in \prod_{i \in I} \mathcal{A}_i^k(f_i^{k-1}(a^{[k-1]})) \right\};$$

- $Z = \left\{z \in \overline{H} : \ell(z) = L\right\}$ is the set of **terminal histories**;
- $H = \overline{H} \setminus Z$ is the set of **non-terminal histories**;
- the set of **personal histories** is

$$\overline{H}_i = \{\left(f_i^0(\varnothing)\right)\} \cup \{(a_i, m_i)^{[k]} \in A_i^{[k]} \times M_i^{[k]} : k \leq L, \exists a_{-i}^{[k]} \in \overline{H}_{a_i^{[k]}}, f_i^{[k]}(a_i^{[k]}, a_{-i}^{[k]}) = m_i^{[k]}\},$$

  and it is paritioned into Z_i and H_i (terminal and non-terminal personal histories).[10]

Personal histories represent what a player is able to observe at each stage given that a certain history has occurred. Whether every player knows her set of personal histories prior to the game

[9] We abuse notation by letting $f_i^0(\varnothing) = \varnothing$.

[10] $\overline{H}_{a_i^{[k]}}$ is the section of $\overline{H}$ at $a_i^{[k]}$.

depends on whether she knows action sets and sequences of constraint correspondences of everybody and incremental feedback function of herself. Whether at a certain stage she is able to use the sequence of own actions and messages she has observed as information depends on her memory. The assumption of *perfect recall* consists thus in saying that, at every stage t, each player i knows her personal history. Given such assumption, personal histories of actions and messages yield a representation of information equivalent to information partitions. To simplify notation, we let $O_i : \overline{H}_i \rightrightarrows \overline{H}$ be the correspondence that assigns to every personal history of i the corresponding information set, i.e., for every $h_i = (\overline{a}_i^{[k]}, m_i^{[k]}) \in \overline{H}_i$:

$$O_i(h_i) = \{g \in \overline{H} : \ell(g) = k, a_i^{[k]}(g) = \overline{a}_i^{[k]}, f_i^{[k]}(a^{[k]}(g)) = m_i^{[k]}\}.$$

Note that we consider players' information also when they are inactive and at the end of the game. The latter information is used at the beginning of the new one-period game when it is repeated. Therefore, there are terminal and non-terminal information sets. Notice that $O_i^{-1} : \overline{H} \to \overline{H}_i$ is a well-defined function assigning to every history the unique corresponding personal history of player i.

We now define strategies. We think of a strategy as the plan in the mind of a player, who can think about the actions she intends to choose at each of her non-terminal personal histories. Formally, she forms a plan $s_i = (s_i(h_i))_{h_i \in H_i} \in S_i = \times_{h_i \in H_i} \mathcal{A}_i^{\ell(h_i)+1}(h_i)$, where for the sake of simplicity we are abusing notation by letting

$$\mathcal{A}_i^{\ell(h_i)+1}(h_i) = \mathcal{A}_i^{\ell(h_i)+1}\left(m_i^{\ell(h_i)}(h_i)\right).$$

From this, it is straightforward to define the induced objective strategy by s_i, based on objective histories: $\varsigma_i = s_i \circ O_i^{-1} \in \times_{h \in H} \mathcal{A}_i^{\ell(h)+1}(O_i^{-1}(h))$. From these elements we can derive the path function $\zeta : S \to Z$ mapping strategies to terminal histories: for every $s \in S = \times_{i \in I} S_i$, $\zeta(s) = (a^k)_{k=1}^L \in Z$, $a^1 = (s_i(\varnothing))_{i \in I}$ and, for every $l \geq 2$, $a^l = ((s_i(O_i^{-1}(a^{[l-1]})))_{i \in I})$.

It is also useful to define, for each player $i \in I$, the sets of profiles that induce, and strategies that allow, some personal history $h_i \in \overline{H}_i$:

- $S(h_i) = \{s \in S : \exists x \in O_i(h_i), x \preceq \zeta(s)\}$ is the set of strategy profiles that induce h_i;
- $S_i(h_i) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, \exists x \in o_i(h_i), x \preceq \zeta(s_{-i}, s_i)\} = proj_{S_i} S(h_i)$ is the set of strategies of i that allow h_i;
- $S_{-i}(h_i) = \{s_{-i} \in S_{-i} : \exists s_i \in S_i, \exists x \in o_i(h_i), x \preceq \zeta(s_{-i}, s_i)\} = proj_{S_{-i}} S(h_i)$ is the set of strategy profiles of opponents that allow h_i.

Informally, *common knowledge of the rules of the game* implies that these correspondences are commonly understood. Analogous definitions holds for "objective" histories $h \in \bar{H}$:

- $S(h) = \{s \in S : h \preceq \zeta(s)\}$;
- $S_i(h) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, h \preceq \zeta(s_{-i}, s_i)\} = proj_{S_i} S(h)$;
- $S_{-i}(h) = \{s_{-i} \in S_{-i} : \exists s_i \in S_i, h \preceq \zeta(s_{-i}, s_i)\} = proj_{S_{-i}} S(h)$.

Intuitively, $S_{-i}(h_i)$ represents the information that h_i reveals to i (assuming that she has perfect recall) concerning the strategies that the co-players are carrying out. Hence, for every i,

$$\mathcal{C}_i = \{S_{-i}(h_i) : h_i \in H_i\}$$

is the collection of observable events concerning the co-players' behavior that will be used to define the set of conditional probability systems (CPSs) of i.

We now define two other useful objects: the set of personal histories consistent with a given strategy, that is, for any $i \in I$ and $s_i \in S_i$,

$$\overline{H}_i(s_i) = \{h_i \in \overline{H}_i : s_i \in S_i(h_i)\}$$

(again, its partition in $H_i(s_i)$ and $Z_i(s_i)$ is straightforward); the set of continuation strategies of i at any given personal history $h_i \in H_i$,

$$S_i^{\succeq h_i} = \times_{g_i \in H_i, g_i \succeq h_i} \mathcal{A}_i^{\ell(g_i)+1}(g_i).$$

Defining a set of outcomes Y, an outcome function $\gamma : Z \to Y$ and von Neuman-Morgenstern utility functions $(v_i : Y \to \mathbb{R})_{i \in I}$, we construct pay-off functions $(u_i = v_i \circ \gamma : Z \to \mathbb{R})_{i \in I}$ attached to terminal histories. From this, we can also conveniently define strategic form utilities $U_i : S \to \mathbb{R}$ such that $U_i = u_i \circ \zeta$, where the strategic form utility of i is defined over strategies acting on personal histories of i.

Then, a multistage game Γ with informational structure $F = (F_i)_{i \in I}$ is defined as

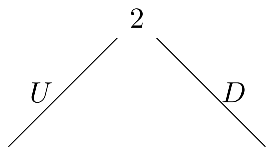
$$\Gamma = \langle I, (A_i, M_i, \mathcal{A}_i, F_i, u_i)_{i \in I} \rangle.$$

To conclude the subsection, we define the notion of **observationally equivalent** strategies, which plays a fundamental role in our analysis.

Definition 3: *Fix $i \in I$, $s_i \in S_i$, and $\overline{s}_{-i} \in S_{-i}$. We say that a profile of strategies $s_{-i} \in S_{-i}$ is observationally equivalent, given s_i, to the profile of strategies $\overline{s}_{-i}$, if $s_{-i} \in S_{-i}(O_i^{-1}(\zeta(s_i, \overline{s}_{-i})))$.*

In words, s_{-i} is observationally equivalent to $\overline{s}_{-i}$, given s_i, if these two profiles, when played along with s_i, induce the same sequence of messages observed by player i, who thus is unable to distinguish between the two profiles. Indeed, we obtain the following formal characterization: for any given $s_i \in S_i$, two pairs of opponents' strategy profiles s_{-i} and $\overline{s}_{-i}$ are observationally equivalent if and only if $O_i^{-1}(\zeta(s_i, \overline{s}_{-i})) = O_i^{-1}(\zeta(s_i, s_{-i}))$.

Example: Consider the following multistage game, which we will adopt as "running example". Payoffs and actions of player **1** are in **bold**, payoffs and actions of player *2* are in *italic*.



1 \ 2	ℓ	c	r
u	**1**, *1*	**0**, *2*	**2**, *0*
m	**1**, *2*	**2**, *0*	**0**, *1*
d	**0**, *1*	**2**, *0*	**0**, *1*

1 \ 2	ℓ	c	r
u	**2**, *0*	**1**, *1*	**3**, *-1*
m	**1**, *3*	**2**, *2*	**0**, *2*
d	**-1**, *3*	**1**, *1*	**-1**, *3*

We do not assume that terminal information fully reveals the realized payoffs. In particular, actions are ex post observable up to the following restrictions: first, independently of initial move by player 2, when playing **m**iddle player **1** can not distinguish between ℓeft and $c$enter, while player *2*, when playing ℓ or c, can not distinguish between **m** and **u**. Moreover, Down has an observational drawback for player *2*: after D, she loses the informativeness of r, which delivers the same message independently of the action chosen by **1**.

The restrictions on feedback imposed can be formally expressed as:

$$f^2_{\mathbf{1}}((x,(\mathbf{m},\ell))) = f^2_{\mathbf{1}}((x,(\mathbf{m},c))), \qquad f^2_{\mathit{2}}((x,(\mathbf{m},\ell))) = f^2_{\mathit{2}}((x,(\mathbf{u},\ell))),$$

$$\forall x \in \{U,D\}, \qquad f^2_{\mathit{2}}((x,(\mathbf{m},c))) = f^2_{\mathit{2}}((x,(\mathbf{u},c))),$$

$$f^2_{\mathit{2}}((D,(\mathbf{u},r))) = f^2_{\mathit{2}}((D,(\mathbf{m},r))) = f^2_{\mathit{2}}((D,(\mathbf{d},r))).$$

▲

3.2 Infinitely repeated interaction

The infinite repetition of the game is itself a multistage game, whose elements are clearly characterized by the elements of the one-period game. For example, the set of feasible actions for player i after personal history $(a_i^{[t]}, m_i^{[t]}) \in A_i^{[t]} \times M_i^{[t]}$ is

$$\begin{cases} \mathcal{A}_i^{l+1}(m_i^t) & l>0 \\ \mathcal{A}_i^1(f_i^0(\varnothing)) & l=0 \end{cases},$$

where $l = \ t \mod L$. For the sake of brevity, we abuse notation using $\mathcal{A}_i^t$ to denote also the feasibility correspondence assigning to sequences of i's actions and messages of length t, with t potentially greater than L, the possible actions she might take at that personal history of the infinite game. Similarly, the message observed by i after history $a^{[t]}$ is

$$\begin{cases} f_i^l((a^k)_{k=t+1-l}^t) & l>0 \\ \varnothing & l=0 \end{cases},$$

where again $l = \ t \mod L$ and we abuse notation denoting $\varnothing$ the message of player i signaling the empty history. For the sake of the analysis, we denote with $\mathbf{f}_i^t$ the feedback function mapping from sequences of action profiles of length t to the corresponding message observed by i.

Starting from these primitive elements, we can define the sets of histories and personal histories of the infinitely repeated game. First, we extend the definition of the informational correspondences

$$\mathbf{O}_i : \quad (A_i \times M_i)^{\leq \mathbb{N}} \quad \rightrightarrows \quad A^{\leq \mathbb{N}}$$

such that

$$\mathbf{O}_i((a_i, m_i)^{[t]}) = \left\{ a^{[t]} : a_i^{[t]} = proj_{A_i^{\leq \mathbb{N}}} a^{[t]}, \mathbf{f}_i^{[t]}(a^{[t]}) = m_i^{[t]} \right\}.$$

Again, $\mathbf{O}_i^{-1}$ is a well-defined function. Then, we define the various sets of histories as done for the one-period game:

- $\overline{\mathbf{H}}$ is the set of all histories;
- $\mathbf{Z} = \overline{\mathbf{H}} \cap A^{\mathbb{N}}$ is the set of terminal histories;
- $\mathbf{H} = \overline{\mathbf{H}} \setminus \mathbf{Z}$ is the set of non-terminal histories;
- $\overline{\mathbf{H}}_i = \mathbf{O}_i^{-1}(\overline{\mathbf{H}})$ is the set of i's personal histories;
- $\mathbf{Z}_i = \mathbf{O}_i^{-1}(\mathbf{Z})$ is the set of i's terminal personal histories;
- $\mathbf{H}_i = \mathbf{O}_i^{-1}(\mathbf{H})$ is the set of i's non-terminal personal histories.

Remark 1: $\mathbf{Z} = Z^{\mathbb{N}}$, $\mathbf{Z}_i = Z_i^{\mathbb{N}}$, $\mathbf{H} = \cup_{n \geq 0} (Z^n \times H)$ *and* $\mathbf{H}_i = \cup_{n \geq 0} (Z_i^n \times H_i)$.

In this context, we still call "strategy" the description of the information-dependent behavior of a player *in a single period*. We instead call "**superstrategy**" the description of the information-dependent behavior of a player in the repeated game; a superstrategy of player i is denoted with $\mathbf{s}_i \in \mathbf{S}_i = \times_{\mathbf{h}_i \in \mathbf{H}_i} \mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)$. Then, one can define the path function:

$$\begin{aligned} \boldsymbol{\zeta} : \quad & \mathbf{S} \quad \rightarrow \quad \mathbf{Z} \\ & \mathbf{s} \quad \mapsto \quad \left(\zeta\left(\prod_{i \in I} \mathbf{s}_i|_{H_i} \right), \zeta\left(\prod_{i \in I} \mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(\zeta(\prod_{i \in I} proj_{H_i} \mathbf{s}_i))\} \times H_i} \right), \ldots \right), \end{aligned}$$

where, for every subset of non-terminal personal histories $\widehat{\mathbf{H}}_i \subseteq \mathbf{H}_i$, $\mathbf{s}_i|_{\widehat{\mathbf{H}}_i}$ is the restriction of $\mathbf{s}_i$ to the desired subset $\widehat{\mathbf{H}}_i$.

Let the following sets of superstrategies be as usual:

- $\mathbf{S}(\mathbf{h}_i) = \{\mathbf{s} \in \mathbf{S} : \exists \mathbf{x} \in \mathbf{O}_i(\mathbf{h}_i),\ \mathbf{x} \prec \boldsymbol{\zeta}(\mathbf{s})\}$;
- $\mathbf{S}_i(\mathbf{h}_i) = proj_{\mathbf{S}_i} \mathbf{S}(\mathbf{h}_i)$;
- $\mathbf{S}_{-i}(\mathbf{h}_i) = proj_{\mathbf{S}_{-i}} \mathbf{S}(\mathbf{h}_i)$;
- $\mathbf{S}(\mathbf{h}) = \{\mathbf{s} \in \mathbf{S} : \mathbf{h} \prec \boldsymbol{\zeta}(\mathbf{s})\}$;
- $\mathbf{S}_i(\mathbf{h}) = proj_{\mathbf{S}_i} \mathbf{S}(\mathbf{h})$;

- $\mathbf{S}_{-i}(\mathbf{h}) = proj_{\mathbf{S}_{-i}}\mathbf{S}(\mathbf{h})$.

Let $\mathbf{H}_i(\mathbf{s}_i) = \{\mathbf{h}_i : \mathbf{s}_i \in \mathbf{S}_i(\mathbf{h}_i)\}$, and $\mathbf{S}_i^{\succeq \mathbf{h}_i}$ be the set of continuation superstrategies of i at personal history $\mathbf{h}_i$.

Under the assumption of perfect recall, the collection of conditioning events for the CPSs of player i is

$$\mathfrak{C}_i = \{\mathbf{S}_{-i}(\mathbf{h}_i) : \mathbf{h}_i \in \mathbf{H}_i\}.$$

In particular, beliefs of player i are represented by elements of $\Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$.

Definition 4: *Fix $i \in I$, $\mathbf{s}_i \in \mathbf{S}_i$, and $\overline{\mathbf{s}}_{-i} \in \mathbf{S}_{-i}$. We say that a profile of superstrategies $\mathbf{s}_{-i} \in \mathbf{S}_{-i}$ is observationally equivalent, given $\mathbf{s}_i$, to the profile of superstrategies $\overline{\mathbf{s}}_{-i}$, if $\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\zeta(\mathbf{s}_i, \overline{\mathbf{s}}_{-i})))$.*

Alternatively, $\mathbf{s}_{-i}$ and $\overline{\mathbf{s}}_{-i}$ are observationally equivalent, given $\mathbf{s}_i$, if $\mathbf{O}_i^{-1}(\zeta(\mathbf{s}_i, \overline{\mathbf{s}}_{-i})) = \mathbf{O}_i^{-1}(\zeta(\mathbf{s}_i, \mathbf{s}_{-i}))$. Every period an outcome from the finite set Y (of the one-period game Γ) is generated. Recall from Section 3.1 that outcomes are determined by a function $g : Z \to Y$, and, for every player i, we define a von Neumann-Morgenstern utility function $v_i : Y \to \mathbb{R}$. Then, we construct the profile of one-period payoff functions $(u_i : Z \to \mathbb{R})_{i\in I}$ and one-period normal-form payoff functions $(U_i = u_i \circ \zeta : S \to \mathbb{R})_{i\in I}$. As we have seen in Remark 1, a terminal history of the repeated game can be seen as a sequence of terminal histories of the one-period game, i.e. $\mathbf{z} = (z^t(\mathbf{z}))_{t=1}^{\infty}$, where $z^t(\mathbf{z}) = (a^l(\mathbf{z}))_{l=(t-1)L+1}^{tL}$, and $z^{[t]} = (z^k(\mathbf{z}))_{k=1}^{t}$. Along the path $\mathbf{z}$, a sequence of outcomes $(y^t)_{t\in\mathbb{N}} = (g(z^t(\mathbf{z})))_{t\in\mathbb{N}}$ is generated. Hence, starting from the one-period payoff functions, we define the intertemporal payoff function of player i over histories as the summation of the discounted accrued one-period payoffs. We endow every individual with a discount factor $\delta_i \in [0, 1)$ representing his intertemporal preferences. We define the payoff function on terminal histories as

$$\begin{aligned} \overline{u}_i : \quad & \mathbf{Z} \to \mathbb{R} \\ & \mathbf{z} \mapsto \overline{u}_i(\mathbf{z}) := \sum_{t=1}^{\infty} \delta_i^{t-1} u_i(z^t(\mathbf{z})). \end{aligned}$$

Hence, we can also define the normal-form payoff function of i:

$$\begin{aligned} \mathbf{U}_i : \quad & \mathbf{S} \to \mathbb{R} \\ & \mathbf{s} \mapsto \overline{u}_i(\boldsymbol{\zeta}(\mathbf{s})). \end{aligned}$$

When players are **impatient**, meaning they have zero discount factor, no payoff can be attached to infinite terminal histories. As we will see, to cope with this issue and generalize over any possible $\delta_i \in [0, 1)$, we directly rely on a form of **sequential rationality** based on **continuation values**, that is, we require players to take, at every personal history, choices that maximize the discounted expected utility computed *at that point in time*. As we show in Appendix A, the behavioral implications of this computation are equivalent to the ones of a computation from an *ex ante perspective* whenever the discount factor is strictly positive.

We have given all the elements to construct the infinite repetition of the multistage game Γ, where the informational structure is given by $F = (F_i)_{i\in I}$, and players have intertemporal preferences represented by $\delta = (\delta_i)_{i\in I}$:

$$\Upsilon(\Gamma, \delta) = \langle I, (A_i, M_i, \mathcal{A}_i, u_i, F_i, \delta_i)_{i\in I}\rangle.$$

4. Rationality and sophisticated reasoning

We begin this section defining continuation values, that is, current discounted expected utilities computed at any personal history, on the basis of the beliefs about the continuation of the game. We define these objects both for the one-period game and for the infinitely repeated interaction. Such continuation values drive the decisions of the players, and thus we will use them to give our representation of Rationality. We connect rationality and one-period rationality, and we characterize and connect the behavioral and first-order belief implications of rationality and common strong belief in rationality (RCSBR) and one-period rationality and common strong belief in rationality.

Definition 5: *Take a player $i \in I$ and a personal history $\mathbf{h}_i \in \mathbf{H}_i$. The continuation value of superstrategy $\mathbf{s}_i \in \mathbf{S}_i$ at $\mathbf{h}_i$, given CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, is the expected value at $\mathbf{h}_i$ of future payoffs, assuming that $\mathbf{s}_i$ is played and μ^i is believed from there onward, i.e.*

$$V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i) = \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)),$$

where $\mathbf{s}_i|\mathbf{h}_i$ is the superstrategy allowing $\mathbf{h}_i$ and playing like $\mathbf{s}_i$ at each personal history that does not (strictly) precede $\mathbf{h}_i$. Similarly, taking a one-period personal history $h_i \in H_i$, the continuation value of strategy s_i at h_i, given one-period CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, is

$$V^{\gamma^i}_{i,h_i}(s_i) = \sum_{s_{-i}\in S_{-i}(h_i)} U_i((s_i|h_i, s_{-i}))\gamma^i(s_{-i}|S_{-i}(h_i)).$$

Continuation values have well known continuity properties (see Appendix A), which prove useful to define rationality and carry out the epistemic analysis of sophisticated strategic reasoning.

4.1 Rational planning

Player i is rational if she play a strategy that satisfies one-step optimality given her CPS. This definition of rationality can be seen as a generalization of folding-back optimality to the infinite horizon case.

Definition 6: *We say that a superstrategy $\mathbf{s}_i$ is one-step optimal in the repeated game given a CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$—written $\mathbf{s}_i \in \mathcal{BR}_i\left(\mu^i\right)$—if, for all $\mathbf{h}_i \in \mathbf{H}_i$,*

$$\mathbf{s}_i(\mathbf{h}_i) \in \arg\max_{a_i\in\mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i|_{\mathbf{h}_i}a_i),$$

where $\mathbf{s}_i|_{\mathbf{h}_i}a_i$ is the superstrategy that allows $\mathbf{h}_i$, plays a_i at $\mathbf{h}_i$ and behaves like $\mathbf{s}_i$ at any other personal history that does not precede $\mathbf{h}_i$. Similarly, a strategy s_i is one-step optimal in the one-period game given a one-period CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$—written $s_i \in \mathcal{OR}_i(\gamma^i)$—if, for all $h_i \in H_i$,

$$s_i(h_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(h_i)+1}(h_i)} V_{i,h_i}^{\gamma^i}(s_i|_{h_i}a_i).$$

If players only care about the present, it follows intuitively that, in every one-period game, they should act so as to maximize their current one-period expected utility. Proposition 1 formalizes this fact.

Proposition 1: *When player i is impatient, a superstrategy $\bar{\mathbf{s}}_i$ is one-step optimal given CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ if and only if, for every period t and path $z^{[t-1]}$, the strategy induced in the corresponding one-period game is one-step optimal given the induced one-period CPS.*

The following remark clarifies that a repeated-game CPS induces in each period a one-period CPS.

Remark 2: *Fix a CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, a period $t \in \mathbb{N}$, and a path $z^{[t-1]}$. Then, for each personal history of the form $\mathbf{h}_i = (\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)$ with $h_i \in H_i$, the marginal of on the co-players' strategies played in the t-th repetition of the one-period game is a CPS of the one-period game.*

In appendix A we show the following existence result.

Corollary of Proposition A.1 and Proposition A.2: *For every player and every possible (one-period) CPS, there always exists a one-step optimal superstrategy (strategy).*

4.2 Strategic thinking and strong rationalizability

Battigalli & Tebaldi (2019, BT) extend the analysis of rationality and common strong belief in rationality of Battigalli & Siniscalchi (2002) to a class of infinite sequential games, which includes the infinite repetition of finite one-period games (simultaneous or sequential). To provide perspective for our results, it is useful to relate to their work. Events about behavior and interactive strategic thinking can be defined within the canonical type structure $\left(\beta_i : \mathbf{T}_i \to \Delta^{\mathbf{H}_i}(\mathbf{S}_{-i} \times \mathbf{T}_{-i})\right)_{i\in I}$ based on the given multistage game, in our case, the infinitely repeated game: $\mathbf{T}_i$ is the space of epistemic types of player i, that is, infinite hierarchies of conditional probability systems based on the countable collection $\{\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}\}_{\mathbf{h}_i \in \mathbf{H}_i}$ of conditioning events corresponding to personal histories; $\beta_i(\mathbf{t}_i) = (\beta_{i,\mathbf{h}_i}(t_i))_{\mathbf{h}_i \in \mathbf{H}_i}$ (with $\beta_{i,\mathbf{h}_i}(t_i) \in \Delta(\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i})$ for each $\mathbf{h}_i \in \mathbf{H}_i$) is the CPS over superstrategies and types of the co-players associated with type (infinite hierarchy) $\mathbf{t}_i$, and function β_i is a homeomorphism.[11] With this,

- an event about player i is a measurable subset of $\mathbf{S}_i \times \mathbf{T}_i$;
- R_i is event "i is rational," that is, the set of i-states $(\mathbf{s}_i, \mathbf{t}_i)$ such that $\mathbf{s}_i$ is one-step/sequentially optimal given the first-order CPS in hierarchy $\mathbf{t}_i$, which is obtained from the marginal of each

[11] What really matters is that the type structure à la Battigalli & Siniscalchi features continuous and onto belief maps.

conditional belief $\beta_{i,\mathbf{h}_i}(\mathbf{t}_i)$ on $\mathbf{S}_{-i}(\mathbf{h}_i)$;

- $\mathrm{SB}_i(E_{-i})$ is the event that i **strongly believes** E_{-i}, that is, CPS $\beta_i(\mathbf{t}_i)$ assigns probability 1 to E_{-i} whenever $E_{-i} \cap (\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}) \neq \emptyset$;
- $R_i^{m+1} = R_i^m \cap \mathrm{SB}_i\left(R_{-i}^{m-1}\right)$, with $R_i^1 = R_i$; for example, R_i^2 is the event that i is rational and strongly believes in the co-players' rationality;
- **rationality and common strong belief in rationality** (RCSBR) is event $\times_{i\in I} R_i^\infty = \times_{i\in I} \cap_{m=1}^\infty R_i^m$;
- finally note that $\cap_{m=1}^\infty \mathrm{SB}_i\left(R_{-i}^{m-1}\right) = \mathrm{SB}_i\left(R_{-i}^\infty\right)$; thus, $(\mathbf{s}_i, \mathbf{t}_i) \in R_i^\infty$ implies that $\beta_{i,\mathbf{h}_i}(\mathbf{t}_i)$ assigns probability 1 to R_{-i}^∞ whenever $R_{-i} \cap (\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}) \neq \emptyset$.

Of course, a similar analysis applies to all finite games (see Battigalli & Siniscalchi 2002), including the one-period games considered here. We are interested in the implications of RCSBR for strategic behavior and beliefs about co-players' behavior (first-order beliefs). Building on BT and adapting their results, one can show that such implications are characterized by the strong rationalizability solution concept defined below.

Definition 7: *For every player $i \in I$, let $\mathbf{\Sigma}_i^0 = \mathbf{S}_i$, $\Sigma_i^0 = S_i$,*

$$\mathbf{\Sigma}_i^1 = \{(\mathbf{s}_i, \mu^i) \in \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) : \mathbf{s}_i \in \mathcal{BR}_i(\mu^i)\},$$

$$\Sigma_i^1 = \{(s_i, \gamma^i) \in S_i \times \Delta^{\mathcal{C}_i}(S_{-i}) : s_i \in \mathcal{OR}_i(\gamma^i)\},$$

and recursively define, for each $k \in \mathbb{N}$,

$$\mathbf{\Sigma}_i^{k+1} = \{(\mathbf{s}_i, \mu^i) \in \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) : \mathbf{s}_i \in \mathcal{BR}_i(\mu^i),\ \forall m \leq k,$$

$$\forall \mathbf{h}_i \in \mathbf{H}_i,\ proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^m \cap \mathbf{S}_{-i}(\mathbf{h}_i) \neq \emptyset \Rightarrow \mu^i(proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^m | \mathbf{S}_{-i}(\mathbf{h}_i)) = 1\},$$

and

$$\Sigma_i^{k+1} = \{(s_i, \gamma^i) \in S_i \times \Delta^{\mathcal{C}_i}(S_{-i}) : s_i \in \mathcal{OR}_i(\gamma^i), \forall m \leq k,$$

$$\forall h_i \in H_i, proj_{S_{-i}}\Sigma_{-i}^m \cap S_{-i}(h_i) \neq \emptyset \Rightarrow \gamma^i(proj_{S_{-i}}\Sigma_{-i}^m | S_{-i}(h_i)) = 1\},$$

where $\mathbf{\Sigma}_{-i}^m = \prod_{j\neq i} \mathbf{\Sigma}_j^m$ and $\Sigma_{-i}^m = \prod_{j\neq i} \Sigma_j^m$. Then let $\mathbf{\Sigma}_i^\infty = \cap_{k\in\mathbb{N}} \mathbf{\Sigma}_i^k$ and $\Sigma_i^\infty = \cap_{k\in\mathbb{N}} \Sigma_i^k$. We say that superstrategy $\mathbf{s}_i$ (strategy s_i) is k-strongly rationalizable if $\mathbf{s}_i \in proj_{\mathbf{S}_i}\mathbf{\Sigma}_i^k$ ($s_i \in \Sigma_i^k$), and that belief μ^i (one-period belief γ^i) is k-strongly rationalizable if $\mu^i \in proj_{\Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})}\mathbf{\Sigma}_i^k$ ($\gamma^i \in proj_{\Delta^{\mathcal{C}_i}(S_{-i})}\Sigma_i^k$). If Σ_i^k and $\mathbf{\Sigma}_i^k$ are substituted with Σ_i^∞ and $\mathbf{\Sigma}_i^\infty$ in the definitions above, we say that the belonging objects are strongly rationalizable.

We can use this characterization result to study the implications of RCSBR in the case of impatient players. The intuition is that the behavior of players satisfies one-period strong rationalizability at every history, because rationality for impatient players is equivalent to one-period rationality, and thus players at the beginning of the game should expect with probability 1 the behavioral implications of one-period rationality and common strong belief in rationality. As long as players carry out strongly rationalizable strategies, common strong belief in rationality implies that they keep assigning probability 1 to the strongly rationalizable strategies of the co-players even if they observe personal histories to which their earlier beliefs assigned probability 0. In other words, players should expect to observe, in the first period, a message consistent with one-period strong rationalizability. Then, as this expectation is confirmed in every period, that is it is not contradicted by the play, it remains the highest possible level of sophistication that can be ascribed to others, and—by the *best rationalization principle* embedded in RCSBR—players continue to believe in one-period strong rationalizability in the following periods. Theorem 1 formalizes this intuition.

Theorem 1: *When players are impatient, rationality and common strong belief in rationality induce strongly rationalizable strategies and one-period CPSs in every period.*

Evidently, one-period RCSBR is not implied after deviations from RCBSR, which are rationalized ascribing to co-players lower levels of sophisticated reasoning.

5. Learning

We now focus on convergence of beliefs. In particular, we start characterizing profiles of superstrategies and CPSs where beliefs have already converged. For this section, let $((\mathbf{s}_i, \mu^i))_{i\in I}$ be the *true state*, that is, the profile of superstrategies played and beliefs actually held by players–equivalently, the true state indicates a profile of superstrategies and types $(\mathbf{s}_i, \mathbf{t}_i)_{i\in I}$ in the canonical type structure. To simplify notation, for every $t \in \mathbb{N}$, we denote with $\mathbf{h}_i^t$ the personal history of i at the beginning of period t induced by the true state, i.e. such that $\mathbf{h}_i^t \prec \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i}))$ and $\ell(\mathbf{h}_i^t) = L(t-1)$.

Definition 8: *A CPS on superstrategies for i, $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, has converged from period T, if for every $t, k \geq T$,*

$$\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = \mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^k)).$$

Suppose, without loss of generality, that $k > t$. For every $E_{-i} \subseteq \mathbf{S}_{-i}(\mathbf{h}_i^k)$, the chain rule implies

$$\mu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = \mu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i^k)) \cdot \mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^k)|\mathbf{S}_{-i}(\mathbf{h}_i^t)).$$

Hence, convergence requires that $\mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^k)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1$. Since this must hold for every $k, t \geq T$, we get the following characterization.

Remark 3: *μ^i has converged from period T if and only if the belief conditional on the observed personal history at T, $\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^T))$, assigns probability one to the set of opponents' superstrategies observationally equivalent, given i's own superstrategy, to the true ones.*

Hence, belief convergence completely characterizes learning, intended as the ability to perfectly forecast the future messages one will observe. Intuitively, a player that is certain, and correct, about the message she will observe at every period, has no reason to change her beliefs about others' behavior. Now, we are interested in understanding when learning takes place.

Comment: Since beliefs are updated according to the chain rule, a sufficient condition for asymptotic convergence is the well known requirement that beliefs assign, at a certain point (finite history) during the game, positive probability to the "true state of the world". In our setting, the "true state of the world" on which a player gradually gains information is the set of opponents' superstrategies observationally equivalent to the true ones, given the player's feedback and her own superstrategy. We call this property "observational grain of truth". Assigning positive probability to the true sequence of messages one's is going to receive may be seen as a strong requirement, or as a weak one, depending on the specific case. Our results are clearly similar to the ones of Kalai & Lehrer (1993), extending them to the case of imperfect monitoring and multistage one-period games. Indeed, restricting beliefs to the σ-algebra generated by the collection $(\mathbf{S}_{-i}(\mathbf{g}_i))_{\mathbf{h}_i \preceq \mathbf{g}_i \preceq \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))}$, whenever player i, at some personal history $\mathbf{h}_i$, assigns positive probability to $\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s})))$, then i's belief is absolutely continuous with respect to the objective distribution (the objective distribution is deterministic). Then, as Kalai and Lehrer have shown, beliefs strongly converge to the objective distribution, asymptotically. In the short/medium run, for every $\varepsilon > 0$, there exists a time starting from which beliefs are "ε-close" to the objective distribution. In our case, "ε-closeness" means that the belief assigns probability at least $1-\varepsilon$ to the set of superstrategies observationally equivalent to the true ones. We will thus show that, by finiteness of the strategy space, from a certain period onward these "ε-close" beliefs cause impatient players to play "ε-versions" of the one-period equilibrium concept attained in the long run, that is, induced strategies are part of such equilibrium, while induced one-period CPSs are ε-close to the corresponding equilibrium ones. Our setting is much less convoluted than the one of Kalai and Lehrer, due mainly to the fact that we do not allow players to adopt behavioral superstrategies (nor strategies). Hence, we can provide a very simple proof of learning.

Definition 9: *Fix a player i, a profile of superstrategies $(\mathbf{s}_i, \mathbf{s}_{-i})$, and a CPS μ^i; we say that μ^i contains an **"observational grain of truth"** given $(\mathbf{s}_i, \mathbf{s}_{-i})$ if there exists a $T \in \mathbb{N}$ such that*

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)) > 0.$$

Observe that, while in our definitions we focus for simplicity on the beliefs held at the beginning of periods, both Definition 8 and Definition 9 can be given equivalently in terms of personal histories, induced by the true state, of general length.

Proposition 2: *If CPS μ^i contains an* "observational grain of truth" *given $(\mathbf{s}_i, \mathbf{s}_{-i})$, then*

$$lim_{t\to\infty}\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1.$$

Thus, if a player's belief contains an observational grain of truth, the player will asymptotically learn to perfectly forecast the future messages she will observe. In the short/medium run, learning has "ε-closeness" implications.

Corollary of Proposition 2: *If CPS μ^i contains an* "observational grain of truth" *given $(\mathbf{s}_i, \mathbf{s}_{-i})$,,* *then, for every $\varepsilon > 0$, there exists a time T such that, for all $t \geq T$*

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1 - \varepsilon.$$

Obviously, observational grain of truth in the present context is a characterization of learning, as its necessity is immediate.

6. Strong rationalizability, learning, and equilibrium

Let us start this section by defining the concepts of **SCE** and **SCE with strongly rationalizable beliefs**, both for the one-period game and its infinite repetition. In the latter case, it is useful to define also a variation of these two concepts, which we label as "*eventual*." This simply requires the characterizing conditions to hold only from a certain period on. Subsequently, we remark that ε-confirmation of a CPS implies ε-confirmation of the induced one-period CPSs which, among other things, helps us connect the concepts of SCE and one-period SCE (and refinements).

Definition 10: *A one-period SCE is a profile of strategies and CPSs pairs $((s_i, \gamma^i))_{i \in I} \in \prod_{i \in I} S_i \times \Delta^{\mathcal{C}_i}(S_{-i})$ such that, for every i, γ^i is confirmed by s and s_i one-step optimal given γ^i, i.e.:*

(i) (confirmation of beliefs) $\gamma^i \left(S_{-i}(o_i^{-1}(\zeta(s)))|S_{-i}\right) = 1$;

(ii) (rationality) $s_i \in \mathcal{OR}_i(\gamma^i)$.

$((s_i, \gamma^i))_{i \in I} \in \prod_{i \in I}$ is a one-period SCE with strongly rationalizable beliefs if, for every i, γ^i is confirmed by s and (s_i, γ^i) are strongly rationalizable, i.e.:

(i) $\gamma^i \left(S_{-i}(o_i^{-1}(\zeta(s)))|S_{-i}\right) = 1$;

(ii) $(s_i, \gamma^i) \in \Sigma_i^\infty$.

Definition 11: *An SCE is a profile of superstrategies and CPSs pairs $((\mathbf{s}_i, \mu^i))_{i \in I} \in \prod_{i \in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ such that, for every i:*

(i) (confirmation of beliefs) $\mu^i \left(\mathbf{S}_{-i}(\mathbf{o}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}\right) = 1$;

(ii) (rationality) $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$.

A profile $((\mathbf{s}_i, \mu^i))_{i \in I} \in \prod_{i \in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ is an SCE with strongly rationalizable beliefs if it is an SCE of the infinitely repeated game $\Upsilon(\Gamma, \delta)$ such that, for every $i \in I$, $(\mathbf{s}_i, \mu^i) \in \boldsymbol{\Sigma}_i^\infty$.

As anticipated, in case condition (i) of Definition 11 is substituted by "(i') there exists some $T \in \mathbb{N}$ such that $\mu^i \left(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)\right) = 1$", we call the corresponding profile $((\mathbf{s}_i, \mu^i))_{i \in I}$ an *eventual* SCE (with strongly rationalizable beliefs).

Remark 4: *For every player $i \in I$ and CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, for every $\varepsilon \geq 0$ and $T \in \mathbb{N}$ such that, for all $t \geq T$,*

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1 - \varepsilon,$$

all one-period CPSs $(\mu_t^i)_{t \geq T}$, induced by μ^i at every period $t \geq T$ starting at the true personal history $\mathbf{h}_i^t$, satisfy

$$\mu_t^i(S_{-i}\left(o_i^{-1}(\zeta(s^t))\right)|S_{-i}) \geq 1 - \varepsilon.$$

Remark 4 allows us to draw many of the conclusions listed in the remainder of this section. To start, we use it to derive the implications of confirmation of beliefs, about the infinite interaction, on confirmation of beliefs about the one-period games induced by the true state, when players are impatient.

Corollary of Proposition 1 and Remark 4: *If players are impatient, every Self-Confirming Equilibrium of the infinite repetition induces a sequence of one-period Self-Confirming Equilibria.*

Corollary of Theorem 1 and Remark 4: *If players are impatient, every Self-Confirming Equilibrium with strongly rationalizable beliefs induces a sequence of one-period Self-Confirming Equilibria with strongly rationalizable beliefs.*

6.1 Equilibrium implications of learning and RCSBR "after" beliefs convergence

In this subsection, we explicit the implications on plays that we are able to derive from the analysis up to this point, in terms of the equilibrium concepts previously defined, considering the case in which players' beliefs have already converged.

As shown in Remark 3, convergence of beliefs is equivalent to learning the exact sequence of future messages one will observe, which is in turn equivalent to confirmation of beliefs of the individual. Think of a **state of the game**, either as a profile of superstrategis and epistemic types $(\mathbf{s}_i, \mathbf{t}_i)_{i \in I}$ in the canonical type structure, or—equivalently for our purposes—a "first-order state" comprising a profile $\left(\mathbf{s}_i, \mu^i\right)_{i \in I}$ of superstrategies and beliefs (CPSs) about the co-players' superstrategies. With this, from the aforementioned equivalence, we can make sense of the following statements.

Remark 5: *A profile $((\mathbf{s}_i, \mu^i))_{i \in I}$ is an eventual Self-Confirming Equilibrium if and only if there exists a period T starting from which all players beliefs have converged (and players are rational).*

Corollary of Remark 4 and Remark 5: *If players satisfy the assumptions of rationality and common strong belief in rationality, and their beliefs converge, then the true state of the game must feature an eventual Self-Confirming Equilibrium with strongly rationalizable beliefs.*

Corollary of Remark 3, Remark 4, Remark 5 and Theorem 1: *If players are impatient, satisfy the assumptions of rationality and common strong belief in rationality, and their beliefs converge from period T, then the true state of the game must induce, from period T onward, a sequence of one-period Self-Confirming Equilibria with strongly rationalizable beliefs.*

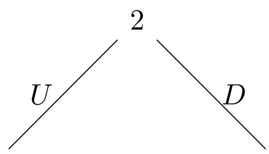
So far, we have fully characterize learning with eventual SCE (Remark 5). Thanks to the following remark, we can provide an equivalent characterization with one-period SCEs, whenever individuals are impatient.

Remark 6: *If at a state, from a certain time T onward (that is, for every $t \geq T$), all induced one-period CPSs of player i are confirmed, then her CPS on superstrategies μ^i has converged from T.*

Remark 6 is, roughly speaking, the inverse of a part of Remark 4.

Corollary of Remark 4, Remark 6, and Proposition 1: *When players are impatient, a profile $((\mathbf{s}_i, \mu^i))_{i\in I}$ induces, starting from some period T, a sequence of one-period Self-Confirming Equilibria, if and only if there exists a period T starting from which all players beliefs have converged and players are rational.*

Example: We report the game graphic representation:



1 \ 2	ℓ	c	r
u	**1**, *1*	**0**, *2*	**2**, *0*
m	**1**, *2*	**2**, *0*	**0**, *1*
d	**0**, *1*	**2**, *0*	**0**, *1*

1 \ 2	ℓ	c	r
u	**2**, *0*	**1**, *1*	**3**, *-1*
m	**1**, *3*	**2**, *2*	**0**, *2*
d	**-1**, *3*	**1**, *1*	**-1**, *3*

Given the informational structure we defined, the only strategy profiles part of an RSCE are

$$\cup_{\mathbf{x}\in\{\mathbf{u},\mathbf{m}\}}\{(\mathbf{m}.\mathbf{x}, U.\ell.c)\},$$

and thus the only terminal history consistent with an RSCE is $(U, (\mathbf{m}, \ell))$, which is the SPE outcome. The strategy profiles part of an SCE with strongly rationalizable beliefs instead are, besides the one above,

$$\cup_{\mathbf{x}\in\{\mathbf{u},\mathbf{m}\}}\{(\mathbf{m}.\mathbf{x}, U.c.c), (\mathbf{u}.\mathbf{x}, U.\ell.c)\},$$

$$\cup_{(\mathbf{x},x)\in\{\mathbf{u},\mathbf{m},\mathbf{d}\}\times\{\ell,c,r\}}\{(\mathbf{x}.\mathbf{m}, D.x.\ell), (\mathbf{x}.\mathbf{u}, D.x.\ell)\},$$

and

$$\cup_{(\mathbf{x},x)\in\{\mathbf{u},\mathbf{m},\mathbf{d}\}\times\{\ell,r\}} \in \{(\mathbf{x}.\mathbf{m}, D.x.c)\},$$

which implies that all terminal histories in

$$\cup_{X\in\{U,D\}}\{(X, (\mathbf{u}, \ell)), (X, (\mathbf{m}, \ell)), (X, (\mathbf{m}, c))\}$$

can be induced by one-period SCEs with strongly rationalizable beliefs. Moreover, the game has an SCE outcome in $(D, (\mathbf{u}, r))$.

Assume individuals are impatient. Instead of completely defining a superstrategy, we focus on paths, that is sequences of actions, induced by a superstrategy. For any fixed path, there is an infinite number of superstrategies inducing it. We say that a path is optimal given a certain "partially defined" CPS (specified only at the personal histories allowed by the path), in the sense of one-step optimality. If a path is optimal, then there exists, as shown in the proof of Proposition A.2, a superstrategy, inducing that path, that is one-step optimal given a CPS whose projection is the above mentioned "partially defined" CPS. Hence, if a certain path is consistent with strong rationalizability then there exists a strongly rationalizable strategy inducing it. To assess whether a path is consistent with strong rationalizability, we take advantage of the "circular property" of strong rationalizablity, derived by the underlying algorithmic procedure: if some path $(a^1, a^2, a^3, \ldots)$ is one-step optimal given a CPS strongly believing in (some generic strategies inducing) a path that, itself, is one-step optimal given a CPS strongly believing in a path that, itself,... in a path that, itself, is one-step optimal given a CPS strongly believing in $(a^1, a^2, a^3, \ldots)$, then all paths listed are consistent with strong rationalizability.

Let $\mathbf{s}^*_\mathbf{1}$ be the superstrategy played by player **1**, and suppose that player **1**'s CPS on *2*'s superstrategies is $\mu^\mathbf{1}$ such that

$$\mu^\mathbf{1}(\mathbf{S}_2(\boldsymbol{\zeta}(\mathbf{s}^*_\mathbf{1}, \mathbf{s}_2))|\mathbf{S}_2(\varnothing)) = \begin{cases} \frac{1}{2} & proj_{A_2^{\mathbb{N}}}\boldsymbol{\zeta}(\mathbf{s}^*_\mathbf{1}, \mathbf{s}_2) = (U, \ell, U, \ell, U, \ell, \ldots) \\ \frac{1}{2^{2t+1}} & proj_{A_2^{\mathbb{N}}}\boldsymbol{\zeta}(\mathbf{s}^*_\mathbf{1}, \mathbf{s}_2) \in \{(U, c), (U, r)\}^t \times \{(U, \ell)\} \times \ldots \end{cases} .$$

Then, it can be seen that $\mathbf{s}^*_\mathbf{1}$ taken such that $\mathbf{s}^*_\mathbf{1}(\mathbf{h_1}) = \mathbf{u}$, for every $\mathbf{h_1}$ allowed by $\mathbf{s}^*_\mathbf{1}$ and any superstrategy in the support of $\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\varnothing))$, at which **1** is active, is one-step optimal given $\mu^\mathbf{1}$. Indeed, at every such personal history,

$$marg_{A_2}\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\mathbf{h_1}))(\ell) = \frac{1}{2}, \qquad marg_{A_2}\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\mathbf{h_1}))(c) = marg_{A_2}\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\mathbf{h_1}))(r) = \frac{1}{4},$$

and thus $\mathbf{u}$ is always optimal. Now we check that the paths in the "support" of $\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\varnothing))$ are consistent with strong rationalizability. There exists a superstrategy $\mathbf{s}_2$ inducing $(U, \ell, U, \ell, U, \ell, \ldots)$ that is one-step optimal given a CPS strongly believing $\mathbf{s_1}$ inducing $(\mathbf{m}, \mathbf{m}, \mathbf{m}, \ldots)$ in the subgame after U, which itself satisfies one-step optimality given a CPS strongly believing in $\mathbf{s}_2$ inducing $(U, \ell, U, \ell, U, \ell, \ldots)$. Thus these paths are consistent with strong rationalizability. As for the superstrategies $\mathbf{s}_2$'s inducing paths in $\{(U, c), (U, r)\}^t \times \{(U, \ell)\} \times \ldots$, the ones involving an initial chain of (U, c)'s followed by (U, ℓ)'s satisfy one-step optimality given a CPS strongly believing in an initial chain of $\mathbf{u}$'s followed by $\mathbf{m}$'s in the subgame at U. Then this latter path satisfies one-step optimality given a CPS strongly believing in $(U, \ell, U, \ell, U, \ell, \ldots)$. Since this last path has already been shown to be consistent with strong rationalizability, so are the two preceding paths. With regards to the paths inducing an initial chain of (U, r)'s followed by (U, ℓ)'s, they are justified by an initial chain of $\mathbf{d}$'s followed by $\mathbf{m}$'s, which is justified by an initial chain of (U, c)'s followed by (U, ℓ)'s. Since this latter is consistent with strong rationalizability, so are the preceding two paths. As for superstrategies $\mathbf{s}_2$'s that induce an initial mixed chain of (U, c)'s and (U, r)'s, they can be seen to be consistent with strong rationalizability with an analogous iterated procedure of justification. Letting the support of $\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\varnothing))$ be included in the set of strongly rationalizable superstrategies, $\mathbf{s}^*_\mathbf{1}$ can be built to be strongly rationalizable (while inducing path $(\mathbf{u}, \mathbf{u}, \ldots)$).

Analogously, let $\mathbf{s}^*_2$ be the superstrategy played by player *2*. Suppose that the CPS over **1**'s superstrategies held by player *2*, μ^2, is such that

$$\mu^2 \left(\{\mathbf{s_1} : proj_{A_\mathbf{1}^{\mathbb{N}}} \boldsymbol{\zeta}(\mathbf{s_1}, \mathbf{s}_2^*) = ((\mathbf{u}^{[n]}), (\mathbf{m}, \mathbf{m}, \ldots)\} | \mathbf{S_1}(\varnothing) \right) = 1,$$

i.e., it assigns, at the beginning of the game, probability 1 to superstrategies inducing a path of an initial chain of $\mathbf{u}$'s of length $n \in \mathbb{N}$, followed by $\mathbf{m}$'s. We have already seen that such path is consistent with strong rationalizability. Hence, $\mathbf{s}_2^*$ can be made strongly rationalizable, and in particular such that, at every personal history allowed by $\mathbf{s}_2^*$ and any superstrategy in the support of μ^2, $\mathbf{s}_2^*$ plays (U, c) in all periods until n, and (U, ℓ) after. Profile $(\mathbf{s}_2^*, \mu^2)$, may be interpreted as a belief of player *2* that player **1** starts the game playing $\mathbf{u}$, and upon repeatedly observing $f_\mathbf{1}^2((U, (\mathbf{u}, c)))$, will eventually be convinced (or "believe to have learned") that *2* is playing the constant superstrategy c. Hence **1** will deviate to $\mathbf{m}$, knowing that player *2* can not observe such deviation when playing (U, c) (nor (U, ℓ)). In particular, *2* believes that **1** will need exactly n observations to change his behavior.

In conclusion, a state of the world satisfying rationality and common strong belief in rationality can be constituted by such $\mathbf{s}_\mathbf{1}^*$, $\mu^\mathbf{1}$, $\mathbf{s}_2^*$, and μ^2. Notice that both $\mu^\mathbf{1}(\cdot|\mathbf{S}_2(\varnothing))$ and $\mu^2(\cdot|\mathbf{S_1}(\varnothing))$ contain a grain of truth. In particular, μ^2 has already converged, since

$$\mathbf{O}_2^{-1}(\boldsymbol{\zeta}(\mathbf{s}_\mathbf{1}^*, \mathbf{s}_2^*)) = \Big(((U,c)^{[n]}, (U,\ell)^{[\infty]}), ([(f_2^1(U), f_2^2((U,\mathbf{u},c)))]^{[n]}, [(f_2^1(U), f_2^2((U,\mathbf{u},\ell)))]^{[\infty]}) \Big) =$$

$$= \Big(((U,c)^{[n]}, (U,\ell)^{[\infty]}), ([(f_2^1(U), f_2^2((U,\mathbf{u},c)))]^{[n]}, [(f_2^1(U), f_2^2((U,\mathbf{m},\ell)))]^{[\infty]}) \Big) = \mathbf{O}_2^{-1}(\boldsymbol{\zeta}(\mathbf{s_1}, \mathbf{s}_2^*))$$

for every $\mathbf{s}_1 \in supp\mu^2(\cdot|\mathbf{S_1}(\varnothing))$. With regards to player **1**,

$$\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{O}_\mathbf{1}^{-1}(\boldsymbol{\zeta}(\mathbf{s}_\mathbf{1}^*, \mathbf{s}_2^*))|\mathbf{S}_2(\varnothing)) = \mu^\mathbf{1}(\{\mathbf{s}_2 : proj_{A_2^{\mathbb{N}}} \boldsymbol{\zeta}(\mathbf{s_1}, \mathbf{s}_2) = (((U,c)^{[n]}), (U,\ell,U,\ell,\ldots)\}|\mathbf{S}_2(\varnothing))$$

$$= \frac{1}{2^{2n+1}} > 0.$$

Let $\mathbf{h}_\mathbf{1}^t$ be the personal history of player **1** of length $2t$ preceding $\mathbf{O}_\mathbf{1}^{-1}(\boldsymbol{\zeta}(\mathbf{s}^*))$. Notice that

$$\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{h}_\mathbf{1}^t)|\mathbf{S}_2(\varnothing)) = \mu^\mathbf{1}(\{\mathbf{s_1} : proj_{A_2^{\mathbb{N}}} \boldsymbol{\zeta}(\mathbf{s_1}, \mathbf{s}_2^*) \in \{(U,c)\}^t \times A_2 \times \ldots\}|\varnothing) =$$

$$= \sum_{k=t}^{\infty} 2^{k-t} \frac{1}{2^{2k+1}} = \frac{1}{2^{t+1}} \sum_{k=t}^{\infty} (2\frac{1}{2^2})^k = \frac{1}{2^{2t+1}} \sum_{k=0}^{\infty} (\frac{1}{2})^k = \frac{1}{2^{2t}}$$

for every $t \leq n$, and thus

$$\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{O}_\mathbf{1}^{-1}(\boldsymbol{\zeta}(\mathbf{s}_\mathbf{1}^*, \mathbf{s}_2^*))|\mathbf{S}_2(\mathbf{h}_\mathbf{1}^t)) = \frac{\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{O}_\mathbf{1}^{-1}(\boldsymbol{\zeta}(\mathbf{s}_\mathbf{1}^*, \mathbf{s}_2^*))|\mathbf{S}_2(\varnothing))}{\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{h}_\mathbf{1}^t)|\mathbf{S}_2(\varnothing))} = \frac{1}{2^{2(n-t)+1}}.$$

Clearly, upon observing the first ℓ, i.e. from $\mathbf{h}_\mathbf{1}^{n+1}$ onward, for every $t \geq n+1$,

$$\mu^\mathbf{1}(\mathbf{S}_2(\mathbf{h}_\mathbf{1}^{-1}(\boldsymbol{\zeta}(\mathbf{s}_\mathbf{1}^*, \mathbf{s}_2^*))|\mathbf{S}_2(\mathbf{h}_\mathbf{1}^t)) = 1,$$

and complete convergence is achieved in discrete time. Hence, the grain of truth leads to learning and, in the end, the play converges to the one-period SCE with rationalizable conjectures $(U, (\mathbf{u}, \ell))$.

Alternation between different one-period SCEs with strongly rationalizable beliefs can clearly be achieved. For simplicity, we know assume that players' beliefs have converged from the beginning. In particular, player **1** strongly believes in *2*'s superstrategies that induce path $(U, c, U, \ell, U, c, U, \ell, \ldots)$, while *2*'s strongly believes in opponent's superstrategies that induce the constant path $(\mathbf{m}, \mathbf{m}, \ldots)$ in the subgame starting at U. Then, **1** adopts a superstrategy that, along path, induces $(\mathbf{m}, \mathbf{u}, \mathbf{m}, \mathbf{u}, \ldots)$ in the subgame starting at U, while *2* plays the constant path $(U.\ell, U.\ell, \ldots)$. Both beliefs are strongly rationalizable, and both are confirmed by the sequence of observed messages. ▲

Battle of the Sexes: As mentioned, the sufficient conditions for learning can be seen as very demanding. Players must assign a positive probability to the set of opponents' superstrategies observationally equivalent to the true ones. Moreover, the definition of beliefs over superstrategies, while formally making sense of individuals' sophistication, can in fact lead to very unlikely situations, as we show here with the Battle of the Sexes. Consider as one-period game the following BoS:

1 \ 2	B	S
B	2, 1	0, 0
S	0, 0	1, 2

Player 1 prefers B to S, and player 2 viceversa. Suppose there are observable actions, or equivalently $f_i = u_i$ for each $i \in \{1, 2\}$. Thus, the set of one-period Self-Confirming Equilibira $(\{(B, B), (S, S)\})$ coincides with the set of one-period SCEs with rationalizable conjectures, the set of one-period RSCEs, and the set of one-period Nash Equilibria. Clearly, there exist superstrategies $\mathbf{s}_1$ and $\mathbf{s}_2$ inducing the alternated path $((B, B), (S, S), (B, B), (S, S), \ldots)$ that are strongly rationalizable, as each of them is a sequential best replies to a CPS strongly believing in the other. Such CPS is clearly confirmed, which means that has converged. Consequently, the above path is consistent with learning and RCSBR.

Now, suppose that player 1 firmly believes that player 2 wants to "cooperate" and play the alternated sequence of equilibria. In particular, 1 is sure that 2 will start playing B. Upon observing S, she then is sure that 2 decided to start with S, but will now play B. Upon observing S again, 1 thinks that 2 was hoping that 1 would have come along, but now has understood 1's intention and will play B. Upon observing S for the fourth time, she makes the same identical reasoning. If 1 goes on with this thinking, then her CPS is such that, at any time along a sequence of (B, S)'s, 1's one-period belief assigns probability 1 to 2 playing B. This indeed is a completely acceptable feature, given the definition of CPS. As a result, 1's conjectures have converged along such path. However, they are clearly not confirmed. Observe that, if 2 follows an analogous reasoning, with S in place of B, then the actual play is exactly an infinite sequence of (B, S)'s, so that beliefs over superstrategies that have not converged, but one-period beliefs that have. ▲

6.2 Sufficient equilibrium conditions for learning and RCSBR

We have shown that playing a sequence of one-period Self-Confirming Equilibria with strongly rationalizable beliefs is a necessary implication of learning, impatience, and RCSBR. Now we show the converse idea, that is, every sequence of one-period Self-Confirming Equilibria with strongly rationalizable beliefs can be the implication of those behavioral assumptions. This fact, stated in Theorem 2, stems from the equivalence of learning and confirmations of beliefs (Remark 4, Remark 5, and Remark 6), and from the next proposition, which formally shows something that might be already intuitively understood from the examples at the end of Section 6.1.

Proposition 3: *Let players be impatient. If a terminal history coincides with a sequence of one-period terminal histories consistent with one-period rationality and common strong belief in rationality, then there exists a profile of strongly rationalizable superstrategies that induces it.*

Proposition 3 tells us that every sequence of one-period terminal histories consistent with one-period strong rationalizability is consistent with strong rationalizability (of the infinite interaction). Then this connection between histories can be translated into a connection between beliefs and between behaviors, and implies the following theorem.

Theorem 2: *Let players be impatient. Every sequence of one-period SCEs with strongly rationalizable beliefs can be induced at some state that features learning and RCSBR.*

The same sufficiency results of course holds also, obviously, for any SCE with strongly rationalizable beliefs of the infinite repetition and possibly patient players.

6.3 Equilibrium implications of learning and RCSBR "during" belief convergence

Proposition 2 provides sufficient conditions for asymptotic learning. As a consequence, if players satisfy the assumptions of RCSBR and impatience, and their beliefs contain an observational grain of truth, then asymptotically they end up playing one-period Self-Confirming Equilibria with strongly rationalizable beliefs. Moreover, Remark 4 tells us that, from certain finite periods onward, beliefs must be at least arbitrarily close to exactly forecasting the future messages. Hence, as the next proposition shows, the fact that the strategy space is finite allows to characterize the behavior of such players, after a finite number of periods, as consistent with one-period SCEs with strongly rationalizable beliefs, where their beliefs are "almost confirmed". Before stating the proposition, we define this concept reflecting an "almost" one-period SCE with strongly rationalizable beliefs.

Definition 12: *A one-period ε-Self-Confirming Equilibrium with strongly rationalizable beliefs is a profile of pairs $((s_i, \gamma^i))_{i \in I} \in \prod_{i \in I} S_i \times \Delta^{\mathcal{C}_i}(S_{-i})$ such that, for some $(\nu^i)_{i \in I} \in \prod_{i \in I} \Delta^{\mathcal{C}_i}(S_{-i})$,*

(i) *$((s_i, \nu^i))_{i \in I}$ is a one-period Self-Confirming Equilibrium with strongly rationalizable beliefs;*

(ii) *for every $i \in I$, there exists $0 < \delta \leq \varepsilon$ such that $\gamma^i(S_{-i}(o_i^{-1}(\zeta(s)))|S_{-i}) = 1-\delta$, and $(s_i, \gamma^i) \in \Sigma_i^{\infty}$.*

As anticipated, the concept of one-period ε-Self-Confirming Equilibrium with rationalizable beliefs indicates a profile of strategies that is part of a one-period Self-Confirming Equilibrium with rationalizable beliefs, and is paired with a profile of beliefs ε-close to the ones part of said equilibrium.

Proposition 4: *Fix $i \in I$, $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, $s_i \in \mathcal{OR}_i(\gamma^i)$, and $z_i \in Z_i$. Suppose that, for some $\varepsilon > 0$ arbitrarily chosen,*

$$\gamma^i(S_{-i}(z_i)|S_{-i}) \geq 1 - \varepsilon.$$

Then, there exist $\bar{\varepsilon}$ *and* $\nu^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, *where*

$$\nu^i(S_{-i}(z_i)|S_{-i}(h_i)) = 1 \quad \forall h_i \preceq z_i,$$

$$supp\nu^i(\cdot|S_{-i}(h_i)) \subseteq supp\gamma^i(\cdot|S_{-i}(h_i)) \quad \forall h_i \in H_i,$$

such that, if $\varepsilon \leq \bar{\varepsilon}$, *then* $s_i \in \mathcal{OR}_i(\nu^i)$.

Proposition 4 allows us to say that, when players are impatient, one-period ε-Self-Confirming Equilibria are repeatedly played from a certain time on, since being a one-period best reply to the ε-close belief actually held implies being a one-period best reply to the corresponding confirmed belief. Putting the pieces together, we can state the following as a corollary.

Corollary of Proposition 1 (Theorem 1), Proposition 2, Remark 4, and Proposition 4: *If the belief of every player contains an observational grain of truth, and players satisfy RCSBR and are impatient, then there exists a period* T *starting from which only one-period* ε*-Self-Confirming Equilibria with strongly rationalizable beliefs are played, and in the long run one-period beliefs converge to confirmed ones.*

Sketch of the proof: Let $\bar{\varepsilon} = min_{i\in I}\bar{\varepsilon}_i$ as defined in Proposition 4, and take any $\varepsilon < \bar{\varepsilon}$. By the corollary of Proposition 2, if all players beliefs contain a grain of truth, there exists $(T_i)_{i\in I}$ such that, for all $t \geq T_i$, the belief along path of i a time t assigns probability at least $1-\varepsilon$ to the set of superstrategies observationally equivalent, given i's own, to the true ones. Thus, by Remark 4, i's induced one-period beliefs assign probability at least $1-\varepsilon$ to strategies observationally equivalent to the truly induced ones. Then, by Proposition 1 the strategy played by i at each period after T_i satisfies is one-step optimal given such induced one-period belief. By Proposition 4, such strategy one-step optimal also given a one-period belief that is confirmed by the true play. Since, for every one-period personal history, the support of such confirmed belief is contained in the support of the truly induced belief, the confirmed belief strongly believes in everything the true one does. Thus, by Theorem 1, both one-period CPSs are strongly rationalizable. When $t \geq T = max_{i\in I}T_i$, this holds for every player. Consequently, for every such t we have a one-period ε-Self-Confirming Equilibrium with strongly rationalizable beliefs. By Proposition 2, in the long-run beliefs strongly converge to the fully confirmed ones.

We can ask ourselves what learning and RCSBR entail in the medium run when players are not impatient. Intuitively, since the space of superstrategies is not finite, what results is a notion of eventual "almost" Self Confirming Equilibrium with strongly rationalizable beliefs of the infinite interaction, where not only beliefs are "almost confirmed", but also superstrategies are "almost optimal" given the equilibrium (confirmed) beliefs. In Remark A.1 we show that optimization at a given personal history with respect to own superstrategies is equivalent to optimization with respect to own continuation superstrategies. Similarly, it should be noticed that, if one were to consider a form of optimality starting only at a certain personal history, only the beliefs held from that personal history onward would matter. Hence, we are able to express the connotation of "eventual", which requires properties to hold from a certain period onward, as requiring the properties to hold only for the continuation superstrategies, and the beliefs about them, at every personal history starting from a particular one.

Now we define these "continuation objects" of interest. Given a player i and a personal history $\mathbf{h}_i$, let

$$\mathfrak{C}_i^{\mathbf{h}_i} = \{\mathbf{S}_{-i}(\mathbf{g}_i) \subseteq \mathbf{S}_{-i} : \mathbf{g}_i \succeq \mathbf{h}_i\}$$

be the set of conditional events that are induced by the sub-tree of $\overline{\mathbf{H}}_i$ with root $\mathbf{h}_i$, and denote by $\Delta^{\mathfrak{C}_i^{\mathbf{h}_i}}(\mathbf{S}_{-i})$ the corresponding set of CPSs. Once $\mathbf{h}_i$ is reached, "optimality from there on" depends only on these "sub-CPSs". In other words, once $\mathbf{h}_i$ is reached, optimality starting at $\mathbf{h}_i$ should be intended as optimality in the "subjective sub-game with root $\mathbf{h}_i$". To substantiate Remark A.1 and the other claim made above, given a continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ and a CPS $\nu^i \in \Delta^{\mathfrak{C}_i^{\mathbf{h}_i}}(\mathbf{S}_{-i})$, the continuation values of the continuation game can be defined without changes. For every $\mathbf{g}_i \succeq \mathbf{h}_i$,

$$V_{i,\mathbf{g}_i}^{\nu^i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}) = \sum_{t=(\ell(\mathbf{g}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{g}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|\mathbf{g}_i, \mathbf{s}_{-i})))\nu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)) =$$

$$= \sum_{t=(\ell(\mathbf{g}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{g}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{g}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)) = V_{i,\mathbf{g}_i}^{\mu^i}(\mathbf{s}_i),$$

where $\boldsymbol{\zeta}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|\mathbf{g}_i, \mathbf{s}_{-i})$ is the terminal history induced by playing continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ after $\mathbf{g}_i$ and superstrategies $\mathbf{s}_{-i}$, while $\mathbf{s}_i$ and μ^i are any superstrategy and any belief such that

$$proj_{\mathbf{S}_i^{\succeq \mathbf{h}_i}} \mathbf{s}_i = \mathbf{s}_i^{\succeq \mathbf{h}_i} \quad \wedge \quad proj_{[\Delta(\mathbf{S}_{-i})]^{\mathfrak{C}_i^{\mathbf{h}_i}}} \mu^i = \nu^i.$$

For fixed superstrategy $\mathbf{s}_i$ and CPS μ^i, denote with $s_i^{\succeq \mathbf{h}_i}$ the continuation strategy induced by $\mathbf{s}_i$, and by $\mu^i_{\succeq \mathbf{h}_i}$ the projection of CPS μ^i over the set of CPSs $\Delta^{\mathfrak{C}_i^{\mathbf{h}_i}}(\mathbf{S}_{-i})$. Then, $\mathbf{s}_i$ is optimal starting at $\mathbf{h}_i$, with respect to μ^i, if $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ is one-step optimal given $\mu^i_{\succeq \mathbf{h}_i}$, that is, for every $\mathbf{g}_i \succeq \mathbf{h}_i$,

$$\mathbf{s}_i^{\succeq \mathbf{h}_i}(\mathbf{g}_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}} V_{i,\mathbf{g}_i}^{\mu^i_{\succeq \mathbf{h}_i}}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|_{\mathbf{g}_i} a_i),$$

where $V_{i,\mathbf{g}_i}^{\mu^i_{\succeq \mathbf{h}_i}}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|_{\mathbf{g}_i} a_i)$ denotes, in the usual way, the continuation value at $\mathbf{g}_i$ of playing continuation like superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ after $\mathbf{g}_i$ and a_i. With an abuse of notation, we denote such optimality property as

$$\mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i(\mu^i_{\succeq \mathbf{h}_i}).$$

An obvious remark is worth stating.

Remark 7: *If a superstrategy is one-step optimal given a certain CPS, then the superstrategy is optimal starting at any personal history given the same CPS.*

Formally, Remark 7 says that, given $\mathbf{s}_i$, $\mathbf{s}_i^{\succeq \mathbf{h}_i}$, μ^i, and $\mu^i_{\succeq \mathbf{h}_i}$ as above, for every $\mathbf{h}_i$,

$$\mathbf{s}_i \in \mathcal{BR}_i(\mu^i) \quad \Rightarrow \quad \mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i(\mu^i_{\succeq \mathbf{h}_i}).$$

Definition 13: *Fix a personal history $\mathbf{h}_i \in \mathbf{H}_i$. We say that a continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathbf{S}_i^{\succeq \mathbf{h}_i}$*

is an ϵ-best reply to a CPS $\nu^i \in \Delta^{\mathfrak{C}_i^{\succeq \mathbf{h}_i}}(\mathbf{S}_{-i})$, *where* $\epsilon > 0$, *if, for any* $\bar{\mathbf{s}}_i^{\succeq \mathbf{h}_i} \in \mathbf{S}_i^{\succeq \mathbf{h}_i}$ *such that* $\bar{\mathbf{s}}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i(\nu^i)$, *and for every* $\mathbf{g}_i \succeq \mathbf{h}_i$,

$$V_{i,\mathbf{g}_i}^{\nu^i}(\bar{\mathbf{s}}_i^{\succeq \mathbf{h}_i}) - V_{i,\mathbf{g}_i}^{\nu^i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}) \leq \epsilon,$$

written

$$\mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i^{\epsilon}(\nu^i).$$

Definition 14: *An eventual ε,ϵ-Self Confirming Equilibrium with strongly rationalizable beliefs is a state* $(\mathbf{s}_i, \mu^i)_{i \in I} \in \prod_{i \in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ *such that, for some* $\mathbf{h} \prec \boldsymbol{\zeta}(\mathbf{s})$ *and some* $(\nu^i)_{i \in I} \in \prod_{i \in I} \Delta^{\mathfrak{C}_i^{\mathbf{O}_i^{-1}(\mathbf{h})}}(\mathbf{S}_{-i})$, *for every* $i \in I$,

(i) *there exists* $\bar{\mathbf{s}}_i^{\succeq \mathbf{O}_i^{-1}(\mathbf{h})} \in \mathbf{S}_i^{\succeq \mathbf{O}_i^{-1}(\mathbf{h})}$ *such that* $(\bar{\mathbf{s}}_i^{\succeq \mathbf{O}_i^{-1}(\mathbf{h})}, \nu^i)_{i \in I}$ *is a Self-Confirming Equilibrium of the sub-game with root* $\mathbf{h}$, *i.e.,* $\bar{\mathbf{s}}_i^{\succeq \mathbf{O}_i^{-1}(\mathbf{h})} \in \mathcal{BR}_i(\nu^i)$ *and, for every* $\mathbf{O}_i^{-1}(\mathbf{h}) \preceq \mathbf{h}_i \prec \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))$,

$$\nu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}(\mathbf{h}_i)) = 1,$$

and $\mathbf{s}_i^{\succeq \mathbf{O}_i^{-1}(\mathbf{h})}$, *which is the continuation strategy induced by* $\mathbf{s}_i$, *is an ϵ-best reply to* ν^i;

(ii) *for every* $\mathbf{O}_i^{-1}(\mathbf{h}) \preceq \mathbf{h}_i \prec \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))$ *and* $E_{-i} \subseteq \mathbf{S}_{-i}$,

$$\nu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) > 0 \Rightarrow \frac{\mu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))}{\nu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))} \geq 1 - \varepsilon,$$

and $(\mathbf{s}_i, \mu^i) \in \boldsymbol{\Sigma}_i^{\infty}$.

Proposition 5: *Fix* $i \in I$, $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$, *and* $\mathbf{z}_i \in \mathbf{Z}_i$. *If for every* $\bar{\varepsilon} > 0$ *there exists* $\mathbf{h}_i^{\bar{\varepsilon}} \prec \mathbf{z}_i$ *such that*

$$\mu^i(\mathbf{S}_{-i}(\mathbf{z}_i)|\mathbf{S}_{-i}(\mathbf{h}_i^{\bar{\varepsilon}})) \geq 1 - \bar{\varepsilon},$$

then, for every $\epsilon > 0$, *there exists* $\mathbf{h}_i^{\varepsilon}$ *and* $\nu^i \in \Delta^{\mathfrak{C}_i^{\mathbf{h}_i^{\varepsilon}}}(\mathbf{S}_{-i})$ *such that*

$$\nu^i(\mathbf{S}_{-i}(\mathbf{z}_i)|\mathbf{S}_{-i}(\mathbf{h}_i)) = 1 \quad \forall \mathbf{h}_i^{\varepsilon} \preceq \mathbf{h}_i \preceq \mathbf{z}_i,$$

and $proj_{\mathbf{S}_i^{\succeq \mathbf{h}_i^{\varepsilon}}} \mathbf{s}_i \in \mathcal{BR}_i^{\epsilon}(\nu^i)$.

We can now put together the pieces and state the following corollary.

Corollary of Proposition 2 and Proposition 5: *Whenever the belief of every player contains an observational grain of truth, and players satisfy RCSBR, the true state must be an eventual ε,ϵ-Self-Confirming Equilibrium with strongly rationalizable beliefs, for every $\varepsilon > 0$ and $\epsilon > 0$. In the long-run, beliefs converge to full confirmation and continuation strategies to "full" sequential best replies.*

In other words, in the medium run individuals play, at every point along path, an "ε,ϵ-Self-Confirming Equilibrium with strongly rationalizable beliefs" of the continuation game. For every pair (ε, ϵ), there exists a discrete time from which the conditions of $\varepsilon-$confirmed and $\epsilon-$best reply begin to be satisfied, After beliefs have converged, the play becomes a Self-Confirming Equilibrium with strongly rationalizable beliefs.

7. Literature review and discussion

In this paper we analyze the limit behavior of strategically sophisticated rational players in infinitely repeated games with imperfect feedback. We model sophisticated strategic thinking by assuming common strong belief in rationality and prove that, under an "observational grain of truth" assumption, players' behavior and first-order beliefs converge to a self-confirming equilibrium (SCE) with strongly rationalizable beliefs of the repeated game. If players are impatient, in the long run they play SCEs with strongly rationalizable beliefs of the one-period game, but the one-period equilibrium may change over time. We also show that our assumptions are tight. We are now in a position to discuss the related literature in detail. While doing this, we consider the limitations and possible extensions of our work.

Drawing on Battigalli (1987), Battigalli & Guaitoli (1988) use the notion of SCE in (strongly) rationalizable beliefs to analyze economic policy in a macroeconomic game with incomplete information. This equilibrium concept is adapted and used by Schipper (2021) to analyze discovery and equilibrium in games with unawareness (lack of conception of some features of the game). Here we provide both an epistemic and a learning foundation to the equilibrium concept. Although we assume complete information, we can easily extend our results to environments with incomplete information about payoff functions, as in the epistemic analysis of Battigalli & Siniscalchi (2002) and Battigalli & Tebaldi (2019). We conjecture that our approach can be extended to analyze processes of learning and discovery as (impatient) agents repeatedly play a game with unawareness, but this is well beyond the scope of this paper.[12]

As mentioned in the Introduction, Fudenberg & Levine (1993) coined the term "self-confirming equilibrium." They put forward a notion of randomized SCE motivated by a population-game scenario whereby agents are drawn from large populations and randomly matched in every period to play a sequential game, so that randomized strategies of the one-period game are interpreted as stable statistical distributions of pure strategies within populations. In this case, belief-confirmation means that each agent assigns probability 1 to the set of co-players' randomized strategies inducing the actual frequency distribution of observations given her (pure) strategy. The large-population scenario also justifies one-period expected payoff maximization despite a positive discount factor, as agents understand that they cannot affect the behavior of future co-players, who are almost certainly different from their current co-players, and—in the long run—they also have no incentive to experiment. We do not consider a population-game scenario for two reasons. First, many recurrent interactions feature a fixed set of players. Second, the analysis would be technically more difficult. We relate to one-period game equilibria by assuming impatient players, while with patient players we obtain convergence to repeated-game SCE (cf. Kalai & Lehrer 1993, 1995). We conjecture that we could cover the case of large but finite populations allowing for chance moves and analyzing the population game as a grand game with (finitely) many agents partitioned according to their role. Another difference with

[12] See the discussion in Schipper (2021), pages 3-4.

Fudenberg & Levine (1993) is that, unlike us, they assume perfect feedback about chosen actions at the end of the one-period game. When the latter is a sequential game, co-players' one-period strategies are nonetheless imperfectly observable, which is what makes their SCE concept different from Nash equilibrium. Note, however, that under perfect feedback pure SCEs in two-person games are realization-equivalent to Nash equilibria.[13] Fudenberg & Kamada (2015, 2018) remove the perfect feedback assumption, positing a terminal information partition for each player.[14]

We explained in the Introduction the main conceptual difference between SCE with rationalizable beliefs and the rationalizable SCE concept of Rubinstein & Wolinsky (1994): unlike the former, the latter postulates common certainty of the confirmation of beliefs. This is argued informally in their paper, and it is formally proved in the epistemic analysis of Esponda (2013), who focuses on games with incomplete information. Another important difference between our work and these papers on rationalizable SCE is that they consider simultaneous-move games. While the SCE concept, which does not presume strategic sophistication, can be meaningfully applied to the strategic form of a sequential game,[15] notions of SCE with strategically sophisticated players must be adapted to take sequential moves into account, because their application to the strategic form of a sequential game with feedback would allow for non-credible threats.[16] Dekel et al (1999) analyze a version of rationalizable SCE for sequential games with perfect feedback. As mentioned above, Fudenberg & Kamada (2015,2018) allow for imperfect feedback. These papers on rationalizable SCE in sequential games feature a weak notion of strategic sophistication, as they assume that there is common certainty of rationality and belief confirmation at the beginning of the game, but not if players are surprised by moves that are compatible with such assumptions. We instead assume common strong belief in rationality. Yet, we do not assume common strong belief in confirmation and we do not allow for randomization; thus, the two concepts are not nested.

The learning aspect of our paper is related to Kalai & Lehrer (1993) who analyze repeated games with perfect monitoring where each player knows her payoff function, and Kalai & Lehrer (1995) on repeated games with imperfect monitoring and imperfect knowledge of one's own payoff function. As in their work, we obtain convergence of beliefs about superstrategies from a kind of "grain of truth" condition. As in Kalai & Lehrer (1995), our condition concerns the personal observations that will be made by each player, rather than the path of play. Furthermore, since we model beliefs as conditional probability systems, we can state this condition as something that holds eventually, that is, we allow for finitely many surprises. The most important difference between our work and these papers is that they do not assume sophisticated strategic thinking, which is the reason why only knowledge of one's own payoff function matters, rather than interactive knowledge about the game.

Finally, we do not model the information structure of the one-period game and of the repeated game by means of information partitions. We represent the flow of information accruing to players

[13] The latter may be partially randomized off path. Cf. Battigalli (1987) and Fudenberg & Levine (1993).

[14] For this reason, they call the equilibrium "partition confirmed." Instead, we keep the same terminology independently of the information/feedback structure.

[15] Provided that also feedback, besides the payoff functions, is accurately represented in strategic form. See the discussion in Battigalli et al (2019), who point out that this is not true when players are ambiguity averse.

[16] Rubinstein & Wolinsky (1994) write that their analysis concerns "normal-form games." They do not clarify whether they mean that the analysis can be meaningfully applied to the normal/strategic form of the given game with feedback. But it is obvious that this is not the case.

between stages and periods by means of feedback functions and thereby comply with the following "separation principle" of Battigalli & Generoso (2021): the description of the rules of the game is independent of players' personal features, such as their mnemonic abilities. [17] Besides this conceptual advantage, our representation allows to seamlessly blend information flows within each one-period game with repeated-game monitoring. To simplify the exposition, we assume a multistage structure (cf. Myerson 1986), but our analysis and results can be extended to more general sequential games represented as in Battigalli & Generoso (2021).

[17] Of course, our analysis of rationality presumes that each player always remembers the sequence of actions she chose and messages she received (personal history). As shown in Battigalli & Generoso (2021), this perfect-memory assumption allows to recover from our "flow representation" of information a "stock representation" with information partitions satisfying perfect recall.

References

[1] ALIPRANTIS, C.D., AND K.C. BORDER (2006): *Infinite Dimensional Analysis. A Hitchhiker's Guide,* third edition, Springer-Verlag, Berlin.

[2] BATTIGALLI, P. (1987): "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali," unpublished thesis, UniversitÃ Bocconi.

[3] BATTIGALLI, P., S. CERREIA-VIOGLIO, F. MACCHERONI, AND M. MARINACCI (2015): "Self-Confirming Equilibrium and Model Uncertainty," *American Economic Review*, 105, 646-677.

[4] BATTIGALLI, P., AND N. GENEROSO (2021): "Information Flows and Memory in Games", Igier Working Paper, no. 678.

[5] BATTIGALLI, P., AND D. GUAITOLI (1988): "Conjectural Equilibria and rationalizability in a Game with Incomplete Information," Quaderni di Ricerca, UniversitÃ Bocconi (published in *Decisions, Games and Markets*, Kluwer, Dordrecht, 97-124, 1997).

[6] BATTIGALLI, P., AND M. SINISCALCHI (2002): "strong belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356-391.

[7] BATTIGALLI, P., AND P. TEBALDI (2019): "Interactive epistemology in simple dynamic games with a continuum of strategies," *Economic Theory*, 68, 737-763.

[8] DEKEL, E., D. FUDENBERG, AND D.K. LEVINE (1999): "Payoff Information and Self-confirming Equilibrium," *Journal of Economic Theory*, 89, 165-185.

[9] DEKEL, E., AND M. SINISCALCHI (2015): "Epistemic Game Theory," in *Handbook of Game Theory with Economic Applications, Volume 4*, ed. by P. Young and S. Zamir. Amsterdam: North-Holland, 619-702.

[10] EASLEY, D., AND N.M. KIEFER (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045-1064.

[11] ESPONDA, I. (2013): "Rationalizable Conjectural Equilibrium: A Framework for Robust Predictions," *Theoretical Economics*, 8, 467-501.

[12] FUDENBERG, D., AND Y. KAMADA (2015): "Rationalizable partition-confirmed equilibrium," *Theoretical Economics*, 10, 775-806.

[13] FUDENBERG, D., AND Y. KAMADA (2018): "Rationalizable partition-confirmed equilibrium with heterogeneous beliefs," *Games and Economic Behavior*, 109, 364-381.

[14] FUDENBERG, D., AND D.M. KREPS (1995): "Learning in extensive-form games I. Self-confirming equilibria," *Games and Economic Behaviour*, 8, 20-55.

[15] GILLI, M. (1999): "Adaptive Learning in Imperfect Monitoring Games," *Review of Economic Dynamics*, 2, 472-485.

[16] KALAI, E., AND E. LEHRER (1993): "Rational Learning leads to Nash Equilibrium," *Econometrica*, 61, 1019-1045.

[17] KALAI, E., AND E. LEHRER (1995): “Subjective Games and Equilibria,” *Games and Economic Behaviour*, 8, 123-163.

[18] RUBINSTEIN, A., AND A. WOLINSKY (1994): “Rationalizable Conjectural Equilibrium: Between Nash and rationalizability,” *Games and Economic Behavior*, 6, 299-311.

[19] SCHIPPER, B. (2021): “Discovery and Equilibrium in Games with Unawareness,” *Journal of Economic Theory*, 198, 105365.

Appendix

A - Comparison of optimality conditions and existence

In the present appendix we compare the one-step optimality with sequential optimality and weak sequential optimality. First we show how, thanks to dynamic consistency of expected utility maximization, whenever players have strictly positive discount factors, the definitions of optimality based on continuation values are equivalent to maximization of expected utility with respect to continuation strategies. This holds because it is possible to attach utilities to terminal histories of the infinite game. Hence, the use of continuation values allows us to extend optimality conditions "computed under this ex-ante perspective" to the case of impatient intertemporal preferences in a multiperiod game. Subsequently, we provide a proof of the **One-Shot Deviation Principle**, which states the equivalence between one-step optimality and sequential optimality. Then, we adapt known arguments to prove the existence of sequentially optimal and weakly sequentially optimal superstrategies (and strategies), and to prove that a superstrategy (strategy) is weakly sequentially optimal if and only if there exists a behaviorally equivalent sequentially optimal superstrategy (strategy). As implication, we obtain that our behavioral characterization of rationality and common strong belief in rationality, and of the corresponding one-period assumptions, is behaviorally equivalent to the one that would be induced by using weak sequential optimality as the representation of rationality, as done by BT.

First of all, we state the continuity properties of value functions.

Lemma A.1: *Let $\Upsilon(\Gamma, \delta)$ be the infinite repetition of the multistage game Γ with discount factors $\delta = (\delta_i)_{i\in I}$. $\Upsilon(\Gamma, \delta)$ satisfies continuity at infinity, i.e.*

$$\forall i \in I, \forall \mathbf{h}_i \in \mathbf{H}_i, \quad lim_{t\to\infty}[sup\{|V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i) - V^{\mu^i}_{i,\mathbf{h}_i}(\bar{\mathbf{s}}_i)| \mathbf{s}_i, \bar{\mathbf{s}}_i \in \mathbf{S}_i, \mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}),$$

$$\forall \mathbf{g}_i \in \mathbf{H}_i, \ell(\mathbf{g}_i) < t,\ \mathbf{s}_i(\mathbf{g}_i) = \bar{\mathbf{s}}_i(\mathbf{g}_i)\}] = 0.$$

Lemma A.2: *For every $i \in I$ and $\mathbf{h}_i \in \mathbf{H}_i$, $V_{i,\mathbf{h}_i} : \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) \to \mathbb{R}$ is jointly continuous. Fix a one-period personal history $h_i \in H_i$, then also $V_{i,h_i} : S_{-i} \times \Delta^{\mathcal{C}_i} \to \mathbb{R}$ is jointly continuous.*

Definition A.1: *A superstrategy $\mathbf{s}^*_i$ is sequentially optimal given a CPS $\mu^i \in \Delta^{\mathfrak{C}_i)}(\mathbf{S}_{-i})$ if, for every $\mathbf{h}_i \in \mathbf{H}_i$,*

$$\mathbf{s}^*_i \in \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i).$$

Similarly, a strategy s^*_i is sequentially optimal given a one-period CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$ if, for every $h_i \in H_i$,

$$s^*_i \in \arg\max_{s_i\in S_i(h_i)} V^{\gamma^i}_{i,h_i}(s_i).$$

Observe that, for every $i \in I$, $\mathbf{h}_i \in \mathbf{H}_i$, and $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, by compactness of $\mathbf{S}_i$ and continuity of $V^{\mu^i}_{i,\mathbf{h}_i}(\cdot)$, $V^{\mu^i}_{i,\mathbf{h}_i}(\cdot)$ admits a maximizer. Obviously, the same holds with respect to the one-period continuation values.

Definition A.2: *A superstrategy* $\mathbf{s}_i^*$ *is weakly sequentially optimal given a CPS* $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ *if, for every* $\mathbf{h}_i \in \mathbf{H}_i(\mathbf{s}_i^*)$,

$$\mathbf{s}_i^* \in \arg\max_{\mathbf{s}_i \in \mathbf{S}_i} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i).$$

Similarly, a strategy s_i^* *is weakly sequentially optimal given a one-period CPS* $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$ *if, for every* $h_i \in H_i(s_i^*)$,

$$s_i^* \in \arg\max_{s_i \in S_i} V_{i,h_i}^{\gamma^i}(s_i).$$

Remark A.1: *By definition of continuation value, optimization with respect to superstrategies (strategies) is equivalent to optimization with respect to continuation superstrategies (strategies), at every personal history and for each (one-period) CPS.*

In our definition of sequential optimality (and its weaker version) and one-step optimality, for the infinite intertemporal interaction, we look at the individual choice not as the choice of a superstrategy to which the player commits ex-ante, but as the repeated decision at each personal history of the immediate action to take, re-planning in her mind the continuation superstrategy. For this reason, the discounted expected utility guiding such decision is the one computed at the current period, not the one computed ex-ante. It is easy to see that, when players are not impatient, our definitions based on continuation values coincide with the traditional ex-ante ones. Instead, obviously, the conditions differ when players are impatient. Hence, our definitions allow to extend traditional optimality conditions to the case of possibly impatient players, when utility can not be attached to terminal histories.

Remark A.2: (Dynamic consistency of expected utility maximization) *When players are not impatient, optimality computed via continuation values is equivalent to optimality computed ex-ante.*

The following two propositions allow us to see that our definition of rationality is equivalent to sequential optimality, and it is hence behaviorally equivalent to the rationality definition of BT, provided that players are not impatient.

Proposition A.1: (One-Shot Deviation Principle) *Fix a player* i*, a superstrategy* $\mathbf{s}_i$ *and a CPS* μ^i *over opponents' superstrategies. Then,* $\mathbf{s}_i$ *is one-step optimal given* μ^i *if and only if* $\mathbf{s}_i$ *is sequentially optimal given* μ^i.

Proposition A.2: *Fix a player* i *and a CPS* μ^i *over opponents' superstrategy profiles. Then, there always exists at least one sequentially optimal superstrategy and one weakly sequentially optimal superstrategy. Furthermore, every superstrategy behaviorally equivalent to a sequentially optimal superstrategy is weakly sequentially optimal. Consequently, there exist infinite weakly sequentially optimal superstrategies. The same results hold with respect to strategies and one-period CPSs.*

Corollary of Proposition A.1 and Proposition A.2: *Fix* $i \in I$*, superstrategy* $\mathbf{s}_i$*, and CPS over* *opponents' superstrategy profiles μ^i. Then $\mathbf{s}_i$ is weakly sequentially optimal given μ^i if and only if there exists a behaviorally equivalent strategy $\overline{\mathbf{s}}_i$ which is one-step optimal given μ^i.*

In conclusion, our representation of rationality is equivalent to sequential optimality, and behaviorally equivalent to weak sequential optimality. This latter, whenever players are not impatient, coincides with the representation of rationality of BT.

Now we define another notion of optimality, namely folding back optimality, which in dynamic finite games (and thus in our one-period game) is equivalent to one-step optimality and sequential optimality. This optimality is based on the definition of *optimal values*, which is done iteratively below.

Basis step: Fix a CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, and consider every pre-terminal personal history $h_i \in H_i$, that is, h_i is such that $\ell(h_i) = L-1$. Then, the optimal value at h_i given γ^i is

$$\widehat{V}_i^{\gamma^i}(h_i) = \max_{a_i \in \mathcal{A}_i^L(h_i)} \sum_{a_{-i} \in \mathcal{A}_{-i}^L(h_i)} \sum_{s_{-i} \in S_{-i}^{a_{-i}}(h_i)} u_i((a^{L-1}(h_i, s_{-i}), (a_i, a_{-i})))\gamma^i(s_{-i}|S_{-i}(h_i)) =$$

$$= \max_{a_i \in \mathcal{A}_i^L(h_i)} \sum_{a_{-i} \in \mathcal{A}_{-i}^L(h_i)} \sum_{h \in o_i(h_i)} u_i((h, (a_i, a_{-i})))\gamma^i(S_{-i}(h, a_{-i})|S_{-i}(h_i)),$$

where $S_{-i}(h_i, a_{-i}) = \{s_{-i} \in S_{-i}(h_i) : s_{-i}(h_i) = a_{-i}\}$, while $a^{L-1}(h_i, s_{-i})$ is the unique sequence of $L-1$ action profiles induced by s_{-i} when personal history h_i has been played.

Recursive step: Fix a CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, and assume that, for some $k \in \mathbb{N}$, for every personal history $h_i \in H_i$ such that $\ell(h_i) > L-k$, $\widehat{V}_i^{\gamma^i}(h_i)$ has been defined. Now take every $g_i \in H_i$ such that $\ell(g_i) = L-k$. For every $a_i \in \mathcal{A}_i^{\ell(g_i)+1}(g_i)$, define

$$\widehat{V}_i^{\gamma^i}(g_i, a_i) = \sum_{s_{-i} \in S_{-i}(g_i)} \widehat{V}_i^{\gamma^i}((g_i, (a_i, f_i^{\ell(g_i)+1}((g_i, (a_i, s_{-i}(g_i))))))\gamma^i(s_{-i}|S_{-i}(g_i)) =$$

$$= \sum_{a_{-i} \in \mathcal{A}_{-i}^{\ell(g_i)+1}(g_i)} \widehat{V}_i^{\gamma^i}((g_i, (a_i, f_i^{\ell(g_i)+1}((g_i, (a_i, a_{-i}))))\gamma^i(S_{-i}^{a_{-i}}(g_i)|S_{-i}(g_i)).$$

Then, define the optimal value at g_i given γ^i as

$$\widehat{V}_i^{\gamma^i}(g_i) = \max_{a_i \in \mathcal{A}_i^{\ell(g_i)+1}(g_i)} \widehat{V}_i^{\gamma^i}(g_i, a_i).$$

Definition A.3: *A strategy s_i is folding back optimal given CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$ if, for every $h_i \in H_i$,*

$$s_i(h_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(h_i)+1}(h_i)} \widehat{V}_i^{\gamma^i}(h_i, a_i).$$

Alternatively, s_i is folding back optimal given γ^i if and only if

$$\widehat{V}_i^{\gamma^i}(h_i, s_i(h_i)) = \widehat{V}_i^{\gamma^i}(h_i).$$

The following proposition is considered well-known, and we do not prove it here.

Proposition A.3: (Folding Back Principle) *Take a strategy s_i and a CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$. Then s_i satisfies one-shot deviation property given γ^i if and only if s_i is folding back optimal given γ^i.*

Corollary of Proposition A.1 and Proposition A.3: *Take a strategy s_i and a CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$. Then s_i is sequentially optimal given γ^i if and only if s_i is folding back optimal given γ^i.*

B - Remarks on game structure

Remark B.1: *If we endow $\overline{H}$ with the natural partial order on sequences $\preceq$, we obtain the objective tree with root $\varnothing$, $(\overline{H}, \preceq)$. Similarly, endowing $\overline{H}_i$ with the product partial order*

$$h_i \preceq g_i \Leftrightarrow a_i^{[\ell(h_i)]}(h_i) \preceq a_i^{[\ell(g_i)]}(g_i) \wedge m_i^{[\ell(h_i)]}(h_i) \preceq m_i^{[\ell(g_i)]}(g_i),$$

where $h_i, g_i \in \overline{H}_i$, $\ell(h_i)$ is the length of h_i (i.e. the cardinality of the sequence of pairs), and $\ell(h_i) \leq \ell(g_i)$, we obtain the subjective tree with root $\varnothing$, $(\overline{H}_i, \preceq)$.

Observe that the $\preceq$ relation can be applied to $\overline{H}_i$ (and $\overline{\mathcal{I}}_i$) in an alternative but equivalent way to the one defined in remark 1, starting from $\overline{H}$: for every $h_i, g_i \in \overline{H}_i$ ($C, D \in \overline{\mathcal{I}}_i$), $h_i \preceq g_i \Leftrightarrow$ exists $h' \in o_i(h_i)$ (C), $g' \in o_i(g_i)$ (D) such that $h' \preceq g'$. Again we abuse notation calling the inherited relation on $\overline{H}_i$ and $\overline{\mathcal{I}}_i$ with the same symbol: $\preceq$. Then, also $(\overline{\mathcal{I}}_i, \preceq)$ is obviously a subjective tree.

Remark B.2: $S(h_i) = \{s \in S : \forall g_i \prec h_i,\ (g_i, (s_i(g_i), f_i^{\ell(g_i)}(s(g_i))) \preceq h_i\}$.

Remark B.3: *$S(h_i) = S_i(h_i) \times S_{-i}(h_i)$ for all $h_i \in H_i$ and $i \in I$, and for every $g_i, h_i \in \overline{H}_i$, $g_i \preceq h_i \Rightarrow S(h_i) \subseteq S(g_i) \Leftrightarrow S_i(h_i) \subseteq S_i(g_i) \wedge S_{-i}(h_i) \subseteq S_{-i}(g_i)$.*

By inspection of the definition, observe that, for every $h_i \in \overline{H}_i$, $S(h_i) = \cup_{h \in o_i(h_i)} S(h)$. Thus, by Remark B.3, $S_{-i}(h_i) = \cup_{h \in o_i(h_i)} S_{-i}(h)$.

Remark B.4: *$(\mathbf{H}, \preceq)$ is an objective tree, $(\mathbf{H}_i, \preceq)$ is a subjective tree, where $\preceq$ is the "prefix of" relation, inherited, respectively, from $A^{\mathbb{N}}$ and $A_i^{\mathbb{N}} \times M_i^{\mathbb{N}}$.*

Remark B.5: $\mathbf{S}(\mathbf{h}_i) = \left\{\mathbf{s} \in \mathbf{S} : \forall \mathbf{g}_i \prec \mathbf{h}_i,\ (\mathbf{g}_i, (\mathbf{s}_i(\mathbf{g}_i), \mathbf{f}_i(\mathbf{s}(\mathbf{g}_i))) \preceq \mathbf{h}_i^{\ell(\mathbf{g}_i)}\right\} = \mathbf{S}_i(\mathbf{h}_i) \times \mathbf{S}_{-i}(\mathbf{h}_i)$. *More-*

over, for every $\mathbf{g}_i, \mathbf{h}_i \in \overline{\mathbf{H}}_i$,

$$\mathbf{g}_i \preceq \mathbf{h}_i \Rightarrow \mathbf{S}(\mathbf{h}_i) \subseteq \mathbf{S}(\mathbf{g}_i) \Leftrightarrow \mathbf{S}_i(\mathbf{h}_i) \subseteq \mathbf{S}_i(\mathbf{g}_i) \wedge \mathbf{S}_{-i}(\mathbf{h}_i) \subseteq \mathbf{S}_{-i}(\mathbf{g}_i).$$

C - Proofs

Proofs for Section 4

Proof of Remark 2: Let

$$marg_{\times_{\mathbf{h}_i \in \left\{\mathbf{O}_i^{-1}(z^{[t-1]})\right\}\times H_i} \mathcal{A}_{-i}^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)} \mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i)) = \mu_t^i(\cdot|S_{-i}(proj_{H_i}\mathbf{h}_i))$$

for every $\mathbf{h}_i \in \left\{\mathbf{O}_i^{-1}(z^{[t-1]})\right\} \times H_i$. Clearly,

$$\left\{\mathbf{O}_i^{-1}(z^{[t-1]})\right\} \times H_i \simeq H_i \simeq \mathcal{C}_i$$

and $\mu_t^i = (\mu_t^i(\cdot|S_{-i}(h_i)))_{h_i \in H_i} \in [\Delta(S_{-i})]^{\mathcal{C}_i}$. We want to show that ν^i has the CPS properties. Notice that indeed

$$\mu_t^i\left(S_{-i}(h_i)|S_{-i}(h_i)\right) = \mu^i\left(\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)\right)|\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)\right)\right) = 1$$

for every $h_i \in H_i$. Then, observe also that, for every $E_{-i} \subseteq S_{-i}$ and $g_i, h_i \in H_i$ such that $h_i \preceq g_i$,

$$\mu_t^i\left(E_{-i} \cap S_{-i}(g_i)|S_{-i}(h_i)\right) = \mu^i\left(\mathbf{S}_{-i}^{E_{-i}}\left(\mathbf{O}_i^{-1}(z^{[t-1]})\right) \cap \mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), g_i)\right)|\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)\right)\right) =$$

$$\mu^i\left(\mathbf{S}_{-i}^{E_{-i}}\left(\mathbf{O}_i^{-1}(z^{[t-1]})\right)|\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), g_i)\right)\right)\mu^i\left(\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), g_i)\right)|\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)\right)\right) =$$

$$= \mu_t^i\left(E_{-i}|S_{-i}(g_i)\right)\mu_t^i\left(S_{-i}(g_i)|S_{-i}(h_i)\right),$$

where

$$\mathbf{S}_{-i}^{E_{-i}}\left(\mathbf{O}_i^{-1}(z^{[t-1]})\right) = \left\{\mathbf{s}_{-i} \in \mathbf{S}_{-i}\left(\mathbf{O}_i^{-1}(z^{[t-1]})\right) : \mathbf{s}_{-i}|_{\left\{\mathbf{O}_i^{-1}(z^{[t-1]})\right\}\times H_i} \in E_{-i}\right\}.$$

■

Proof of Proposition 1: Take a period t and an objective history $z^{[t-1]}$. Take any personal history $\mathbf{g}_i = \left(\mathbf{O}_i^{-1}(z^{[t-1]}), g_i\right) \in \left\{\mathbf{O}_i^{-1}(z^{[t-1]})\right\} \times H_i$, i.e., any personal history which is part of i's subjective one-period game that follows $z^{[t-1]}$ (and any history observationally equivalent to $z^{[t-1]}$). Observe that, for every $\mathbf{s}_i \in \mathbf{S}_i$,

$$V_{i,\mathbf{g}_i}^{\mu^i}(\mathbf{s}_i) = \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i\left(z^t\left(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{g}_i, \mathbf{s}_{-i})\right)\right)\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)) =$$

$$\int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i\left(\zeta((\mathbf{s}_i|\mathbf{g}_i)|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i}, \mathbf{s}_{-i}|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})\right)\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)).$$

Moreover, notice that

$$(\mathbf{s}_i|\mathbf{g}_i)|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i} = (\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i.$$

Let

$$\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i) = \left\{\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{g}_i) : \mathbf{s}_{-i}|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i} = s_{-i}\right\},$$

we have that

$$V_{i,\mathbf{g}_i}^{\mu^i}(\mathbf{s}_i) = \sum_{s_{-i}\in S_{-i}} \int_{\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i)} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= \sum_{s_{-i}\in S_{-i}} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\int_{\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i)} \mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= \sum_{s_{-i}\in S_{-i}} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\mu^i(\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= \sum_{s_{-i}\in S_{-i}} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\mu^i(\mathbf{S}_{-i}^{s_{-i}}(\mathbf{O}_i^{-1}(z^{[t-1]}))|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= \sum_{s_{-i}\in S_{-i}} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\mu_t^i(s_{-i}|g_i)$$

$$= \sum_{s_{-i}\in S_{-i}(g_i)} u_i\left(\zeta((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i, s_{-i})\right)\mu_t^i(s_{-i}|g_i)$$

$$= V_{i,g_i}^{\mu_t^i(\cdot|g_i)}((\mathbf{s}_i|_{\{\mathbf{O}_i^{-1}(z^{[t-1]})\}\times H_i})|g_i),$$

where the first three equalities follows from the definition of $\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i)$ and μ^i, and the subsequent ones from the definition of μ_t^i and of continuation value. Since for every $g_i \in H_i$ there is $\mathbf{g}_i = (\mathbf{O}_i^{-1}(z^{[t-1]}), g_i)$, and for every $\mathbf{g}_i \in \mathbf{H}_i$ there is $g_i \in H_i$ and $z^{[t-1]} \in Z^{t-1}$ (with $t-1$ being the quotient between $\ell(\mathbf{g}_i)$ and L) such that $\mathbf{g}_i = (\mathbf{O}_i^{-1}(z^{[t-1]}), g_i)$, we conclude that, for every $g_i \in H_i$ and $a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}(\mathbf{g}_i) = \mathcal{A}_i^{\ell(g_i)+1}(g_i)$,

$$V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t) = V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i) \geq V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i|_{\mathbf{g}_i}a_i) = V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t|_{g_i}a_i),$$

and that, for every $\mathbf{g}_i \in \mathbf{H}_i$ and $a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}(\mathbf{g}_i) = \mathcal{A}_i^{\ell(g_i)+1}(g_i)$,

$$V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i) = V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t) \geq V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t|_{g_i}a_i) = V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i|_{\mathbf{g}_i}a_i).$$

■

Proof of Theorem 1: Let $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I}\mathbf{\Sigma}_i^\infty$. We want to show that, for all $t \in \mathbb{N}$, in the one-period game starting at $z^{[t-1]} = z^{[t-1]}(\boldsymbol{\zeta}(\mathbf{s}))$ (notation used henceforth), the induced strategy $s_i^t(\mathbf{s}_i)$ and the induced one-period CPS μ_t^i are strongly rationalizable, i.e. $(s_i^t(\mathbf{s}_i), \mu_t^i) \in \Sigma_i^\infty$.

First, we prove by induction the following claim.

Claim: For all $k \in \mathbb{N}$, if $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{\Sigma}_i^k$ then, for all $t \in \mathbb{N}$, in the one-period game starting at $z^{[t-1]}$,

$$S_{-i}(h_i) \cap proj_{S_{-i}} \Sigma_{-i}^m \neq \emptyset \Rightarrow \mu_t^i(proj_{S_{-i}} \Sigma_{-i}^m | S_{-i}(h_i)) = 1,$$

for all $m < k$ and for all $h_i \in H_i$.

The claim implies that $\mu_t^i \in proj_{\Delta^{\mathcal{C}_i}(S_{-i})} \Sigma_i^k$. Since $s_i^t(\mathbf{s}_i) \in \mathcal{OR}_i(\mu_t^i)$ (by Proposition 1), then $(s_i^t(\mathbf{s}_i), \mu_t^i) \in \Sigma_i^k$. To do so, we prove contemporarily that, for every $k \in \mathbb{N}$, for every $i \in I$ and for every $t \in \mathbb{N}$, given some $h_i \in H_i$, for every profile of strategies $s_{-i} \in S_{-i}(h_i) \cap proj_{S_{-i}} \Sigma_{-i}^{k-1}$ (provided this last intersection is not empty), there exists $\mathbf{s}_{-i} \in proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^{k-1} \cap \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}))$ such that $s_{-i}^t(\mathbf{s}_{-i}) = s_{-i}$.

Basis step: We start with $k = 2$. By Proposition 1, for every $i \in I$ and $t \in \mathbb{N}$, if $(\mathbf{s}_i, \mu^i) \in \mathbf{\Sigma}_i^1$ then $s_i^t(\mathbf{s}_i) \in \mathcal{OR}_i(\mu_t^i)$, which implies that $s_i^t(\mathbf{s}_i) \in \Sigma_i^1$. Now suppose that $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{\Sigma}_i^2$. Take all $h_i \in H_i$ such that $S_{-i}(h_i) \cap proj_{S_i} \Sigma_{-i}^1 \neq \emptyset$. We want to show that, for any profile of strategies $s_{-i} = (s_j)_{j\neq i} \in S_{-i}(h_i) \cap proj_{S_i} \Sigma_{-i}^1$, there exists at least a profile of superstrategies $\mathbf{s}_{-i} \in proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}$ such that $s_{-i}^t(\mathbf{s}_{-i}) = s_{-i}$. To see this, consider the case $t = 1$. Suppose $s_j \in \mathcal{OR}_j(\gamma^j)$ for some $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$. Then, we can find some CPS $\nu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ such that, for all $h_i \in H_i$ and $E_{-i} \subseteq S_{-i}$,

$$\nu^i(\mathbf{S}_{-i}^{E_{-i}} | \mathbf{S}_{-i}(h_i)) = \gamma^i(E_{-i}|S_{-i}(h_i)).$$

Indeed, let $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ be a CPS such that, for all $h_j \in H_j$ and for every $s_{-j} \in supp\gamma^j(\cdot|S_{-j}(h_j)) \setminus \cup_{g_j \prec h_j} supp\mu^j(\cdot|g_j)$,

$$\nu^j(s_{-j}^{\mathbb{N}} | \mathbf{S}_{-j}(h_j)) = \gamma^j(s_{-j}|S_{-j}(h_j)),$$

where $s_{-j}^{\mathbb{N}}$ is the superstrategy playing like s_{-j} in every period. Indeed, the above condition does not contradict ν^j being a CPS. Clearly, by Proposition 1, there exists $\mathbf{s}_j \in \mathcal{BR}_j(\nu^j)$ such that $s_j^1(\mathbf{s}_j) = s_j$, and hence $\mathbf{s}_{-i} \in proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^1 \cap \mathbf{S}_{-i}(h_i)$.

Suppose now that $t \geq 2$. There exists $\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}))$ such that $\mathbf{s}_{-i} \in proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^1$. For every $h_i \in H_i$ for which it is possible, take some $s_{-i} \in S_{-i}(h_i) \cap proj_{S_{-i}} \Sigma_{-i}^1 \neq \emptyset$. We want to show that there exists $\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{h}_i) \cap proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^1$, where $\mathbf{h}_j = (\mathbf{O}_j^{-1}(z^{[t-1]}), h_j)$. For every $j \neq i$, let $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$ such that $s_j \in \mathcal{OR}_j(\gamma^j)$. Then, let $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ such that:

(i) $\mathbf{s}_j \in \mathcal{BR}_j(\nu^j)$;

(ii) for all $h_j \in H_j$ and $s_{-j} \in supp\gamma^j(\cdot|S_{-j}(h_j)) \setminus \cup_{g_j \prec h_j} supp\gamma^j(\cdot|g_j)$,

$$\nu^j \left(\mathbf{S}_{-j}^{s_{-j}^{\mathbb{N}}} \left((\mathbf{O}_j^{-1}(z^{[t-1]}), h_j) \right) | \mathbf{S}_{-j} \left((\mathbf{O}_j^{-1}(z^{[t-1]}), h_j) \right) \right) = \gamma^j(s_{-j}|S_{-j}(h_j));$$

(iii) $\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}))$.

Conditions (i) and (iii) can clearly coexist. Condition (ii), as before, does not contradict the fact that ν^j is a CPS, nor can it prevent the superstrategy to satisfy one-shot deviation in the previous periods, as that only depends on the past induced one-period CPSs, not modified by this requirement. Indeed, the second condition only affects beliefs about continuation strategies from $\mathbf{O}_j^{-1}(z^{[t-1]})$ onward.

Thus, it allows $s_j^t(\mathbf{s}_j) = s_j$. Consequently, there exists $\mathbf{s}_{-i} \in \mathbf{S}_{-i}((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^1$. Hence, for every $i \in I$, $t \in \mathbb{N}$, and $h_i \in H_i$

$$S_{-i}(h_i) \cap proj_{S_{-i}}\Sigma_{-i}^1 \neq \emptyset \Rightarrow \mathbf{S}_{-i}((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^1 \neq \emptyset \Rightarrow$$

$$\Rightarrow \mu^i(proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^1 | \mathbf{S}_{-i}((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i))) = 1 \Rightarrow \mu_t^i(proj_{S_{-i}}\Sigma_{-i}^1 | S_{-i}(h_i)) = 1$$

where the last implication follows from the fact that, for every $j \in I$ and $t \in \mathbb{N}$,

$$\mathbf{s}_j \in proj_{\mathbf{S}_j}\mathbf{\Sigma}_j^1 \Rightarrow s_j^t(\mathbf{s}_j) \in proj_{S_j}\Sigma_j^1.$$

Hence, $\mu_t^i \in \Sigma_i^2$. Then again,

$$\mathbf{s}_i \in \mathcal{BR}_i(\mu^i) \Rightarrow s_i^t(\mathbf{s}_i) \in \mathcal{OR}_i(\mu_t^i)$$

for every $i \in I$, which implies that $((s_i^t(\mathbf{s}_i), \mu_t^i))_{i \in I} \in \Sigma^2$.

Inductive step: Suppose that, for some $k \in \mathbb{N}$, for every $m \geq k$ and every $t \in \mathbb{N}$, if $((\mathbf{s}_i, \mu^i))_{i \in I} \in \prod_{i \in I} \mathbf{\Sigma}_i^m$ then, for every $i \in I$, $((s_i^t(\mathbf{s}_i), \mu_t^i))_{i \in I} \in \Sigma^m$. Suppose also that there exists $\mathbf{s}'_{-i} \in proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^\ell \cap \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{t-1}))$ such that $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$, for every $\ell < k$, for every given $h_i \in H_i$ such that $S_{-i}(h_i) \cap proj_{S_{-i}}\Sigma_{-i}^\ell \neq \emptyset$, and for every $s_{-i} \in S_{-i}(h_i) \cap proj_{S_{-i}}\Sigma_{-i}^\ell$.

First, we want to show that, given any suitable h_i, for every $s_{-i} \in S_{-i}(h_i) \cap proj_{S_{-i}}\Sigma_{-i}^k$ there exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]})) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k$ such that $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$. Clearly, $\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]})) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k \neq \emptyset$, because it contains $\mathbf{s}_{-i}$. Conditions $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}))$ and $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$ do not contrast one another. Most importantly, neither are $\mathbf{s}'_{-i} \in proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k$ and $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$. Indeed, there exist $\mathbf{s}'_i$ such that $\mathbf{s}' = (\mathbf{s}'_i, \mathbf{s}'_{-i}) \in proj_{\mathbf{S}} \prod_{i \in I} \mathbf{\Sigma}_i^k \cap \mathbf{S}(\mathbf{O}_i^{-1}(z^{[t-1]}))$, which implies that $s_{-i}^t(\mathbf{s}'_{-i}) \in proj_{S_{-i}}\Sigma_{-i}^k$. In particular, we want to build $\mathbf{s}'_{-i}$ such that $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$.

Take any $j \neq i$. Let $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$ be a one-period CPS of interest justifying $s_j = proj_{S_j} s_{-i}$, that is, γ^j strongly believes $proj_{S_{-j}}\Sigma_{-j}^{k-1}$ and $s_j \in \mathcal{OR}_j(\gamma^j)$. Hence, by inductive assumption, for every h_j such that $S_{-j}(h_j) \cap proj_{S_{-j}}\Sigma_{-j}^{k-1}$, for every $s_{-j} \in supp\gamma^j(\cdot|S_{-j}(h_j))$, there exists $\mathbf{s}''_{-j} \in proj_{\mathbf{S}_{-j}}\mathbf{\Sigma}_{-j}^{k-1} \cap \mathbf{S}_{-j}(o_j^{-1}(z^{[t-1]}))$ such that $s_{-j}^t(\mathbf{s}''_{-j}) = s_{-j}$. By defining $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ as a CPS strongly believing in those superstrategies $\mathbf{s}''_{-j}$, and such that $\nu_t^j = \gamma^j$, it immediately follows that there exists $\mathbf{s}'_j \in \mathcal{BR}_j(\nu^j)$ such that $\mathbf{s}'_j \in proj_{\mathbf{S}_j}\mathbf{\Sigma}_j^k \cap \mathbf{S}_j(o_j^{-1}(z^{[t-1]}))$ and $s_j^t(\mathbf{s}'_j) = s_j$. Letting $\mathbf{s}'_j \in \mathbf{S}_j(o_j^{-1}(z^{[t-1]}))$ is clearly possible because $o_j^{-1}(z^{[t-1]})$ is consistent with strong belief of level k in rationality, and thus any superstrategy that is a sequential best reply to a CPS assigning probability one to $o_j^{-1}(z^{[t-1]})$ can allow it without loss of generality, and independently of the subsequent choices (since the player is impatient and the personal history is terminal for period $t-1$).

This holds for every $j \neq i$. Observe that, while in general, for any $\mathbf{h}_i \in \mathbf{H}_i$, $\mathbf{S}_{-i}(\mathbf{h}_i) \neq \prod_{j \neq i} \mathbf{S}_j(\mathbf{h}_i)$, instead for any $\mathbf{h} \in \mathbf{H}$ it holds that

$$\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\mathbf{h})) \supseteq \mathbf{S}_{-i}(\mathbf{h}) = \prod_{j \neq i} \mathbf{S}_j(\mathbf{O}_j^{-1}(\mathbf{h})).$$

Hence, we have shown that there exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]})) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k$ such that $s_{-i}^t(\mathbf{s}'_{-i}) = s_{-i}$, and thus $\mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)\right) \cap proj_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k \neq \emptyset$. Assume now that $((\mathbf{s}_i, \mu^i))_{i \in I} \in \mathbf{\Sigma}^{k+1}$. Consequently, for every $i \in I$, $t \in \mathbb{N}$, $m \neq k$, and $h_i \in H_i$

$$S_{-i}(h_i) \cap proj_{S_i} \Sigma^m_{-i} \neq \emptyset \Rightarrow \mathbf{S}_{-i}((\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)) \cap proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}^m_{-i} \neq \emptyset \Rightarrow$$

$$\Rightarrow \mu^i(proj_{\mathbf{S}_{-i}} \mathbf{\Sigma}^m_{-i} | \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}), h_i)) = 1 \Rightarrow \mu^i_t(proj_{S_{-i}} \Sigma^m_{-i} | S_{-i}(h_i)) = 1,$$

where the last implication follows from the inductive assumption. Hence, $\mu^i_t \in proj_{\Delta^{\mathcal{C}_i}(S_{-i})} \Sigma^{k+1}_i$. Then again,

$$\mathbf{s}_i \in \mathcal{BR}_i(marg_{\mathbf{S}_{-i}} \mu^i) \Rightarrow s^t_i(\mathbf{s}_i) \in \mathcal{OR}_i(\mu^i_t)$$

for every $i \in I$, that is, $((s^t_i(\mathbf{s}_i), \mu^i_t))_{i \in I} \in \Sigma^{k+1}$.

If $((\mathbf{s}_i, \mu^i))_{i \in I} \in \mathbf{\Sigma}^\infty$, then for every $t \in \mathbb{N}$ and $i \in I$ it holds that, for every $h_i \in H_i$ and $k \in \mathbb{N}$,

$$S_{-i}(h_i) \cap proj_{S_{-i}} \Sigma^\infty_{-i} \neq \emptyset \Rightarrow \mu^i_t(proj_{S_{-i}} \Sigma^k_{-i} | S_{-i}(h_i)) = 1 \Rightarrow \mu^i_t(proj_{S_{-i}} \Sigma^\infty_{-i} | S_{-i}(h_i)) = 1,$$

where the last implication follows from continuity of measures and the fact that $\Sigma^\infty_{-i} = \cap_{k \geq 0} \Sigma^k_{-i}$. Hence, $((s^t_i(\mathbf{s}_i), \mu^i_t))_{i \in I} \in \Sigma^\infty$. ∎

Proofs for Section 5

Proof of Remark 3: Observe that $\mu^i(\cdot | \mathbf{S}_{-i}(\mathbf{h}^t_i)) = \mu^i(\cdot | \mathbf{S}_{-i}(\mathbf{h}^k_i))$ if and only if

$$\mu^i(\mathbf{S}_{-i}(\mathbf{h}^k_i) | \mathbf{S}_{-i}(\mathbf{h}^t_i))$$

for every $k \geq t \geq T$, where the if part is immediate. Then this requires that, at T, for all $t \geq T$,

$$\mu^i(\mathbf{S}_{-i}(\mathbf{h}^t_i) | \mathbf{S}_{-i}(\mathbf{h}^T_i)) = 1,$$

which happens if and only if

$$\mu^i(\cap_{t \geq T} \mathbf{S}_{-i}(\mathbf{h}^t_i) | \mathbf{S}_{-i}(\mathbf{h}^T_i)) = 1,$$

i.e.

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i}))) | \mathbf{S}_{-i}(\mathbf{h}^T_i)) = 1.$$

Once this holds for T, it clearly holds for every $t \geq T$. In words, if the belief of player i over opponents superstrategies has converged starting from T, then at every $t \geq T$ player i believes with certainty in opponents' superstrategy profiles that are observationally equivalent, given i's own superstrategy, to the true ones. Clearly, also the the reverse implications hold. ∎

Proof of Proposition 2: Observe that the sequence $(\mathbf{S}_{-i}(\mathbf{h}^t_i))_{t \in \mathbb{N}}$ is decreasing, and such that $\mathbf{S}_{-i}(\mathbf{h}^t_i) \downarrow \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))) = \cap_{t \in \mathbb{N}} \mathbf{S}_{-i}(\mathbf{h}^t_i)$, where this last equality holds by Remark B.5. Hence, by continuity of measures, for every $k \in \mathbb{N}$,

$$lim_{t \to \infty} \mu^i(\mathbf{S}_{-i}(\mathbf{h}^t_i) | \mathbf{S}_{-i}(\mathbf{h}^k_i)) = \mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i}))) | \mathbf{S}_{-i}(\mathbf{h}^k_i)).$$

If there exists a $T \in \mathbb{N}$ such that $\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)) > 0$, then, for all $\ell \geq t \geq T$,

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) > 0 \quad \wedge \quad \mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^\ell)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) > 0,$$

by chain rule. Hence, for all $\ell \geq t \geq T$, again applying the chain rule, it is true that

$$\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^\ell)) = \frac{\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))}{\mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^\ell)|\mathbf{S}_{-i}(\mathbf{h}_i^t))}.$$

Taking the limit for ℓ, we obtain that

$$lim_{\ell\to\infty}\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^\ell)) = \frac{\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))}{\mu^i(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))} = 1,$$

proving the claim. ■

Proofs for Section 6

Proof of Remark 4: For every $t \in \mathbb{N}$, it holds that

$$1 \geq \mu_t^i(S_{-i}\left(o_i^{-1}(\zeta(s^t))\right)|S_{-i}) = \mu^i(\mathbf{S}_{-i}\left(\left(\mathbf{h}_i^t, o_i^{-1}(\zeta(s^t))\right)\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq$$

$$\geq \mu^i(\mathbf{S}_{-i}\left(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)),$$

where the first inequality is by definition of probability measure and the second by the fact that $\left(\mathbf{h}_i^t, o_i^{-1}(\zeta(s^t))\right) \preceq \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))$. Hence, for all $t \in \mathbb{N}$ and $\varepsilon \geq 0$,

$$\mu^i(\mathbf{S}_{-i}\left(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1-\varepsilon \Rightarrow \mu_t^i(S_{-i}\left(o_i^{-1}(\zeta(s^t))\right)|S_{-i}) \geq 1-\varepsilon.$$

■

Proof of Remark 6: For every $t \geq T$, let $\mathbf{h}_i^t \prec \mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))$ be such that $\ell(\mathbf{h}_i^t) = L(t-1)$. To play one-period Self-Confirming Equilibria, it must hold that

$$\mu^i(\mathbf{S}_{-i}\left(\mathbf{h}_i^{t+1}\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1.$$

Since this holds for every $t \geq T$, by chain rule it can be shown, by induction, that for every $k \in \mathbb{N}$,

$$\mu^i(\mathbf{S}_{-i}\left(\mathbf{h}_i^{T+k}\right)|\mathbf{S}_{-i}(\mathbf{h}_i^T)) = 1.$$

Hence, as shown in the proof of Proposition 2, we can conclude that

$$\mu^i(\mathbf{S}_{-i}\left(\mathbf{O}_i^{-1}(\boldsymbol{\zeta}(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i^T)) = 1.$$

■

Proof of Proposition 3: We have shown in the proof of Theorem 1 that, when players are impatient, for every $k \in \mathbb{N}$, for every $t \in \mathbb{N}$, for every $i \in I$, for every $z^{[t]} \in Z^t$ such that $\mathbf{S}(z^{[t]}) \cap proj_{\mathbf{S}} R^k \neq \emptyset$, and for every $s_i \in proj_{S_i} SR^k$, there exists $\mathbf{s}_i \in proj_{\mathbf{S}_i} R_i^k \cap \mathbf{S}_i(z^{[t]})$ such that $s_i^t(\mathbf{s}_i) = s_i$.

Consequently, since $\mathbf{S}(\mathbf{h}) = \prod_{i \in I} \mathbf{S}_i(\mathbf{h})$ for every $\mathbf{h} \in \overline{\mathbf{H}}$, it follows that, for every $t \in \mathbb{N}$, for every $\mathbf{h} \in Z^t$ such that $\mathbf{S}(\mathbf{h}) \cap proj_{\mathbf{S}} R^\infty \neq \emptyset$, for every $z \in Z$ with $proj_S SR^\infty \cap S(z) \neq \emptyset$, and for every $k \in \mathbb{N}$, there exists $\mathbf{s}_k \in proj_{\mathbf{S}} R^k \cap \mathbf{S}((\mathbf{h}, z))$, i.e. $proj_{\mathbf{S}} R^k \cap \mathbf{S}((\mathbf{h}, z)) \neq \emptyset$. Since $proj_{\mathbf{S}} R^k$ is closed and $\mathbf{S}((\mathbf{h}, z))$ is clopen, $proj_{\mathbf{S}} R^k \cap \mathbf{S}((\mathbf{h}, z))$ is closed, and thus compact. By finite intersection property of compact sets, and because for every $m < k$ we know that $proj_{\mathbf{S}} R^m \cap proj_{\mathbf{S}} R^k = proj_{\mathbf{S}} R^k$, it holds that $\cap_{k \in \mathbb{N}} \left(\mathbf{S}((\mathbf{h}, z)) \cap proj_{\mathbf{S}} R^k\right) = \mathbf{S}((\mathbf{h}, z)) \cap proj_{\mathbf{S}} R^\infty \neq \emptyset$. In other words, there exists $\mathbf{S} \in \mathbf{S}(\mathbf{h}) \cap proj_{\mathbf{S}} R^\infty$ such that $\zeta(s^t(\mathbf{s})) = z$.

Finally, by induction on t, we show that, if $\mathbf{z} \in Z^\infty$ is such that $z^t(\mathbf{z})$ is consistent with one-period RCSBR, i.e., there exists $s \in SR^\infty$ such that $\zeta(s) = z^t(\mathbf{z})$, for every t, then there exists $\mathbf{s} \in proj_{\mathbf{S}} R^\infty$ such that $\boldsymbol{\zeta}(\mathbf{s}) = \mathbf{z}$. Indeed, for $t = 0$, for any $z \in Z$ consistent with one-period RCSBR, there exists $\mathbf{s} \in proj_{\mathbf{S}} R^\infty$ such that $z^1(\boldsymbol{\zeta}(\mathbf{s})) = z$. If for some $t \in \mathbb{N}$, it holds that, for every $\mathbf{h} \in Z^t$ with $z^k(\mathbf{h})$ consistent with one-period RCSBR, for $k \leq t$, there exists $\mathbf{s} \in proj_{\mathbf{S}} R^\infty$ such that $\mathbf{s} \in \mathbf{S}(\mathbf{h}) \cap proj_{\mathbf{S}} R^\infty$, then for every $z \in Z$, consistent with one-period RCSBR, there exists $\mathbf{s} \in \mathbf{S}((\mathbf{h}, z)) \cap proj_{\mathbf{S}} R^\infty$. Hence, if $\mathbf{z} \in Z^\infty$ is such that, for every $t \in \mathbb{N}$, $z^t(\mathbf{z})$ is consistent with one-period RCSBR, then, for every t, $\mathbf{S}(z^{[t]}(\mathbf{z})) \cap proj_{\mathbf{S}} R^\infty \neq \emptyset$. Similarly as before, $\mathbf{S}(z^{[t]}(\mathbf{z})) \cap proj_{\mathbf{S}} R^\infty$ is close, and thus compact, for each t. Moreover, $\mathbf{S}(z^{[t]}(\mathbf{z})) \cap \mathbf{S}(z^{[\ell]}(\mathbf{z})) = \mathbf{S}(z^{[t]}(\mathbf{z}))$ whenever $\ell \leq t$. Thus, by finite intersection property of compact sets, $\cap_{t \in \mathbb{N}} \left(\mathbf{S}(z^{[t]}(\mathbf{z})) \cap proj_{\mathbf{S}} R^\infty\right) = \mathbf{S}(\mathbf{z}) \cap proj_{\mathbf{S}} R^\infty \neq \emptyset$. ■

Proof of Theorem 2: Let $\mathbf{z} \in Z^\infty$ be the terminal history induced by the sequence of strategy and one-period CPS profiles $\left(((s_i^t, \gamma_t^i))_{i \in I}\right)_{t \in \mathbb{N}} \in \left[\prod_{i \in I} \left(S_i \times \Delta^{\mathcal{C}_i}(S_{-i})\right)\right]^{\mathbb{N}}$. For every t, $((s_i^t, \gamma_t^i))_{i \in I}$ is a one-period SCE with strongly rationalizable profiles. For every $i \in I$, define $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ such that, for every $t \in \mathbb{N}$, for every $h_i \in H_i$, and for every $s_{-i} \in S_{-i}$,

$$\gamma_t^i(s_{-i}|S_{-i}(h_i)) = \mu^i \left(\mathbf{S}_{-i}^{s_{-i}}(\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z}))) \cap proj_{\mathbf{S}_{-i}} R_{-i}^\ell | \mathbf{S}_{-i}((\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z})), h_i))\right),$$

where

$$\mathbf{S}_{-i}^{s_{-i}}(\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z}))) = \{\mathbf{s}_{-i} \in \mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z}))) s_{-i}^t(\mathbf{s}) = s_{-i}\},$$

and

$$\ell = \sup\{k \in \mathbb{N} : \mathbf{S}_{-i}^{s_{-i}}(\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z}))) \cap proj_{\mathbf{S}_{-i}} R_{-i}^k \neq \emptyset\}.$$

Since, for every t, $z^{[t-1]}(\mathbf{z})$ is consistent with strong rationalizability, by Proposition 3, $\ell = \sup\{k \in \mathbb{N} s_{-i} \in proj_{S_{-i}} SR^k\}$. Since γ_t^i strongly believes $(SR^k)_{k=1}^\infty$, then, by definition and by Proposition 3, $proj_{\times_{t \in \mathbb{N}, h_i \in H_i} \mathbf{S}_{-i}\left((\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z})), h_i)\right)} \mu^i$ strongly believes $(R^k)_{k=1}^\infty$. Observe that, by definition, μ^i assigns initial probability one to the collection of sets $(\mathbf{S}_{-i}(\mathbf{O}_i^{-1}(z^{[t]}(\mathbf{z}))))_{t \in \mathbb{N}}$, and thus is confirmed by personal histories preceding $\mathbf{o}_i^{-1}(\mathbf{z})$. Furthermore, let μ^i strongly believe in $(proj_{\mathbf{S}_{-i}} R_{-i}^k)_{k=0}^\infty$ (which consists in imposing constraints at personal histories outside $\cup_{t \in \mathbb{N}} \left(\{\mathbf{O}_i^{-1}(z^{[t-1]}(\mathbf{z}))\} \times H_i\right)$). Then, there exists $\mathbf{s}_i \in \mathbf{S}_i$ such that $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$, and $s_i^t(\mathbf{s}_i) = s_i^t$. Therefore, there exists $((\mathbf{s}_i, t_i))_{i \in I} \in R^\infty$ such that $\boldsymbol{\zeta}(\mathbf{s}) = \mathbf{z}$ and $\beta_i(t_i)$ has converged for every player. ■

Proof of Proposition 4: Fix $i \in I$. Let $z_i \in Z_i$ and $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$. Assume that, for some $\varepsilon > 0$ and for every $h_i \preceq z_i$,

$$\gamma^i \left(S_{-i}(z_i)|S_{-i}(h_i)\right) = 1 - \delta_{h_i} \geq 1 - \varepsilon.$$

Define $\nu^i_\varepsilon \in [\Delta(S_{-i})]^{\mathcal{C}_i}$ in the following way:

$$\forall\, h_i \preceq z_i,\ \forall E_{-i} \subseteq S_{-i} \quad \nu^i \left(E_{-i}|S_{-i}(h_i)\right) = \frac{\gamma^i \left(E_{-i} \cap S_{-i}(z_i)|S_{-i}(h_i)\right)}{1 - \delta_{h_i}}$$

$$\forall\, h_i \npreceq z_i,\ \forall E_{-i} \subseteq S_{-i}, \quad \nu^i \left(E_{-i}|S_{-i}(h_i)\right) = \gamma^i \left(E_{-i}|S_{-i}(h_i)\right).$$

It can be checked that ν^i_ε is a CPS, i.e. $\nu^i_\varepsilon \in \Delta^{\mathcal{C}_i}(S_{-i})$. Moreover, for all $h_i \in H_i$, $\nu^i_{\varepsilon,h_i} = \nu^i_\varepsilon(\cdot|S_{-i}(h_i))$ is absolutely continuous with respect to $\gamma^i_{h_i} = \mu^i(\cdot|S_{-i}(h_i))$, written $\nu^i_{\varepsilon,h_i} << \mu^i_{h_i}$. Then, the Radon-Nikodym derivative is

$$\frac{\mathrm{d}\nu^i_{\varepsilon,h_i}}{\mathrm{d}\gamma^i_{h_i}} = \frac{1}{1 - \delta_{h_i}} \mathbb{1}_{S_{-i}(z_i)}$$

whenever $h_i \preceq z_i$, and simply 1 otherwise. Take any measurable function u, then

$$\int_{S_{-i}} u \mathrm{d}\nu^i_{\varepsilon,h_i} = \int_{S_{-i}} u \frac{\mathrm{d}\nu^i_{\varepsilon,h_i}}{\mathrm{d}\gamma^i_{h_i}} \mathrm{d}\gamma^i_{h_i}$$

becomes, when $h_i \preceq z_i$, in

$$\int_{S_{-i}(z_i)} u \mathrm{d}\gamma^i_{h_i} = (1 - \delta_{h_i}) \int_{S_{-i}} u \mathrm{d}\nu^i_{\varepsilon,h_i}.$$

Hence, for every such h_i and any $\bar{s}_i \in S_i$,

$$\mathbb{E}_{\gamma^i_{h_i}}[U_i(\bar{s}_i, \cdot)|h_i] = \sum_{s_{-i} \in S_{-i}(h_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i})$$

$$= (1 - \delta_{h_i}) \sum_{s_{-i} \in S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \nu^i_{\varepsilon,h_i}(s_{-i}) + \sum_{s_{-i} \in S_{-i}(h_i) \setminus S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i})$$

$$= (1 - \delta_{h_i}) \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_i(\bar{s}_i, \cdot)|h_i] + \sum_{s_{-i} \in S_{-i}(h_i) \setminus S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i}).$$

Let

$$n_i = min_{s \in S} U_i(s), \quad N_i = max_{s \in S} U_i(s), \quad \kappa_i = min_{w,v \in U_i(S), w \neq v} |w - v|.$$

Suppose by contradiction that, for every $\varepsilon > 0$ and some $h_i \preceq z_i$, there exists $\widehat{s}_i$ such that $\mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,\widehat{s}_i}] > \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,s_i}]$. Then

$$0 \geq \mathbb{E}_{\gamma^i_{h_i}}[U_{i,\widehat{s}_i}] - \mathbb{E}_{\gamma^i_{h_i}}[U_{i,s_i}] = (1 - \delta_{h_i}) \left[\mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,\widehat{s}_i}] - \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,s_i}]\right] +$$

$$+ \sum_{s_{-i} \in S_{-i}(h_i) \setminus S_{-i}(z_i)} [U_i(\widehat{s}_i|h_i, s_{-i}) - U_i(s_i|h_i, s_{-i})] \cdot \gamma^i_{h_i}(s_{-i})$$

$$\geq (1 - \delta_{h_i})\kappa_i - \delta_{h_i}(M_i - n_i).$$

Thus the inequality is satisfied only if

$$\varepsilon \geq \delta_{h_i} \geq \frac{\kappa_i}{\kappa_i + M_i - N_i} \in (0,1).$$

Then, there exists $\overline{\varepsilon} < \frac{\kappa_i}{\kappa_i+M_i-N_i}$ such that a contradiction is reached. Since, for every $\bar{s}_i \in S_i$,

$$\mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,s_i}] = \mathbb{E}_{\gamma^i_{h_i}}[U_{i,s_i}] \geq \mathbb{E}_{\gamma^i_{h_i}}[U_{i,\bar{s}_i}] = \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,\bar{s}_i}]$$

when $h_i \not\preceq z_i$, then the statement is satisfied. ■

Proofs for Appendix A

Proof of Lemma A.1: Let

$$M = max_{(s'_i,s''_i,s_{-i})\in S_i\times S_i\times S_{-i}}[u_i(\zeta(s'_i,s_{-i})) - u_i(\zeta(s''_i,s_{-i}))].$$

Fix $\mathbf{h}_i \in \mathbf{H}_i$ and $\epsilon > 0$, and let $t > \ell(\mathbf{h}_i)$ be arbitrary. Then, for every $\mathbf{s}_i, \bar{\mathbf{s}}_i \in \mathbf{S}_i$ such that $\mathbf{s}_i(\mathbf{g}_i) = \bar{\mathbf{s}}_i(\mathbf{g}_i)$ at every personal history $\mathbf{g}_i$ with $\ell(\mathbf{g}_i) < t$, and for every $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$,

$$|V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i) - V^{\mu^i}_{i,\mathbf{h}_i}(\bar{\mathbf{s}}_i)| =$$

$$= |\sum_{k=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{k-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} \left[u_i(z^k(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i,\mathbf{s}_{-i}))) - u_i(z^k(\boldsymbol{\zeta}(\bar{\mathbf{s}}_i|\mathbf{h}_i,\mathbf{s}_{-i})))\right]$$

$$\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(h_i))| = |\sum_{k=(t \mod L)}^{\infty} \delta_i^{k-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)}$$

$$\left[u_i(z^k(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i,\mathbf{s}_{-i}))) - u_i(z^k(\boldsymbol{\zeta}(\bar{\mathbf{s}}_i|\mathbf{h}_i,\mathbf{s}_{-i})))\right] \mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(h_i))| \leq$$

$$\leq \left|\sum_{k=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{k-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} M\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(h_i))\right| =$$

$$= M \cdot \sum_{k=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{k-(\ell(\mathbf{h}_i) \mod L)-1} = \frac{M}{1-\delta_i} \cdot \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1},$$

where the first equality follows by definition, the second by the superstrategies prescribing same behavior before t, the inequality by definition of M, and the last two equalities by definition of μ and convergence of the geometric series. Since $\delta_i \in [0,1)$, there exists t large enough such that

$$\delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} < \frac{\epsilon(1-\delta_i)}{M}.$$

Furthermore, this t is independent of μ^i and $\mathbf{s}_i, \bar{\mathbf{s}}_i$. ■

Proof of Lemma A.2: For the continuation values of the one-period game, the property is immediate.

For the continuation values of the infinitely repeated interaction, the property stems from continuity at infinity (Lemma 1) and continuity of the discounted summation of one-period utilities. Take a sequence $((\mathbf{s}_i^n, \mu_n^i))_{n\in\mathbb{N}} \in \left(\mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})\right)^{\mathbb{N}}$ such that $(\mathbf{s}_i^n, \mu_n^i) \to (\mathbf{s}_i, \mu^i)$, and fix a personal history $\mathbf{h}_i \in \mathbf{H}_i$. We want to show that $V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i^n) \to V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i)$. Since $\mathbf{s}_i^n \to \mathbf{s}_i$, for all $t \in \mathbb{T}$ there exists a $n_t \in \mathbb{N}$ such that, for every $n \geq n_t$, $\mathbf{s}_i^n(\mathbf{g}_i) = \mathbf{s}_i(\mathbf{g}_i)$, for all $\mathbf{g}_i \in \mathbf{H}_i$ with $\ell(\mathbf{g}_i) < t$. By continuity at infinity, for every $\varepsilon > 0$, there exists m such that, for all $n \geq m$, $|V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i) - V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i^n)| < \frac{\varepsilon}{2}$.

Fix $t \in \mathbb{N}$ and $\mathbf{s}_i \in \mathbf{S}_i$, and let $\hat{U}_{i,\mathbf{s}_i}^t : \mathbf{S}_{-i} \to \mathbb{R}$ be such that, for every $\mathbf{s}_{-i} \in \mathbf{S}_{-i}$,

$$\hat{U}_{i,\mathbf{s}_i}^t(\mathbf{s}_{-i}) = u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i}))).$$

Then, for every t and $\mathbf{s}_i$, $\hat{U}_{i,\mathbf{s}_i}^t$ is continuous and bounded. Thus, $(\sum_{t=1}^n \delta_i^{t-1}\hat{U}_{i,\mathbf{s}_i}^t)_{n\in\mathbb{N}}$ is a sequence of continuous functions such that $\sum_{t=1}^n \delta_i^{t-1}\hat{U}_{i,\mathbf{s}_i}^t \to \sum_{t=1}^\infty \delta_i^{t-1}\hat{U}_{i,\mathbf{s}_i}^t$. Moreover, such convergence is uniform, which implies that $\sum_{t=1}^\infty \delta_i^{t-1}\hat{U}_{i,\mathbf{s}_i}^t$ is continuous. Also, it is clearly bounded between $\frac{1}{1-\delta_i}\min_{s\in S} U_i(s)$ and $\frac{1}{1-\delta_i}\max_{s\in S} U_i(s)$. The same properties apply for

$$\sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1}\hat{U}_{i,\mathbf{s}_i}^t.$$

Therefore, by convergence of μ_n^i, for every $\mathbf{s}_i \in \mathbf{S}_i$, there exists $k_{\mathbf{s}_i}$ such that, for all $n \geq k_{\mathbf{s}_i}$,

$$\left|V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i) - V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i)\right| = \left|\int_{\mathbf{S}_{-i}} \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1}\hat{U}_{i,\mathbf{s}_i}^t \mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(h_i)) + \right.$$

$$\left. - \int_{\mathbf{S}_{-i}} \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1}\hat{U}_{i,\mathbf{s}_i}^t \mu_n^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(h_i))\right| < \frac{\varepsilon}{2}.$$

Hence, for all $n \geq max\{m, k_{\mathbf{s}_i}\}$,

$$|V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i) - V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i^n)| \leq |V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i) - V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i)| + |V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i) - V_{i,\mathbf{h}_i}^{\mu_n^i}(\mathbf{s}_i^n)| < \varepsilon.$$

■

Proof of Remark A.2: The traditional definitions of sequential optimality and one-shot deviation property are respectively

$$\forall \mathbf{h}_i \in \mathbf{H}_i, \quad \mathbf{s}_i^* \in \arg\max_{\mathbf{s}_i \in \mathbf{S}_i} \int_{\mathbf{S}_{-i}} \mathbf{U}_i(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))$$

$$\mathbf{s}_i^*(\mathbf{h}_i) \in \arg\max_{a_i \in \mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)} \int_{\mathbf{S}_{-i}} \mathbf{U}_i(\mathbf{s}_i^*|_{\mathbf{h}_i} a_i, \mathbf{s}_{-i})\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

We want to show that these definitions coincide with the ones based on continuation values, provided that it is possible to attach payoffs at terminal histories, i.e. $\delta_i > 0$ for every i. Indeed notice that, for every $\mathbf{h}_i \in \mathbf{H}_i$ and $\mathbf{s}_i \in \mathbf{S}_i$,

$$\int_{\mathbf{S}_{-i}} \mathbf{U}_i(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) =$$

$$= \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} \sum_{t=1}^{(\ell(\mathbf{h}_i) \mod L)} \delta_i^{t-1} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))+$$

$$+ \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-1} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

Then, $\mathbf{s}_i^*$ is sequentially optimal given μ^i in the ex-ante computation if and only if

$$\forall \mathbf{h}_i \in \mathbf{H}_i, \quad \forall \mathbf{s}_i \in \mathbf{S}_i,$$

$$\sum_{t=1}^{(\ell(\mathbf{h}_i) \mod L)} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i^*|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))+$$

$$+ \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i^*|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) \geq$$

$$\geq \sum_{t=1}^{(\ell(\mathbf{h}_i) \mod L)} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))+$$

$$+ \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

By the fact that, taking any strategy in $\mathbf{S}_i$, for every $t \leq (\ell(\mathbf{h}_i) \mod L)$, $z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i}))$ depends solely on $\mathbf{s}_{-i}$, the above inequality holds if and only if

$$\sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i^*|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) \geq$$

$$\geq \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

Now it is easy to see that, when the definition of payoffs on terminal history is possible, i.e. when $\delta_i > 0$, the inequality is equivalent to

$$V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i^*) = \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i^*|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) \geq$$

$$\geq \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) = V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i).$$

Hence, the two types of sequential optimality are equivalent. A similar argument can be made for one-shot deviation, and for weak sequential optimality.

In conclusion, our definitions of optimality coincide with the traditional ones when payoffs on terminal histories can be defined, while allowing us to treat also the case in which players are impatient, i.e. $\delta_i = 0$ for all i. ∎

Proof of Proposition A.1: The if part is immediate. It follows from the fact that a sequentially optimal superstrategy is immune to deviations. Indeed, if exists $\mathbf{h}_i \in \mathbf{H}_i$ and $a_i \in \mathfrak{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)$, such that

$$V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i|_{\mathbf{h}_i}a_i) > V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i|_{\mathbf{h}_i}s_i(\mathbf{h}_i)) = V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i),$$

then the superstrategy $\mathbf{s}_i|_{\mathbf{h}_i}a_i$ obtained by substituting a_i to $\mathbf{s}_i(\mathbf{h}_i)$ and keeping fixed all other moves is "sequentially better" than $\mathbf{s}_i$, which can thus not be sequentially optimal given μ^i.

The converse implication is obtained because the extensive form Υ of the game is continuous at infinity (Lemma 1), and because the One-Shot Deviation Principle holds for the finite horizon case, which can be shown by induction. To be used as finite horizon case, we define the truncated game. Given a period T, a superstrategy $\mathbf{s}_i$ and a CPS μ^i, define the game truncated at T as

$$\Upsilon^{T,\mathbf{s}_i,\mu^i} = \langle I, (A_i, \mathcal{A}_i^T(\cdot), M_i, F_i, u_i, \delta_i)_{i\in I}\rangle,$$

where $\mathcal{A}_i^{T,\ell(\mathbf{h}_i)+1}(\mathbf{h}_i) = \mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)$ if $\ell(\mathbf{h}_i) < LT$, and $\mathcal{A}_i^{T,\ell(\mathbf{h}_i)+1}(\mathbf{h}_i) = \emptyset$ if $\ell(\mathbf{h}_i) \geq LT$, so that $\overline{\mathbf{H}}^T = \{\mathbf{h} \in \mathbf{h} : \ell(\mathbf{h}) \leq LT\}$ and $\mathbf{Z}^T = \{\mathbf{h} \in \mathbf{H} : \ell(\mathbf{h}) = LT\}$; then $\overline{\mathbf{H}}_i^T = \{\mathbf{h}_i \in \mathbf{H}_i : \mathbf{O}_i^{-1}(\mathbf{h}_i) \subseteq \overline{H}^T\}$ and $Z_i^T = \{h_i \in H_i : o_i^{-1}(h_i) \subseteq Z^T\}$. u_i is the one-stage payoff function of Γ, M_i is the usual set of possible messages i can observe, F_i the usual one-period incremental feedback function, δ_i the discount factor of i and A_i the set of possible actions. Then we can define

- the set of truncated superstrategies $\mathbf{S}_i^T = \times_{\mathbf{h}_i\in\mathbf{H}_i^T}\mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)$, and for every personal history of the truncated game $\mathbf{h}_i \in \overline{\mathbf{H}}_i^T$, the set of truncated own superstrategies that allow it $\mathbf{S}_i^T(\mathbf{h}_i) = \{\mathbf{s}_i^T \in \mathbf{S}_i^T : \exists \mathbf{x} \in \mathbf{O}_i(\mathbf{h}_i), \exists \mathbf{s}_{-i} \in \mathbf{S}_{-i}, \mathbf{x} \preceq \boldsymbol{\zeta}^T(\mathbf{s}_i^T, \mathbf{s}_{-i})\}$. $\boldsymbol{\zeta}^T(\mathbf{s}_i^T, \mathbf{s}_{-i}) = z^{[T]}(\boldsymbol{\zeta}((\mathbf{s}_i|\mathbf{s}_i^T), \mathbf{s}_{-i}))$, where $(\mathbf{s}_i|\mathbf{s}_i^T) \in \mathbf{S}_i$ is the superstrategy playing like $\mathbf{s}_i^T$ at every personal history shorter than LT, and like $\mathbf{s}_i$ every where else.;
- for any truncated superstrategy $\mathbf{s}_i^T$ and strategy $\mathbf{s}_{-i} \in \mathbf{S}_{-i}$, the truncated strategic form payoff function
$$\mathbf{U}_i^T(\mathbf{s}_i^T, \mathbf{s}_{-i}) = \overline{u}_i((\mathbf{s}_i|\mathbf{s}_i^T), \mathbf{s}_{-i}).$$

Under an ex-ante computation, a truncated superstrategy $\widehat{\mathbf{s}}_i^T$ is sequentially optimal given a CPS $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ in the game truncated at T if, for every $\mathbf{h}_i \in \mathbf{H}_i^T$,

$$\widehat{\mathbf{s}}_i^T \in \arg\max_{\mathbf{s}_i^T\in\mathbf{S}_i^T(\mathbf{h}_i)} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} \mathbf{U}_i^T(\mathbf{s}_i^T|\mathbf{h}_i, \mathbf{s}_{-i})\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)),$$

whereas, it satisfies one-shot deviation property given μ^i in the truncated game if, for every $\mathbf{h}_i \in \mathbf{H}_i^T$,

$$\widehat{\mathbf{s}}_i^T(\mathbf{h}_i) \in \arg\max_{a_i\in\mathfrak{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} \mathbf{U}_i^T(\mathbf{s}_i^T|_{\mathbf{h}_i}a_i, \mathbf{s}_{-i})\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

Hence, with an analogous argument to the one used in remark A.2, we can derive the continuation value of playing truncated superstrategy $\mathbf{s}_i^T$ at personal history $\mathbf{h}_i \in \mathbf{H}_i^T$, given μ^i, and equivalently express the two optimality conditions in terms of such values. In particular,

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T) =$$

$$= \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\mathbf{s}_i^T)|\mathbf{h}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) =$$

$$= V_{i,\mathbf{h}_i}^{\mu^i}((\mathbf{s}_i|\mathbf{s}_i^T)).$$

The definition of continuation value in the truncated game clarifies that if a superstrategy is sequentially optimal given μ^i then its truncation at T is sequentially optimal given μ^i in the truncated game after which the superstrategy itself is the continuation. In the same way, if a superstrategy is one-step optimal given μ^i, then its truncation satisfies it in said truncated game.

Now we show via induction how in any finite (truncated) game one-shot deviation property implies sequential optimality. Suppose $\mathbf{s}_i^T$ is one-step optimal given μ^i.

Basis step: Fix some $\mathbf{h}_i \in \mathbf{H}_i^T$, with $\ell(\mathbf{h}_i) = LT-1$. Take any truncated superstrategy $\overline{\mathbf{s}}_i^T \in \mathbf{S}_i^T$, then

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T) = V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\widehat{\mathbf{s}}_i^T|_{\mathbf{h}_i}\overline{\mathbf{s}}_i^T(\mathbf{h}_i)).$$

for any $\widehat{\mathbf{s}}_i^T \in \mathbf{S}_i^T$. This holds because the truncated strategy does not matter anymore after $\overline{\mathbf{s}}_i^T(\mathbf{h}_i)$ has been played, that is, the choice at $\mathbf{h}_i$ is the last one, in the truncated game, that matters. Hence, by one-shot deviation,

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T) = V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T|_{\mathbf{h}_i}\mathbf{s}_i^T(\mathbf{h}_i)) \geq V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T|_{\mathbf{h}_i}\overline{\mathbf{s}}_i^T(\mathbf{h}_i)) = V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)$$

for every $\overline{\mathbf{s}}_i^T \in \mathbf{S}_i^T(\mathbf{h}_i)$.

Inductive step: Suppose that, for some $k \in \{1,\dots LT-1\}$, for all $\ell \in \{1,\dots,k\}$ and for all $\mathbf{h}_i \in \mathbf{H}_i^T$ such that $\ell(\mathbf{h}_i) = LT-\ell$, $\mathbf{s}_i^T \in \arg\max_{\overline{\mathbf{s}}_i^T \in \mathbf{S}_i^T(\mathbf{h}_i)} V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)$. Let $\mathbf{h}_i \in \mathbf{H}_i^T$ with $\ell(\mathbf{h}_i) = LT-k-1$. Take any truncated superstrategy $\overline{\mathbf{s}}_i^T \neq \mathbf{s}_i^T$, then

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T) =$$

$$= \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i))$$

$$= \sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \sum_{\mathbf{g}_i \in \mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\, \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\, \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0}$$

$$\int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{g}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)).$$

Observe the following. Let

$$\begin{aligned} \widehat{u}_{i,\mathbf{s}_i}^{t,\mathbf{g}_i} : \quad \mathbf{S}_{-i} &\to \mathbb{R}_0^+ \\ \mathbf{s}_{-i} &\mapsto u_i(z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{g}_i,\mathbf{s}_{-i}))). \end{aligned}$$

Then, for every $\mathbf{s}_i$, $\bar{\mathbf{s}}_i^T$, $t \in \mathbb{N}$, and $\mathbf{g}_i \in \mathbf{H}_i$,

$$\int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\bar{\mathbf{s}}_i^T)|\mathbf{g}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)) = \int_{\mathbf{S}_{-i}} u^{t,\mathbf{g}_i}_{i,(\mathbf{s}_i|\bar{\mathbf{s}}_i^T)} \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)).$$

See that $u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}$ is clearly continuous. Indeed, for every $(a, b) \subseteq \mathbb{R}$ and $t \in \mathbb{N}$,

$$(u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i})^{-1}((a, b)) = \{\mathbf{s}_{-i} \in \mathbf{S}_{-i} : z^t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{g}_i, \mathbf{s}_{-i})) \in u_i^{-1}((a, b))\} =$$

$$\cup_{z^{[t-1]}\in Z^{t-1}\cap \mathbf{H}_i(\mathbf{s}_i|\mathbf{g}_i)} \cup_{z\in u_i^{-1}((a,b)):(z^{[t-1]},z)\in \mathbf{H}_i(\mathbf{s}_i|\mathbf{g}_i)} \cap_{h\preceq z}\mathbf{S}_{-i}((z^{[t-1]}, h))$$

where $\mathbf{S}_{-i}((z^{[t-1]}, h))$ is a clopen set. Hence $(u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i})^{-1}((a, b))$ is open and $u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}$ is continuous.

Since $u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}$ is a non-negative Borel-measurable (by continuity) function,

$$\int_{\mathbf{S}_{-i}} u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i} \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)) =$$

$$sup\{\int_{\mathbf{S}_{-i}} f \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)) : f \ simple \ function,\ 0 \le f \le u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}\}.$$

For every such f simple (or step) function, there is $n_f \in \mathbb{N}$, $(x_k)_{k=1}^{n_f} \in (\mathbb{R}_0^+)^{n_f}$, and a partition (in measurable sets) $(A_k)_{k=1}^{n_f}$ of $\mathbf{S}_{-i}$, such that

$$f = \sum_{k=1}^{n_f} x_k \cdot \mathbb{1}_{A_k}, \qquad \int_{\mathbf{S}_{-i}} f\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)) = \sum_{k=1}^{n_f} x_k \cdot \mu^i(A_k|\mathbf{S}_{-i}(\mathbf{g}_i)),$$

$$f \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)} = \sum_{k=1}^{n_f} x_k \cdot \mathbb{1}_{A_k\cap \mathbf{S}_{-i}(\mathbf{g}_i)},$$

$$\int_{\mathbf{S}_{-i}} f \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)) = \sum_{k=1}^{n_f} x_k \cdot \mu^i(A_k \cap \mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{g}_i)).$$

By the chain rule of probability, for any A_k and $\mathbf{g}_i \succ \mathbf{h}_i$ such that $\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i)) > 0$,

$$\mu^i(A_k \cap \mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{g}_i)) = \frac{\mu^i(A_k \cap \mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))}{\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))}.$$

Then, for such $\mathbf{g}_i$ and $\mathbf{h}_i$,

$$\int_{\mathbf{S}_{-i}} u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i} \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= sup\{\sum_{k=1}^{n_f} x_k \cdot \frac{\mu^i(A_k \cap \mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))}{\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))} : f \ simple \ function,\ 0 \le f \le u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}\}$$

$$\frac{1}{\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))} \cdot sup\{\sum_{k=1}^{n_f} x_k \cdot \mu^i(A_k \cap \mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i)) : f \ simple \ function,\ 0 \le f \le u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}\}$$

$$= \frac{1}{\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))} \cdot sup\{\int_{\mathbf{S}_{-i}} f \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i)) : f \ simple \ function,\ 0 \le f \le u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i}\}$$

$$= \frac{1}{\mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))} \cdot \int_{\mathbf{S}_{-i}} u^{t,\mathbf{g}_i}_{i,\mathbf{s}_i} \cdot \mathbb{1}_{\mathbf{S}_{-i}(\mathbf{g}_i)}\mathrm{d}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{g}_i)).$$

We can now go back to $V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)$, and see that

$$\sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0}$$

$$\int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{g}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) =$$

$$\sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))$$

$$\sum_{t=(\ell(\mathbf{h}_i) \mod L)+1}^{\infty} \delta_i^{t-(\ell(\mathbf{h}_i) \mod L)-1} \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^t(\boldsymbol{\zeta}((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{g}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i))$$

$$= \sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))V_{i,\mathbf{g}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T),$$

provided that $\ell(\mathbf{g}_i) \mod L = \ell(\mathbf{h}_i) \mod L$. Otherwise (i.e. if $\ell(\mathbf{g}_i) \mod L = (\ell(\mathbf{h}_i) \mod L)+1$), we obtain that

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T) = \sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))\ \delta_i V_{i,\mathbf{g}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)+$$

$$+\int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(z^{(\ell(\mathbf{g}_i) \mod L)}(\boldsymbol{\zeta}((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{g}_i,\mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)),$$

where the second term of the summation is independent of $(\mathbf{s}_i|\overline{\mathbf{s}}_i^T)$ (though indirectly depends on it as it determines the possible $\mathbf{g}_i$'s). By inductive assumption, observe that

$$\sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))V_{i,\mathbf{g}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)$$

$$\leq \sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))V_{i,\mathbf{g}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T).$$

Hence, whenever $\ell(\mathbf{g}_i) \mod L = \ell(\mathbf{h}_i) \mod L$,

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T) \leq$$

$$\leq \sum_{\mathbf{g}_i\in\mathbf{H}_i((\mathbf{s}_i|\overline{\mathbf{s}}_i^T)|\mathbf{h}_i):\ \ell(\mathbf{g}_i)=\ell(\mathbf{h}_i)+1,\ \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))>0} \mu^i(\mathbf{S}_{-i}(\mathbf{g}_i)|\mathbf{S}_{-i}(\mathbf{h}_i))V_{i,\mathbf{g}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T) =$$

$$= V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T|_{\mathbf{h}_i}\overline{\mathbf{s}}_i^T(\mathbf{h}_i)) \leq V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T|_{\mathbf{h}_i}\mathbf{s}_i^T(\mathbf{h}_i)) = V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T),$$

where the last inequality follows from one-shot deviation property. It can be checked that the same holds in case $\ell(\mathbf{g}_i) \mod L = (\ell(\mathbf{h}_i) \mod L)+1$. Since these hold for every $\overline{\mathbf{s}}_i^T \in \mathbf{S}_i^T$, then $\mathbf{s}_i^T \in \arg\max_{\overline{\mathbf{s}}_i^T\in\mathbf{S}_i^T(\mathbf{h}_i)} V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T)$.

Now suppose $\mathbf{s}_i$, which satisfies one-shot deviation property in the full game given μ^i, is not sequentially optimal, i.e.

$$\exists \widehat{\mathbf{s}}_i \in \mathbf{S}_i,\ \exists \mathbf{h}_i \in \mathbf{H}_i,\quad V_{i,\mathbf{h}_i}^{\mu^i}(\widehat{\mathbf{s}}_i) - V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i) = \epsilon > 0.$$

Let $\mathbf{s}_i^T$ and $\widehat{\mathbf{s}}_i^T$ be the truncations at period T of, respectively, $\mathbf{s}_i$ and $\widehat{\mathbf{s}}_i$, and let $\overline{\mathbf{s}}_i = (\mathbf{s}_i|\widehat{\mathbf{s}}_i^T) \in \mathbf{S}_i$ be the strategy playing like $\widehat{\mathbf{s}}_i$ at all personal histories with length less than LT, and like $\mathbf{s}_i$ everywhere else (with $\widehat{\mathbf{s}}_i^T$ being its truncation at T). By continuation at infinity (lemma 1), there exists a $T > (\ell(\mathbf{h}_i) \mod L) + 1$ such that

$$V_{i,\mathbf{h}_i}^{\mu^i}(\overline{\mathbf{s}}_i) \geq V_{i,\mathbf{h}_i}^{\mu^i}(\widehat{\mathbf{s}}_i) - \epsilon > V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i).$$

Hence, in the truncated game $\Upsilon^{T,\mathbf{s}_i,\mu^i}$,

$$V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\overline{\mathbf{s}}_i^T) = V_{i,\mathbf{h}_i}^{\mu^i}(\overline{\mathbf{s}}_i) > V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i) = V_{i,\mathbf{h}_i}^{T,\mathbf{s}_i,\mu^i}(\mathbf{s}_i^T),$$

that is, $\mathbf{s}_i^T$ is not sequentially optimal in the truncated game, implying it does not satisfy the one-shot deviation property in the truncated game, which is a contradiction.

An analogous, simpler proof can be provided for the one-period case. ■

Proof of Proposition A.2: Fix $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$. By compactness of $\mathbf{S}_i$ and continuity of $V_{i,\mathbf{h}_i}^{\mu^i}(\cdot)$, we know that for all $\mathbf{h}_i \in \mathbf{H}_i$, $\arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i)$ is non-empty and closed (hence compact). To see closedness, take a sequence $(\mathbf{s}_i^n)_{n\in\mathbb{N}} \in \left(\arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i)\right)^{\mathbb{N}}$ such that $\mathbf{s}_i^n \to \overline{\mathbf{s}}_i$, then

$$\forall n \in \mathbb{N}, \forall \mathbf{s}_i \in \mathbf{S}_i,\quad V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i^n) \geq V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i).$$

Taking the limit for n,

$$\forall \mathbf{s}_i \in \mathbf{S}_i,\quad V_{i,\mathbf{h}_i}^{\mu^i}(\overline{\mathbf{s}}_i) = lim_{n\to\infty} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i^n) \geq V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i),$$

where the first equality follows from continuity of $V_{i,\mathbf{h}_i}^{\mu^i}(\cdot)$.

Let $\mathcal{W}^0 = \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V_{i,\varnothing}^{\mu^i}(\mathbf{s}_i)$. Then, for every $t \in \mathbb{N}$, pick a superstrategy $\mathbf{s}_i^t \in \mathcal{W}^{t-1}$ and let

$$\mathcal{W}^t = \left\{\mathbf{s}_i \in \mathcal{W}^{t-1} : \forall \mathbf{g}_i \in \mathbf{H}_i(\mathbf{s}_i^t),\ \ell(\mathbf{g}_i) < t,\ \mathbf{s}_i(\mathbf{g}_i) = \mathbf{s}_i^t(\mathbf{g}_i)\right\} \cap$$

$$\cap \left(\cap_{\mathbf{h}_i\in\mathbf{H}_i(\mathbf{s}_i^t):\ \ell(\mathbf{h}_i)=t} \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i)\right),$$

where $\left\{\mathbf{s}_i \in \mathcal{W}^{t-1} : \forall \mathbf{g}_i \in \mathbf{H}_i(\mathbf{s}_i^t),\ \ell(\mathbf{g}_i) < t,\ \mathbf{s}_i(\mathbf{g}_i) = \mathbf{s}_i^t(\mathbf{g}_i)\right\}$ is the subset of $\mathcal{W}^{t-1}$ consisting in strategies behaviorally equivalent to $\mathbf{s}_i^t$ at all personal histories with length less than t. Such set is trivially closed and non-empty. Then, $\mathcal{W}^t \neq \emptyset$, because, as highlighted in Remark A.1, for the optimization at each different personal history $\mathbf{h}_i$, the only thing that matters is the continuation superstrategy, not what has happened in the past (dynamic consistency of expected utility maximization). Hence, a superstrategy which "continues optimally" at each personal history of any given length can be always constructed. Also, $\mathcal{W}^t$ is closed and hence compact. Since in the decreasing sequence of compact sets $(\mathcal{W}^t)_{t=0}^{\infty}$ every finite intersection is non-empty (for all finite subsequences, the intersection equals the smallest set), then we can apply the finite intersection property of compact sets, and state that

$\cap_{t\geq 0}\mathcal{W}^t \neq \emptyset$. By inspection of the definitions, $\cap_{t\geq 0}\mathcal{W}^t$ is a subset of the set of all weakly sequentially optimal strategies given μ^i, implying that this latter is non-empty. 1

Similarly, let $\mathcal{S}^0 = \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\varnothing}(\mathbf{s}_i)$. Then, for every $t\in\mathbb{N}$, let

$$\mathcal{S}^t = \mathcal{S}^{t-1} \cap (\cap_{\mathbf{h}_i\in\mathbf{H}_i:\ell(\mathbf{h}_i)=t} \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i)).$$

Clearly, $\cap_{\mathbf{h}_i\in\mathbf{H}_i:\,\ell(\mathbf{h}_i)=t} \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i)$ is non-empty, by the same argument made above. Hence, a superstrategy which "continues optimally" at each personal history of any given length can be always constructed. Again, if $\mathcal{S}^{t-1}\neq\emptyset$, $\mathcal{S}^t\neq\emptyset$. Also, $\mathcal{S}^t$ is closed and hence compact, and $\mathcal{S}^t\subseteq\mathcal{W}^t$ by inspection of the definitions. Then, since in the decreasing sequence of compact sets $(\mathcal{S}^t)_{t=0}^{\infty}$ every finite intersection is non-empty (for all finite subsequences, the intersection equals the smallest set), then we can again apply the finite intersection property of compact sets, and state that $\cap_{t\geq 0}\mathcal{S}^t$ is non-empty, closed and compact. By inspection of the definitions, $\cap_{t\geq 0}\mathcal{S}^t$ is the set of all sequentially optimal superstrategies given μ^i.

The last claim is obvious. To prove it formally, one can think of the following "non traditional" argument. Pick $\bar{\mathbf{s}}_i\in\cap_{t\geq 0}\mathcal{S}^t$, and use it for the iterative construction of the sequence of sets $(\overline{\mathcal{W}}^t)_{t=0}^{\infty}$ as before, that is, $\overline{\mathcal{W}}^0 = \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\varnothing}(\mathbf{s}_i)$ and, for all $t\in\mathbb{N}$,

$$\overline{\mathcal{W}}^t = \left\{\mathbf{s}_i\in\overline{\mathcal{W}}^{t-1} : \forall\mathbf{g}_i\in\mathbf{H}_i(\bar{\mathbf{s}}_i),\ \ell(\mathbf{g}_i)<t,\ \mathbf{s}_i(\mathbf{g}_i)=\bar{\mathbf{s}}_i(\mathbf{g}_i)\right\}\cap$$

$$\cap\left(\cap_{\mathbf{h}_i\in\mathbf{H}_i(\bar{\mathbf{s}}_i):\,\ell(\mathbf{h}_i)=t} \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i)\right).$$

Such construction is possible since, by definition, $\bar{\mathbf{s}}_i\in\overline{\mathcal{W}}^t$ for every t. Then, as seen in the first construction above, $\cap_{t\geq 0}\overline{\mathcal{W}}^t$, which is non-empty, closed and compact, is a set of weakly sequentially optimal superstrategies. Furthermore, by inspection of the definitions, for all $t\in\mathbb{N}$ and all $\mathbf{g}_i\in\mathbf{H}_i(\bar{\mathbf{s}}_i)$ such that $\ell(\mathbf{g}_i)<t$, for all $\mathbf{s}_i\in\cap_{t\geq 0}\overline{\mathcal{W}}^t$, $\mathbf{s}_i(\mathbf{g}_i)=\bar{\mathbf{s}}_i(\mathbf{g}_i)$, i.e. $\mathbf{s}_i(\mathbf{g}_i)=\bar{\mathbf{s}}_i(\mathbf{g}_i)$ for all $\mathbf{g}_i\in\mathbf{H}_i(\bar{s}_i)$. In addition, every strategy $\mathbf{s}_i\in\mathbf{S}_i$ such that, for all $\mathbf{g}_i\in\mathbf{H}_i(\mathbf{s}_i)$, $\mathbf{s}_i(\mathbf{g}_i)=\bar{\mathbf{s}}_i(\mathbf{g}_i)$, belongs, for all $t\in\mathbb{N}$, both to $\{\mathbf{s}_i\in\overline{\mathcal{W}}^{t-1} : \forall\mathbf{g}_i\in\mathbf{H}_i(\bar{\mathbf{s}}_i),\ \ell(\mathbf{g}_i)<t,\ \mathbf{s}_i(\mathbf{g}_i)=\bar{\mathbf{s}}_i(\mathbf{g}_i)\}$ and $\cap_{\mathbf{h}_i\in\mathbf{H}_i(\bar{\mathbf{s}}_i):\,\ell(\mathbf{h}_i)=t} \arg\max_{\mathbf{s}_i\in\mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i)$, and consequently belongs to $\overline{\mathcal{W}}^t$. Hence, $\cap_{t\geq 0}\overline{\mathcal{W}}^t$ coincides with the set of superstrategies behaviorally equivalent to $\bar{\mathbf{s}}_i$, which are thus all weakly sequentially optimal.

An identical proof can be provided for the one-period case. ■

Proofs for Appendix B

Proof of Remark B.2: (and first part of **Remark B.5**) Let

$$C(h_i) = \{s\in S : \forall g_i\prec h_i, (g_i,(s_i(g_i), f_i(s(g_i))))\preceq h_i\}.$$

First, take any $s\in C(h_i)$. By hypothesis, $(a_i^{\ell([h_i])}(\zeta(s)), f_i^{\ell([h_i])]}(a^{\ell([h_i])}(\zeta(s)))) = h_i$, which implies that $a^{\ell([h_i])}(\zeta(s))\in o_i(h_i)$. Since $a^{\ell([h_i])}(\zeta(s))\preceq\zeta(s)$ by definition of $\zeta(s)$, we have shown that $C(h_i)\subseteq S(h_i)$.

For the other inclusion, take $s \in S(h_i)$. Suppose, by contradiction, that there is some $g'_i \prec h_i$ such that $(g'_i, (s_i(g'_i), f_i(s(g'_i))) \not\preceq h_i$. Let g_i be the shortest of such personal histories, i.e., for every $g''_i \prec g_i$, $(g''_i, (s_i(g''_i), f_i(s(g''_i))) \preceq h_i$. By definition, $s \in C((g'_i, (s_i(g'_i), f_i(s(g'_i))))$, and thus $s \in S((g'_i, (s_i(g'_i), f_i(s(g'_i))))$. Then, there is $y \in o_i((g'_i, (s_i(g'_i), f_i(s(g'_i))))$ such that $y \preceq \zeta(s)$. Since $(g'_i, (s_i(g'_i), f_i(s(g'_i))) \not\preceq h_i$, for every $x \in o_i(h_i)$ and $y \in o_i((g'_i, (s_i(g'_i), f_i(s(g'_i))))$, $y \not\preceq x$. Consequently $x \not\preceq \zeta(s)$ and a contradiction is reached. ■

Proof of Remark B.3: (and second part of **Remark B.5**) A proof "by brute force" of the first statement can be given by induction. Let $s' = (s'_i, s'_{-i}), s'' = (s''_i, s''_{-i}) \in S(h_i)$.

Basis step: For $t = 1$, $s'_i(\varnothing) = a^1_i(h_i) = s''_i(\varnothing)$. Since

$$f_i(s'_i(\varnothing), s'_{-i}(\varnothing)) = f_i(a^1_i(h_i), s'_{-i}(\varnothing)) = m^1_i(h_i) = f_i(s''_i(\varnothing), s''_{-i}(\varnothing)) = f_i(a^1_i(h_i), s''_{-i}(\varnothing)),$$

then

$$f_i(s'_i(\varnothing), s''_{-i}(\varnothing)) = f_i(s''_i(\varnothing), s'_{-i}(\varnothing)) = m^1_i(h_i).$$

Hence, $(s'_i, s''_{-i}), (s''_i, s'_{-i}) \in S(a^{[1]}_i(h_i), m^{[1]}_i(h_i))$.

Inductive step: Suppose that for some $k \in \{1, \ldots, \ell(h_i) - 1\}$,

$$(s'_i, s''_{-i}), (s''_i, s'_{-i}) \in S(a^{[k]}_i(h_i), m^{[k]}_i(h_i)).$$

Clearly, because the choice of i depends solely on his personal history, and $(a^{[k]}_i(h_i), m^{[k]}_i(h_i)) \prec h_i$,

$$s'_i(a^{[k]}_i(h_i), m^{[k]}_i(h_i)) = a^{k+1}_i(h_i) = s''_i(a^{[k]}_i(h_i), m^{[k]}_i(h_i))$$

by characterization of $S(h_i)$. Also, since, by the same reasons,

$$f_i(s'(a^{[k]}_i(h_i), m^{[k]}_i(h_i))) = f_i(a^{[k+1]}_i(h_i), s'_{-i}(a^{[k]}_i(h_i), m^{[k]}_i(h_i))) = m^{k+1}_i(h_i) =$$

$$= f_i(s''(a^{[k]}_i(h_i), m^{[k]}_i(h_i))) = f_i(a^{[k+1]}_i(h_i), s''_{-i}(a^{[k]}_i(h_i), m^{[k]}_i(h_i))).$$

Then

$$f_i(s'_i(a^{[k]}_i(h_i), m^{[k]}_i(h_i)), s''_{-i}(a^{[k]}_i(h_i), m^{[k]}_i(h_i))) =$$

$$= f_i(s''_i(a^{[k]}_i(h_i), m^{[k]}_i(h_i)), s'_{-i}(a^{[k]}_i(h_i), m^{[k]}_i(h_i))) = m^{k+1}_i(h_i),$$

and thus

$$(s'_i, s''_{-i}), (s''_i, s'_{-i}) \in S(a^{[k+1]}_i(h_i), m^{[k+1]}_i(h_i)).$$

□

The property just proven reflects the very simple intuition that every $s_i \in S_i(h_i)$ is behaviorally equivalent along (and until) h_i, which also implies that

$$a^{[\ell(h_i)]}(\zeta(s'^{[\ell(h_i)]}(\zeta(s''_i, s'_{-i})) = a^{[\ell(h_i)]}(\zeta(s''^{[\ell(h_i)]}(\zeta(s'_i, s''_{-i})).$$

Now we prove the second statement. The fact that $g_i \preceq h_i$ implies the existence of $g' \in o_i(g_i)$, $h' \in o_i(h_i)$ such that $g' \preceq h'$. This is equivalent to: for every $h'' \in o_i(h_i)$ exists $g'' \in o_i(g_i)$ such that $g'' \preceq h''$. This equivalence follows from $(\overline{H}_i, \preceq)$ being a tree, which implies that each personal history has a unique predecessor for any given length. Indeed, intuitively, if there were two predecessors of, say, length l, then two histories part of the same personal history would have predecessors belonging to different personal histories, and thus such personal histories would be different, making it impossible for two longer histories to belong to the same personal history.

Given the above equivalence, for every strategy $s \in S(h_i)$, there exists $h'' \in o_i(h_i)$ such that s is inducing h'' ($h'' \preceq \zeta(s)$). Then there exists $g'' \in o_i(g_i)$, where $g''^{[\ell(g_i)]}(h'')$, such that s is reaching g'', which implies that $s \in S(g_i)$. □

Clearly, the relation $\supseteq$ on $\mathcal{C}_i$ inherited from the relation on personal histories makes $(\supseteq, \mathcal{C}_i)$ a tree, since such is $(\overline{H}_i, \preceq)$. ■