ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

LIMITS TO RATIONAL LEARNING

Yehuda John Levy

Number 731
November 2014

Manor Road Building, Manor Road, Oxford OX1 3UQ

Limits to Rational Learning

Yehuda John Levy*†

November 5, 2014

Abstract

A long-standing open question raised in the seminal paper of Kalai and Lehrer (1993) is whether or not the play of a repeated game, in the rational learning model introduced there, must eventually resemble play of *exact* equilibria, and not just play of *approximate* equilibria as demonstrated there. This paper shows that play may remain distant - in fact, mutually singular - from the play of any equilibrium of the repeated game. We further show that the same inaccessibility holds in Bayesian games, where the play of a Bayesian equilibrium may continue to remain distant from the play of any equilibrium of the true game.

1 Introduction

The premise of *rational learning* is that decision-making agents update their beliefs about what other agents will do based on the actions that they have observed. The seminal work [Blackwell and Dubins (1962)] shows that when a single agent learns rationally in this way, even if his prior beliefs about the process are incorrect but do contain a minimal amount of truth, his posterior beliefs will eventually lead to true beliefs about the process. The staple work in game theory which incorporates this paradigm into the multi-agent setting is [Kalai and Lehrer (1993)], in which agents both learn and also try to maximise their utility in a repeated game. Their work studies the question of whether the beliefs of the agents not only merge, but converge to beliefs induced by the *Nash equilibria* of the repeated game. As it turns out, the answer is, not necessarily.

[Kalai and Lehrer (1993)] was not the only work at that time to explore questions of convergence under rational learning; in fact, other influential works that both support it and contrast with it were carried out around the same time and in the following years. [Jordan (1991)], for example, shows that under appropriate assumptions, rational learning converges to the set of stage game Nash

*Nuffield College and Department of Economics, Oxford University, john.calculus@gmail.com

†The author expresses thanks to Peyton Young for his helpful remarks and for bringing to the author's attention the conjecture this paper address, and to Ron Peretz for his help with the probabilistic tools this paper utilizes.

equilibria. However, [Kalai and Lehrer (1993)] struck on a property of learning discussed very little at the time: the convergence of the agents' strategies to (approximate) equilibria of the *repeated game*, not of the strategic game that is being repeated. While their assumptions that the players' beliefs all contain a grain of truth[1] have been scrutinised as both an over-demanding coordination requirement (e.g., [Miller and Sanchirico (1999)]) and a highly non-generic condition (e.g., [Miller and Sanchirico (1997)]), the resulting body of literature extending, discussing, and contrasting [Kalai and Lehrer (1993)] "is, in many respects, a natural successor to the earlier literature on learning rational expectations" as "both literatures address the question of whether decision-makers can, through repeated experience, learn to make optimal or equilibrium decisions", [Jordan (1993)].

Two main veins of subsequent work exist. One direction generalises the results of [Kalai and Lehrer (1993)], some works by weakening the absolutely continuity assumptions on the beliefs, as in [Sandroni (1998)] and [Norman (2012)], others by weakening assumptions on the players' knowledge, as in [Kalai and Lehrer (1995)], [Jordan (1995)], and [Nyarko (1998)], and still other variations, e.g., [Gilli (2001)] and the references there. Another direction, however, was to point out the limitations of the assumptions and results, as in the papers [Miller and Sanchirico (1997)] and [Miller and Sanchirico (1999)], [Nachbar (1997)], and [Foster and Young (2001)].

The contribution of this paper is to answer a long-standing open question raised in [Kalai and Lehrer (1993), Sec 7.1]. The classical result of that paper ensures convergence of the play to the set of *approximate* equilibria (i.e., ε-equilibria) of the repeated game. (It is clear that one cannot in general expect convergence to a specific equilibrium or approximate equilibrium, as players may, for example, rotate among different equilibria.) The authors raise, but leave open, the question of whether the play must converge to the set of *exact* equilibria? In this paper we show by example that this need not be the case. Furthermore, not only does convergence fail to occur, but the play induced by any Nash equilibrium remains far - in fact, mutually singular - from actual play.

A related problem, discussed in [Kalai and Lehrer (1993), Sec 6], is learning in Bayesian games. In these games, each player is privately assigned a type - types are chosen independently - and a player's payoff may depend on his own type. Players can then condition their actions both on the public play so far and on their type. [Kalai and Lehrer (1993), Sec 6] shows that as time goes by, play must converge to the set of the *approximate* equilibria of the true game. In other words, players begin to play as we would expect in ε-equilibria *as if they were learning the types of the other players*, even though they need not be.[2] We also show in our paper that this result cannot be strengthened to deduce convergence to the set of plays induced by equilibria of the true game.

[1] Or somewhat weaker absolute continuity conditions.

[2] E.g., in the case that a player does not condition his actions on his type, the others would never actually garner any information about his type.

The rational learning model is presented formally in Section 2. A brief informal overview of the construction is given in Section 3. The stage game of our example is presented in Section 4, while the strategies and beliefs are given in Section 5. In order to prove that convergence to Nash equilibrium does not occur, Section 6 contains a preliminary result, while the proof itself is given in Section 7. Section 8 presents the Bayesian game. Section 9 presents a slightly different example, with its own virtues and disadvantages. Some probabilistic tools appear in Appendix A, while proofs from Section 6 appear in Appendix B.

2 The Rational Learning Model

Let $\mathcal{P}$ be a set of players with finite action spaces $(A^k)_{k\in\mathcal{P}}$, and let G be a strategic game on this set of players with payoff functions $r = (r^k)_{k\in\mathcal{P}}$. Let $\overline{A} = \prod_{k\in\mathcal{P}} A^k$ be the set of pure action profiles - hence $r^k : \overline{A} \to \mathbb{R}$ for each $k \in \mathcal{P}$ - and for $T = 0, 1, 2, \ldots, \infty$, let $H_{T+1} = \overline{A}^T$ be the collection of histories of the T-stage repeated game[3] (with $H_1 = \{\emptyset\}$). Denote $H_* = \underset{t<\infty}{\cup} H_t$.

A *behavioural strategy* for Player $k \in \mathcal{P}$ is a mapping[4] $\sigma^k : H_* \to \Delta(A^k)$. A profile of strategies $\sigma = (\sigma^k)_{k\in\mathcal{P}}$ induces a measure P_σ (and associated expectations operator E_σ) in the T-stage repeated games in the usual way, i.e., on H_{T+1}, for each $T = 0, 1, 2, \ldots, \infty$, defined by $P_\sigma(a_1, \ldots, a_t) = \prod_{s<t} \sigma(a_1, \ldots, a_{s-1})[a_s]$. The payoff in the infinitely repeated game G^∞ is given by

$$\overline{r}(a_1, a_2, \ldots) = \sum_{t=1}^{\infty} \beta^{t-1} r(a_t)$$

where $0 < \beta < 1$ is a fixed discount factor. For $\varepsilon \geq 0$, a strategy profile σ of G^∞ is an ε-*equilibrium* (or just *equilibrium* when $\varepsilon = 0$) if for each player $k \in \mathcal{P}$ and each strategy τ of Player k,

$$E_\sigma[\overline{r}^k] + \varepsilon \geq E_{(\tau,(\sigma^j)_{j\neq k})}[\overline{r}^k]$$

For each $h = (a_1, \ldots, a_T) \in H_*$ and $t \leq T$, denote $h|_t = (a_1, \ldots, a_{t-1})$, and for $k \in \mathcal{P}$ and strategy σ^k of Player k, let σ^k_h be the strategy defined by $\sigma^k_h(h') = \sigma^k(h, h')$, where for $h' = (a'_1, \ldots, a'_t)$ we have $(h, h') = (a_1, \ldots, a_T, a'_1, \ldots, a'_t)$, and similarly for profiles of strategies. To recall the results of [Kalai and Lehrer (1993)] and to state our results clearly, we introduce the following concepts:

Definition 2.1. *Let $(\tau^{j,k})_{j,k\in\mathcal{P}}$ be a $\mathcal{P} \times \mathcal{P}$ collection of strategies, where $\tau^{j,k}$ is a strategy of Player k (interpreted as the belief of Player j about Player k's strategy). We say that in $(\tau^{j,k})_{j,k\in\mathcal{P}}$ each player is best-replying to his beliefs*

[3]It is convenient to denote histories of the T-stage game as H_{T+1}, since we will view this set as the space of plays proceeding stage $T + 1$.

[4]For a set X, $\Delta(X)$ denotes the space of probability measures on X.

if for each $j \in \mathcal{P}$, $\tau^{j,j}$ is a best reply to $(\tau^{j,k})_{k\neq j}$, i.e., for each strategy π of Player j,

$$E_{(\tau^{j,k})_{k\in\mathcal{P}}}[\overline{r}^j] \geq E_{(\pi,(\tau^{j,k})_{k\neq j})}[\overline{r}^j]$$

Given such $\tau = (\tau^{j,k})_{j,k\in\mathcal{P}}$, by P_τ we mean the distribution induced by the diagonal $(\tau^{j,j})_{j\in\mathcal{P}}$, interpreted as the strategies which are actually played (since each player knows his own strategy).

For a measurable space $(\Omega, \mathcal{B})$, the set of probability measures on Ω is denoted $\Delta(\Omega)$, and for each $\mu, \nu \in \Delta(\Omega)$, the *total variation distance* between μ, ν is:

$$||\mu - \nu|| = 2 \cdot \sup_{A\in\mathcal{B}} |\mu(A) - \nu(A)|$$

μ is *absolutely continuous w.r.t.* ν, denoted $\mu \ll \nu$, if $\nu(A) = 0$ implies $\mu(A) = 0$. μ and ν are mutually singular if there is $A \in \mathcal{B}$ such that $\mu(A) = \nu(\Omega\backslash A) = 1$. Observe that $\mu\perp\nu$ implies $||\mu - \nu|| = 2$.

Definition 2.2. *Let P, Q be probability measures (on the same space); then Q contains a grain of truth of P if $P \ll Q$ and the Radon-Nikodym derivative $\frac{dP}{dQ}$ is bounded; equivalently, if for some $0 < \lambda < 1$ and some probability measure P', $Q = \lambda P + (1-\lambda)P'$.*

If Ω is any set and $f : \Omega \to \mathbb{R}^N$, denote $||f||_\infty = \sup_{\omega\in\Omega, n\leq N} |f^n(\omega)|$. We use this notation also for $p, q \in \Delta(\Omega)$ for finite Ω: $||p - q||_\infty = \max_{\omega\in\Omega} |p[\omega] - q[\omega]|$.[5]

The following is the main result of [Kalai and Lehrer (1993)]:

Theorem 2.3. *Let $\tau = (\tau^{j,k})_{j,k\in\mathcal{P}}$ be such that everyone is best-replying to their beliefs. Denote $\tilde{\tau}^j = (\tau^{j,k})_{k\in\mathcal{P}}$, the beliefs of Player j. Assume that for each $j \in \mathcal{P}$, $P_\tau \ll P_{\tilde{\tau}^j}$. Then for P_τ-a.e. $h \in H_\infty$ and any $\varepsilon > 0$, there exists $T \in \mathbb{N}$ such that for all $t \geq T$, there is an ε-equilibrium[6] σ of G^∞ such that $||P_{\tau_{h|_t}} - P_\sigma|| < \varepsilon$.*

It should be emphasised that this result does not say that there exists any particular ε-equilibrium σ such that from *any* late enough time period t, the play induced in the game after t stages is close to the play induced by σ.[7] The result says that the distribution of play induced after enough time is *close to the set* of possible distributions of play induced by all ε-equilibria. That is, at some periods, the induced play may be close to the play induced by one ε-equilibrium, but at other periods, close to the play induced by another.

In Section 7.1 of [Kalai and Lehrer (1993)], the following question is posed: Can Theorem 2.3 be strengthened to require that the process converges to the

[5]Hence, on $\Delta(\Omega)$ for finite Ω, we have both the supremum norm $||\cdot||_\infty$ and the total variation norm $||\cdot||$.

[6]The choice of ε-equilibrium σ may depend on t.

[7]Indeed, if the stage game possesses multiple equilibria and if the strategies alternate between these equilibria, we will would never get such convergence.

set of *exact equilibria* (that is, 0-equilibria) of the G^∞, and not just the set of ε-equilibria? This question has remained open until now. The purpose of this paper is to present a counter-example, showing that we may not, in general, be able to guarantee that play of the game will eventually resemble the play of an *exact* equilibrium. To be more precise:

We construct a game G and $(\tau^{j,k})_{j,k\in\mathcal{P}}$ in which everyone is best-replying to their beliefs, such that for each $j \in \mathcal{P}$, $P_\tau \ll P_{\tilde{\tau}^j}$ - in fact, the beliefs contain a grain of truth - and such that for P_τ-a.e. $h \in H_\infty$, for each $t \in \mathbb{N}$, and each equilibrium σ of G^∞, $P_{\tau_{h|_t}} \perp P_\sigma$.

A particular type of beliefs arise from *repeated Bayesian games* (a.k.a. *games of incomplete information*): Each player has a discrete type space, $(T^k)_{k\in\mathcal{P}}$. At the beginning of play, types are chosen by Nature before play begins independently - i.e., via a commonly known product distribution $\mu = \prod_{k\in\mathcal{P}} \mu_k$. Each player is informed of his own type: Hence, a strategy for Player k is then a mapping $\eta^k : H_* \times T^k \to \Delta(A^k)$, and a profile $\eta = (\eta^k)_{k\in\mathcal{P}}$ of such strategies together with the prior μ induces a measure P_η on $H_\infty \times \prod_{k\in\mathcal{P}} T^k$. The payoff of Player k may include dependence on the type of Player k (and only his type), and hence is a function $r^k : T^k \times \overline{A} \to \mathbb{R}$. Let $G^\infty(\mu)$ denote this game of incomplete information, and for each profile of types $(t^k)_{k\in\mathcal{P}} \in \prod_{k\in\mathcal{P}} T^k$, let $G((t^k)_{k\in\mathcal{P}})$ be the strategic game resulting from this selection of types (i.e., the one-shot game in which the types $(t^k)_{k\in\mathcal{P}}$ are selected and made public), and let $G^\infty((t^k)_{k\in\mathcal{P}})$ be the game in which the strategic-form game $G((t^k)_{k\in\mathcal{P}})$ is repeated infinitely many times. Defining the payoffs also to be the discounted sum of the stream of the stage payoffs, Bayesian equilibrium is defined in the same way - if no player can increase his expected payoff by deviating.[8]

Theorem 2.4. *[[Kalai and Lehrer (1993)] , Sec. 6] Let $\eta = (\eta^j)_{j\in\mathcal{P}}$ be a Bayesian equilibrium of $G^\infty(\mu)$. Then for each $\varepsilon > 0$, for μ-a.e. choice of types $(t^k)_{k\in\mathcal{P}} \in \prod_{k\in\mathcal{P}} T^k$ of Nature, for $P_\eta(\cdot \mid (t^k)_{k\in\mathcal{P}})$-a.e. play $h \in H_\infty$, there is a time T such that for each $t \geq T$, there is an ε-equilibrium $\sigma = (\sigma^j)_{j\in\mathcal{P}}$ of $G^\infty((t^k)_{k\in\mathcal{P}})$, such that $||P_{\eta_{h|_t}}(\cdot \mid (t^k)_{k\in\mathcal{P}}) - P_\sigma|| < \varepsilon$.*

In this paper:

We construct a repeated Bayesian game $G^\infty(\mu)$ with finite type spaces $(T^k)_{k\in\mathcal{P}}$ and a Bayesian equilibrium $(\eta^k)_{k\in\mathcal{P}}$ of $G^\infty(\mu)$, such that for some choice of types $(t^k)_{k\in\mathcal{P}} \in \prod_{k\in\mathcal{P}} T^k$ by Nature and $P_\eta(\cdot \mid (t^k)_{k\in\mathcal{P}})$-a.e. play $h \in H_\infty$, for each $t \in \mathbb{N}$, and each equilibrium σ of $G^\infty((t^k)_{k\in\mathcal{P}})$, $P_{\eta_{h|_t}}(\cdot \mid (t^k)_{k\in\mathcal{P}}) \perp P_\sigma$.

We raise some questions worthy of further investigation but that we shall not attempt to answer here. The first issue is that in the examples presented in

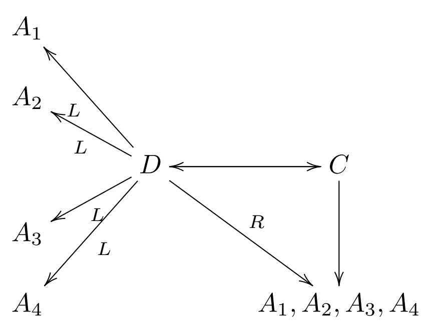
[8]Since the type space is discrete, we needn't differentiate between ex-ante deviations - i.e., deviations proceeding the selections of types - or ex-post deviations - i.e., deviations following the selection of types. If the type space were more general, technical measurability issues arise. However, Theorem 2.4 does not remain correct in such a general framework anyway.

this paper, the payoffs will be seen to be highly non-generic, and their success in evading rational learning converging to equilibria is dependent precisely on how they are defined. Hence, it is natural to inquire: Does convergence to the set of equilibria occur for generic games, i.e., that the set of stage games for which this strengthened convergence does not hold is small in an appropriate sense (first category, measure zero, etc)?

A second issue concerns our Bayesian game, presented in Section 8, in which there may not be convergence to the set of equilibria of the true game. The lack of convergence in that example is only for some of the type profiles that can arise. One can ask if this limitation of our example is fundamental: Must convergence in Bayesian games occur for a set of type profiles of positive probability, or is it possible to find examples where convergence does not occur for almost any type profiles?

3 Informal Construction Overview

- The game has six players, $\mathcal{P} = \{A_1, A_2, A_3, A_4, C, D\}$. Each player has two actions, L and R.
- C and D play a coordination game between them. When D plays L, everyone is incentivised to play L as well. When D plays R, A_1, A_2, A_3, A_4 play a sort of anti-coordination game among themselves - which also depends on C - and it is a mixed equilibrium of this game we will focus on. The equilibrium in which all play L is far preferred by all players.
- The influence of the players on each other can be viewed via the following graph:



- When C and D coordinate on (R, R) for the n-th time, players A_1, A_2, A_3, A_4 will incorrectly put some small probability δ_n^3 on C playing L.[9] As a result, equilibria for A_1, A_2, A_3, A_4 are mixed. We will concentrate on the equilibrium where each plays L with probability $\frac{1}{2} + \delta_n$.

[9] This can be justified, as we will do later in the framework of Bayesian games, by believing that with small probability C is actually indifferent between his actions and that he puts positive probability on both.

- Although the beliefs about C's action contain only a small error - on the order of δ_n^3 the n-th time C and D coordinate on R - this leads to a much larger - order of δ_n - deviation in equilibrium strategies for A_1, A_2, A_3, A_4. Although (δ_n) will converge to zero quickly enough to guarantee that the beliefs contain a grain of truth, they will shrink slowly enough so that if an equilibrium *of the stage game* is used *at each stage*, the play of the game must be far (in fact, mutually singular) to the true play.

- Thus, to avoid equilibria in the repeated game in which players do not play equilibria in the stage game - by implementing intertemporal dependence and coordination - we have in our strategies and beliefs in between such stages at which C and D coordinate on (R, R) long periods in which they coordinate on (L, L), hence incentivising A_1, A_2, A_3, A_4 to play L as well. Hence, at those stages in which C and D coordinate on (R, R), play must be very close to an equilibrium of the stage game, since all know that the next time any player has an incentive to play anything other than L - which, recall, is a much-preferred outcome for all - will be so far into the future and hence so discounted, no player has any hope that a deviation could lead to a future profit.

4 The Stage Game

The stage game G has six players, which we denote $\mathcal{P} = \{A_1, A^2, A^3, A^4, C, D\}$. Each player has action set $A = \{L, R\}$. Let x be a mixed action profile. For $k \in \mathcal{P}$, we will write x^k instead of $x^k[L]$ (the probability that k plays L). For $j = 1, 2, 3, 4$, denote $\Delta^{-j}(x) = \prod_{k \neq j}(x^{A_k} - \frac{1}{2})$.

The payoff to Players C, D depends only on each other, and is a coordination game given by:

$r^{C,D}(\cdot, a^{-\{C,D\}}) =$

$C \backslash D$	L	R
L	$2,2$	$0,0$
R	$0,0$	$1,1$

For $j = 1, 2, 3, 4$, the payoff to A_j is given by:

$r^{A_j}(\cdot, x^{-A_j}) =$

	$a^D = L$	$a^D = R$
L	2	$x^C - \Delta^{-j}(x)$
R	0	0

This gives a well-defined payoff since the payoff for A^j depends in a multi-linear way on the other's actions.

Lemma 4.1. *For each $0 < \delta < 1$, there exists $\eta > 0$ such that there is no η-equilibrium z of G satisfying:*

- $z^C = z^D = R$.

- *For $j = 1, 2, 3, 4$, $\frac{1}{2} + \frac{1}{2}\delta \leq z^{A^j}$.*

Proof. Fix $0 < \delta < 1$. Suppose for each $\eta > 0$, there was an η-equilibrium z_η satisfying these conditions. Taking a limit of a subnet of $(z_\eta)_{\eta>0}$ would give an equilibrium y of G satisfying these properties, and in particular, for all $j = 1,2,3,4$, $z^C - \Delta^{-j}(z) < -\frac{1}{8}\delta^3 < 0$; hence, since $z^C = z^D = R$, we would have $z^{A^j} = R$ for all $j = 1,2,3,4$, a contradiction. □

It is immediate to verify:

Lemma 4.2. *The profile $\overline{L}$ with $\overline{L}^k = L$ for each $k \in \mathcal{P}$ is an equilibrium of G.*

5 The Strategies and Beliefs

Let $(\delta_n)_{n=1}$ be any positive sequence with

$$\sup_n \delta_n < \min[\frac{1}{4}, \varepsilon, \zeta^{-1}], \ \sum_{n=1}^{\infty} \delta_n^3 < \infty, \ \sum_{n=1}^{\infty} \delta_n^2 = \infty \tag{5.1}$$

where $\varepsilon, \zeta > 0$ are specified later in Proposition 6.1. Let $L_n = T_0(\delta_n)$, with $T_0(\cdot)$ also defined later in Proposition 6.1. Denote $S_n = 1 + \sum_{k<n} L_n$ (note that $S_1 = 1$). We define the following array of strategies $(\tau^{j,k})_{j,k\in\mathcal{P}}$; they are history-independent and hence we will not reference the history explicitly in the definition, i.e., we will denote $\tau_t^{j,k}$ instead of $\tau^{j,k}(h|_t)$. Recall also that we denote τ^k instead of $\tau^{k,k}$.

- $$\tau_t^C = \tau_t^D = \begin{cases} R & \text{if } \exists n \in \mathbb{N}, t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

- For each $j = 1,2,3,4$,

$$\tau_t^{A_j} = \begin{cases} \frac{1}{2} + \delta_n & \text{if } t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

- All players have correct beliefs about A_1, A_2, A_3, A_4, D: $\tau^{k,A_j} = \tau^{A_j}$, $\tau^{k,D} = \tau^D$, for $k \in \mathcal{P}$ and $j = 1,2,3,4$.

- However, for all $k \in \mathcal{P}$, $k \neq C$,

$$\tau^{k,C} = \begin{cases} \delta_n^3 & \text{if } t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

Recall that we denote the beliefs of Player $j \in \mathcal{P}$ by $\tilde{\tau}^j := (\tau^{j,k})_{k\in\mathcal{P}}$, and similarly we denote $\tilde{\tau}_t^j := (\tau_t^{j,k})_{k\in\mathcal{P}}$.

Lemma 5.1. *In τ, each player is best-replying to his beliefs.*

Proof. Because of the history independence, it's enough to check that for each $t \in \mathbb{N}$, each $k \in \mathcal{P}$, τ_t^k is a best-reply to his beliefs $\tilde{\tau}_t^k$ in the stage game G. If $\forall n, t \neq S_n$, then τ_t is just the equilibrium $\overline{L}$ from Lemma 4.2 and all players have the correct beliefs of the other players' action at stage t, i.e, $\tau_t^{j,k} = \tau_t^k$. Suppose $t = S_n$. We contend that the (A_j) are all indifferent; indeed,

$$r^{A_j}(R, (\tilde{\tau}_{S_n}^{A_j,k})_{k \neq A_j}) = \tau_{S_n}^{A_j,C} - \Delta^{-j}(\tilde{\tau}_{S_n}^{A_j}) = \delta_n^3 - ((\frac{1}{2} + \delta_n) - \frac{1}{2})^3 = 0$$

Furthermore, since $y^D = R$, C prefers R; and it's easy to see that D prefers R since $\tau^{D,C} = \delta_n^3 < \frac{1}{3}$. □

Lemma 5.2. *For each player $k \in \mathcal{P}$, $P_\tau \ll P_{\tilde{\tau}^k}$, and in fact, $P_{\tilde{\tau}^k}$ contains a grain of truth of P_τ.*

Proof. C has correct beliefs about all players; if $k \neq C$, then denoting

$$u_n^C = \delta_n^3, u^D = R, \forall j,\ u_n^{A_j} = \frac{1}{2} + \delta_n$$

Then one verifies

$$\tilde{\tau}_t^{A_j} = \tilde{\tau}_t^D = \begin{cases} u_n & \text{if } t = S_n \\ \overline{L} & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

where $\overline{L}$ is the equilibrium of the stage game given by $\overline{L}^k = L$ for all $k \in \mathcal{P}$. Let v_n denote the profile

$$v_n^C = v_n^D = R, \forall j,\ v_n^{A_j} = u_n^{A_j} = \frac{1}{2} + \delta_n \tag{5.2}$$

Then

$$\tau_t = \begin{cases} v_n & \text{if } t = S_n \\ \overline{L} & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

Since $\forall a \in \overline{A}$, $u_n[a] = 0$ implies $v_n[a] = 0$, by Corollary 10.3 and Proposition 10.4 it suffices to check that

$$\sum_{n=1}^{\infty} ||u_n - v_n||_\infty < \infty$$

and that there is some $\alpha > 0$ such that for all $t \in \mathbb{N}$ and all profile $a \in \overline{A}$ if $\tau_t[a] > 0$, then $\tau_t[a] \geq \alpha$. For the latter claim, take[10] $\alpha = (\frac{1}{4})^4$. For the former, observe that $||u_n - v_n||_\infty \leq \sum_{k \in \mathcal{P}} |u_n^k - v_n^k| = \delta_n^3$, and by assumption, $\sum_{n=1}^{\infty} \delta_n^3 < \infty$. □

[10] Since $\delta_n < \frac{1}{4}$.

6 Preliminary Repeated Game

Fix a discount factor $\beta > 0$. For $T \in \mathbb{N}$ and $g : H_{T+1} \to \mathbb{R}^{\mathcal{P}}$, let $\Gamma(T, g)$ be the game with payoff from play $h = (a_1, \ldots, a_T) \in H_{T+1}$ given by:

$$\overline{r}_g(h) = \sum_{t=1}^{T} \beta^{t-1} r(a_t) + \beta^T \cdot g(h)$$

I.e., the players play T rounds, and then at stage $T + 1$, get payoff $g(h)$ (in addition to the accrued payoffs) and the game is over; payoffs are discounted. g will be referred to as a *continuation payoff*. Also, for $t \leq T$,

$$\overline{r}_t[h] = \overline{r}_t[h|_{t+1}] = \sum_{s \leq t} \beta^{s-1} r(a_s) \tag{6.1}$$

Denote

$$H_\alpha = \{h = (a_1, \ldots, a_T) \in H_{T+1} \mid a_1^C = a_1^D = R \text{ and } \forall k \in \mathcal{P}, 2 \leq t \leq T, a_t^k = L\} \tag{6.2}$$

i.e., those histories in which Players C, D plays R in the first round, and after the first round, everyone plays only L. Note that H_α is finite, $|H_\alpha| = 2^4$. The following is the main proposition that we will require:

Proposition 6.1. *There exists $\varepsilon > 0$ such that for each $0 < \delta < \frac{1}{2}$, there exists $T_0 = T_0(\delta) \in \mathbb{N}$ such that if $T \geq T_0$ and g is a continuation payoff satisfying $||g||_\infty \leq \frac{||r||_\infty}{1-\beta}$, then there does not exist an equilibrium[11] profile σ of $\Gamma(T, g)$ which satisfies:*

- *For all $j = 1, 2, 3, 4$, $\frac{1}{2} + \frac{1}{2}\delta \leq \sigma^{A_j}(\emptyset)$.*
- *$P_\sigma(H_\alpha) > 1 - \varepsilon$ and $P_\sigma(h) > \frac{1}{2}\zeta^{-1}$ for each $h \in H_\alpha$, where $\zeta = 4^4$.*

Specifically, let $\varepsilon > 0$ and $T_1 \in \mathbb{N}$ satisfy:

$$\varepsilon < \frac{1}{8},\ \zeta\varepsilon(2 + ||r||_\infty)\frac{1}{1-\beta} < \frac{1}{4} \tag{6.3}$$

and

$$\beta^{T_1} \frac{||r||_\infty}{1-\beta} \leq \min[\frac{\eta}{2}, \frac{1}{8}] \tag{6.4}$$

where η corresponds to δ as in Lemma 4.1, and set $T_0 := 2 \cdot T_1$.

Suppose σ were such an equilibrium (for some $T \geq T_0$ and some continuation payoff g). The following lemmas are proved in Appendix B:

Lemma 6.2. *For all $h \in H_\alpha$ and all $2 \leq t \leq T - T_1$ and each player $k \in \mathcal{P}$, $\sigma^k(h|_t) = L$.*

[11] The notion of equilibrium extends to $\Gamma(T, g)$, with $\overline{r}_g$ replacing $\overline{r}$.

Lemma 6.3. *If $T \geq T_1$, then $\sigma^C(\emptyset) = \sigma^D(\emptyset) = R$.*

Lemma 6.4. *If $T \geq T_0$, then $\sigma(\emptyset)$ is a η-equilibrium of the stage game G, where η corresponds to δ as in Lemma 4.1.*

(Proof of Proposition 6.1). By Lemma 6.4, $z := \sigma(\emptyset)$ is an η-equilibrium of the stage game. Hence, we cannot have both the conditions listed in Lemma 4.1 holding. However, Lemma 6.3 shows indeed that $z^C = z^D = R$, while by assumption, $z^{A_j} \geq \frac{1}{2} + \frac{1}{2}\delta$ for each $j = 1, 2, 3, 4$. □

7 Proof of No Convergence to Equilibrium

Proposition 7.1. *For any equilibrium σ of G^∞, $P_\tau \perp P_\sigma$.*

Proof. For each $n \in \mathbb{N}$, let $X_n = (\{L, R\}^{\mathcal{P}})^{L_n}$ the possible sequences of play in the L_n-repeated game, where recall that $S_n = 1 + \sum_{k<n} L_k$. Then clearly $H_\infty = X := \prod_{n \in \mathbb{N}} X_n$, while $H_{S_n} = \overline{X}_n := \prod_{k<n} X_k$. (Essentially, we've partitioned the stages into blocks $X_1, X_2, \ldots$) We use the following notations:

For a strategy profile τ and its induced measure $P_\tau \in \Delta(X)$, let $(P_\tau)_n$ denote the marginal of P_τ on $\overline{X}_{n+1} = H_{S_{n+1}}$, and let $(P_\tau)_n[\cdot|\cdot]$ be the conditional on X_n w.r.t. $\overline{X}_n$: I.e., $(P_\tau)_n[\cdot \mid \overline{x}_n]$ is the distribution on X_n given $\overline{x}_n = (x_1, \ldots, x_{n-1}) \in \overline{X}_n$; if $P_\tau(\overline{x}_n) = 0$, then these conditional distributions are arbitrary, and any specifications of it are called a *version* of $(P_\tau)_n[\cdot \mid \cdot]$. (Referring back to our game, $(P_\tau)_n$ is the distribution on the L_n-block of the repeated game, given the play of the proceeding blocks of the sizes $L_1, \ldots, L_{n-1}$.)

Now in our case, for our specific strategy profile τ we have the version of $(P_\tau)_n[\cdot \mid \cdot]$ given by $(P_\tau)_n[\cdot \mid \cdot] \equiv P_{\rho_n}$, where ρ_n is a strategy profile in L_n stage game where $\rho_n(\emptyset) = v_n$ and $\rho_n(h) = \overline{L}$ if $h \neq \emptyset$, with v_n defined in (5.2), and $\overline{L}^k = L$ for all $k \in \mathcal{P}$; i.e., $P_{\rho_n} = v_n \otimes_{k=2}^{L_n} \overline{L}$.

On the other hand, for $\overline{x}_n \in \overline{X}_n = \prod_{k<n} X_k = H_{S_n}$, letting σ be an equilibrium of G^∞, we have the version of $(P_\sigma)_n[\cdot \mid \overline{x}_n]$ given by:

- If $P_\sigma(\overline{x}_n) > 0$, then $(P_\sigma)_n[\cdot \mid \overline{x}_n] = P_{\pi\{\overline{x}_n\}}$, where $\pi\{\overline{x}_n\}$ is the strategy profile in the L_n-stage game defined by $\pi\{\overline{x}_n\}(h) = \sigma_{\overline{x}_n}(h)$ for $h \in \bigcup_{t=1}^{L_n} H_t$.
- If $P_\sigma(\overline{x}_n) = 0$, $\pi\{\overline{x}_n\}$ is chosen to be $\pi\{\overline{y}_n\}$ for some $\overline{y}_n$ with $P_\sigma(\overline{y}_n) > 0$.

Hence, for all $\overline{x}_n \in \overline{X}_n$, $\pi\{\overline{x}_n\}$ is an equilibrium[12] in the game $\Gamma(L_n, g_{\overline{x}_n})$ for the continuation payoff $g_{\overline{x}_n}(h) = E_{\sigma(\overline{x}_n, h)}[\overline{r}]$, which satisfies $||g_{\overline{x}_n}||_\infty \leq ||\overline{r}||_\infty \leq \frac{||r||_\infty}{1-\beta}$.

[12]Indeed, if σ is an equilibrium and $h \in H_*$ with $P_\sigma(h) > 0$, then σ_h is also an equilibrium.

Fix $\overline{x}_n \in \overline{X}_n$, and denote π for brevity instead of $\pi\{\overline{x}_n\}$. Proposition 6.1 and the fact that $L_n = T_0(\delta_n)$ then implies that one of the following holds:

(i) For some $j = 1, 2, 3, 4$, $\pi_n^{A_j}(\emptyset) < \frac{1}{2} + \frac{1}{2}\delta_n$.

(ii) $P_\pi(H_\alpha) \leq 1 - \varepsilon$, where H_α is defined just before Proposition 6.1 (for the L_n-stage game).

(iii) $P_\pi(h) \leq \frac{1}{2}\zeta^{-1}$ for some $h \in H_\alpha$.

Going over these case-by-case:

(i) In this case,

$$|\pi^{A_j}(\emptyset) - \rho_n^{A_j}(\emptyset)| = |\pi^{C_j}(\emptyset) - (\frac{1}{2} + \delta_n)| \geq \frac{\delta_n}{2}$$

so $||P_{\rho_n} - P_\pi|| \geq \frac{\delta_n}{2}$.

(ii) $P_\pi(H_\alpha) \leq 1 - \varepsilon$ while $P_{\rho_n}(H_\alpha) = 1$, hence $||P_{\rho_n} - P_\pi|| \geq 2\varepsilon \geq \delta_n$.

(iii) $P_\pi(h) \leq \frac{1}{2}\zeta^{-1}$, while $P_{\rho_n}(h) \geq \zeta^{-1}$; hence $||P_{\rho_n} - P_\pi|| \geq \zeta^{-1} \geq \delta_n$.

Hence, $||P_{\rho_n} - P_\pi|| \geq \frac{\delta_n}{2}$. Since $P_{\rho_n}(H_\alpha) = 1$ and $|H_\alpha| = 2^4$, we have[13]

$$||(P_\sigma)_n[\cdot|\overline{x}_n] - (P_\tau)_n[\cdot|\overline{x}_n]||_\infty = ||P_{\rho_n} - P_\pi||_\infty \geq \frac{1}{2 \cdot |H_\alpha|}||P_{\rho_n} - P_\pi|| \geq \frac{\delta_n}{2^5}$$

This was for any $\overline{x}_n \in X^*$. Hence, by Theorem 10.7, since $\sum_{n=1}^{\infty} \delta_n^2 = \infty$, $P_\tau \perp P_\sigma$. □

Corollary 7.2. *For P_τ-a.e. $h \in H_\infty$, for all $t \in \mathbb{N}$, $P_{\tau_{h|_t}} \perp P_\sigma$ for any equilibrium σ of G^∞.*

Proof. Fix $h \in H_\infty$ and $t \in \mathbb{N}$ with $P_\tau(h|_t) > 0$ and an equilibrium σ of G^∞. Denote $\tau' = \tau_{h|_{S_n}}$. We will show that $P_{\tau'} \perp P_\sigma$.

First take the case that $t = S_n$ for some n. In this case, one simply observes that $P_{\tau'}$ is induced by the sequence $(L_n, L_{n+1}, \ldots)$ in the same way that τ is induced by the sequence $(L_1, L_2, \ldots)$, and applies Proposition 7.1 (since the sequence $(\delta_n, \delta_{n+1}, \ldots)$ also satisfies (5.1), and $L_n = T_0(\delta_n)$).

[13] If $\mu, \nu \in \Delta(A)$ with A finite and ν supported on a set B of size M, then $||\mu - \nu||_\infty \geq \frac{1}{2M}||\mu - \nu||$. This is because

$$\mu(A \backslash B) = |\mu(A \backslash B) - \nu(A \backslash B)| = |\mu(B) - \nu(B)| \leq \sum_{a \in B} |\mu[a] - \nu[a]| \leq M \cdot ||\mu - \nu||_\infty$$

so

$$||\mu - \nu|| = \sum_{a \in A} |\mu[a] - \nu[a]| = \sum_{a \in B} |\mu[a] - \nu[a]| + \mu(A \backslash B) \leq 2M||\mu - \nu||_\infty$$

Suppose now $S_n < t < S_{n+1}$. Denote $T = S_{n+1} - t$. Fix some $h' \in H_{T+1}$ (hence, $(h|_t, h') \in H_{S_{n+1}}$) such that $P_{\tau'}(h') > 0$ and also $P_\sigma(h') > 0$. Let $\tau'' = \tau'_{h'}$, $\sigma' = \sigma_{h'}$; then σ' is an equilibrium of G^∞ since $P_\sigma(h') > 0$, and τ'' is the induced by the sequence $(L_{n+1}, L_{n+2}, \ldots)$ in the same way that τ is induced by the sequence $(L_1, L_2, \ldots)$. Hence, like the case above, $P_{\tau''} \perp P_{\sigma'}$, i.e. $P_{\tau'_h} \perp P_{\sigma_h}$, and therefore $P_{\tau'}(\cdot \mid h') \perp P_\sigma(\cdot \mid h')$. To sum up, for each $h' \in H_{T+1}$ such that $P_{\tau'}(h') > 0$ and also $P_\sigma(h') > 0$, we have $P_{\tau'}(\cdot \mid h') \perp P_\sigma(\cdot \mid h')$. Hence, from Proposition 10.8, $P_{\tau'} \perp P_\sigma$, as required. □

8 The Bayesian Game

Relying on the payoffs r defined in Section 4 and the strategies and beliefs $\tau = (\tau^{j,k})_{j,k \in \mathcal{P}}$ defined in Section 5, we now define a Bayesian game:

- First, let $(\delta_n)_{n=1}^\infty$ be a positive sequence satisfying, in addition to (5.1), the condition
$$\prod_{k=1}^{\infty} \frac{1}{1 - \delta_k^3} \leq \frac{4}{3} \tag{8.1}$$
(E.g., take $\delta_n = \frac{1}{\sqrt{n+M}}$, for large enough $M > 0$.)

- The set of players $\mathcal{P} = \{C, D, A^1, A^2, A^3, A^4\}$ is the same as before, each with actions $\{L, R\}$.

- Player C has two types, which we denote $t^C =\uparrow$ or $t^C =\downarrow$, and which are chosen with equal likelihood. All other players can each be of one type only, and hence we drop reference to their types.

- For a player $k \neq C$, the payoff ρ^k satisfies $\rho^k = r^k$, while
$$\rho^C(\uparrow, \cdot) = r^C(\cdot), \;\; \rho^C(\downarrow, \cdot) \equiv 0$$

- Define a positive sequence $(p_n)_{n=1}^\infty$ by:
$$p_{n+1} = \begin{cases} \frac{1}{2} & \text{if } n = 0 \\ \frac{p_n}{1-\delta_n^3} & \text{if } n > 0 \end{cases} \tag{8.2}$$

- Now, define the strategies $(\eta^j)_{j \in \mathcal{P}}$ in the following way: $\eta^k = \tau^k$ for $k \neq C$, while
$$\eta_t^{A_j}(t^{A_j}, h|_t) = \begin{cases} L & \text{if } \forall n, t \neq S_n \\ R & \text{if } t = S_n, t^{A_j} =\uparrow \\ \frac{\delta_n^3}{1-p_n} & \text{if } t = S_n, t^{A_j} =\downarrow \end{cases} \tag{8.3}$$

Lemma 8.1. $\frac{1}{2} \leq p_n \leq \frac{2}{3}$ *for all* $n \in \mathbb{N}$.

Hence, since $\delta_n^3 < \frac{1}{3}$, the strategies given by (8.3) are well-defined.

Proof. This follows by (8.1), since one shows inductively that:

$$\frac{1}{2}\prod_{k\leq n}\frac{1}{1-\delta_k^3}=p_{n+1}$$

□

Lemma 8.2. *$p_n = P_\eta(t^C =\uparrow \mid R_n)$ where*

$$R_n = \{\forall m < n, a^C_{S_m} = R \textit{ and } \forall t < S_n \textit{ s.t. } \forall m, t \neq S_m, a^C_t = L\} \tag{8.4}$$

I.e., the probability that the other players associate with the type of C being $\uparrow$, given that only R has been played by him at each stage of the form S_k for $k < n$ and L at all other stages, is p_n. Note that $R_1 = \{\emptyset\}$.

Proof. Observe that by the definition of η,

$$P_\eta(R_{n+1} \mid t^C =\downarrow \cap R_n) = 1 - \frac{\delta_n^3}{1-p_n},\ P_\eta(R_{n+1} \mid t^C =\uparrow \cap R_n) = 1$$

Assume inductively $p_n = P_\eta(t^C =\uparrow \mid R_n)$: This clearly holds for $n = 1$, since $P_\eta(t^C =\uparrow \mid \{\emptyset\}) = P_\eta(t^C =\uparrow) = \frac{1}{2}$. Using Bayes rule, since $R_{n+1} \subseteq R_n$,

$$\begin{aligned} &P_\eta(t^C =\uparrow \mid R_{n+1}) \\ &\quad= \frac{P_\eta(R_{n+1} \mid t^C =\uparrow \cap R_n)\cdot P_\eta(t^C =\uparrow \mid R_n)}{P_\eta(R_{n+1} \mid t^C =\uparrow \cap R_n)\cdot P_\eta(t^C =\uparrow \mid R_n) + P_\eta(R_{n+1} \mid t^C =\downarrow \cap R_n)\cdot P_\eta(t^C =\downarrow \mid R_n)} \\ &\quad= \frac{1\cdot p_n}{1\cdot p_n + (1-\frac{\delta_n^3}{1-p_n})\cdot(1-p_n)} = \frac{p_n}{1-\delta_n^3} = p_{n+1} \end{aligned}$$

as required. □

Lemma 8.3. *For all n, and all player $k \neq C$, $\tau^{k,C}_{S_n} = P_\eta(a^k_{S_n} = L \mid R_n)$, where R_n is defined in (8.4), and if $t \in \mathbb{N}$ is not in $(S_n)_{n\in\mathbb{N}}$, then $\tau^{k,C}_t = P_\eta(a^k_{S_n} = L \mid h)$ for any $h \in H_t$.*

(Recall that τ is history-independent.)

Proof. Indeed, by the previous lemma and the definition of $\tau^{k,C}_{S_n}$,

$$\begin{aligned} P_\eta(a^C_{S_n} = L \mid R_n) &= p_n \cdot \eta^C_{S_n}(\uparrow) + (1-p_n)\eta^C_{S_n}(\downarrow) \\ &= p_n \cdot 0 + (1-p_n)\frac{\delta_n^3}{1-p_n} = \delta_n^3 = \tau^{k,C}_{S_n} \end{aligned}$$

The second part of the lemma is clear from (8.3). □

Corollary 8.4. *$(\eta^k)_{k\in\mathcal{P}}$ is a Bayesian equilibrium, and P_τ is precisely the marginal of $P_\eta(\cdot \mid\uparrow)$ on H_∞.*

Proof. The first part follows by combining Lemma 8.3 with the fact that under $(\tau^{j,k})_{j,k}$ each player was best-replying to his beliefs, while the second part follows by the definition of η. $\square$

Given what we have already proven for τ in the game with payoff r in Section 7:

Corollary 8.5. *$P_\eta(\cdot \mid\uparrow)$-a.e. $h \in H_\infty$, for all $t \in \mathbb{N}$, and for all equilibrium σ of $G^\infty(\uparrow)$, we have $P_{\eta_{h|_t}}(\cdot \mid\uparrow) \perp P_\sigma$.*

Indeed, σ must also be an equilibrium of G^∞, for G defined in Section 4, since $G = G(\uparrow)$.

9 An Alternative Example

We present now an alternate example. The advantage of this example is that it is simpler in some respects, although shares a similar theme. A disadvantage of this example is that it cannot be modelled into the frame of Bayesian games, as we had done for the other construction in Section 8.[14] We do not present a complete proof that convergence to equilibria does not occur, rather we state parallels of Lemma 4.1 and Proposition 6.1, and the proof from there continues very similarly.

There are seven player, $\mathcal{P} = \{A, B_1, B_2, C_1, C_2, C_3, D\}$. Each has actions L, R. The payoffs are given by:

- $r^{C_1} \equiv r^{C_2} \equiv r^{C_3} \equiv r^A \equiv 0$. Denote $\Delta^0(x) = \prod_{j=1,2,3}(x^{C_j} - \frac{1}{2})$, $\Delta^A(x) = x^A - \frac{1}{2}$.
- For $j = 1, 2$,

$r^{B_j}(\cdot, x^{-B_j}) =$

	$x^D = L$	$x^D = R$
L	2	$\Delta^0(x) + (-1)^j \Delta^A(x)$
R	0	0

- $r^D(\cdot, x^{-D}) =$

L	$4x^A - 3$
R	0

Again, the profile $\overline{L}$ in which $\overline{L}^k = L$ for all $k \in \mathcal{P}$ is an equilibrium. The parallel of Lemma 4.1 is:

Lemma 9.1. *For each $0 < \delta < \frac{1}{4}$, there exists $\eta > 0$ such that there is no η-equilibrium z of G satisfying:*

[14] The reason is that, since types must be independent, it is always true that any two players must, at any given time, have the same belief about the type (and hence the actions) of any third player. In this example, this is not the case, and this is crucial for the success of the example.

- *For* $j = 1, 2, 3$, $\frac{1}{2} + \frac{1}{2}\delta \le z^{C_j} \le \frac{5}{8}$.
- $z^{B_1}, z^{B_2} \in [\delta, 1 - \delta]$.

Proof. As in the proof of Lemma 4.1, we would get an equilibrium y satisfying these conditions. Then $y^D = R$, and we have $\Delta^0(y) \neq 0$, so B_1, B_2 cannot be both indifferent between L, R. $\square$

Similar to Section 5, let $(\delta_n)_{n=1}$, $(L_n)_{n=1}^{\infty}$ be a sequences satisfying (5.1) and the definition after it, except with ε,ζ and the function T_0 being specified by Proposition 9.2 below instead of Proposition 6.1. Define:

- $$\tau_t^{C_1} = \tau_t^{C_2} = \tau_t^{C_3} = \begin{cases} \frac{1}{2} + \delta_n & \text{if } t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$
- $$\tau_t^D = \begin{cases} R & \text{if } \exists n \in \mathbb{N},\ t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$
- $$\tau^A = \begin{cases} \frac{1}{2} & \text{if } \exists n \in \mathbb{N},\ t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$
- If $k \neq B_1, B_2$ or $m \neq A$, $\tau^{k,m} = \tau^m$.
- However, for $j = 1, 2$,

$$\tau_t^{B_j, A} = \begin{cases} \frac{1}{2} - (-1)^j \delta_n^3 & \text{if } t = S_n \\ L & \text{if } \forall n \in \mathbb{N}, t \neq S_n \end{cases}$$

We leave it to the reader to show the parallels of Lemmas 5.1 and 5.2 - that is, that each player is best-replying to his beliefs in each stage, and that the beliefs possess grain of truth. Now, if one defines the T-stage version of this game with continuation payoffs as we did in Section 6, one can show in a similar manner that if we denote

$$H_\alpha = \{h = (a_1, \ldots, a_T) \in H_{T+1} \mid a_1^D = R \text{ and } \forall k \in \mathcal{P}, 2 \le t \le T, a_t^k = L\} \tag{9.1}$$

then:

Proposition 9.2. *There exists* $\varepsilon > 0$ *such that for each* $0 < \delta$ *small enough, there exists* $T_0 = T_0(\delta) \in \mathbb{N}$ *such that if* $T \ge T_0$ *and* g *is a continuation payoff satisfying* $||g||_\infty \le \frac{||r||_\infty}{1-\beta}$*, then there does not exist an equilibrium profile* σ *of* $\Gamma(T, g)$ *which satisfies:*

- *For all* $j = 1, 2, 3$, $\frac{1}{2} + \frac{1}{2}\delta \le \sigma^{C_j}(\emptyset) \le \frac{5}{8}$.
- $P_\sigma(H_\alpha) > 1 - \varepsilon$ *and* $P_\sigma(h) > \frac{1}{2}\zeta^{-1}$ *for each* $h \in H_\alpha$*, where* $\zeta = 4^6$.

Indeed, one proves parallels of Lemma 6.2 (only for the non-indifferent players D, B_1, B_2), Lemma 6.3 (only for Player D), and Lemma 6.4 (using η from Lemma 9.1 instead of Lemma 4.1); one then observes that if δ is small enough, $P_\sigma(h) > \frac{1}{2}\zeta^{-1}$ for each $h \in H_\alpha$ implies that $\sigma^{B_1}(\emptyset), \sigma^{B_2}(\emptyset) \in [\delta, 1-\delta]$, and then derives a contradiction using Lemma 9.1 just as in the proof of Proposition 6.1.

Following the proof of this proposition, the proof that convergence to equilibria does not occur proceeds in the same manner as in Section 7.

10 Appendix A: Probabilistic Tools

Let[15] $X_1, X_2, \dots$ be finite, $X = \prod_{n\in\mathbb{N}} X_n$, $\overline{X}_n = \prod_{k<n} X_k$. For a measure[16] $P \in \Delta(X)$, let P_n denote the marginal of P on $\overline{X}_{n+1}$. For $P, Q \in \Delta(X)$, we will say that P is *locally absolutely continuous w.r.t.* Q if $P_n \ll Q_n$ for all n. For each n, let $P_n[\cdot|\cdot]$ be the conditional on X_n w.r.t. $\overline{X}_n$: I.e., $P_n[\cdot \mid \overline{x}_n]$ is the distribution on X_n given $\overline{x}_n \in \overline{X}_n$. The distribution $P_n[\cdot \mid \overline{x}_n]$ is uniquely defined[17] if $P(\overline{x}_n) > 0$. A *version* of $P_n[\cdot \mid \cdot]$ is such a conditional distribution with $P[\cdot \mid \overline{x}_n]$ defined in this unique way when $P(\overline{x}_n) > 0$, and an arbitrary element of $\Delta(X_n)$ if $P(\overline{x}_n) = 0$.

The following is known as the *Kakutani dichotomy* (the version for products of finite spaces); see [Shiryaev (1995), Sec. VII.6], Theorem 4, which generalises [Kakutani (1948)].

Theorem 10.1. *Let $P, Q \in \Delta(X)$, and for each n, such that P is locally absolutely continuous w.r.t. Q. Then*

$$P \ll Q \iff P\Big[\sum_{n=1}^{\infty}\big(1 - \sum_{a\in X_n}\sqrt{P_n[a \mid \overline{x}_n]\cdot Q_n[a \mid \overline{x}_n]}\big) < \infty\Big] = 1$$

and

$$P \perp Q \iff P\Big[\sum_{n=1}^{\infty}\big(1 - \sum_{a\in X_n}\sqrt{P_n[a \mid \overline{x}_n]\cdot Q_n[a \mid \overline{x}_n]}\big) = \infty\Big] = 1$$

Note that the summand is well-defined - since if $Q(\overline{x}_n) = 0$ then $P(\overline{x}_n) = 0$ - and is seen to be non-negative.

Lemma 10.2. *For any finite set A, and any $p, q \in \Delta(A)$,*

$$\frac{1}{8}||p-q||_\infty^2 \le 1 - \sum_{a\in A}\sqrt{p[a]\cdot q[a]} \le \frac{1}{2}|A|\cdot||p-q||_\infty$$

[15] Some of the notation here is a repeat of Section 7.
[16] X is endowed with the Borel σ-algebra induced by the Tychonoff topology.
[17] For all $\overline{x}_n = (x_1, \dots, x_{n-1}) \in X^*$, $P(\overline{x}_n) = \prod_{k<n} P[x_k \mid \overline{x}_k]$.

Proof. The right-hand inequality follows since for $0 \le x, y$,

$$(\sqrt{y} - \sqrt{x})^2 = \Big(\frac{\sqrt{|y-x|}}{\sqrt{y}+\sqrt{x}}\Big)^2 \cdot |y-x| \le |y-x|$$

and therefore, since $\sum_{a\in A} p[a] = \sum_{a \in A} q[a] = 1$,

$$1 - \sum_{a\in A} \sqrt{p[a]\cdot q[a]} = \frac{1}{2}\sum_{a\in A}(\sqrt{p[a]} - \sqrt{q[a]})^2 \le \frac{1}{2}\sum_{a\in A}|p[a]-q[a]| \le \frac{1}{2}|A|\cdot||p-q||_\infty$$

For the left-hand inequality, for $0 \le x, y \le 1$ we have

$$|\sqrt{y} - \sqrt{x}| = |y-x|\frac{1}{\sqrt{y}+\sqrt{x}} \ge \frac{1}{2}\cdot|y-x|$$

and hence

$$1 - \sum_{a\in A} \sqrt{p[a]\cdot q[a]} = \frac{1}{2}\sum_{a\in A}(\sqrt{p[a]} - \sqrt{q[a]})^2 \ge \frac{1}{8}\sum_{a\in A}(p[a]-q[a])^2 \ge \frac{1}{8}||p-q||_\infty^2$$

□

Corollary 10.3. *Let $X_n \equiv A$ for all $n \in \mathbb{N}$, for some finite set A. Let $(p_t)_{t=1}^\infty$, $(q_t)_{t=1}^\infty$ be sequences in $\Delta(A)$, such that for all $t \in \mathbb{N}$, $p_t \ll q_t$. Let $P = \bigotimes_{t\in\mathbb{N}} p_t$, $Q = \bigotimes_{t\in\mathbb{N}} q_t$. If $\sum_{t=1}^\infty ||p_t - q_t||_\infty < \infty$, then $P \ll Q$.*

Proof. Clearly local absolute continuity of P w.r.t. Q holds. One applies the right-hand inequality of Lemma 10.2 to the criteria for absolute continuity given in Theorem 10.1. □

Proposition 10.4. *If, in Corollary 10.3, it in addition holds that for some $\alpha > 0$, for all t, $p_t[x] > 0$ implies $p_t[x] \ge \alpha$, then Q contains a grain of truth of P (i.e., $\frac{dP}{dQ}$ is bounded).*

Proof. For each element $\overline{x} \in X$ s.t. $P(\overline{x}_n) > 0$ for all n (which implies $Q(\overline{x}_n) > 0$ for all n), denote

$$T_n(\overline{x}) = \frac{P(\overline{x}_n)}{Q(\overline{x}_n)} = \prod_{k\le n}\frac{p_k[x_k]}{q_k[x_k]}$$

Then for Q-a.e. $\overline{x}$, $\frac{dP}{dQ}(\overline{x}) = \lim_{n\to\infty} T_n(\overline{x})$. Since $||p_n - q_n||_\infty \to 0$, by shrinking α and possibly disregarding finitely many t, we may assume that $p_t[x] > 0$ implies both $p_t[x] \ge \alpha$ and $q_t[x] \ge \alpha$. Then, for Q-a.e. $\overline{x}$,

$$\begin{aligned}\ln(T_n(\overline{x})) &\le \sum_{t=1}^\infty \sup_{x\in X_t, p_t[x]>0} \max(\ln(p_t[x]) - \ln(q_t[x]), 0)\\ &\le \frac{1}{\alpha}\sum_{t=1}^\infty \sup_{x\in X_t, p_t[x]>0}\max(p_t[x]-q_t[x], 0) \le \frac{1}{\alpha}\sum_{t=1}^\infty ||p_t - q_t||_\infty < \infty\end{aligned}$$

□

Corollary 10.5. *Let $P, Q \in \Delta(X)$, and for each n, let $P_n[\cdot|\cdot], Q_n[\cdot \mid \cdot]$ be versions of the marginals and suppose that for all $x \in X$,*

$$\sum_{n=1}^{\infty}\big(1 - \sum_{a\in X_n}\sqrt{P_n[a \mid \overline{x}_n]\cdot Q_n[a \mid \overline{x}_n]}\big) = \infty$$

Then $P \perp Q$.

Note that the power that Corollary 10.5 adds to the Kakutani criterion for the divergent case is that P need not be locally absolutely continuous w.r.t. Q.

Proof. For each n, let $\mu_n \in \Delta(X_n)$ have full support. Define $\psi_n : [0,1] \times (\Delta(X_n))^2 \to \mathbb{R}$ by

$$\psi_n(\varepsilon, p, q) = 1 - \sum_{a\in X_n}\sqrt{(1-\varepsilon)p[a] + \varepsilon\mu_n[a])((1-\varepsilon)q[a] + \varepsilon\mu_n[a])}$$

ψ_n is clearly continuous. Let ε_n be such that for all $\varepsilon \leq \varepsilon_n$, $\sup_{p,q\in\Delta(X_n)} |\psi_n(\varepsilon, p, q) - \psi_n(0, p, q)| < \frac{1}{2^n}$, and denote $\overline{\varepsilon} = (\varepsilon_n)$. Define new marginals $P'_n[\cdot \mid \cdot], Q'_n[\cdot \mid \cdot]$ by

$$P'_n[\cdot \mid \cdot] = (1-\varepsilon_n)P_n[\cdot \mid \cdot] + \varepsilon_n\mu_n,\ Q'_n[\cdot \mid \cdot] = (1-\varepsilon_n)Q_n[\cdot \mid \cdot] + \varepsilon_n\mu_n$$

and let P', Q' be the induced distributions on X with these conditional distributions. If needed, further shrink ε_n such that $\varepsilon_n \leq \frac{\varepsilon_1}{4^n|X_n|^2}$ - call such a sequence *quickly decreasing*. Observe that, since $|\sqrt{a} - \sqrt{b}| \leq \sqrt{|a-b|}$ and $|c\cdot d - c'\cdot d'| \leq d\cdot|c-c'| + c'\cdot|d-d'|$,we have

$$\Big|\sum_{a\in X_n}\sqrt{(P')_n[a \mid \overline{x}_n]\cdot(Q')_n[a \mid \overline{x}_n]} - \sum_{a\in X_n}\sqrt{P_n[a \mid \overline{x}_n]\cdot Q_n[a \mid \overline{x}_n]}\Big|$$
$$\leq \sum_{a\in X_n}\sqrt{|(P')_n[a \mid \overline{x}_n]\cdot(Q')_n[a \mid \overline{x}_n] - P_n[a \mid \overline{x}_n]\cdot Q_n[a \mid \overline{x}_n]|} \leq |X_n|\cdot\sqrt{2\varepsilon_n} \leq \frac{\sqrt{2\varepsilon_1}}{2^n}$$

Hence, for all $x \in X$,

$$\sum_{n=1}^{\infty}\big(1 - \sum_{a\in X_n}\sqrt{(P')_n[a \mid \overline{x}_n]\cdot(Q')_n[a \mid \overline{x}_n]}\big) = \infty$$

As such, $||P' - P|| \leq \varepsilon_1$, $||Q' - Q|| \leq \varepsilon_1$ by Lemma 10.6 below. Hence, taking a sequence $(\overline{\varepsilon}^n)_{n\in\mathbb{N}}$ of quickly decreasing sequences satisfying $\overline{\varepsilon}^n_1 \to 0$, we have $P' \to P$, $Q' \to Q$, and clearly, each pair P' and Q' are locally absolutely continuous w.r.t. each other. Applying the Kakutani criterion, we have $P' \perp Q'$, and hence[18] $P \perp Q$. □

[18]Indeed, this implies that for each $\varepsilon > 0$, there are disjoint A, B with $P(B) < \varepsilon$, $Q(A) < \varepsilon$; this easily implies $P \perp Q$.

Lemma 10.6. *Let $P, Q \in \Delta(X)$ such that for some versions of P_n, Q_n, for each $n \in \mathbb{N}$, $||P_n[\cdot \mid \cdot] - Q_n[\cdot \mid \cdot]|| < \zeta_n$. Then $||P - Q|| < \sum_n \zeta_n$.*

Proof. Inductively, for each $n \in \mathbb{N}$, it can be shown via standard arguments[19] that

$$||P_n - Q_n|| \leq ||P_{n-1} - Q_{n-1}|| + \sup_{\overline{x}_n \in \overline{X}_n} ||P_n[\cdot|\overline{x}_n] - Q_n[\cdot|\overline{x}_n]||$$

and hence $||P_n - Q_n|| < \sum_{k \leq n} \zeta_k \leq \sum_k \zeta_k$ for each n, and hence[20] $||P - Q|| = \lim_{n\to\infty} ||P_n - Q_n|| \leq \varepsilon$. □

Theorem 10.7. *Let $P, Q \in \Delta(X)$ and suppose that for some versions of $P_n[\cdot \mid \cdot]$, $Q_n[\cdot \mid \cdot]$, and all $x \in X$,*

$$\sum_{n=1}^{\infty} ||P_n[\cdot \mid \overline{x}_n] - Q_n[\cdot \mid \overline{x}_n]||_\infty^2 = \infty \tag{10.1}$$

Then $P \perp Q$.

Proof. This follows from Corollary 10.5 and the left-hand inequality of Lemma 10.2. □

We will also make use of the following:

Proposition 10.8. *Let $P, Q \in \Delta(X)$ and $T \in \mathbb{N}$. Denote $P_+ = \{\overline{x}_T \in \overline{X}_T | P(\overline{x}_T) > 0\}$ and similarly define Q_+. Suppose for each $\overline{x}_T \in P_+ \cap Q_+$, $P(\cdot \mid \overline{x}_T) \perp Q(\cdot \mid \overline{x}_T)$. Then $P \perp Q$.*

Proof. For each $\overline{x}_T \in P_+ \cap Q_+$, let $S_P(\overline{x}_T), S_Q(\overline{x}_T) \subseteq \prod_{n \geq T} X_n$ be disjoint such that $P(\overline{x}_T \times S_P(\overline{x}_T) \mid \overline{x}_T) = 1$, $Q(\overline{x}_T \times S_Q(\overline{x}_T) \mid \overline{x}_T) = 1$. For $\overline{x}_T \notin P_+ \cap Q_+$, take $S_P(\overline{x}_T) = S_Q(\overline{x}_T) = \prod_{n \geq T} X_n$. Then define,

$$S_P = \bigcup_{\overline{x}_T \in \overline{X}_T} (\overline{x}_T \times S_P(\overline{x}_T)), \quad S_Q = \bigcup_{\overline{x}_T \in \overline{X}_T} (\overline{x}_T \times S_Q(\overline{x}_T))$$

Then it is easy to check that

$$P(S_P) = Q(S_Q) = 1 - Q(S_P) = 1 - P(S_Q) = 1$$

□

[19] Indeed, given transition kernels $\eta(\cdot \mid \cdot)$, $\zeta(\cdot \mid \cdot)$ from a Borel space X to a Borel space Y, and measures $\mu, \nu \in \Delta(X)$, letting $\eta(\cdot \mid \mu)$ and $\zeta(\cdot \mid \nu)$ be the induced measures on $X \times Y$, we have

$$||\eta(\cdot \mid \mu) - \zeta(\cdot \mid \nu)|| \leq \sup_{x \in X} ||\eta(\cdot|x) - \zeta(\cdot|x)|| + ||\mu - \nu||$$

[20] In general, for any measure λ on a measurable space $(\Omega, \mathcal{B})$, and any filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$, if $\lambda_n = \lambda|\mathcal{B}_n$, we have $||\lambda_n|| \to ||\lambda||$, where $||\cdot||$ denotes the total variation norm.

11 Appendix B: Proofs from Section 6

Lemma 11.1. *For each $h \in H_\alpha$ and each $2 \le t \le T$,*

$$P_\sigma(H_\alpha \big| \{h|_t\}) > 1 - 2\zeta\varepsilon$$

Proof. If not, let $h \in H_\alpha$ and $2 \le t \le T$ such that $P_\sigma(H_{T+1} \backslash H_\alpha \big| \{h|_t\}) \ge 2\zeta\varepsilon$.

$$P_\sigma(H_{T+1} \backslash H_\alpha) \ge P_\sigma(H_{T+1} \backslash H_\alpha \big| \{h|_t\}) \cdot P_\sigma(h|_t) \ge (2\zeta\varepsilon) \cdot (\frac{1}{2}\zeta^{-1}) = \varepsilon$$

and therefore $P_\sigma(H_\alpha) \le 1 - \varepsilon$, a contradiction. □

Proof. (Proof of Lemma 6.2) Suppose not; let $h = (a_1, \ldots, a_T) \in H_\alpha$, $2 \le t \le T - T_1$, $k \in \mathcal{P}$ be such that $\sigma^k(h|_t) < 1$. Let σ_L (resp. σ_R) be strategy profiles such that $\sigma_L^k(h|_t) = L$ (resp. $\sigma_R^k(h|_t) = R$) and agrees with σ otherwise.[21] We have (recall the notation of (6.1)),

$$E_{\sigma_R}[\overline{r}_g^k | \{h|_t\}] = \overline{r}_{t-1}^k(h|_t) + E_{\sigma_R}\big[\sum_{s=t}^{T} r^k(a_s)\beta^{s-1} + \beta^T g^k(h) \mid \{h|_t\}\big]$$
$$\le \overline{r}_{t-1}^k(h|_t) + 1 \cdot \beta^{t-1} + 2 \cdot \sum_{s=t+1}^{T} \beta^{s-1} + \beta^T ||g||_\infty$$

where we have used the fact that $r^k(R, x^{-k}) \le 1$ for any profile x, and $r^k \le 2$. On the other hand, since for $h = (a_1, \ldots, a_T) \in H_\alpha$, $r^k(a_t) = 2$ for all $2 \le t \le T$, and $r^k \ge -||r||_\infty$,

$$E_{\sigma_L}[\overline{r}_g^k | \{h|_t\}] = \overline{r}_{t-1}^k(h|_t) + E_{\sigma_L}\big[\sum_{s=t}^{T} r^k(a_s)\beta^{s-1} + \beta^T g^k(h) \mid \{h|_t\}\big]$$
$$\ge \overline{r}_{t-1}^k(h|_t) + \big(2 \cdot P_{\sigma_L}\big(H_\alpha \big| \{h|_t\}\big) - ||r||_\infty \cdot (1 - P_{\sigma_L}(H_\alpha \big| \{h|_t\}))\big) \sum_{s=t}^{T} \beta^{s-1} - \beta^T ||g||_\infty$$
$$\ge \overline{r}_{t-1}^k(h|_t) + (2(1 - 2\zeta\varepsilon) - 2\zeta\varepsilon ||r||_\infty) \sum_{s=t}^{T} \beta^{s-1} - \beta^T ||g||_\infty$$

where we have used Lemma 11.1, since $P_{\sigma_L}(H_\alpha \big| h|_t) \ge P_\sigma(H_\alpha \big| h|_t)$. In order to have $E_{\sigma_R}[\overline{r}_g^k] < E_{\sigma_L}[\overline{r}_g^k]$ (which implies $E_\sigma[\overline{r}_g^k] < E_{\sigma_L^k}[\overline{r}_g]$ and gives the desired contradiction to $\sigma^k(h|_t) < 1$), it suffices to have,

$$1 \cdot \beta^{t-1} + \sum_{s=t+1}^{T} 2\beta^{s-1} + \beta^T ||g||_\infty < (2 - 2\zeta\varepsilon(2 + ||r||_\infty)) \sum_{s=t}^{T} \beta^{s-1} - \beta^T ||g||_\infty$$

[21] That is, $\sigma_L^m, \sigma_R^m = \sigma^m$ for $m \neq k$, and $\sigma_L^k(q) = \sigma_R^k(q) = \sigma^k(q)$ for any $q \in \cup_{t \le T} H_t$ except for the one case $q = h|_t$.

or equivalently,

$$2\beta^T ||g||_\infty < \beta^{t-1} - 2\zeta\varepsilon(2 + ||r||_\infty) \sum_{s=t}^{T} \beta^{s-1}$$

Since $||g||_\infty \le \frac{||r||_\infty}{1-\beta}$ and $\sum_{s=t}^{T} \beta^{s-1} \le \frac{\beta^{t-1}}{1-\beta}$, it suffices to have

$$\beta^{T-t+1} \frac{||r||_\infty}{1-\beta} + \zeta\varepsilon(2 + ||r||_\infty)\frac{1}{1-\beta} < \frac{1}{2}$$

For this to hold for $t \le T - T_1$, it is enough to require that the following pair hold:

$$\beta^{T_1+1} \frac{||r||_\infty}{1-\beta} < \frac{1}{4} \text{ and } \zeta\varepsilon(2 + ||r||_\infty)\frac{1}{1-\beta} < \frac{1}{4}$$

which follow from (6.4) and (6.3). □

Proof. (Proof of Lemma 6.3) We deal with Player D; Player C follows similarly. Like above, let σ_L (resp. σ_R) be the strategy profile such that $\sigma_L^D(\emptyset) = L$ (resp. $\sigma_R^D(\emptyset) = R$) and agrees with σ otherwise. Observe that $P_\sigma(H_\alpha) > 1 - \varepsilon$ implies $\sigma^C(\emptyset)[R] > 1 - \varepsilon$, hence

$$\begin{aligned} E_{\sigma_L}[r^D(a_1)] &= E_{\sigma_L}[r^D(a_1)|H_\alpha] P_{\sigma_L}(H_\alpha) + E_{\sigma_L}[r^D(a_1)|H_T \backslash H_\alpha](1 - P_{\sigma_L}(H_\alpha)) \\ &\le E_{\sigma_L}[r^D(a_1)|H_\alpha] \cdot 0 + (2 \cdot \sigma^C(\emptyset)[L] + 0 \cdot \sigma^C(\emptyset)[R]) \cdot (1 - 0) \le 2\varepsilon < \frac{1}{4} \end{aligned}$$

by (6.3) and since $r^D \le 2$. Hence, also since $r^D \le 2$, we have

$$E_{\sigma_L}[\overline{r}_g^D] = E_{\sigma_L}[\sum_{s=1}^{T} r^D(a_s)\beta^{s-1} + \beta^T g(h)] \le \frac{1}{4} + 2\sum_{s=2}^{T} \beta^{s-1} + \beta^T \frac{||r||_\infty}{1-\beta}$$

Similarly, since $P_{\sigma_R}(H_\alpha) \ge P_\sigma(H_\alpha) > 1 - \varepsilon$ and $r^D \ge 0$,

$$E_{\sigma_R}[r^D(a_1)] \ge E_{\sigma_R}[r^D(a_1)|H_\alpha] P_{\sigma^D}(H_\alpha) \ge \sigma^C(\emptyset)[R](1-\varepsilon) \ge (1-\varepsilon)^2 > \frac{3}{4}$$

Therefore,

$$\begin{aligned} E_{\sigma_R}[\overline{r}_g^D] &= E_{\sigma_R}[\sum_{s=1}^{T} r^D(a_s)\beta^{s-1} + \beta^T g^D(h)] \\ &\ge \frac{3}{4} + (2 \cdot P_{\sigma_R}(H_\alpha) - (1 - P_{\sigma_R}(H_\alpha)) \cdot ||r||_\infty) \sum_{s=2}^{T} \beta^{s-1} - \beta^T \frac{||r||_\infty}{1-\beta} \\ &\ge \frac{3}{4} + (2(1-\varepsilon) - ||r||_\infty \varepsilon) \sum_{s=2}^{T} \beta^{s-1} - \beta^T \frac{||r||_\infty}{1-\beta} \end{aligned}$$

where we have used again the fact that $P_{\sigma_R}(H_\alpha) \geq P_\sigma(H_\alpha) > 1 - \varepsilon$. In order to have $E_{\sigma_R}[\overline{r}_g^D] > E_{\sigma_L}[\overline{r}_g^D]$ it is sufficient to require

$$\frac{3}{4} - (2 + ||r||_\infty)\varepsilon \sum_{s=2}^{T} \beta^{s-1} - \beta^T \frac{||r||_\infty}{1-\beta} > \frac{1}{4} + \beta^T \frac{||r||_\infty}{1-\beta}$$

which holds if

$$\varepsilon(2 + ||r||_\infty) \sum_{s=2}^{T} \beta^{s-1} + 2\beta^T \frac{||r||_\infty}{1-\beta} < \frac{1}{2}$$

and since $T \geq T_1$, this holds if

$$\varepsilon(2 + ||r||_\infty) \frac{1}{1-\beta} + 2\beta^{T_1} \frac{||r||_\infty}{1-\beta} < \frac{1}{2}$$

This will hold if

$$2\beta^{T_1} \frac{||r||_\infty}{1-\beta} < \frac{1}{4} \text{ and } \varepsilon(2 + ||r||_\infty) \frac{1}{1-\beta} < \frac{1}{4}$$

which follow from (6.4) and (6.3). □

Proof. (Proof of Lemma 6.4) For each $u \in \{L, R\}^{\overline{\mathcal{P}}}$, where $\overline{\mathcal{P}} = \{A_1, A_2, A_3, A_4\}$, let σ_u be the strategy profile in the T-stage game in which $\sigma_u^{\overline{\mathcal{P}}}(\emptyset) = u$, and σ_u agrees with σ otherwise.[22] $z = \sigma^{\overline{\mathcal{P}}}(\emptyset)$ must then be an equilibrium of the game with payoff

$$E_{\sigma_u}[\overline{r}^{\overline{\mathcal{P}}}] = r^{\overline{\mathcal{P}}}[u] + \xi(u)$$

where

$$\xi(u) = E_{\sigma_u}[\sum_{s=2}^{T} r^{\overline{\mathcal{P}}}(a_s)\beta^{s-1} + \beta^T g^{\overline{\mathcal{P}}}(h)]$$

If we can show that $||\xi(u) - \xi(v)||_\infty < \eta$ for any two $u, v \in \{L, R\}^{\overline{\mathcal{P}}}$, then $x = \sigma(\emptyset)$ will be an η-equilibrium in the stage game $r(\cdot)$, since $\sigma^{C,D}(\emptyset) = (R, R)$ by Lemma 6.3.

We have by Lemmas 6.2 and 6.3,

$$P_\sigma(\forall 2 \leq s \leq T - T_1, a_s = \overline{L}) = 1$$

and hence for $2 \leq s \leq T - T_1$,

$$P_\sigma(r^{\overline{\mathcal{P}}}(a_s) \equiv 2) = 1$$

Furthermore, by the final condition assumed in Propositon 6.1, $\sigma(\emptyset)[u] > 0$ and $\sigma(\emptyset)[v] > 0$ for any $u, v \in \{L, R\}^{\overline{\mathcal{P}}}$, hence $P_{\sigma_u} \ll P_\sigma$ and $P_{\sigma_v} \ll P_\sigma$. Hence, for $2 \leq s \leq T - T_1$,

$$E_{\sigma_u}[r^{\overline{\mathcal{P}}}(a_s)] = E_{\sigma_v}[r^{\overline{\mathcal{P}}}(a_s)] \equiv 2$$

[22] That is, $\sigma_u^{C,D}(\emptyset) = \sigma^{C,D}(\emptyset)$ and if $q \in \cup_{t=2}^{T} H_t$, $\sigma_u(q) = \sigma(q)$.

Hence, to show that $||\xi(u) - \xi(v)||_\infty < \eta$, it's enough to require that

$$\sum_{s=T-T_1+1}^{T} ||r||_\infty \beta^{s-1} + \beta^T \frac{||r||_\infty}{1-\beta} < \frac{\eta}{2}$$

or equivalently,

$$\frac{||r||_\infty \cdot \beta^{T_1}}{1-\beta} < \frac{\eta}{2}$$

which follows from (6.4). □

References

[Ash and Sexton (1985)] A Surface With One Local Minimum, *Mathematical Association of America*, 58, 147–149.

[Blackwell and Dubins (1962)] Merging of Opinions with Increasing Information, *The Annals of Mathematical Statistics*, Vol. 33, 882–886.

[Foster and Young (2001)] On the Impossibility of Predicting the Behavior of Rational Agents, *Proc. Nat. Acad. Sci.*, 98, 12848–12853.

[Gilli (2001)] A General Approach to Rational Learning in Games, *Bull. Econ. Res.*, 53, 275–303.

[Jordan (1991)] Bayesian Learning in Normal Form Games, *Games Econ. Behav.*, 3, 60–81.

[Jordan (1993)] The Problems in Learning Mixed–Strategy Nash Equilibria, *Games. Econ. Behav.*, 5, 368–386.

[Jordan (1995)] Bayesian Learning in Repeated Games, Games Econ. Behav., 9, 8–20.

[Kakutani (1948)] On Equivalence of Infinite Product Measures, *Annals of Mathematics*, 49, 214–224.

[Kalai and Lehrer (1993)] Rational Learning Leads to Nash Equilibrium, *Econometrica*, 61, 1021–1045.

[Kalai and Lehrer (1995)] Subjective Games and Equilibria, *Games. Econ. Behav.*, 8, 123–63.

[Miller and Sanchirico (1997)] Almost Everybody Disagrees Almost All the Time: The Genericity of Weakly Merging Nowhere, Columbia University Discussion Paper No. 9697–25.

[Miller and Sanchirico (1999)] The Role of Absolute Continuity in Merging of Opinions and Rational Learning, *Games Econ. Behav.* 29, 170–190.

[Nachbar (1997)] Prediction, Optimization, and Learning in Repeated Games, *Econometrica*, 65, 275–309.

[Norman (2012)] Almost-Rational Learning of Nash Equilibrium without Absolutely Continuity, *Economics Series Working Papers*, 602, Univ. Oxford, Dep. Econ.

[Nyarko (1998)] Bayesian Learning and Convergence to Nash Equilibria without Common Priors, *Econ. Theory*, 11, 643–55.

[Sandroni (1998)] Necessary and Sufficient Conditions for Convergence to Nash Equilibrium: The Almost Absolute Continuity Hypothesis, *Games Econ. Behav.*, 22, 121–147.

[Sandroni and Smorodinsky (1999)] The Speed of Rational Learning, *Int. J. Game Theory*, 28, 199–210.

[Shiryaev (1995)] *Probability*, 2nd Ed., Springer.