A simple axiomatics of dynamic play in repeated games

Laurent Mathevet

university of texas - austin

17. January 2012

A SIMPLE AXIOMATICS OF DYNAMIC PLAY IN REPEATED GAMES†

LAURENT MATHEVET*

ABSTRACT. This paper proposes an axiomatic approach to study two-player infinitely repeated games. A solution is a correspondence that maps the set of stage games into the set of infinite sequences of action profiles. We suggest that a solution should satisfy two simple axioms: *individual rationality* and *collective intelligence*. The paper has three main results. First, we provide a classification of all repeated games into families, based on the strength of the requirement imposed by the axiom of collective intelligence. Second, we characterize our solution as well as the solution payoffs in all repeated games. We illustrate our characterizations on several games for which we compare our solution payoffs to the equilibrium payoff set of Abreu and Rubinstein (1988). At last, we develop two models of players' behavior that satisfy our axioms. The first model is a refinement of subgame-perfection, known as renegotiation proofness, and the second is an aspiration-based learning model.

Keywords: Axiomatic approach, repeated games, classification of games, learning, renegotiation.

JEL Classification: C71, C72, C73.

I INTRODUCTION

What may be reasonable to observe in a repeated interaction? Equilibrium analyses of repeated interactions propose variants of Nash or subgame perfect Nash equilibrium. Non-equilibrium analyses typically propose learning models in which players follow various behavioral rules (e.g. Kalai and Lehrer (1993) and Fudenberg and Levine (1998)). Both classes of analysis deliver paths of play or distributions on paths of play. Rather than starting with

Date: January 17, 2012.
†Special thanks are due to Philippe Solal, Max Stinchcombe and Tom Wiseman for many discussions and suggestions. I also wish to thank Matthias Blonski, David Frank, Drew Fudenberg, Sahotra Sarkar and Dale Stahl for their helpful comments.
*Dept. of Economics, University of Texas at Austin, 1 University Station, C3100, Austin, TX 78712, U.S.A.; lmath@mail.utexas.edu.

a model of behavior and deriving what will be observed, this paper formulates an answer to the original question in terms of axioms that the dynamic paths should satisfy.[1]

This paper proposes an axiomatic approach to study two-player infinitely repeated games. A repeated game is an infinite sequence of repetitions of a stage game. A solution is a description of the sequences of play that can arise in a repeated game given its stage game. In technical terms, a solution is a correspondence that maps the set of stage games into the set of infinite sequences of action profiles. The case of infinitely patient players and the discounted case are treated separately. In both cases, we suggest that a solution should satisfy two simple axioms.[2] Our first axiom captures a basic notion of individual rationality. A solution should exclude all sequences for which a player receives a strictly smaller payoff than her maxmin level. In many instances, this axiom is uncontroversial, because all players should be able to secure this minimal payoff. Most equilibrium concepts satisfy it. Our central axiom captures a notion of collective intelligence. We say that a sequence demonstrates that the players could make themselves strictly better off if it contains a cyclic subsequence that Pareto improves on the sequence. The collective intelligence axiom is that we should not observe dynamic paths of play that demonstrate that the players could make themselves strictly better off. Most learning models satisfy this axiom in certain games.

The departure of our approach from the standard methodology confers on it different virtues. Traditionally, the analyst presupposes a behavior for the players through an equilibrium concept or learning dynamics, and this behavior produces sequences of play. Our approach takes the sequences of play as starting point, which has several benefits. First, our axioms contain enough information to provide a classification of all repeated games according to the restrictiveness of our theory. This classification helps delineate the domain of applicability of our theory. Second, this approach has an advantage at the characterization stage, because the axioms apply directly to the sequences of play and not to the behaviors that produce them. Finally, it fosters a unifying view of repeated games. In dominance solvable games, for example, non-equilibrium theory often predicts a unique outcome, while

[1]This is similar in spirit to the difference between preference-based demand theory and revealed preference analysis. In the latter, the analyst formulates axioms on observables, i.e. on the consumer's choices.

[2]The following version of the axioms corresponds to infinitely patient players. Weaker versions are provided for the discounted case.

equilibrium theory often predicts an extreme form of multiplicity. The axiomatic method help reconcile both sides by grouping theories according to their observable properties.

Among the main results, we classify *all* repeated games into families, based on the strength of the requirement imposed by the axiom of collective intelligence. For each game, it allows to appreciate the reasonableness of the axiom — under a certain criterion. In game theory, most equilibrium and non-equilibrium theories are silent about the games to which they are more likely to apply. By default, the assumptions that the analyst makes about players seem to apply equally to all games. For example, the statement that the structure of the game and the players' rationality are common knowledge is assumed to hold for all games.[3] Similarly, the statement that players use Bayesian learning or adaptive dynamics is assumed to hold for all games. This — involuntary or deliberate — form of universality is appealing, but it makes it difficult to appreciate whether a particular theory is more plausible in certain games. Harsanyi and Selten (1988, p.8) express a related concern: "although classical game theory offers a number of alternative solution concepts for cooperative games, it fails to provide a clear criterion as to *which* solution concept is to be employed in analyzing any real-life social situation." In our approach, the axiom of collective intelligence contains enough information to suggest a natural criterion to develop a classification of games. Precisely, each repeated game belongs to a family $\mathscr{F}_n$ for some $n \in \mathbb{N}\setminus\{0\}$, and collective intelligence is weaker in $\mathscr{F}_n$ than in $\mathscr{F}_{n+1}$. Thus, the characterizations are more reasonable for games in $\mathscr{F}_n$ than in $\mathscr{F}_{n+1}$. Our classification criterion will also be used for a given game (instead of across games) to classify the payoff areas where the axiom is more likely to hold.

We continue the analysis by characterizing the solution, i.e. the set of sequences of play that satisfy our axioms, in all repeated games. This result leads to a complete characterization of the solution payoffs. The characterizations are simple and general. In the discounted case, they take an *eventual* form. In 2×2 games, the solution payoffs are reminiscent of the equilibrium payoffs of Abreu and Rubinstein (1988), and thus we compare both characterizations in several examples. In Battle of the Sexes, for instance, both characterizations are the same. Like Abreu and Rubinstein (1988), our results are in sharp contrast with

[3]For example, the fact that a game is dominance solvable — and thus common knowledge of rationality and of the structure of the game gives a unique rationalizable equilibrium (see e.g. Brandenburger (1992)) — does not make it a situation where the common knowledge assumption is more likely to hold.

traditional folk theorems (e.g. Aumann and Shapley (1994)) since our solution payoff set is substantially smaller. This reduction is dramatic for common interest games for which each player has a strictly dominant strategy and not playing it results in very low payoffs. In those games, our characterization delivers a unique prediction. This suggests that an axiomatic approach may offer selection arguments to equilibrium analysis (see Blonski et al. (2011)). In general, the axioms eliminate the Nash equilibria for which an alternative way of playing the profiles played on the path would benefit both players. In other words, in those Nash equilibria, each player believes that there is no hope to teach the other player to respond to a deviation that could be jointly beneficial other than by getting punished, although the other player plays appropriately against it on the equilibrium path.

So far the results characterize the solution as well as the associated payoffs, and the classification of games gives perspective. However, this does not explain *how* players may produce sequences that satisfy the axioms. This leads to our final contribution.

We provide two models of players' behavior that support our axioms. The first model is a refinement of subgame perfection, known as renegotiation proofness (Farrell and Maskin (1989), Bernheim and Ray (1989)). The idea behind renegotiation proofness is that players do not want to renegotiate the existing terms of an equilibrium at any point in time. We show that every such equilibrium generates a sequence of play that satisfies the axioms. Many definitions of renegotiation-proofness are available in the literature, ranging from a weak requirement (e.g. Farrell and Maskin (1989)) to a strong requirement (e.g. Ray (1994)). Our definition is intermediate. Our second model is an aspiration-based learning model. In every period, each player has an implicit aspiration level. If a player's current average payoff, over the recent past, does not meet her aspiration, then she switches with some probability to another course of actions. This process produces the desired sequences with probability one, at least when players consider increasingly long past histories. The result re-emphasizes the game classification.

The axiomatic method in economics and the literature on repeated games have led to more achievements than a literature review can cover. We refer the reader to standard texts, such as Thomson (2001) and Mailath and Samuelson (2006). To the best of our knowledge, the axiomatic method has rarely been used in repeated games, so it is hoped to provide new perspectives. A recent paper by Blonski et al. (2011) uses an axiomatic method for equilibrium selection in the infinitely repeated Prisoners' Dilemma. They define a solution as a subset of the (stage game) payoffs and discount factors for which some cooperation should be expected on the path. Their axioms nicely characterize a unique solution. Earlier contributions used an axiomatic approach in extensive form games (Abreu and Pearce (1984) and Kohlberg and Mertens (1986)) to study the stability of Nash equilibrium.

The paper is organized as follows. We present the repeated game model in the next section. Section III introduces the axiomatic approach. Section IV provides the classification of games. Section V contains the characterizations and the examples. Section VI describes the model of renegotiation-proofness and states the corresponding result. Section VII gives learning arguments to support the axioms but we relegate the formal treatment of our learning model to the appendix. Section VIII extends our results to sophistication constraints and discounted payoffs. Finally we conclude.

II The Model

Let $G = (A_1, A_2, u_1, u_2)$ be a two-person game in normal form. A_i is player i's finite set of actions and $A = A_1 \times A_2$ is the set of action profiles. Define $u_i : A \to \mathbb{R}$ to be player i's payoff function. Let

$$\Pi(G) = \{\pi \in \mathbb{R}^2 : \exists a \in A \text{ s.t. } \pi = (u_1(a), u_2(a))\}$$

be the set of pure payoff vectors in G. Denote by $\mathrm{Co}(X)$ the convex hull of $X \subset \mathbb{R}^2$. The set of feasible payoffs in G is $\mathrm{Co}(\Pi(G))$. Player i's maxmin payoff is

$$\underline{u}_i(G) = \max_{a_i \in A_i} \min_{a_j \in A_j} u_i(a_i, a_j).$$

A maxmin action for player i is any action a_i such that $\min_{a_j \in A_j} u_i(a_i, a_j) = \underline{u}_i(G)$. Let $\underline{u}(G) = (\underline{u}_1(G), \underline{u}_2(G))$ be the maxmin payoff vector. A payoff vector π is individually rational if $\pi_i \geq \underline{u}_i(G)$ for all i. The set of individually rational payoffs is

$$\Pi^{IR}(G) = \{\pi \in \mathrm{Co}(\Pi(G)) : \pi_1 \geq \underline{u}_1(G) \text{ and } \pi_2 \geq \underline{u}_2(G)\}.$$

The repeated game with stage game G consists of an infinite sequence of repetitions of G at discrete time periods $t = 1, 2, \ldots$ At period t, the players make simultaneous moves, denoted by $a_i^t \in A_i$, that become common knowledge (forever). The history of play up to time t, denoted by h_t, is the sequence of action profiles $h_t = (a^1, \ldots, a^t)$. Let H_t be the set of histories of length t and let $H = \cup H_t$ be the set of all (finite) histories.

A repeated game strategy for player i is a function $\sigma_i : H \to A_i$ that maps each history to an action in A_i. Starting from any history $h_t \in H_t$, the continuation game is the infinitely repeated game that begins in period t. For any strategy profile σ, player i's continuation strategy induced by h_t, denoted $\sigma_i|_{h_t}$, is given by $\sigma_i|_{h_t}(h) = \sigma_i(h_t h)$ for all $h \in H$, where $h_t h$ is the concatenation of history h_t followed by history h.

Repeated game payoffs are determined by the inferior limit of means. We study discounted payoffs in Section VIII. Let $S(G) = A^\infty$ be the set of all infinite sequences of play in the repeated game with stage game G. For notational ease, we will use S to refer to $S(G)$ when there is no confusion. An element $s \in S$ is written as $s = (s^1, s^2, \ldots)$. For $s \in S$, player i's payoff is given by

$$\pi_i(s) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i(s^t).$$

A pair of strategies $\sigma = (\sigma_1, \sigma_2)$ produces a sequence of play $\boldsymbol{a}(\sigma) \in S$. Thus, each player i receives utility $\pi_i(\boldsymbol{a}(\sigma))$ from strategy profile σ. For any $\pi, \pi' \in \mathbb{R}^2$, $\pi \gg \pi'$ means $\pi_1 > \pi'_1$ and $\pi_2 > \pi'_2$, i.e. π (*strictly*) Pareto dominates π'. Notation $\pi > \pi'$ means that $\pi_1 \geq \pi'_1$ and $\pi_2 \geq \pi'_2$ and at least one inequality holds strictly, i.e. π *weakly* Pareto dominates π. Let

$$\mathscr{P}(X) = \{\pi \in X : \nexists x \in X \text{ s.t. } x \gg \pi\}$$

be the set of (weakly) Pareto efficient elements in X relative to X.

III The Axiomatic Approach

Preliminaries

We study the class of two-player infinitely repeated games with finite stage games. Let $\mathscr{G}$ be the set of all (finite two-person) stage games. A solution $\mathscr{S}$ is a function that assigns to each stage game $G \in \mathscr{G}$ a subset $\mathscr{S}(G) \subset S(G)$. We think of $\mathscr{S}(G)$ as a description of the set of infinite sequences that can arise in the repeated game with stage game G.

Before stating our axioms, we define the notion of cycle. A sequence $s \in S$ has a cycle if there exist T and ℓ such that $s^t = s^{t+\ell}$ for all $t \geq T$ and there is no $\ell' < \ell$ for which $s^t = s^{t+\ell'}$ for all $t \geq T$. The length ℓ of a cycle is equal to the number of profiles that form the cycle. A sequence with a cycle is called a cyclic sequence. These sequences play an important role in repeated games. They appear in folk theorems (Aumann and Shapley (1994)). They emerge naturally in repeated games with boundedly rational players who use automaton strategies (Abreu and Rubinstein (1988); see the appendix). Automata are sometimes interpreted as mental systems with psychological states (Compte and Postlewaite (2010)). From this point of view, a cycle captures a notion of joint abilities of the players. We also know from Kalai and Stanford (1988) that all subgame perfect equilibria can be approximated by profiles that generate cyclic sequences. Finally, our characterization results are unchanged if we adopt a more general definition of cycle that includes the above as a special case.

The Axioms

We provide two simple axioms that capture sensible properties of a solution. Section VIII contains the version of the axioms for the discounted case.

Axiom 1. (Individual Rationality) For any stage game $G \in \mathscr{G}$, if $s \in S(G)$ is such that $\pi_i(s) < \underline{u}_i(G)$ for some player i, then $s \notin \mathscr{S}(G)$.

This axiom eliminates all sequences for which there exists a player whose dynamic payoff is strictly lower than her maxmin level. In many situations, Axiom 1 should be uncontroversial. If we restrict attention to cyclic sequences, then any player prefers to switch to her maxmin action forever rather than obtaining the sequence s described in the axiom. Indeed, a player can, at any time, decide to play her maxmin action forever. For all periods that follow this choice, she will receive per-period payoffs that weakly exceed $\underline{u}_i(G)$, hence her dynamic payoff will be weakly larger than $\underline{u}_i(G)$.

For arbitrary sequences, the axiom deserves further explanation. Let $\epsilon \in (0,1)$, and consider a game where $u = 3(1-\epsilon)$ is the payoff of some action profile and $\underline{u}_1(G) = 1$. Let s be a sequence that gives the following payoffs to player 1:

$$u, u, 0, 0, 0, 0, u, u, u, u, u, u, u, u, \ldots \tag{1}$$

In (1), every finite string of u's (0's) is followed by twice as many 0's (u's). It is easy to see that $\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}u_1(s^t)$ does not exist and $\pi_1(s)=\frac{1}{3}u<\underline{u}_1(G)$. Axiom 1 says that the solution should exclude sequence s as a possible outcome. One argument is that player 1 may induce a sequence that she prefers to s by playing her maxmin action. Given two arbitrary sequences, in particular without cycle, it is not obvious to define preferences between them based on payoffs. Under the limit-of-means criterion (Aumann and Shapley (1994)), a player prefers s' to s'' if $\liminf(1/T)\sum u_i(s'^t)\geq \limsup(1/T)\sum u_i(s''^t)$. Another criterion may express a preference for stable payoffs: s' is strictly preferred to s'' whenever $u_i(s'^t)>\pi_i(s'')$ for all t large enough. Under this criterion, each player prefers a stable payoff flow to a more chaotic path that may, at times, pay more. Each player can then induce a sequence that she prefers to s by switching to her maxmin action.

We now present our second axiom.

Axiom 2. (Collective Intelligence) For any stage game $G\in\mathscr{G}$, if $s\in S(G)$ has a cyclic subsequence s' such that $\pi(s')\gg\pi(s)$, then $s\notin\mathscr{S}(G)$.[4]

One reading of this axiom is that if a sequence of play contains an alternative "scenario" (i.e. a subsequence) that benefits both players, then they should find it and the original sequence should not materialize. s' is required to be cyclic to exclude alternative scenarios of extreme complexity; since they may be much more complex than the sequence from which they are extracted, it is unclear that players may find them (see Section IV).

Let us make a key observation: a sequence s cannot be eliminated by using cycles made of actions that do not appear in s or that appear only finitely often in s. This means that Axiom 2 applies *if a sequence of play carries in itself* forever *an alternative way of playing that benefits both players*. The first part of the key observation follows from the definition of a subsequence. All elements of s' are contained in s, hence players do not play new profiles in s'. What explains $\pi(s')\gg\pi(s)$ is that s adds poor choices in between the elements of s'; see the proposition below. The second part of the observation follows from the definition of π as a limit of means. If an element only appears finitely many times in a sequence, then it does not affect the payoffs. Thus, if there were action profiles in s yielding high payoffs for both players, and if they were only played finitely many times in s, then they could only be

[4]The discounted-payoff analog of our axioms is given in Section VIII.

played finitely many times in s'. As a result, $\pi(s') \gg \pi(s)$ could not hold. For $\pi(s') \gg \pi(s)$ to hold, s must feature "good" profiles played infinitely many times, yet these profiles are not given a more prominent part of the interaction.

Most learning models lead to satisfy collective intelligence in certain games (Section VII). Therefore, denying the axiom as a whole contradicts these models. For example, all (finite-time) convergent learning dynamics satisfy collective intelligence, because all subsequences of a convergent sequence give the same dynamic payoffs as the sequence. In (finite) dominance solvable games, where most learning models predict convergence, the acceptance of these models is tied to the acceptance of the axiom.

The importance of the experimental literature in repeated games[5] raises the question: is collective intelligence corroborated by experiments? The question actually applies to the discounted-payoff version of the axiom. A rigorous answer is beyond the scope of this paper. We will simply say that some data are encouraging, at least in certain games. In Arifovic et al. (2006), for example, many pairs of subjects end up playing the efficient equilibrium in Chicken and Stag Hunt, and alternating between the equilibria in Battle of the Sexes, and playing a seemingly convergent path, cooperative or not, in the Prisoners' Dilemma.

Before concluding this section, we reformulate the axiom in terms of inferior choices. The next proposition shows that the existence of an improving subsequence is equivalent to the existence of an inferior subsequence. Interpreted along these lines, Axiom 2 requires players to eventually abandon inferior choices.

Proposition 1. A cyclic sequence s has a cyclic subsequence s' such that $\pi(s') \gg \pi(s)$, if and only if, it has a cyclic subsequence s'' such that $\pi(s'') \ll \pi(s)$.

Proofs omitted in the main text are relegated to the appendix.

Although collective intelligence implies some cooperation between the players along certain sequences, it is compatible with non-cooperative game theory. Indeed, collective intelligence only deals with joint improvements that come from within a sequence. Therefore, the axiom allows for total cooperation failure, which occurs when players revert to their

[5]The recent literature includes, among others, Dal Bo (2005), Arifovic et al. (2006), Duffy and Ochs (2009), Blonski et al. (2011), and Dal Bo and Frechette (2011). For a more comprehensive literature review, especially of the earlier literature (e.g. Roth and Murnighan (1978), Holt (1985), etc), we refer the reader to the recent experimental papers.

maxmin action forever. Likewise, the axiom allows for the existence of punishments, even if they harm both players. However, phases of mutually harmful behaviors cannot co-exist indefinitely with instances of cooperation. In that case, cooperation would take over, possibly after an arbitrary long time, i.e. after arbitrarily many cooperation failures or punishments. This line of reasoning is defended by Axelrod (1984)'s works on the emergence of cooperation in non-cooperative settings. Overall, our axiom has a flavor of cooperative competition (Brandenburger and Nalebuff (1999)).

We will provide a formal defense of our axioms in later sections. For any game G, let $\mathscr{C}(G)$ be the set of sequences that satisfy the axioms, i.e. all sequences s that do not have any cyclic subsequence s' such that $\pi(s') \gg \pi(s)$ and for which $\pi_i(s) \geq \underline{u}_i(G)$ for all i.

IV Axioms and Complexity

In this section, we propose a criterion to assess the reasonableness of collective intelligence and we use it to classify *all* stage games, and indirectly all repeated games, into families.

The main hurdle to the axiom lies in the players' difficulty of noticing and implementing an improving subsequence. Our criterion is based on the comparison between the complexity of a sequence and the complexity of some improving subsequence. It seems reasonable to assume that if players conform to some convention, then they should be able to conform to simpler conventions that do not add new observations. That is, if players succeed in playing a cycle, then they should be able to play another cycle of shorter or equal length that uses the same action profiles. This is especially plausible because, once the players play a cycle, they will have infinitely many opportunities to observe it, question it, and eventually change it. Therefore, one of the main reasons why players may not play another cycle, although it uses the same profiles and benefits both players, must be that it is longer and thereby requires more sophistication. But suppose that it can never happen. Suppose that any cyclic sequence s that contains an improving subsequence also contains one, call it s'', that is simpler than s, i.e. whose cycle is smaller than that of s.[6] Then the axiom should be rather reasonable, because along s the players will not only experience its cycle arbitrarily

[6]A subsequence could be more complex, i.e. have a longer cycle, than the sequence from which it is extracted. For example, $s = (a, b, a, b, \ldots)$ has a cycle of length 2 and $s' = (a, b, b, a, b, b, \ldots)$, constructed by removing one a out of two from s, has a cycle of length 3.

many times, allowing them to detect s'', but a joint improvement only requires the players to implement a simpler cycle than their current one. To summarize, in these situations, they should find an improving subsequence. The main result shows that there are games that satisfy this property. Unfortunately, for some games, if an improving subsequence exists, then it may require a longer cycle than the original sequence; thus, the players may never reach the level of sophistication necessary to improve themselves. Yet the axiom may still hold, depending on the capabilities that the players gain over time. After all, the players have an infinite horizon to learn to coordinate on increasingly complex cycles.

Our main result classifies stage games into families. In each family, an improving subsequence cannot exceed a certain complexity in comparison to the original sequence.

Defining Family $\mathscr{F}_n$

This section presents the formal definition of family $\mathscr{F}_n$ as well as technical material. The next section contains the main result. Some readers may prefer reading the next section and its main result before covering the technicalities.

A triangle is a convex set $\Delta = \mathrm{Co}(\{u^1, u^2, u^3\})$ where u^1, u^2 and u^3 are three non-collinear vectors in $\Pi(G)$,[7] called *vertices*. Let $V(\Delta)$ be the set of vertices of Δ. Let $\mathscr{N}(\Delta) = \{u \in V(\Delta) : u \in \mathscr{P}(V(\Delta))\}$ be the set of Pareto undominated vertices within $V(\Delta)$. Let $\mathscr{T}(G)$ be the set of all triangles in game G.

Given a stage game, the next definition classifies the different areas of the feasible payoffs into families, and this classification will determine the family to which the game belongs. Let $\mathbb{Q}[0,1]$ be the set of rational numbers between 0 and 1. Given a rational number q, $d(q)$ denotes its denominator and $\nu(q)$ its numerator.

Definition 1. Given $G \in \mathscr{G}$, a feasible payoff π is in $\mathscr{U}_k$, $k \in \mathbb{N}^*$, if for aany $\Delta \in \mathscr{T}(G)$ such that $\pi \in \Delta$ and $|\mathscr{N}(\Delta)| = 3$, Δ can be written as $\mathrm{Co}(\{u^1, u^2, u^3\})$ such that, for some $q \in \mathbb{Q}[0,1]$, either $(u^2 - u^1) + q(u^2 - u^3) \gg 0$ and $d(q) \leq k$, or $(u^1 - u^2) + q(u^3 - u^2) \gg 0$ and $d(q) + \nu(q) \leq k$.[8]

[7]Three points $x, y, z \in \mathbb{R}^2$ are collinear if $\frac{y_2 - x_2}{y_1 - x_1} = \frac{z_2 - x_2}{z_1 - x_1}$.

[8]The requirement $q \in [0,1]$ is inconsequential because we can relabel the vertices of a triangle: if $(u^2 - u^1) + 2(u^2 - u^3) \gg 0$, then $\frac{1}{2}(u^2 - u^1) + (u^2 - u^3) \gg 0$, and thus, exchanging the labels of u^1 and u^3, we have $(u^2 - u^1) + \frac{1}{2}(u^2 - u^3) \gg 0$.

A feasible payoff profile is in $\mathscr{U}_k$ if every triangle that contains it, and whose vertices are pairwise Pareto unordered, can be written such that a movement along one of its edges, followed by a fraction q of a movement along another, leads to a Pareto improvement, where q is a rational number satisfying certain properties with respect to k. By convention, if there is no Δ containing π such that $|\mathcal{N}(\Delta)| = 3$, then $\pi \in \mathscr{U}_1$. Our main definition is the following.

Definition 2. A game G is in family $\mathscr{F}_n$ if for every $\pi \in \mathrm{Co}(\Pi(G))$, $\pi \in \cup_{k=1}^{n} \mathscr{U}_k$.

A game is in family $\mathscr{F}_n$ if any feasible payoff is in $\mathscr{U}_k$ for some $k \leq n$. Every stage game is in $\mathscr{F}_n$ for some $n \geq 1$. Let us illustrate the definition in the Prisoners' Dilemma with the payoffs given in Figure II. There are four possible triangles, as shown in the next figure.

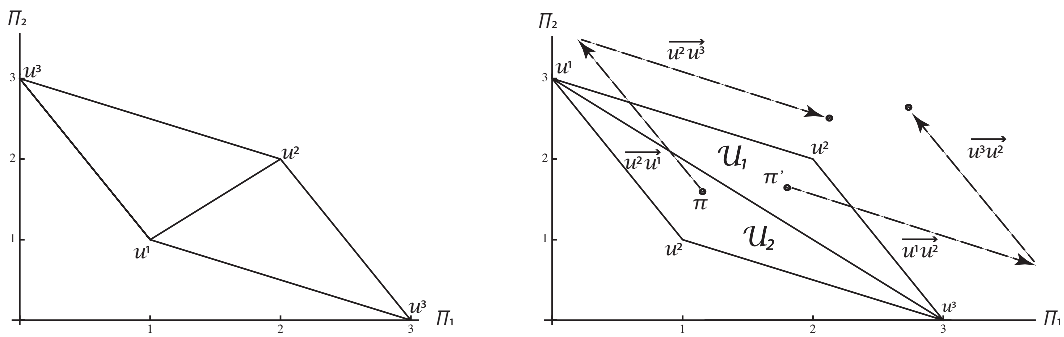


FIGURE I. Triangles in the Prisoners' Dilemma

For every triangle Δ in the left-hand side panel, $|\mathcal{N}(\Delta)| = 2$ and thus these triangles are not subject to the requirement of Definition 1. Triangles in the right-hand side panel, however, are subject to the requirement, because $|\mathcal{N}(\Delta)| = 3$. In the bottom triangle, a movement along vector $(u^1 - u^2) + (u^3 - u^2)$, where $u^\ell - u^k = \overrightarrow{u^k u^\ell}$, gives a Pareto improvement over any π. Therefore, $q = 1$ and $d(q) + \nu(q) = 2$. In the top triangle, a movement along vector $(u^2 - u^1) + (u^2 - u^3)$ gives a Pareto improvement over any π'. Again, $q = 1$ and hence $d(q) = 1$. Putting these observations together, we obtain $G \in \mathscr{F}_2$. The reason why the Prisoners' Dilemma is in $\mathscr{F}_2$ is intuitive. In the bottom triangle (right panel), a Pareto improvement requires to re-allocate the weight from u^2 onto two different profiles, which may require an increased level of coordination from the players.

We show that Battle of the Sexes and Chicken are in $\mathscr{F}_1$. Consider Figure III in Section V. There is only one triangle and it satisfies $|\mathcal{N}(\Delta)| = 2$. Therefore Battle of the Sexes is in

$\mathscr{F}_1$. In Chicken, Figure IV in Section V, only one triangle satisfies $|\mathscr{N}(\Delta)| = 3$, and we can label its vertices so that $(u^2 - u^1) + (u^2 - u^3) \gg 0$. Therefore, Chicken is in $\mathscr{F}_1$.

Complexity and Classification of Games

The complexity of a (cyclic) sequence of play is measured by the length of its cycle. Under automaton strategies, this measure gives a limit on the number of states that a pair of automata can have in order to generate the desired sequence. Alternatively, if a cycle results from an agreement by the players to correlate their actions via a public signaling device,[9] then the complexity measure gives an indication on the complexity of the device. Given a cyclic sequence s, say that a sequence s'' is at most n times more complex than s if s'' has a cycle whose length is smaller than or equal to n times the length of the cycle of s. When $n = 1$, sequence s'' is said to be simpler than s.

We first state and discuss our result, and then sketch the intuition of the proof.

Proposition 2. For any game $G \in \mathscr{F}_n$, if sequence s has a cycle and if it contains a subsequence s' such that $\pi(s') \gg \pi(s)$, then there is a subsequence s'' of s such that s'' is at most n times more complex than s and $\pi(s'') \gg \pi(s)$.

If an improving subsequence exists, then there must be one whose complexity cannot exceed a certain limit compared to the original sequence. In virtue of this proposition, families $\mathscr{F}_n$ indicate how plausible it is that the players implement a jointly improving scenario. As n increases, it becomes less plausible, because all improving subsequences may be much more complex than what there is evidence of the players' capability (in s). That said, the axiom may continue to hold for large n if the players gain sophistication over time. The proposition implies that if we reject the axiom for $G \in \mathscr{F}_n$, then we accept the fact that two players can play a cycle of length ℓ forever, without ever realizing that a (sub)cycle of length at most $n\ell$ would give them both strictly higher payoffs, or without ever being able to implement it. Thus, collective intelligence is particularly compelling in $\mathscr{F}_1$, because if an improving subsequence exists, then there must also be a simpler one than the original sequence.

[9]Take an urn with balls of different colors. When a ball is publicly drawn, each player behaves according to the color. The number of balls, which indicates the complexity of the device, is the length of the cycle.

We now sketch the intuition behind the proposition. Consider a sequence with cycle $\{(C,D)(C,C)(D,C)\}$ in the Prisoners' Dilemma. A look at Figure I shows that the resulting payoffs are π'. If players eventually modified the cycle by replacing (D,C) and (C,D) by (C,C), then their payoffs would change from $\frac{1}{3}(u^1+u^2+u^3)$ to

$$\frac{1}{3}(u^1+u^2+u^3+(u^2-u^1)+(u^2-u^3)) = u^2. \tag{2}$$

Figure I shows that a movement along vector $\overrightarrow{u^1u^2}+\overrightarrow{u^3u^2}$, where $\overrightarrow{u^ku^\ell} = u^\ell - u^k$, gives a Pareto improvement from π'. The modification in (2) is equivalent to a change in payoffs along $\frac{1}{3}(\overrightarrow{u^1u^2}+\overrightarrow{u^3u^2})$, thus it must be a Pareto improvement. In other words, any subsequence that eventually selects (C,C) only from the above cycle is simpler than and improves on the starting sequence. More generally, when players reach a cycle whose associated payoffs lie in triangle Δ, each vertex of Δ corresponds to the payoffs of an action profile from the cycle. Definition 1 identifies the different areas from the feasible payoff set that lend themselves to Pareto improvements by moving along vectors, where these vectorial movements represent substitutions between the elements of the cycle.[10]

We end this section with two remarks:

Instead of classifying games into families, we could classify areas of the feasible payoffs for a given game; this is Definition 1. The main interest is to develop a theory subject to sophistication constraints. For example, an analyst may be studying a game in $\mathscr{F}_2$, while she only believes that the players can implement an improving subsequence if it is simpler than the starting sequence. This is analyzeed in Section VIII.

The definition of $\mathscr{F}_n$ imposes strong conditions. Albeit not necessary, these conditions seem to be minimally sufficient to establish Proposition 2. Consider the Prisoners' Dilemma with utilities $u(D,D) = (\frac{1}{2},\frac{1}{2})$, $u(C,D) = (-1,3)$, $u(D,C) = (3,-1)$, and $u(C,C) = (2,2)$. This game is in $\mathscr{F}_2$. Take a sequence whose cycle (of length 7) has one (D,D), four (C,D)'s and two (D,C)'s. This sequence admits no improving subsequence with a cycle of length $\ell \leq 7$, confirming $G \notin \mathscr{F}_1$. The shortest subcycle that dominates the above has length 11.

[10]For example, if $|\mathscr{N}(\Delta)| = 3$ and $(u^2-u^1)+(u^2-u^3) \gg 0$, then players should drop some weight allocated to (profiles giving) u^1 and u^3 and transfer it onto u^2.

V Characterizations

In this section, we characterize our solution, i.e. the set of sequences that are compatible with our axioms, in all repeated games. We use this result to completely characterize the solution payoffs. Our results are illustrated on several examples. In 2×2 games, our solution payoffs are reminiscent of the basic structure of the equilibrium payoff set of Abreu and Rubinstein (1988). Both characterizations are compared in the examples.

While studying our characterizations, it is important to keep in mind the classification of games, as the characterizations may be more reasonable in certain games than others.

Before proceeding to the results, we introduce several definitions. The line segment between any u^k and u^ℓ in $\Pi(G)$ (not necessarily different) is defined as $\overline{u^k u^\ell} = \mathrm{Co}(\{u^k, u^\ell\})$. Let $L(G)$ be the set of all segments in game G with generic element g. For any sequence $s \in S$, the recurrent payoff set is defined as

$$R(s) = \{u \in \Pi(G) : \exists a \in A \text{ s.t. } u = u(a) \text{ and } (\forall T \in N)(\exists t \geq T) \text{ s.t. } s^t = a\}.$$

Each sequence s induces an infinite sequence of payoffs and $R(s)$ is the set of payoffs that appear infinitely many times in that sequence.[11]

The Results

Two notions play a central role in our characterization results, individual rationality and internal efficiency.

Definition 3. A sequence s is individually rational if $\pi_i(s) \geq \underline{u}_i$ for all i.

Definition 4. A sequence s is internally efficient if $\pi(s) \in \mathscr{P}(\mathrm{Co}(R(s)))$.

The first notion is obvious. The second notion is interesting for its normative and positive content. A sequence is internally efficient if it generates a payoff vector that lies on the weak-Pareto frontier of (the convex hull of) its *own* set of recurrent payoffs. It is socially desirable that players generate the largest surplus from their interaction. Nonetheless, it may be too demanding to assume Pareto efficiency, because it does not take into account players'

[11]This interpretation of $R(s)$ is accurate, because A is finite. If A were infinite, then it could be that a payoff vector appears infinitely many times in s and yet is not in $R(s)$.

capabilities. They may be unable to cooperate to generate the maximal surplus. They may not even be aware of the actions to take to produce the maximal surplus. Internal efficiency, however, requires that among the recurrent profiles, there cannot be a strict Pareto improving arrangement. If players are sufficiently patient and if they observe the same profiles forever, then they may arrive at an efficient arrangement of these profiles. The concept of internal efficiency provides an answer to the criticisms of the adaptive learning literature, according to which adaptive players can play a repeated pattern forever without becoming aware of it, i.e. without noticing the (joint) opportunities for profit (e.g. Sonsino (1997)).

Theorem 1. A solution $\mathscr{S}$ satisfies Axioms 1 and 2, if and only if, $\mathscr{S}(G)$ is a subset of the set of internally efficient and individually rational sequences for all $G \in \mathscr{G}$.

Proof. Suppose that solution $\mathscr{S}$ satisfies the axioms.

Step 1. Internal efficiency. Take any G and any sequence $s \in \mathscr{S}(G)$. For every $\pi \in \mathrm{Co}(R(s))$, there is a rational convex combination $\sum_k u^k \alpha_k$, with $\{u^k\} \subset R(s)$ and $\{\alpha_k\} \subset \mathbb{Q}$, that is arbitrarily close to π. For every such rational convex combination, we can build a cyclic sequence s' whose elements $\{s'^t\}$ are all in $R(s)$ and such that $\pi(s') = \sum_k u^k \alpha_k$. By construction, s' is a subsequence of s. By way of contradiction, suppose that sequence s is not internally efficient, i.e. $\pi(s) \notin \mathscr{P}(\mathrm{Co}(R(s)))$. There are two possible cases: either (i) $\pi(s) \notin \mathrm{Co}(R(s))$ or (ii) $\pi(s) \in \mathrm{Co}(R(s))$. In case (i), it must be that sequence $\{z^T\}$, where $z^T = \frac{1}{T}\sum_{t=1}^{T} u(s^t)$, does not converge. Since the payoffs are given by $\liminf z^T$, there exists $\pi \in \mathrm{Co}(R(s))$ such that $\pi \gg \pi(s)$. This fact obviously holds in case (ii). From the previous argument, there is a rational convex combination of elements of $R(s)$ that is arbitrarily close to π, and thus there is a cyclic subsequence s' of s such that $\pi(s') \gg \pi(s)$. Axiom 2 is violated. This is a contradiction.

Step 2. Individual rationality. This part follows immediately.

Suppose now that $\mathscr{S}(G)$ is a subset of the set of internally efficient and individually rational sequences for all G. Take any G and any sequence $s \in \mathscr{S}(G)$. Then $\pi(s) \in \mathscr{P}(\mathrm{Co}(R(s)))$. For every cyclic subsequence s' of s, $\pi(s') \in \mathrm{Co}(R(s))$. Thus there cannot be a subsequence s' of s such that $\pi(s') \gg \pi(s)$, for otherwise there would be a convex combination from $R(s)$ that is a Pareto improvement on $\pi(s)$. Therefore, Axiom 2 holds. Axiom 1 is trivially satisfied. □

If players can implement improving subsequences, then a sequence that survives the axiom of collective intelligence cannot offer strict Pareto-improvement opportunities within its recurrent payoffs, and thus it must be internally efficient. We now use Theorem 1 to derive a characterization of the solution payoffs in all repeated games.

Theorem 2. For every game $G \in \mathscr{G}$, $\pi(\mathscr{C}(G))$ is dense in $(\cup_{g \in L(G)} \mathscr{P}(g)) \cap \Pi^{IR}(G)$.

In every repeated game, the solution payoffs must be individually rational and lie on the Pareto frontier of a segment formed by some feasible payoffs. This representation as a union of segments dramatically reduces the set of possible payoffs compared to standard folk theorems (e.g. Aumann and Shapley (1994)).

The axioms contain additional information omitted by our characterizations. The axioms produce stable sequences in most games, where stability means convergence of the average payoffs. This is the next result.

Definition 5. A sequence s is payoff-convergent if $\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(s_t)$ exists.

In a payoff convergent sequence, although players may not play a cycle, it is as if they did from a payoff perspective, because the limit of means exists. The trajectory of play along such sequences is not too chaotic. For example, it could be generated by players who play a cycle and depart from it from time to time, but these departures have vanishing frequencies.

In the axioms, the players express a form of preference for stability and recurrence. In particular, collective intelligence suggests an environment where players prefer a sequence to another if the former is cyclic and pays more, which justifies discarding a sequence with an improving cyclic subsequence. The next result is a manifestation of this intuition, as chaotic sequences will be discarded in most games.

Let $\mathscr{G}'$ be the family of stage games for which no two distinct action profiles yield either player the same payoff. That is, for games in $\mathscr{G}'$, if $u_i(a) = u_i(a')$, then $a = a'$. This condition holds for almost all games.

Proposition 3. For every game $G \in \mathscr{G}'$, if a solution $\mathscr{S}$ satisfies Axioms 1 and 2, then $\mathscr{S}(G)$ is a subset of the set of payoff-convergent sequences.

Examples and Uniqueness Result

We demonstrate the conclusions of Theorem 2 on several examples. We also describe a class of common interest games for which our axioms give a unique prediction. Abreu and Rubinstein (1988) study infinitely repeated games under bounded rationality and the left figures represent the Nash equilibrium payoff set of their machine game. The right figures represent our solution payoffs (i.e. $\mathscr{C}(G)$). The intersection of both figures correspond to the Nash equilibrium payoffs among boundedly rational and collectively intelligent players. All the games below are in $\mathscr{F}_1$, except the Prisoners' Dilemma, which is in $\mathscr{F}_2$.

1 \ 2	C	D
C	2, 2	0, 3
D	3, 0	1, 1

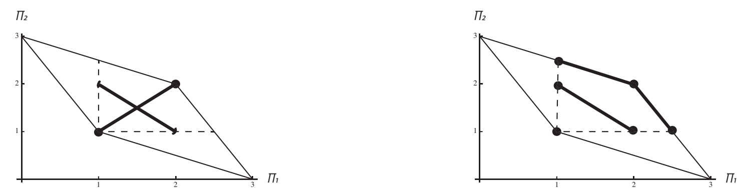


FIGURE II. The Prisoners' Dilemma

1 \ 2	C	D
C	2, 1	0, 0
D	0, 0	1, 2

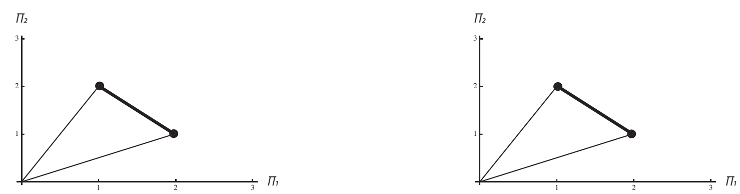


FIGURE III. Battle of the Sexes

1 \ 2	C	D
C	4, 4	2, 5
D	5, 2	0, 0

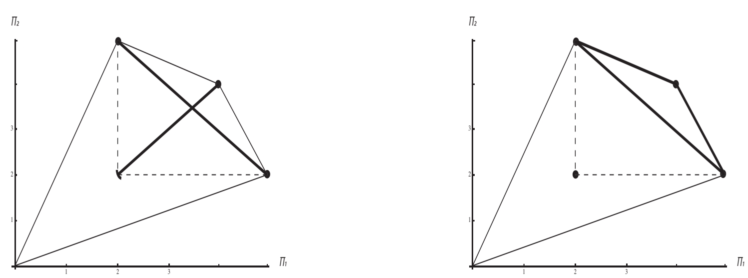


FIGURE IV. Chicken

Abreu and Rubinstein (1988) insist on the dramatic effect of their structure theorem (Theorem 1*, p.1271) on the set of equilibrium payoffs in 2×2 games. In more general games, the set of equilibrium payoffs of the machine game may lose its representation as a union of segments, while this continues to hold under our axioms by Theorem 2. For instance, in

1 \ 2	C	D
C	4, 4	0, 3
D	3, 0	2, 2

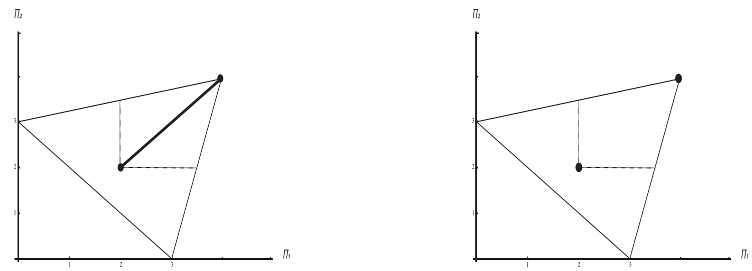


FIGURE V. Stag Hunt

1 \ 2	C	D	E
C	3, 3	0, 0	1, 2
D	1, 2	1, 2	4, 1
E	0, 0	0, 0	4, 1

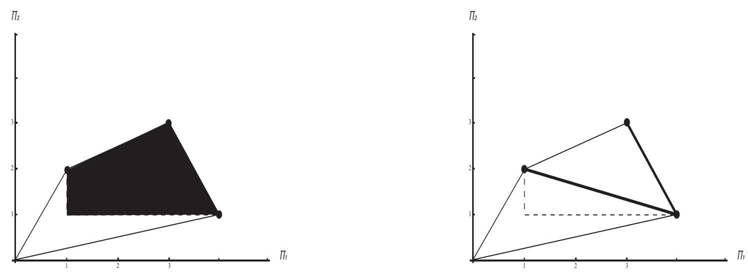


FIGURE VI. A Two-player Three-action Game

the game of Figure VI, the equilibrium payoff set of the machine game is the same as in the traditional folk theorem.[12]

Figures VII and VIII depict two common interest games. Let $u_i^* = \max_{a \in A} u_i(a)$ for $i = 1, 2$. A two-player game is a common interest game if $u_i(a') = u_i^*$ for some player i implies $u_i(a') = u_j^*$ for player j. The next proposition identifies a class of common interest games for which our solution predicts a unique payoff.

1 \ 2	C	D
C	4, 4	2, 3
D	3, 2	0, 0

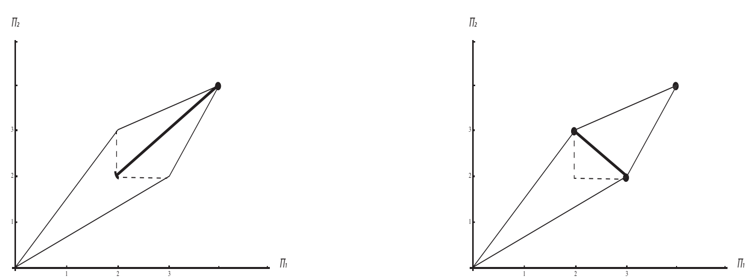


FIGURE VII. A Common Interest Game

[12]All the convex combinations of payoff vectors (3, 3), (1, 2) and (4, 1) can be supported in equilibrium, because each vector corresponds to a static Nash equilibrium payoff. Moreover, an equilibrium pair of automata can be constructed to generate any convex combination of payoff vectors (1, 2), (4, 1) and (0, 0).

1 \ 2	C	D
C	4, 4	2, 0
D	0, 2	1, 1

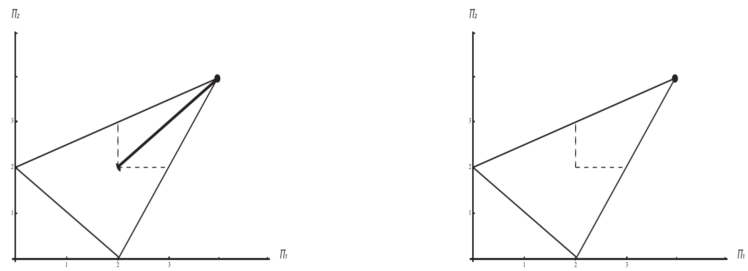


FIGURE VIII. A Common Interest Game

Proposition 4. In game G, if there is $\pi^* \in \Pi(G)$ such that either (i) $\pi^* \gg \underline{u}(G) \geq \pi$ for all $\pi \in \Pi(G)\backslash\{\pi^*\}$ and $\underline{u}(G) \notin \Pi(G)$, or (ii) $\pi^* \geq \underline{u}(G) \gg \pi$ for all $\pi \in \Pi(G)\backslash\{\pi^*\}$ and $\underline{u}(G) \in \Pi(G)$ implies $\underline{u}(G) = \pi^*$, then $\pi(\mathscr{C}(G)) = \{\pi^*\}$.

The first condition implies that playing the profile of maxmin actions secures maximal payoff π^*. In a common interest game where $\underline{u}(G) \geq \pi$ for all $\pi \in \Pi(G)\backslash\{\pi^*\}$, the players are bound to experience payoffs below their maxmin level if they do not play their maxmin actions. Since playing the maxmin actions yields payoffs π^*, we suspect that π^* should be the unique outcome of such games. The game in Figure VII violates both conditions of the proposition and our prediction is not unique. The game in Figure VIII satisfies the first condition, hence our prediction is unique.

VI AXIOMS AND RENEGOTIATION

The previous sections do not explain *h*ow players can produce sequences of play that satisfy the axioms. In this section, we suggest a theory, known as renegotiation-proofness, that also clarifies the relationship of our approach with equilibrium analysis.

A standard definition of renegotiation-proofness is due to Farrell and Maskin (1989).[13] Players agree ex-ante to play a subgame-perfect equilibrium, but they are able to renegotiate the continuation play after every period. In this case, as the authors write, "players are unlikely to play, or to be deterred by, a proposed continuation equilibrium (on or off the equilibrium path) that is strictly Pareto dominated by another equilibrium that they believe is available to them." Therefore, a subgame-perfect equilibrium σ is renegotiation-proof if there do not exist continuation equilibria $\sigma|_h$ and $\sigma|_{h'}$ such that $\sigma|_h$ (strictly) Pareto

[13] Bernheim and Ray (1989) independently provided an equivalent definition.

dominates $\sigma|_{h'}$. This definition imposes restrictions between the equilibrium, σ (i.e. $\sigma|_{\emptyset}$), and its continuation profiles. If σ (strictly) Pareto dominates some continuation equilibrium $\sigma|_h$, then the previous argument applies. Conversely, if a continuation profile $\sigma|_h$ (strictly) Pareto dominates equilibrium profile σ, then presumably the players would renegotiate the equilibrium path and play $\sigma|_h$ instead. Our definition extends the idea of renegotiating the equilibrium path. To avoid fruitless technical complications, we assume that players use strategies representable by finite automata (see Section A in appendix). For any sequence $s \in S$, let

$$R^*(s) = \{a \in A : (\forall T \in N)(\exists t \geq T) \text{ s.t. } s^t = a\} \tag{3}$$

be the set of recurrent action profiles in s, and let $J(s) = \{a \in A : \exists t \text{ s.t. } s^t = a\}$ be the set of all profiles played in s.

Definition 6. A subgame-perfect equilibrium σ is renegotiation-proof if *(i)* there do not exist continuation equilibria $\sigma|_h$ and $\sigma|_{h'}$ of σ such that $\sigma|_h$ strictly Pareto dominates $\sigma|_{h'}$ and *(ii)* there is no subgame-perfect equilibrium σ' that strictly Pareto dominates σ and satisfies *(i)*, $R^*(\boldsymbol{a}(\sigma')) \subset R^*(\boldsymbol{a}(\sigma))$ and $\cup_h J(\boldsymbol{a}(\sigma'|_h)) \subset \cup_h J(\boldsymbol{a}(\sigma|_h))$.

Part (i) is the definition by Farrell and Maskin (1989). Part (ii) reflects our position on renegotiation: players are able to revise the existing terms of an equilibrium but they cannot introduce new ones. If players can renegotiate σ in favor of one of its continuations, as assumed in (i), then they may also renegotiate σ in favor of another equilibrium σ' that uses *on* (and *off*) its equilibrium path the *same* action profiles played *on* (and *off*) the equilibrium path of σ. The players constantly observe the recurrent action profiles of the path, so they have infinitely many opportunities to propose an alternative arrangement of these profiles. For example, if players have been playing the cycle $((C,D)(C,C)(D,C))$ in the Prisoners' Dilemma, eventually they could say: "Why are we repeatedly playing (C,D) and (D,C)? Why not renegotiate and just play (C,C)?" The requirement of part (ii) applies to off-path profiles as well. The players cannot enforce their new equilibrium-path arrangement with profiles that are not played in σ. Therefore, all the continuation profiles of σ', in particular off-path, must be played under σ, $\cup_h J(\boldsymbol{a}(\sigma'|_h)) \subset \cup_h J(\boldsymbol{a}(\sigma|_h))$. Another argument in favor of our definition is that, for games in $\mathscr{F}_1$, players may have a double incentive to renegotiate according to Definition 6: not only this may increase their payoffs but this may also decrease their complexity costs.[14]

Trivially, there is always at least one renegotiation-proof equilibrium. Let $a = (a_1, a_2)$ be a Nash equilibrium of the one-shot game. Define the repeated-game strategy σ_i^* for player i that always plays a_i for every history. Then $\sigma^* = (\sigma_1^*, \sigma_2^*)$ is a subgame-perfect equilibrium. It has no other continuation equilibrium other than itself, $R^*(\boldsymbol{a}(\sigma)) = J(\boldsymbol{a}(\sigma|_h)) = \{a\}$ for all h. Any σ' such that $R^*(\boldsymbol{a}(\sigma')) \subset R^*(\boldsymbol{a}(\sigma))$ cannot Pareto dominate σ, hence it satisfies our definition. We state our result and return to the existing literature.

Proposition 5. For any stage game G, every renegotiation-proof equilibrium generates a path of play in $\mathscr{C}(G)$ (i.e. it satisfies individual rationality and collective intelligence).

The result relies on the specification of payoffs as limits of means. When a subgame-perfect equilibrium σ violates collective intelligence, it is possible to build another equilibrium σ' whose path rewards both players more than σ. What is less obvious is that σ' can always use some of the continuation profiles of σ as punishments to enforce its equilibrium path. Although these punishments may be inefficient in the short run, and thus might a priori violate (i), they are efficient in the long run if the punishment phase only lasts finitely many periods. This shows that the requirement $\cup_h J(\boldsymbol{a}(\sigma'|_h)) \subset \cup_h J(\boldsymbol{a}(\sigma|_h))$, which is very demanding, may become too restrictive for discounted payoffs.[15]

There are several definitions of renegotiation-proofness in the repeated game literature. The definition by Farrell and Maskin (1989) imposes a weak requirement, and thus in many games it does not restrict the set of equilibrium payoffs. In response, stronger definitions have been provided, for example by Bernheim and Ray (1989) and Ray (1994). In Ray (1994), the set of equilibrium payoffs consists of singletons and the Pareto frontier. Our definition is intermediate. The main difference with the stronger definitions is that we exclude renegotiation towards profiles (hence payoffs) that have not been played under σ.

[14]In Abreu and Rubinstein (1988), players prefer machines with fewer states (see Section A). If renegotiation leads to implement a simpler subsequence, then they may obtain larger payoffs *and* spare states.

[15]Under discounted payoffs, the result should hold if we restrict attention to a class of equilibria. Consider the class of games studied in Farrell and Maskin (1989): there are profiles $a^1, a^2 \in A$ such that for all $v \in \mathrm{Co}(\Pi(G))$ with $v_i > \underline{u}_i$ for all i, we have $\max_{a_i} u_i(a_i, a_j^i) < v_i$ and $u_j(a^i) > v_j$ for each i. In these games, we could restrict attention to those equilibria whose punishment phases only use some a^1 and a^2.

Since there is no evidence that players can play or are aware of these profiles, we believe that it can be a sensible choice.

VII Axioms and Learning

In this section, we explain *h*ow learning behaviors can lead players to produce the desired sequences.

Observation. *If every $s \in \mathscr{S}^{learn}(G)$ has at most two recurrent profiles, $|R^*(s)| \leq 2$, whose payoffs are not weakly Pareto ordered, then $\mathscr{S}^{learn}$ satisfies collective intelligence for G.*

Most learning models are a solution $\mathscr{S}^{\text{learn}}$ that satisfies the above condition in certain games. For example, a solution that only keeps (individually rational and) convergent sequences satisfies both axioms, because all subsequences of a convergent sequence give the same dynamic payoffs as the sequence. In dominance solvable games, most adaptive and sophisticated learning dynamics converge to the unique equilibrium (Milgrom and Roberts (1990, 1991)). In those games, models of reinforcement learning (Er'ev and Roth (1995)), pattern recognition (Sonsino (1997)), and Bayesian learning (Fudenberg and Levine (1998)) also predict convergence to the unique equilibrium. In acyclic games, which include coordination games and common interest games, the adaptive learning model with sampling by Young (1993) predicts a.s. convergence to a strict Nash equilibrium. In games with strategic complementarities, potential functions, and bandwagon effects, the probabilistic learning model of Sanchirico (1996) predicts that play will remain almost always in a Nash equilibrium. If we agree that these learning models provide acceptable predictions — albeit maybe incomplete since players are often myopic — then the axioms should be acceptable in certain games.

Furthermore, we propose an aspiration-based learning model inspired from Karandikar et al. (1998). For the sake of continuity, we develop it in the appendix. Under some assumption, it produces paths of play that satisfy both axioms with probability one. This offers an (non-equilibrium) alternative to renegotiation proofness. The contribution is partly conceptual, because our notion of aspiration differs from the existing ones. Besides, the model re-emphasizes the dichotomy between families $\mathscr{F}_1$ and $\{\mathscr{F}_n\}_{n\geq 2}$. In $\{\mathscr{F}_n\}_{n\geq 2}$, the process only satisfies collective intelligence if evolution is very slow and the memory size very large.

This allows players to exploit past information to eliminate improving subsequences. In contrast, for games in $\mathscr{F}_1$, evolution can be fast, and the memory limited, yet the path of play will satisfy the axioms with probability one.

VIII Extensions

Sophistication Constraint

An analyst may have some opinion about the players' ability to detect and implement an improving subsequence. The analyst may believe that an improving subsequence can be found (by the players) only if it is no more than $K < n$ times more complex than the starting sequence, where K is a subjective bound. Nevertheless, the analyst may be studying a game in $\mathscr{F}_n$. How can we study this scenario?

We model this situation by imposing an alternative version of Axiom 2:

Axiom 3. For any stage game $G \in \mathscr{G}$, if $s \in S$ has a cyclic subsequence s' such that $\pi(s') \gg \pi(s)$, and if $\pi(s) \in \cup_{k \le K} \mathscr{U}_k$, then $s \notin \mathscr{S}(G)$.

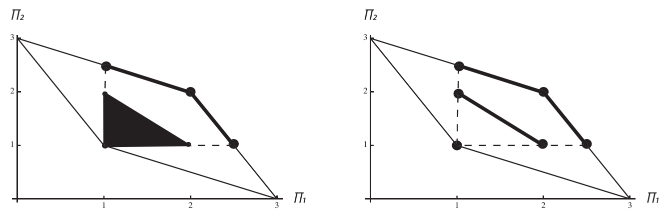
By virtue of Proposition 2, this axiom has the desired implication: if sequence s has a cycle and contains a subsequence s' such that $\pi(s') \gg \pi(s)$, and if $\pi(s) \in \cup_{k \le K} \mathscr{U}_k$, then there is a subsequence s'' of s such that s'' is at most K times more complex than s and $\pi(s'') \gg \pi(s)$.

We can easily modify our characterization theorems to apply them to this situation. Theorem 1 becomes: a solution $\mathscr{S}$ satisfies Axioms 1 and 3, if and only if, $\mathscr{S}(G)$ is a subset of the set of individually rational sequences that are also internally efficient whenever their payoff is in $\cup_{k \le K} \mathscr{U}_k$. Quite simply, internal efficiency only applies to those sequences that are guaranteed to meet the sophistication constraint. Let $\mathscr{C}^K(G)$ be the set of sequences that satisfy (i.e. do not violate) Axioms 1 and 3. We have the following result:

Theorem 3. For every game $G \in \mathscr{G}$, $\pi(\mathscr{C}^K(G))$ is dense in

$$\Pi^{IR}(G) \bigcap \left(\bigcup_{k > K} \mathscr{U}_k \bigcup \left(\bigcup_{g \in L(G)} \mathscr{P}(g) \bigcap \left(\bigcup_{k \le K} \mathscr{U}_k \right) \right) \right).$$

The new axioms and the solution do not say anything about the areas of the feasible payoffs outside of $\cup_{k \le K} \mathscr{U}_k$. In $\cup_{k \le K} \mathscr{U}_k$, however, internal efficiency implies that payoffs



(A) (B)

FIGURE IX. Solution-$\mathscr{C}^K$ payoffs in the Prisoners' Dilemma for (A) $K = 1$, and (B) $K \geq 2$.

must lie on segments. The Prisoners' Dilemma gives a nice illustration. The analyst knows this game is in $\mathscr{F}_2$ but she may have some opinion about K (see Figure IX).

Discounted Payoffs

Part of the repeated game literature is concerned with discounted payoffs. We adapt Axioms 1 and 2 to the discounted case. This gives us long-run characterizations. Cripps et al. (2004) emphasize the importance of long-run characterizations. First, an analyst may be studying an on-going relationship whose starting date she does not know. In those cases, long-run characterizations may provide useful indications about current behavior. Moreover, it is common to be interested in the limit points of a learning process, or in the steady states of a model, again directing attention to long-run behaviors.

For a sequence $s \in S$, player i's (average) discounted payoff is given by

$$\pi_i(s) = \frac{1-\delta}{\delta} \sum_{t=1}^{\infty} \delta^t u_i(s^t)$$

where $\delta \in (0,1)$ is the discount factor. Let us analyze the current axioms in this context.

Under discounting, Axioms 1 and 2 do not forgive "errors" in the early rounds of interaction. Take the following sequences from the Prisoners' Dilemma, $\{(C,D)(D,D)\ldots\}$, $\{(D,D)(D,D)(C,C)\ldots\}$ and $\{(C,D)(D,C)(C,C)\ldots\}$. In the first sequence, player 1 cooperates in period 1 and then both players play D forever. This sequence violates Axiom 1 for all $\delta < 1$. Player 1 should not have cooperated in the first period. The latter two sequences violate Axiom 2.[16] The players coordinate on the cooperative outcome after two periods, which

[16]For the last sequence, Axiom 2 is violated for all $\delta > .85$.

is a reasonably good collective outcome. According to the axiom, however, the players should not have failed to cooperate in the first two periods.

Under discounted payoffs, weaker versions of the axioms should apply to the tail of sequences. Our original motivation was concerned with eventual behaviors. Therefore, we choose to impose axioms on a tail that contains the recurrent profiles, because these profiles are observed infinitely many times. Recall the definition of recurrent profiles, $R^*(s)$, defined in (3). For any sequence $s \in S$, let $T(s) = \{\hat{s} \in S : (\exists N \in \mathbb{N})(\forall t \in \mathbb{N} \setminus \{0\}),\ \hat{s}^t = s^{t+N} \text{ and } \hat{s}^t \in R^*(s)\}$ be the set of continuations of s that only contain recurrent profiles.

Axiom 4. For any stage game $G \in \mathscr{G}$, if $s \in S$ is such that, for every $\hat{s} \in T(s)$, $\pi_i(\hat{s}) < \underline{u}_i(G)$ for some player i, then $s \notin \mathscr{S}(G)$.

Axiom 5. For any stage game $G \in \mathscr{G}$, if $s \in S$ is such that, for every $\hat{s} \in T(s)$, $\hat{s}$ has a cyclic subsequence s' such that $\pi(s') \gg \pi(\hat{s})$, then $s \notin \mathscr{S}(G)$.

Our characterization results, Theorems 1 and 2, *eventually* apply to each sequence. Precisely, if Axioms 4 and 5 hold, then $s \in \mathscr{S}(G)$ must eventually become (i.e. some tail $\hat{s}$ is) individually rational and internally efficient. The argument goes as follows. For each s, each payoff in the convex hull of the recurrent set, $\mathrm{Co}(R(s))$, can be generated as the discounted payoff $\pi(s')$ of a subsequence s' of any $\hat{s} \in T(s)$. The absence of improving subsequence must imply internal efficiency of the tail of s.

IX Conclusion

Most studies of repeated games postulate a behavior for the players via an equilibrium concept or learning dynamics. While these postulates may not be observable, they lead to infinite sequences of play that are, in principle, observable.

This paper suggests starting with the infinite sequences of play and imposing axioms on them. This methodology is similar in spirit to the revealed preference approach of consumer theory, inasmuch as the primitives of the model are, in principle, observable. Although the standard repeated game approach gives invaluable insights on strategic and learning behaviors, the axiomatic method may bring new perspectives. From the perspective of equilibrium theory, it may present valuable selection arguments where folk theorems often predict multiplicity. From the perspective of adaptive learning, it may imply a greater sophistication of the players, especially their awareness of being involved in a repeated game. The results may unify both views of repeated interactions. In dominance solvable games, for example, non-equilibrium theory often predicts a unique outcome, while equilibrium theory often predicts an extreme multiplicity.

This paper has proposed two simple axioms from which a solution follows as well as solution payoffs. Other axioms are likely to be relevant. Fortunately, our axioms are accompanied by a classification of games according to which their performance can be evaluated. This classification identifies families of games in which our theory and its predictions are more likely to apply than in others. Beyond the particular classification, we believe that the meaning of it is important. Consider, for example, the theory according to which players play iteratively undominated strategies. While this theory gives sharp predictions in dominance-solvable games, it does not mean that dominance-solvable games are games where it is more likely to apply. In general, game theory has proceeded in a seemingly universal fashion: many theories are silent about the games to which they are supposed to apply, and by default they seem to apply equally to all games. This makes it difficult to appreciate whether a particular theory is more plausible in certain games. Admittedly, our classification is only one possible classification that makes conceptual sense. Ultimately, the question of classifying theories according to contexts may be an empirical question. This seems to be an important topic in game theory.

One obvious direction for future research is experimental testing. The nature of our theory makes it an appropriate candidate. In this respect, our classification of games comes in handy. A natural starting point is the class $\mathscr{F}_1$. One way to proceed is to pursue with "higher" families, such as $\mathscr{F}_2$, etc. Section VIII offers another perspective. Games in $\{\mathscr{F}_n\}_{n\geq 2}$ may have payoff regions with the same nice properties as games in $\mathscr{F}_1$. We would expect collective intelligence to hold at least within these regions.

APPENDIX

Appendix A Complexity

An automaton is a behavior model composed of finitely many states, transitions between those states, and an action is played in each state. The definitions below are given for convenience; see Abreu and Rubinstein (1988) and Kalai and Stanford (1988) for details.

Definition 7. A (finite) automaton M_i for player i is a tuple $(Q_i, q_i^1, \lambda_i, \mu_i)$ where $Q_i \subseteq \mathbb{N}$ is a (finite) set of states; $q_i^1 \in Q_i$ is the initial state; $\mu_i : Q_i \times A \longrightarrow Q_i$ is the transition function; $\lambda_i : Q_i \longrightarrow A_i$ is the behavior function.[17]

A pair of finite machines $M = (M_1, M_2)$ induces deterministically a sequence of action profiles. The definition implies that this sequence must have a cycle. In Abreu and Rubinstein (1988), players have preferences over the nature of the automaton. Complexity enters players' utility lexicographically, giving priority to payoffs and then breaking ties according to the number of states (fewer states are preferred).

Appendix B Aspiration-Based Learning

In every period, each of two satisficing players has an (implicit) individual aspiration level. If a player's current payoff is below her aspiration level, then the player is not satisfied, and she changes her course of actions with some probability. Players' aspirations evolve in time and are determined by individual and group-related performance indicators. First, player i's aspiration level is always larger than $\underline{u}_i(G)$, for obvious reasons. Second, a player observes the recent history, and if *both* she and the other player could have done strictly better, simply by duplicating some profiles from the recent past and dropping others, then her aspiration rises above her current average payoff. The player considers that an alternative and jointly-beneficial play was within reach and yet it did not happen. Therefore the player seeks improvement.

There are many ways in which aspirations may be defined. In our model, there is a sense in which players are not as naive (or optimistic) as in earlier models. In Karandikar

[17] This definition allows a player to react to his own deviations.

et al. (1998), the future aspiration is a convex combination of current aspiration and current payoff. Therefore, if the current payoff is high, then a player's aspiration increases, whether or not this might be due to a mistake by the other player that cannot be durable. A similar remark applies to Hart and Mas-Colell (2000). The authors present a no-regret learning model. Think of regret as a manifestation of unfulfilled aspiration. In that paper, a player experiences regret at period t, if the action that she played last could have been replaced, every time that it was played in the past, by another action that would have yielded a higher average payoff, assuming that other players would have played the same actions. In the Prisoners' Dilemma, for example, player 1 experiences regret at history $((C,C),\dots,(C,C))$, because she thinks that action D would have produced $((D,C),(D,C),\dots,(D,C))$. This view is naive, because it does not take into account others' reactions to these hypothetical changes. Our model suggests that a player should experience regret for not obtaining larger payoffs only when doing so was compatible with her opponent also receiving larger payoffs.

The game G is played repeatedly through time $t = 1,2,\dots$ Players only consider the last $K \geq 3$ periods of their interaction. This may be due to memory size or computational power.[18] Abusing notation, player i's payoff from any $a = (a^1,\dots,a^K)$ is defined as $\pi_i(a) = \frac{1}{K}\sum_{n=1}^{K} u_i(a^n)$. Denote the history at time t by $\boldsymbol{a}(t) = (a^{t-K+1},\dots,a^t)$.

Definition 8. A sequence of action profiles $b = (b^1,\dots,b^K)$ is an internal improvement on sequence $a = (a^1,\dots,a^K)$, if $b^k \in \{a^1,\dots,a^K\}$ for all $k = 1,\dots,K$ and $\pi(b) \gg \pi(a)$.

An internal improvement on a is a history b that only differs from a in that it duplicates some profiles from a and drop others, and both players strictly prefer b to a.

We study the following stochastic process. Recall that $A = A_1 \times A_2$ is the set of action profiles. A state θ is a history of length K. Let A^K be the set of states. At each time $t \in \{K, 2K, 3K, \dots\}$, a player (say i) who is at state $\boldsymbol{a}(t)$ computes the set of internal improvements on $\boldsymbol{a}(t)$, possibly empty, and then her play over the next K periods is governed by individual probabilistic transitions. These one-step individual probabilities determine a K-step joint transition matrix, $Q^t_{\theta,\theta'}$, that represents the probability to have $\boldsymbol{a}(t+K) = \theta'$ given

[18]The players will compute the image of a correspondence taking the last K-period history as input. There is a sense in which larger inputs require more computational power.

$\boldsymbol{a}(t) = \theta$. We specify rules on the joint transition matrices $\{Q^t\}$. After each rule, we give an explanation based on individual transitions.

Rule 1. For all t, if $u_i(\boldsymbol{a}(t)) < \underline{u}_i(G)$ for some i, then $Q^t_{\boldsymbol{a}(t),\theta} \geq \epsilon_t$ for some θ such that $u(\theta) \geq \underline{u}(G)$.

If player i receives strictly lower payoffs than her maxmin at t, then she might play her maxmin action in the following K periods. Player j does not exclude this event and thus j might play a best-response to i's maxmin action. There is a probability of at least ϵ_t that both players exceed their maxmin levels K periods later.

Rule 2. For all t, if b is an internal improvement on $\boldsymbol{a}(t)$, then $Q^t_{\boldsymbol{a}(t),b} \geq \delta_t$.

For every internal improvement b on $\boldsymbol{a}(t)$ (there are finitely many), each player will play at $t+1$ the action prescribed to her by the first profile, b^1, with a probability of at least $\delta_t^{\frac{1}{2K}}$. If b^1 is realized at $t+1$, which occurs with probability at least $\delta_t^{\frac{1}{K}}$, then each player interprets this as a sign that her opponent may be willing to play b. Then each player will play the action prescribed by b^2 with a probability of at least $\delta_t^{\frac{1}{2K}}$ at $t+2$, and so on. Therefore, given $\boldsymbol{a}(t)$, there is a probability of at least δ_t that K periods later the history will be b.

Rule 3. For all t, if $u(\boldsymbol{a}(t)) \geq \underline{u}(G)$ and $\boldsymbol{a}(t)$ has no internal improvement, then $Q^t_{\boldsymbol{a}(t),\boldsymbol{a}(t)} \geq \mu_t$.

An individually rational state $\boldsymbol{a}(t)$ with no internal improvement is left with probability at most $1-\mu_t$.

For $\epsilon_t, \delta_t > 0$, we think of this stochastic process as a reduced form of aspiration-based learning. In every period, each player i has an individual aspiration level exceeding $\underline{u}_i(G)$. If a player's payoff does not meet her aspiration, then she switches with some probability to another course of actions for the next K periods. K periods later, she re-evaluates the situation. Precisely, either a player does not achieve her maxmin payoff, in which case she might resort to her maxmin action to secure it (rule 1), or there exist internal improvements (rule 2). In the latter, *both* she and her opponent notice that they could have done strictly better, simply by duplicating profiles from the recent past and dropping others. Therefore, they may attempt to realize these internal improvements. Rule 3 applies when a player's aspiration is met. Before presenting our result, we introduce some notation.

Assumption 1. Let $\beta_{nK} = \min\{\delta_{nK}, \epsilon_{nK}\delta_{(n+1)K}\}$. Assume

$$\lim_{T\to\infty}\sum_{\tau\geq T}\left\{\beta_{\tau K}\prod_{T\leq n\leq \tau-1}(1-\beta_{nK})\left(\prod_{\ell\geq\tau}\mu_{\ell K}\right)\right\} = 1.$$ [19]

The term in brackets bounds the probability to reach an individually rational state without internal improvement for the first time at period n, starting in period T, and to stay in this state hereafter. Let $\theta^\infty = (\theta,\theta,\ldots)$ be the infinite concatenation of history $\theta \in A^K$. Define

$$H(K,G) = \{(a^t)_{t=1}^\infty : \text{there exist } T \text{ and } \theta\in A^K \text{ such that } (a^t)_{t\geq T} = \theta^\infty, \quad \text{where } \theta \text{ admits no internal improvement and } u_i(\theta)\geq \underline{u}_i(G) \text{ for all } i\} \tag{4}$$

to be the set of infinite sequences whose cycle (of length K) is individually rational and has no internal improvements. The distance between two sets of infinite sequences, S_1 and S_2, is given by

$$d_h(S_1,S_2) = \max\left\{\sup_{s_1\in S_1}\inf_{s_2\in S_2}||\pi(s_1)-\pi(s_2)||, \sup_{s_2\in S_2}\inf_{s_1\in S_1}||\pi(s_1)-\pi(s_2)||\right\},$$

where $||\cdot||$ is the Euclidean distance. Under metric d_h, the distance between two sets of sequences is the Hausdorff distance between the set of payoffs that they generate.

Proposition 6. If Assumption 1 holds, then for all $G\in\mathscr{G}$, the adaptive procedure produces an infinite path in $H(K,G)$ with probability one, $\text{Prob}((a^t)_{t=1}^\infty \in H(K,G)) = 1$. If $G\in\mathscr{F}_1$, then $H(K,G)\subset\mathscr{C}(G)$ for all K. If $G\in\mathscr{F}_n$ with $n\geq 2$, then $\lim_{K\to\infty} d_h(H(K,G),\mathscr{C}(G)) = 0$.

The first part of the result is not surprising. The assumption implies that the probability to stay in an individually rational state with no internal improvement is 1 in the limit. In turn, this implies that the probability to reach an individually rational state with no internal improvement, starting from period T, tends to 1 as $T\to\infty$. Therefore, the process must eventually find and stay in a desired state. This happens, for example, if $\mu_t = 1$ and $\epsilon_t = \delta_t = \underline{\epsilon} > 0$ for all t, as the process is "bounded" by an absorbing Markov chain.

Although the learning model is artificial, it draws an interesting dichotomy between $\mathscr{F}_1$ and $\{\mathscr{F}_n\}_{n\geq 2}$. To see this, let $G\in\mathscr{F}_2$ be the Prisoners' Dilemma and $K=3$. History $\theta = \{(D,D),(C,D),(D,C)\}$ has no internal improvement, yet its infinite concatenation θ^∞ violates

[19] When $\tau = T$, assume $\prod_{T\leq n\leq\tau-1}(1-\beta_{nK}) = 1$.

Axiom 2 — there is an improving subsequence. As a result, $H(3,G) \not\subset \mathscr{C}(G)$. However, for games in $\{\mathscr{F}_n\}_{n\geq 2}$, collective intelligence tends to be satisfied as $K \to \infty$.[20] A large K slows down the evolution process, since it postpones the next opportunity for revision, and in addition, it increases the "memory" size. This allows players to exploit past information to eliminate improving subsequences. In contrast, for games in $\mathscr{F}_1$, evolution can be fast and the memory or computational abilities limited (K small), yet the path of play will satisfy the axioms with probability 1.

Since Assumption 1 is minimal to satisfy both axioms, the learning model suggests a "basic plot" so that the play satisfies both axioms. Players can play freely for an arbitrary amount of time, such as trying to extract as much as possible from the interaction. But there must be a time from which they are inclined to settle. However, none of them accepts to settle, i.e. to reproduce the recent history as a cycle, if they think that they can do better. Each player thinks that she can do better if the recent history admits an internal improvement (rule 2), or if she does not secure her maxmin payoff (rule 1). In all other cases, after some arbitrarily long time, each player thinks that challenging the current history cannot lead to a durably better outcome (rule 3).

Appendix C Proofs

Proposition 1

Proof. Suppose that there is a cyclic subsequence s' of s such that $\pi(s') \gg \pi(s)$. Clearly, this implies $R(s) > 1$. If $R(s) = 2$, then it must be that $R(s) = \{u, u'\}$ where $u \gg u'$. To see why, suppose that it were not the case, i.e. u and u' are Pareto unranked. Then, sequence s would yield payoff $\pi(s) \in \mathscr{P}(\mathrm{Co}(\{u, u'\}))$ because it has a cycle. This would violate the existence of a subsequence s' with the above properties. As a result, $u \gg u'$, and if we define s'' as a subsequence that only plays u', then it satisfies $\pi(s'') \ll \pi(s)$. Given that s has a cycle, if $R(s)$ contains at least three non-collinear points (otherwise the previous argument would apply),[21] then $\pi(s)$ lies in the interior of $\mathrm{Co}(R(s))$. Therefore there must be a rational convex

[20] The process is parameterized by time t and step size K, $\{\boldsymbol{a}_K(t)\}$. Our analysis studies $\lim_{T\to\infty} \lim_{K\to\infty} \boldsymbol{a}_K(t)$.
[21] See Footnote 7 for a definition of collinearity.

combination p of elements of $R(s)$ such that $p \ll \pi(s)$. This implies the existence of a cyclic subsequence s'' of s such that $\pi(s'') = p$, establishing the claim.

Suppose now that the converse holds: there is a cyclic subsequence s'' of s such that $\pi(s'') \ll \pi(s)$. This implies $R(s) > 1$ and the same arguments as the above apply. $\square$

Proposition 2

Proof. Suppose sequence s has a cycle. Consider two cases: (a) $R(s) \leq 2$, (b) $R(s) \geq 3$.

Case (a). If $R(s) = 1$, there is no subsequence s' of s such that $\pi(s') \gg \pi(s)$ and this violates the assumption. If $R(s) = 2$, then either the elements of $R(s)$ are not Pareto ranked, i.e. $R(s) = \{u, u'\}$ with $u \not\gg u'$ and $u' \not\gg u$, in which case the assumption is violated because there is no subsequence s' of s such that $\pi(s') \gg \pi(s)$, or these elements are Pareto ranked. Say $u \gg u'$ without loss of generality. Then the subsequence s'' that always plays the action profile yielding u is simpler than s (because it has a cycle of length 1) and gives each player a strictly higher payoff than s.

Case (b). Suppose $R(s) \geq 3$. If $\pi(s) \in \mathscr{P}(\text{Co}(R(s)))$, then there is no subsequence s' of s such that $\pi(s') \gg \pi(s)$, which violates the assumption of the proposition. Therefore, assume $\pi(s) \notin \mathscr{P}(\text{Co}(R(s)))$. Then there must be a triangle $\Delta = \text{Co}(\{u^1, u^2, u^3\})$ such that (*i*) $\pi(s) \in \Delta$, (*ii*) $\{u^1, u^2, u^3\} \subset R(s)$, and (*iii*) $\pi(s) \notin \mathscr{P}(\Delta)$. There are two sub-cases, depending on the cardinality of $\mathscr{N}(\Delta)$.

Case (b1). Suppose $|\mathscr{N}(\Delta)| \leq 2$. This means that a vertex of Δ is Pareto dominated. Therefore, there exist $u, u' \in R(s)$ such that $u' \gg u$. Recall that sequence s has a cycle and let a' be the action profile played in the cycle for which $u(a') = u'$. We construct subsequence s'' of s as follows: s'' plays the same cycle as s, except that each time an action profile yielding payoff u is played in the cycle of s, that profile is replaced with a'. Subsequence s'' generates a cycle whose length is smaller than or equal to that of s, and it satisfies $\pi(s'') \gg \pi(s)$.

Case (b2). Suppose $|\mathscr{N}(\Delta)| = 3$. Since s has a cycle, we can write $\pi(s)$ as a convex combination of the elements played in its cycle

$$\pi(s) = \sum_{m=1}^{M} \frac{\alpha_m}{\ell} u^m$$

where $M \geq 3$, $\{u^m\} \subset R(s)$, each $\alpha_m \in \mathbb{N}\backslash\{0\}$ and $\ell = \sum \alpha_m$ is the length of the cycle of s. Recall that $\pi(s) \in \Delta = \text{Co}(\{u^1, u^2, u^3\})$. For any $\alpha^* > 0$, define

$$\pi^* = \pi(s) + \frac{\alpha^*}{\ell}((u^2 - u^1) + q(u^2 - u^3)) \tag{5}$$

and

$$\pi^{**} = \pi(s) + \frac{\alpha^*}{\ell}((u^1 - u^2) + q(u^3 - u^2)), \tag{6}$$

where $q \in \mathbb{Q}[0,1]$ is written as $q = \nu(q)/d(q)$. Since G is in family $\mathscr{F}_n$, either $(u^2 - u^1) + q(u^2 - u^3) \gg 0$ and $d(q) \leq n$, or $(u^1 - u^2) + q(u^3 - u^2) \gg 0$ and $d(q) + \nu(q) \leq n$.[22] Therefore, we know that for all $\alpha^* > 0$, either $\pi^* \gg \pi(s)$ or $\pi^{**} \gg \pi(s)$ holds. We must verify that, in each case, we can choose α^* to obtain a well-defined convex combination. For $\alpha^* = \min\{\alpha_1, \alpha_3\}$, (5) gives

$$\pi^* = \sum_{m \neq 1,2,3} \frac{\alpha_m d(q)}{\ell d(q)} u^m + \frac{(\alpha_1 - \alpha^*) d(q)}{d(q)\ell} u^1 + \frac{\alpha_2 d(q) + \alpha^*(\nu(q) + d(q))}{d(q)\ell} u^2 + \frac{\alpha_3 d(q) - \alpha^* \nu(q)}{d(q)\ell} u^3,$$

which is a well-defined convex combination implementable by a sequence of length at most $d(q)\ell \leq n\ell$. For $\alpha^* = \alpha_2/(1+q)$, (6) gives

$$\begin{aligned} \pi^{**} = \sum_{m \neq 1,2,3} \frac{\alpha_m (d(q) + \nu(q))}{\ell(d(q) + \nu(q))} u^m + u^1 \left(\frac{\alpha_1 (d(q) + \nu(q)) + d(q)\alpha_2}{\ell(d(q) + \nu(q))} \right) \\ + u^3 \left(\frac{\alpha_3 (d(q) + \nu(q)) + \nu(q)\alpha_2}{\ell(d(q) + \nu(q))} \right) \end{aligned} \tag{7}$$

which is a well-defined convex combination implementable by a sequence of length at most $(d(q) + \nu(q))\ell \leq n\ell$. □

Theorem 2

Proof. Take any $\pi \in \pi(\mathscr{C}(G))$. Then there is a sequence $s \in \mathscr{C}(G)$ such that $\pi(s) = \pi$. Sequence s must be individually rational by Theorem 1, hence $\pi \in \Pi^{IR}(G)$. By the same theorem, s must also be internally efficient. Recall the definition of recurrent action profiles, (3). If $R^*(s) = \{a\}$ for some profile a, then $\pi \in \mathscr{P}(g)$ where $g = \{u(a)\}$. If $R^*(s) = \{a, a'\}$, then internal efficiency implies $\pi \in \mathscr{P}(g)$ where $g = \text{Co}(\{u(a), u(a')\})$. Suppose now that $R^*(s) \geq 3$.

[22] We actually know that either $(u^2 - u^1) + q(u^2 - u^3) \gg 0$ and $d(q) \leq k$, or $(u^1 - u^2) + q(u^3 - u^2) \gg 0$ and $d(q) + \nu(q) \leq k$, for some $k \leq n$.

Every point in $\mathscr{P}(\mathrm{Co}(R(s)))$, π in particular, must lie on the boundary of (i.e. a segment in) $\mathrm{Co}(R(s))$, for if π lied in the interior, then there would be a Pareto improvement over π (and s would not be internally efficient). Since π must lie on a segment $g \subset \mathrm{Co}(R(s))$, it must be that $\pi \in \mathscr{P}(g)$. Now take any $\pi \in (\cup_{g \in L(G)} \mathscr{P}(g)) \cap \Pi^{IR}(G)$. We show that we can build a sequence that satisfies both axioms and generates payoffs arbitrarily close to π. By definition, there exist $u^1, u^2 \in \Pi(G)$ such that $g = \mathrm{Co}(\{u^1, u^2\})$ and $\pi \in \mathscr{P}(g)$. If $u^1 \gg u^2$ (or vice versa), then $\mathscr{P}(g) = \{u^1\}$, hence the sequence that only plays the profile yielding payoff $u^1 = \pi$ is internally efficient. If u^1 and u^2 are not Pareto ordered, then it is possible to build a cyclic sequence s such that $R(s) = \{u^1, u^2\}$ and whose payoff $\pi(s)$ is arbitrarily close to π. Such a sequence must be internally efficient, because π^1 and u^2 are not Pareto ordered and s is cyclic. Thus, $(\cup_{g \in L(G)} \mathscr{P}(g)) \cap \Pi^{IR}(G)$ is included in the closure of $\pi(\mathscr{C}(G))$. $\square$

Proposition 3

Proof. Recall that for every $\pi \in \mathrm{Co}(R(s))$, there is a rational convex combination $\sum_k u^k \alpha_k$, where $\{u^k\} \subset R(s)$ and $\{\alpha_k\} \subset \mathbb{Q}$, that is arbitrarily close to π. By way of contradiction, suppose that sequence s satisfies the axioms but $\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(s_t)$ does not exist. Define

$$z_T = \frac{1}{T} \sum_{t=1}^{T} u(s_t).$$

By assumption, $\{z_T\}$ is divergent and since it lives in a compact space, it has at least two (distinct) cluster points z' and z'' in $\mathrm{Co}(R(s))$. By definition of the payoffs, $\underline{z} = (\min\{z'_1, z''_1\}, \min\{z'_2, z''_2\}) \geq \pi(s)$. First, consider the case where $\{z', z''\} \not\subset \mathscr{P}(\mathrm{Co}(R(s)))$. Clearly $\underline{z} \notin \mathscr{P}(\mathrm{Co}(R(s)))$, and thus there is a cyclic subsequence s' of s such that $\pi(s') \gg \underline{z}$. This violates Axiom 2. Second, consider the case where $\{z', z''\} \subset \mathscr{P}(\mathrm{Co}(R(s)))$. Then there exist $a, a' \in R^*(s)$ such that $u(a), u(a') \in \mathscr{P}(\mathrm{Co}(R(s)))$, and both z' and z'' lie on the line segment $\overline{u(a)u(a')}$. By assumption, the stage game is such that no two distinct action profiles yield either player the same payoff, $u_1(a) \neq u_1(a')$ and $u_2(a) \neq u_2(a')$. Therefore, $\underline{z} \notin \mathscr{P}(\mathrm{Co}\{R(s)\})$. There is a rational convex combination yielding $\pi \gg \underline{z}$, and thus there is a cyclic subsequence s' of s such that $\pi(s') = \pi$. This is a violation of Axiom 2. $\square$

Proposition 4

Proof. Suppose first that there exists $\pi^* \in \Pi(G)$ such that $\pi^* \gg \underline{u}(G) \geq \pi$ for all $\pi \in \Pi(G)\setminus\{\pi^*\}$ and $\underline{u}(G) \notin \Pi(G)$. By Theorem 2, $\pi(\mathscr{C}(G)) \subset (\cup_{g\in L(G)}\mathscr{P}(g)) \cap \Pi^{IR}(G)$. Since $\underline{u}(G) \notin \Pi(G)$ and $\pi^* \gg \pi$ for all $\pi \in \Pi(G)\setminus\{\pi^*\}$, the only segments $g \in L(G)$ such that $g \cap \Pi^{IR}(G) \neq \emptyset$ must be such that $\pi^* \in g$. For all segments g with $\pi^* \in g$, we have $\mathscr{P}(g) = \{\pi^*\}$. Therefore, $\pi(\mathscr{C}(G)) \subset \{\pi^*\}$. Notice that the sequence that always plays the action profile a^* with $u(a^*) = \pi^*$ is in $\mathscr{C}(G)$. Thus $\mathscr{C}(G)$ is nonempty and $\pi(\mathscr{C}(G)) = \{\pi^*\}$. Suppose now that there is $\pi^* \in \Pi(G)$ such that $\pi^* \geq \underline{u}(G) \gg \pi$ for all $\pi \in \Pi(G)\setminus\{\pi^*\}$. If $\underline{u}(G) \notin \Pi(G)$, then $\pi^* \gg \pi$ for all $\pi \in \Pi(G)\setminus\{\pi^*\}$, and the first argument of the proof applies. If $\underline{u}(G) \in \Pi(G)$, then $\underline{u}(G) = \pi^*$ by assumption, so that $L(G) \cap \Pi^{IR}(G) = \{\pi^*\}$. Applying Theorem 2 and building a sequence that always produces payoff π^* show that $\mathscr{C}(G)$ is nonempty. Hence $\pi(\mathscr{C}(G)) = \{\pi^*\}$. □

Proposition 5

Proof of Proposition 5. Take any renegotiation-proof equilibrium σ. Trivially, $\pi_i(\sigma) \geq \underline{u}_i(G)$ for all i, for otherwise some player would not be playing a best-response. By way of contradiction, suppose that $\boldsymbol{a}(\sigma)$ violates Axiom 2. Then there exists a cyclic subsequence s' of $\boldsymbol{a}(\sigma)$ such that $\pi(s') \gg \pi(\boldsymbol{a}(\sigma))$. We want to build a subgame-perfect equilibrium σ' that has s' as its equilibrium path. Therefore, σ' will satisfy $R^*(\boldsymbol{a}(\sigma')) \subset R^*(\boldsymbol{a}(\sigma))$ and will Pareto dominate σ. Since σ is a subgame-perfect equilibrium, for each player i there must be a profile $a^i \in \cup_h J(\boldsymbol{a}(\sigma|_h))$ such that $\max_{a_i} u_i(a_i, a^i_j) \leq \pi_i(\boldsymbol{a}(\sigma))$. If this were not the case, then σ_i would not be a best-response to σ_j because $\max_{a_i} u_i(a_i, \sigma_j(h)) > \pi_i(\boldsymbol{a}(\sigma))$ for all $h \in H$.[23] Therefore, for each player i, there must be such a profile a^i in $\cup_h J(\boldsymbol{a}(\sigma|_h))$. Of course this implies $u_i(a^i) \leq \pi_i(\boldsymbol{a}(\sigma))$ hence $u_i(a^i) < \pi_i(\boldsymbol{a}(\sigma'))$. From here, the proof will follow the equilibrium construction from Rubinstein (1994) (Proposition 146.2, p. 147) and rely on one critical observation. We construct a pair of automaton strategies that produces (the cycle of) sequence s' on its equilibrium path. We also specify the strategies of the players so that if player i deviates from the equilibrium path, and thus deserves to be punished, then player j plays a^i_j and player i responds by playing a^i_i. Any punishment begins in the period that

[23] Consider for example $\hat{\sigma}_i : H \to S_i$ defined as $\hat{\sigma}_i(h) = br_i(\sigma_j(h))$ for all $h \in H$, the pointwise best-response to σ_j. Since $u_i(\hat{\sigma}_i(h), \sigma_j(h)) > \pi_i(\boldsymbol{a}(\sigma))$ for all h, then $\pi_i(\boldsymbol{a}(\hat{\sigma}_i, \sigma_j)) > \pi_i(\boldsymbol{a}(\sigma))$.

follows the completion of the cycle. A deviation is punished for enough periods to cancel out any possible gain for the deviator, but then players return to the equilibrium path, starting at the beginning of the cycle. As a result, after every possible history h, the payoff profile from $\sigma'|_h$, $\pi(\boldsymbol{a}(\sigma'|_h))$, is equal to $\pi(\boldsymbol{a}(\sigma'))$. Thus, σ' satisfies requirement (i) from Definition 6. The critical observation is that player i's (repeated game) strategy is optimal, although player i plays a_i^i in her punishment phase and a_i^i may not be a best-response to a_j^i. This is because the punishment phase is finite and player i could not do better than $\pi_i(\boldsymbol{a}(\sigma'))$ in the long-run by best-responding to a_j^i in the punishment phase. Since σ' only uses a^1 and a^2 off-path, it is clear that $\cup_h \boldsymbol{J}(\boldsymbol{a}(\sigma'|_h)) \subset \cup_h \boldsymbol{J}(\boldsymbol{a}(\sigma|_h))$. This is a contradiction of the renegotiation-proofness of σ. $\square$

Proposition 6

Proof of Proposition 6. The proof consists of three lemmas:

Lemma 1. If $G \in \mathscr{F}_1$, then $H(K,G) \subset \mathscr{C}(G)$ for every $K \geq 1$.

Proof. Take any sequence $(a^t)_{t=1}^{\infty} \in H(K,G)$ whose cycle is denoted by $\theta = (\theta^1,\ldots,\theta^K)$. Suppose by way of contradiction that $(a^t)_{t=1}^{\infty} \notin \mathscr{C}(G)$. Since $\pi_i((a^t)_{t=1}^{\infty}) \geq \underline{u}_i$ for all i by definition of $H(K,G)$ (see (4)), it must be that there exists a cyclic subsequence s' of $(a^t)_{t=1}^{\infty}$ such that $\pi(s') \gg \pi((a^t)_{t=1}^{\infty})$. Since $G \in \mathscr{F}_1$, Proposition 2 implies that there is a cyclic subsequence s'' of $(a^t)_{t=1}^{\infty}$ such that s'' is simpler than $(a^t)_{t=1}^{\infty}$ and $\pi(s'') \gg \pi((a^t)_{t=1}^{\infty})$. Letting θ'' be the cycle of s'', we have $\pi(\theta'') \gg \pi(\theta)$. Therefore, θ'' is an internal improvement on a, a contradiction. $\square$

Lemma 2. If $G \in \mathscr{F}_n$ with $n \geq 2$, then $\lim_{K\to\infty} d_h(H(K,G),\mathscr{C}(G)) = 0$.

Proof. Take any sequence $(a^t)_{t=1}^{\infty} \in H(K,G)$ with cycle $a = (a^1,\ldots,a^K)$. Consider the set $R_a = \{a^1,\ldots,a^K\}$. Suppose first that $\mathscr{P}(R_a)$ has only one point (say a^1). Then it must be that $a = (a^1,\ldots,a^1)$, for otherwise a would have an internal improvement. Therefore there is a time after which $(a^t)_{t=1}^{\infty}$ only plays a^1, so $(a^t)_{t=1}^{\infty} \in \mathscr{C}(G)$. Suppose now that $\mathscr{P}(\text{Co}(R_a))$ has at least two points. This implies that $\mathscr{P}(\text{Co}(R_a))$ consists of $M_a \geq 1$ segments (see Section V for a definition). Let $\{g_m\}_{m=1}^{M_a}$ be the set of segments such that $\cup_{m=1}^{M_a} g_m = \mathscr{P}(R_a)$. Take any segment g_m, cut it up into K pieces of equal length, and collect the $K+1$ extremities of those

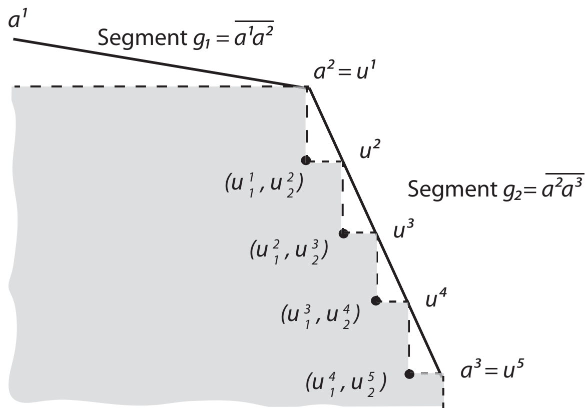


FIGURE X. $\overline{D}(E_a^2)$ is the union of the white triangles along segment g_2

K pieces into set $E_a^m = \{u^n\}_{n=1}^{K+1}$ where the u^n's are labeled in decreasing order w.r.t. their second coordinate (See Figure X for an illustration where g_2 is divided into $K = 4$ pieces). Each element of E_a^m represents a payoff that can be generated by a finite sequence of action profiles with K elements from R_a.[24] For each $u \in E_a^m$, consider the set of payoffs that are strictly smaller than u: $D(u) = \{w \in \mathbb{R}^2 : w_1 < u_1 \text{ and } w_2 < u_2\}$. Let $\overline{D}(E_a^m) = \mathrm{Co}(R_a) \cap (\cup_{n=1}^{K+1} D(u^n))^c$ — c stands for complement — be the set of payoffs that are (strictly) undominated by any of E_a^m; Figure X depicts $\overline{D}(E_a^2)$. Using our labeling assumption, we have

$$
\begin{aligned}
\sup_{u \in \overline{D}(E_a^m)} \inf_{w \in g_m} \|u - w\| &= \max_{n=1,\dots,K} \inf_{w \in g_m} \|(u_1^n, u_2^{n+1}) - w\| \\
&= \sup_{w \in g_m} \inf_{u \in \overline{D}(E_a^m)} \|u - w\|,
\end{aligned} \tag{8}
$$

for each $a \in A^K$ with no internal improvement and segment $g_m \in \mathscr{P}(R_a)$. This equation simply says that the furthest point in $\overline{D}(E_a^m)$ from g_m must take the form (u_1^n, u_2^{n+1}); see Figure X. For every $m = 1, \dots, M_a$, it holds that

$$
\lim_{K \to \infty} \max_{n=1,\dots,K} \inf_{w \in g_m} \|(u_1^n, u_2^{n+1}) - w\| = 0. \tag{9}
$$

Moreover it must be that $\pi(a) \in \overline{D}(E_a^m)$, for otherwise an alternative cycle of length K would generate payoff $u^n \gg \pi(a)$ for some $u^n \in E_a^m$. In that case, $(a^t)_{t=1}^{\infty}$ would not be in $H(K,G)$, a contradiction. Therefore, $\pi(a)$ and any payoff attached to a sequence in $H(K,G)$ must be in

[24] In Figure X, u^4 can be generated by a cycle of length $K = 4$ by playing three times a^3 and one time a^2.

some $\overline{D}(E_a^m)$. It follows from (8) and (9) that $\overline{D}(E_a^m)$ converges to $g_m \subset \mathscr{C}(G)$ for all m, and thus $\lim_{K\to\infty} d_h(H(K,G),\mathscr{C}(G)) = 0$. □

Lemma 3. $\text{Prob}((a^t)_{t=1}^{\infty} \in H(K,G)) = 1$.

Proof. The process $\{\boldsymbol{a}(nK)\}_{n=1}^{\infty}$ is a non-stationary Markov chain with set of states A^K and transition matrices $\{Q_{\theta,\theta'}^{nK}\}$. Let $\phi : A^K \twoheadrightarrow A^K$ be the correspondence that assigns to each $\theta \in A^K$ the set of internal improvements on θ. The set of states can be partitioned into three sets: Θ_1 denotes the set of states $\theta \in A^K$ for which there is i with $\pi_i(\theta) < \underline{u}_i$; Θ_2 denotes the set of states θ such that $\phi(\theta) \neq \emptyset$ and $\pi_i(\theta) \geq \underline{u}_i$ for all i; Θ^* consists of all the other states, i.e. the states for which $\phi(\theta) = \emptyset$ and $\pi_i(\theta) \geq \underline{u}_i$ for all i. We will prove that under Assumption 1,

$$\lim_{n\to\infty} \text{Prob}\left(\boldsymbol{a}(\tau K) = \theta \in \Theta^* \text{ for all } \tau \geq n\right) = 1. \tag{10}$$

Note that for all finite T,

$$\text{Prob}((a^t)_{t=1}^{\infty} \in H(K,G)) = \text{Prob}((a^t)_{t=1}^{\infty} : (a^t)_{t=T}^{\infty} \in H(K,G)),$$

and thus $\text{Prob}((a^t)_{t=1}^{\infty} \in H(K,G)) = 1$ whenever equality (10) holds. Under the parameter restrictions, the learning dynamic satisfies

$$\text{Prob}(\boldsymbol{a}((n+2)K) \in \Theta^* | \boldsymbol{a}(nK) \in \Theta_1) \geq \epsilon_{nK}\delta_{(n+1)K}, \tag{11}$$

$$\text{Prob}(\boldsymbol{a}((n+1)K) \in \Theta^* | \boldsymbol{a}(nK) \in \Theta_2) \geq \delta_{nK}, \tag{12}$$

and $\text{Prob}(\boldsymbol{a}((n+1)K) = \theta | \boldsymbol{a}(nK) = \theta \in \Theta^*) \geq \mu_{nK}$. According to Assumption 1,

$$\lim_{T\to\infty} \sum_{n>T} \left\{ \beta_{nK} \prod_{T\leq m<n} (1-\beta_{mK}) \left(\prod_{\ell \geq n+1} \mu_{\ell K} \right) \right\} = 1,$$

where $\beta_{nK} = \min\{\delta_{nK}, \epsilon_{nK}\delta_{(n+1)K}\}$. Consider sequence $\{p_n\}$ where p_n is the product $\prod_{\ell\geq n+1} \mu_{\ell K}$. By way of contradiction, suppose that $\lim_{n\to\infty} p_n \neq 1$. Since sequence $\{p_n\}$ is monotone, $\lim_{n\to\infty} p_n \neq 1$ implies the existence of $\overline{b} < 1$ such that $p_n < \overline{b}$ for all n. Then, for all T,

$$\sum_{n>T} \left\{ \beta_{nK} \prod_{T\leq m<n} (1-\beta_{mK}) \left(\prod_{\ell\geq n+1} \mu_{\ell K} \right) \right\} \leq \overline{b} \sum_{n>T} \left\{ \beta_{nK} \prod_{T\leq m<n} (1-\beta_{mK}) \right\} \leq \overline{b},$$

which is in contradiction with Assumption 1. Thus, $\lim_{n\to\infty} p_n = 1$. Define

$$\mathscr{A}_T^n = \left\{\boldsymbol{a}(nK) = \theta \in \Theta^* \text{ and } \boldsymbol{a}(mK) \notin \Theta^* \text{ for all } T \le m < n\right\}$$

to be the event that the stochastic process reaches Θ^* at time nK for the first time since TK. For large n, if the process reaches Θ^*, i.e. $\boldsymbol{a}(nK) \in \Theta^*$, then it never leaves it because $\lim_{n\to\infty} p_n = 1$. Therefore,

$$\begin{aligned} \lim_{n\to\infty} \text{Prob}\left(\boldsymbol{a}(\tau K) = \theta \in \Theta^* \text{ for all } \tau \ge n\right) &= \lim_{T\to\infty} \text{Prob}\left(\cup_{n>T} \mathscr{A}_T^n\right) \\ &\ge \lim_{T\to\infty} \sum_{n>T} \left\{\beta_{nK} \prod_{T\le m<n} (1-\beta_{mK})\right\}, \end{aligned} \tag{13}$$

where the inequality is a consequence of (11) and (12). The rhs of (13) is equal to 1 for otherwise Assumption 1 would be violated given $\lim_{n\to\infty} p_n = 1$. We conclude that (10) holds. □

This completes the proof of Proposition 6. □

References

Abreu, D., Pearce, D. G., 1984. On the inconsistency of certain axioms on solution concepts for non-cooperative games. Journal of Economic Theory 34 (1), 169–174.

Abreu, D., Rubinstein, A., 1988. The structure of nash equilibrium in repeated games with finite automata. Econometrica 56 (6), 1259–81.

Arifovic, J., McKelvey, R. D., Pevnitskaya, S., 2006. An initial implementation of the turing tournament to learning in repeated two-person games. Games and Economic Behavior 57 (1), 93–122.

Aumann, R., Shapley, L., 1994. Long-term competition—a game-theoretic analysis. In: Megiddo, N. (Ed.), Essays in Game Theory. Springer-Verlag, New York, pp. 1–15.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New-York.

Bernheim, D. B., Ray, D., 1989. Collective dynamic consistency in repeated games. Games and Economic Behavior 1 (4), 295–326.

Blonski, M., Ockenfels, P., Spagnolo, G., 2011. Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. American Economic

Journal: Microeconomics 3 (3), 164–192.

Brandenburger, A., 1992. Knowledge and equilibrium in games. Journal of Economic Perspectives 6 (4), 83–101.

Brandenburger, A. M., Nalebuff, B. J., 1999. Coopetition: A Revolution Mindset That Combines Competition and Cooperation, 1st Edition. Currency Doubleday, New York.

Compte, O., Postlewaite, A., 2010. Plausible cooperation, department of Economics, University of Pennsylvania.

Cripps, M., Mailath, G., Samuelson, L., 2004. Imperfect monitoring and impermanent reputations. Econometrica 72 (2), 407–432.

Dal Bo, P., 2005. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. American Economic Review 95 (5), 1591–1604.

Dal Bo, P., Frechette, G., 2011. The evolution of cooperation in infinitely repeated games: Experimental evidence. American Economic Review 101 (1), 411–29.

Duffy, J., Ochs, J., 2009. Cooperative behavior and the frequency of social interaction. Games and Economic Behavior 66 (2), 785–812.

Er'ev, I., Roth, A. E., 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8, 164–212.

Farrell, J., Maskin, E., 1989. Renegotiation in repeated games. Games and Economic Behavior 1 (4), 327–360.

Fudenberg, D., Levine, D. K., 1998. The Theory of Learning in Games, Cambridge, Massachusetts.

Harsanyi, J. C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press Books. The MIT Press.

Hart, S., Mas-Colell, A., 2000. A simple adaptive procedure leading to correlated equilibrium. Econometrica 68 (5), 1127–1150.

Holt, C. A., 1985. An experimental test of the consistent-conjectures hypothesis. American Economic Review 75 (3), 314–25.

Kalai, E., Lehrer, E., 1993. Rational learning leads to nash equilibrium. Econometrica 61 (5), 1019–45.

Kalai, E., Stanford, W., 1988. Finite rationality and interpersonal complexity in repeated games. Econometrica 56 (2), 397–410.

Karandikar, R., Mookherjee, D., Ray, D., Vega-Redondo, F., 1998. Evolving aspirations and cooperation. Journal of Economic Theory 80 (2), 292–331.

Kohlberg, E., Mertens, J.-F., 1986. On the strategic stability of equilibria. Econometrica 54 (5), 1003–37.

Mailath, G., Samuelson, L., 2006. Repeated Games and Reputations: Long-Run Relationships. Oxford University Press, New-York.

Milgrom, P., Roberts, J., 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica 58 (6), 1255–1277.

Milgrom, P., Roberts, J., 1991. Adaptive and sophisticated learning in normal form games. Games and Economic Behavior 3 (1), 82–100.

Ray, D., 1994. Internally renegotiation-proof equilibrium sets: Limit behavior with low discounting. Games and Economic Behavior 6, 162–177.

Roth, A., Murnighan, J. K., 1978. Equilibrium behavior and repeated play of the prisoners' dilemma. Journal of Mathematical Psychology 39 (1), 83–96.

Rubinstein, A., 1994. A Course in Game Theory, 1st Edition. MIT Press, Cambridge, Massachusetts.

Sanchirico, C. W., 1996. A probabilistic model of learning in games. Econometrica 64 (6), 1375–93.

Sonsino, D., 1997. Learning to learn, pattern recognition, and nash equilibrium. Games and Economic Behavior 18 (2), 286–331.

Thomson, W., 2001. On the axiomatic method and its recent applications to game theory and resource allocation. Social Choice and Welfare 18 (2), 327–386.

Young, H. P., 1993. The evolution of conventions. Econometrica 61 (1), 57–84.