An Axiomatization of Plays in Repeated Games*

Laurent Mathevet †

May 17, 2017

Abstract

Suppose that, in repeated games, players eventually engage in a pattern of action profiles, which we call a *convention*. Do some conventions seem more plausible than others? We answer axiomatically based on the principles of *stability* and *efficient simplicity*. The main solution says that conventions should be constant repetitions of a static Nash equilibrium, or such that players switch between two Pareto unranked profiles (across which they each change action). In some repeated games, this reduces the multiplicity of outcomes and even leads to uniqueness. The paper also reports experimental evidence that supports our findings.

Keywords: conventions, axioms, pattern mining, complexity, stability, evolution, equilibrium selection.

JEL Classification: C72, C73, D83, C92.

*This paper was started while the author was a faculty member at the University of Texas at Austin. I am especially grateful to Aditya Kuvalekar, Jonathan Lhost, David Pearce, Antonio Penta, Julian Romero, Philippe Solal, Dale Stahl, Max Stinchcombe, Ina Taneva, and Tom Wiseman for many discussions and suggestions. I also wish to thank an associate editor and an anonymous referee for their helpful comments and suggestions, as well as Matthias Blonski, In-Koo Cho, David Frank, Guillaume Frechette, Drew Fudenberg, Herve Moulin, Mike Peters, Ariel Rubinstein, Sahotra Sarkar, Rick Szostak, and Peyton Young, and seminar/conference participants at NYU, Purdue, Rice, UBC, UCR, Columbia, UW Madison, UPenn, the 2012 North American Summer Meeting of the Econometric Society, and the 2013 Choice Symposium.

†Dept. of Economics, New-York University, 19 W. 4th Street, 6FL, New York, NY 10012, U.S.A. (lmath@nyu.edu)

1 Introduction

Repeated interactions are often governed by a form of convention, at least in the long run. Colleagues who regularly go to lunch together alternate who pays the bill; pedestrians in a city choose to walk on the same side of the footpath every day (to pass each other); and non-violence can settle in over long periods between enemy soldiers at war (Axelrod (1984)). Which conventions do we expect to emerge and stay? On the one hand, the standard theory of repeated games is not very discriminating in this matter, as patient players would do anything to avoid future punishments. On the other hand, the learning literature offers a plethora of learning models and of answers.[1] This severe multiplicity creates a lack of perspective, as if conventions obeyed no general principles. A closer look at the literature reveals that common forces are at work in the construction of social norms: *efficiency* (Ray (1994)); *complexity* (Abreu and Rubinstein (1988)); *justifiability* (Spiegler (2005)); etc. The time may be ripe for a different perspective that puts these forces at the center of the analysis.

In this paper, we introduce a framework—not explicitly based on standard equilibrium or learning analysis—in which we can express, combine, and compare various principles to study conventions. Suppose that players eventually enter into a pattern of action profiles, which we call a "convention." Do some conventions seem more plausible than others? This question refers to norms that we might expect players to arrive at and settle on. To answer this question, we propose axioms that conventions might satisfy and characterize those that satisfy all axioms.

This framework generalizes axiomatic bargaining (Nash (1950), Kalai and Smorodinsky (1975), etc.) and Harsanyi and Selten (1988, Chapter 3). Since conventions induce distributions of action profiles that, in turn, produce payoffs, any axiom on payoffs (as in axiomatic bargaining) or on distributions over action profiles (as in Harsanyi and Selten) can be expressed as an axiom on conventions. This matters because payoffs have no strategic content without actions and distributions have no time structure. We use this framework in environments with two completely patient players (i.e., whose payoff functions are given by the limit of means) and perfect monitoring. We illustrate the framework with a few axioms, the implications of which are mathematically straightforward but thought-provoking.[2]

[1] When we vary the initial conditions of learning models, allow different players to follow different learning rules, and allow learning to change over time, the set of all learning predictions explodes.

[2] Among other things, the framework shows that extreme multiplicity is sometimes questionable.

We also include experimental evidence from Mathevet and Romero (2012) that supports our solution. The data consist of over 400 sequences generated by human subjects, who played eight different stage games for more than 100 periods on average. We used pattern detection algorithms to extract all sequences that exhibit a pattern, which represents about 2/3 of the data. Overall, at least 80% of the observed patterns satisfy the combination of our strictest axioms.

Our particular axioms tell the following story: players try creating surplus from their interaction, but they also fight over the distribution of the surplus. Consider the following examples:

	S	H
S	4,4	0,3
H	3,0	2,2

Stag Hunt

	$\overline{\text{P}}$	P
$\overline{\text{E}}$	2,3	0,2
E	3,1	1,4

Dinner Game

In Stag Hunt, two individuals go out on a hunt every day. Each can choose to hunt a stag or a hare. Assume that the history of play becomes common knowledge after each period and that players are completely patient. Consider the convention $((S,H)(H,S)\ldots)$, in which players miscoordinate forever. The axiom of *individual rationality* will eliminate this convention because each player earns strictly less, on average, than what she could secure by playing action H every period.

Now consider the convention $((H,H)(S,S)\ldots)$, in which players alternate forever between hunting a hare and a stag. The axiom of *efficient simplicity* will eliminate this convention. The idea is simple: players would not increase the complexity of their conventions to both earn less. In the above, the alternation not only is more complicated (than just playing the good outcome), but it also harms both players. In other words, players complexify their social norms only if doing so benefits at least one of them; otherwise, they would not do it.

Multiplicity is important to explain the richness of outcomes in the observed world, but 'extreme' multiplicity—according to which nearly everything is possible—denies regularities in human behavior. In this framework, many axioms and solutions can be thought of and hence there is also a form of multiplicity. However, this criticism is of orders of magnitude less here, because to each given set of axioms on conventions (observables) we can build infinitely many models (based on unobservable constructs) satisfying them.

While simplicity is a feature of our social norms, instability can result from excessive simplicity. Consider the dinner game. The row player is a child and the column player is the parent, and they play the game every day before dinner. The child chooses to eat a cookie (E) or to not eat it ($\bar{E}$). The parent decides to punish (P) or to not punish ($\bar{P}$) the child. The child has a dominant action in eating the cookie, and the parent prefers to punish the child only if she eats the cookie. Consider the convention $((E, \bar{P})(\bar{E}, \bar{P})\ldots)$, in which the child eats the cookie one time out of two and the parent never punishes the child. The axiom of *stability* eliminates this convention. The idea is that a player may behave myopically unless she has reason to believe that her behavior might get punished.[3] In the above, the child never suffers the negative consequence of eating the cookie and, thus, she may start eating it more often and destabilize the convention.

As shown above, patterns of play are selected based on what happens in the pattern itself. This emphasizes observed events and the self-sustainability of conventions. Note that the framework is agnostic about whether the outcomes come from a naive or a sophisticated reasoning process. In particular, we do not explicitly model players' off-path behavior. By standard folk theorems, however, one can think of conventions as equilibrium paths of play, as long as they produce individually rational payoffs. In this perspective, the idea is that on-path considerations drive equilibrium selection and discipline what happens off-path implicitly (see Footnote 3 for a concrete example). This kind of consideration already appears in Rubinstein (1986) and his notion of semi-perfect equilibrium, where threats should be played on the path or else disappear from a player's machine to save on the maintenance costs. It also appears in Spiegler (2005) and his notion of Nash equilibrium with tests, where past events observed on the path justify current responses on the path.

In the paper, we first introduce stability and then propose a weak and a strong axiom of efficient simplicity. Then we characterize our main solution: the theorem states that conventions should be constant repetitions of a static Nash equilibrium or such that players switch between two Pareto unranked profiles across which they both change actions. This implies that solution payoffs are line segments, which is reminiscent of Abreu and Rubinstein (1988). We document the difference in examples. In some games, as in Stag Hunt, the only conventions that survive our two weakest axioms are constant

[3] Although we do not model what happens off-path, the axiom of stability could be seen as requiring that players assign zero probability to all actions of their opponent that they have not observed in a very long time. That is, in a standard equilibrium analysis, a player would not believe that her opponent can do off-path what she is not doing on path.

repetitions of a static Nash equilibrium. Thus, to obtain more outcomes, players must either be individually irrational or increase the complexity of their social norms to both earn less. In the Supplement, we also classify all stage games into families, based on geometric properties of their feasible payoffs, and mathematically tie each family to the degree of players' sophistication required by the strong axiom of efficient simplicity.

This paper connects different branches of the literature. Efficient simplicity is in the spirit of renegotiation proofness (Bernheim and Ray (1989), Farrell and Maskin (1989)), and stability is in the spirit of Spiegler (2005)'s Nash equilibrium with tests. Our characterizations are also reminiscent of the literature on bounded rationality (Abreu and Rubinstein (1988)). The axiomatic treatment of repeated games is novel, with the exception of the recent paper by Blonski, Ockenfels, and Spagnolo (2011) in which the authors propose an axiomatic framework for selecting equilibria in the repeated Prisoners' Dilemma. They define a solution as a subset of the (stage game) payoffs and discount factors for which some cooperation should arise on the path. Their axioms characterize a unique solution.

The paper is organized as follows. Section 2 presents the framework. Sections 3 and 4 introduce the axioms. Section 5 contains the characterizations. Section 6 studies basic properties of our solutions: payoff invariance and existence. Section 7 discusses the interpretation of the approach and offers experimental evidence. Section 8 concludes. Proofs are in the Appendix.

2 Preliminaries

2.1 The Model

We consider repeated interactions between two players. Let $G = (A_1, A_2, u_1, u_2)$ be a finite two-person game in normal form, where A_i is player i's finite action set; $A = A_1 \times A_2$ is the set of action profiles; and $u_i : A \to \mathbb{R}$ is i's utility function. Let Σ_i be i's set of mixed actions with typical element α_i. The utility functions are extended to mixed actions by taking expectations. Let

$$\underline{u}_i(G) = \max_{\alpha_i \in \Sigma_i} \min_{a_j \in A_j} u_i(\alpha_i, a_j) \tag{1}$$

denote player i's maxmin payoff. In stage game G, we assume that action profiles give rise to non-collinear payoffs and that $\underline{u}_i(G) > u_i(a)$ whenever a player i obtains the same utility $u_i(a) = u_i(a')$ across two distinct action profiles a and a'.[4]

A repeated interaction consists of an infinite sequence of repetitions of a (commonly known) stage game G at discrete time periods $t = 1, 2, \ldots$ In every period t, the players make simultaneous moves—denoted by $a_i^t \in A_i$—that become common knowledge. These choices can be realizations of randomization devices.

A player's payoff will be evaluated by her average utility. Given a sequence $s = (s^1 s^2 \ldots)$, where $s^t = (a_1^t, a_2^t) \in A$, let $\pi_i(s) = \liminf_{T\to\infty} 1/T \sum_{t=1}^{T} u_i(s^t)$.[5] This specification assumes that players are completely patient. There are two reasons for studying this case. The first one is simplicity—for now, we want to abstract away from patience issues. Second, since complete patience is associated with severe equilibrium multiplicity, selection is especially useful in this case.

2.2 Conventions

This section introduces the notion of *convention*. A sequence $s = (s^1 s^2 \ldots)$ has convergent frequencies if

$$\lim_{T\to\infty} \sum_{t=1}^{T} \mathbf{1}_{\{s^t = a\}}/T \in [0, 1]$$

for all $a \in A$. A sequence s is *cyclic* if there exist a time T and a non-negative number $\ell(s)$ (called the length of the cycle) such that $s^t = s^{t+\ell(s)}$ for all $t \geq T$, and there is no $\ell < \ell(s)$ for which $s^t = s^{t+\ell}$ for all $t \geq T$. That is, there is a time after which some fixed pattern repeats itself forever.

Definition 1. *A convention in a repeated game is a sequence with convergent frequencies such that any action profile that appears once appears infinitely often.*

Axiomatized conventions will represent the plausible outcomes of the repeated game. This interpretation does not prevent a normative use of our framework, in which conventions would be interpreted as normative outcomes, but it is not our point of view in this paper. See Section 4.2 for a discussion.

[4] These assumptions are generic. Recall that non-collinearity means that for any a^1, a^2, a^3 in A, $u(a^1)$, $u(a^2)$, and $u(a^3)$, where $u(a) = (u_1(a), u_2(a))$, must not lie on a line.

[5] The liminf is the infimum of the cluster points of $(\overline{s}^t)$, the sequence of average payoffs from 1 to t.

Some action profiles can have zero frequency in a convention and yet appear infinitely often. These profiles or actions are called *occasional* and are played less and less frequently as time passes. As such, they are a basic expression of learning within a convention and describe an evolution of habits in time (for example, players could punish each other less and less in time out of resignation or unnecessity). Note also that players are allowed to play mixed actions, in which case a convention is the realization of random play. Mixed strategies, public randomization devices and regular Markov chains all produce conventions.

As is well-known, automaton strategies produce cyclic conventions (Abreu and Rubinstein (1988)). The fact that a fixed pattern repeats itself forever gives a time structure to cyclic conventions that can be exploited by the axioms. When axioms do not exploit the time structure of cyclic conventions, they are equivalent to axioms on distributions over action profiles. See Section 7.2 for a discussion.

Note also that (non-pure) mixed action cycles typically will induce non-cyclic conventions. This is also true for automata with probabilistic state transitions.

In what follows, fix a stage game G and let $\pi \gg \pi'$ mean that π strictly Pareto dominates π', i.e., $\pi_i > \pi'_i$ for all i.

3 Stability

For any convention s, define

$$\begin{aligned} R(s) &= \{a \in A : s^t = a \text{ for some } t\}, \quad (2)\\ R_j(s) &= \{a_j \in A_j : \exists a_i \in A_i \text{ s.t. } a \in R(s)\} \\ \Gamma(s) &= \{(u_1(a), u_2(a)) : a \in R(s)\} \end{aligned}$$

to be the set of action profiles, actions of player j, and payoff profiles that appear in convention s.[6] Consider the following axiom:

Axiom 1. *s is such that no i has a mixed action $\alpha_i \in \Sigma_i$ that satisfies $u_i(\alpha_i, a_j) > \pi_i(s)$ for all $a_j \in R_j(s)$.*

If a player can do strictly better by playing a mixed action—against every action played by the other player in the convention—than by following the convention, then

[6]By definition of a convention, if a profile appears once, it appears infinitely often.

she might play that mixed action and destabilize the convention. Quite simply, if a player has such a mixed action, then she may want to try it.

This axiom is especially plausible in a world with bounded memory: if a convention is to stay in place forever, then players may eventually forget what happened in the early stages, and so every player i may eventually believe that j's actions are confined to $R_j(s)$ (that is, every player may eventually assign zero probability to all actions that her opponent has not played in a very long time). This condition on beliefs is enough to imply the axiom in a standard equilibrium analysis. As a consequence, players must display *in* their conventions the arguments deterring one another from deviating.[7]

Take the Prisoners' Dilemma (on p.22), and let $a = (C, C)$ and $b = (D, C)$. The fully cooperative convention $s = (aa\ldots)$ is discarded by the axiom because $u_1(D, C) = 4 > 2 = \pi_1(s)$, and the convention $s' = (ab\ldots)$ is also discarded for a similar reason. In both s and s', player 2 never retaliates, so that player 1 might want to try defecting more often. In this game, Axiom 1 eliminates all cyclic conventions that give a payoff profile on the boundary of the feasible payoffs. For cyclic conventions s that generate an interior payoff profile, Axiom 1 requires nothing more than individual rationality (i.e, $\pi(s) \geq (2, 2)$), because, for such s, $R_i(s) = A_i$ for all i. In the Prisoners' Dilemma, therefore, the payoff set associated with cyclic conventions that survive this axiom is the set of individually rational payoffs in $\mathbb{Q}^2$ excluding the Pareto frontier.

While the above argument applies to cyclic conventions, moving beyond cyclic conventions can get us to the Pareto frontier. Imagine if the players play D for every prime t. Then, D occurs occasionally, and hence leaves the payoffs unchanged. At the same time, now D no longer does strictly better against every observed action of the opponent. Therefore, we can sustain cooperation and hence the frontier. This argument and the benefit of occasional strategies are developed in Section 5.

[7] This is related to justifiability à la Spiegler (2005). In Spiegler's equilibrium concept, players play, on the equilibrium path, myopic best responses to what they observe until they get punished for it (on the path), and only the presence of punishments on the path allows them to justify to themselves why they should play non-myopic actions and dynamically optimize.

4 Efficient Simplicity

4.1 Axioms

Most notions of complexity agree that playing a single action profile is simpler than playing it among others. Thus, players would rather play a single action profile than play it as part of a more elaborate convention that decreases their payoffs:

Axiom 2. *s is such that no $a \in R(s)$ satisfies $u(a) \gg \pi(s)$.*[8]

Think of choosing which side of the footpath to walk on everyday. Most pedestrians choose the same side every day to pass each other, left or right, while changing sides depending on the day would make matters more complicated and provide no benefit. Note that humans did not adopt complex pedestrian norms, and then cooperated to move to simpler ones. It is just that complex ones never occurred in the first place. Following this logic, a group may first consider simple conventions, constant in particular, and only complexify them if this benefits at least one member of the group. In this sense, any s having $a \in R(s)$ such that $u(a) \gg \pi(s)$ would not occur in the first place. If one agrees with this logic, there is no learning needed to transition from s to a, since the former never occurs. A related but different justification of this axiom is dynamic in nature: if a group makes an error and adopts s having $a \in R(s)$ such that $u(a) \gg \pi(s)$, then it is plausible that the group would eventually think of a and coordinate on it by some adjustment process (unspecified here).

On a different note, observe that, unlike this axiom, Pareto efficiency ($\pi(s)$ lies on the Pareto frontier of the feasible payoffs) says nothing about complexity.

Apart from repeating one profile, there could be other social arrangements that are simpler than a given convention. If it is simpler to play one profile than to alternate between two, then perhaps it is simpler to play shorter cycles in general. And so, we could use the length $\ell(\cdot)$ of a cycle as complexity measure.[9]

Now, using measure $\ell(\cdot)$, let us imagine that players can complexify their social norms up to a certain limit when this benefits them. This is Axiom 3(n).

[8]This axiom will be invoked in the context of cyclic conventions.

[9]This is a natural but coarse measure. Given a convention of length ℓ, there always exists a pair of automata with, at most, ℓ states that produce this convention. Automata also specify off-equilibrium behaviors that are not taken into account here. Of course, there exist other measures of complexity, based, for example, on data compression, such as Kolmorogov complexity. Similar axioms can be defined, and other solutions derived, under these other measures.

Definition 2. *A convention s' is at most $n \in \mathbb{N}$ times more complex than s if $\ell(s') \leq n\ell(s)$ and $R(s') \subseteq R(s)$.*[10]

Axiom 3(n). s is such that no s' that is at most n times more complex than s satisfies $\pi(s') \gg \pi(s)$.

For $n = 1$, this axiom says that players would not increase the complexity $\ell(\cdot)$ of their conventions—and play s rather than s'—to both earn less. For $n > 1$, however, players can complexify their conventions "by a factor of at most n"[11] when this benefits them. To avoid characterizing a solution for every n, we focus on Axiom 3(∞) instead, defined as follows:

Axiom 3(∞). s is such that no s' satisfying $R(s') \subseteq R(s)$ also satisfies $\pi(s') \gg \pi(s)$.

This axiom drops the complexity requirement and allows players to form cyclic conventions of any complexity or even "non-cyclic" conventions, as long as they play action profiles only from s. Importantly, in the Supplement to this paper, we classify all stage games into families $\{n : n \in \mathbb{N}\}$ such that, in family n, s survives Axiom 3(n) if and only if it survives Axiom 3(∞). This result tells us that characterizations involving Axiom 3(∞) are identical to those we would obtain with the much weaker Axiom 3(n) provided G is in family n.

We now reformulate Axiom 3(∞) in the lemma below.

Definition 3. *A sequence s' is a subsequence of s if there exists a strictly increasing function $h : \mathbb{N} \to \mathbb{N}$ such that $s'^t = s^{h(t)}$ for all t.*

Axiom 3. *s is such that no subsequence s' of s satisfies $\pi(s') \gg \pi(s)$.*

Lemma 1. *Any s survives Axiom 3(∞) if and only if s survives Axiom 3.*

Proof. Since the requirement that $R(s') \subseteq R(s)$ is equivalent to the requirement that s' is a subsequence of s, both axioms are equivalent. □

If a convention s contains another convention s' that strictly benefits both players, then Axiom 3 rules s out, the idea being that players would eventually play the latter

[10] Here, we do not even rely on profiles played outside s, especially since players have not shown the ability to do so.

[11] Some Pareto improvements might not be achieved by any s' such that $\ell(s') \leq n\ell(s)$ and require s' with $\ell(s') \geq (n + 1)\ell(s)$. So, the complexity bound has bite.

and not s. For stage games in family n, it follows from the lemma that s survives Axiom 3(n) if and only if it survives Axiom 3.

Take the example of convention $(abb\ldots)$ in the Prisoners' Dilemma (on p.22), where $a = (D, C)$ and $b = (C, C)$. This is a strictly Pareto-improving subsequence of $s = (abac\ldots)$, given $c = (C, D)$, and hence s is discarded by the axiom.[12]

Axiom 3 allows for inefficiencies, such as 'always defect' in the Prisoners' Dilemma, but it excludes conventions in which mutually harmful behaviors co-exist indefinitely with good ones. Finally, Pareto efficiency implies Axiom 3: if a convention is Pareto efficient (i.e., $\pi(s)$ lies on the Pareto frontier of the feasible payoffs), then it cannot contain a Pareto improving subsequence, and thus it survives Axiom 3.

4.2 Normative or Positive?

The axioms from the previous section impose different forms of collective rationality. In a model where communication is not modeled explicitly, and where there is no explicit centralized process, do these axioms seem plausible?

First, the degree of collective rationality varies greatly from Axiom 2 to Axiom 3. There is a whole sequence of Axioms 3 (n) in between the two, in which collective rationality increases with $n \in \mathbb{N}$. In Axiom 2 and Axiom 3 (1), the degree of collective rationality is so weak that a group can plausibly meet the requirement without explicit communication.

Second, Section 7.3 reports basic results from laboratory experiments that replicated the conditions of standard repeated games, *without* communication between subjects or centralized process. These results provide evidence in support of our axioms, Axiom 3 included. Most of the games tested in the experiments belong to "family 1,"[13] a family in which s satisfies Axiom 3 if and only if it satisfies the much weaker Axiom 3 (1). This reinforces our previous point.

Finally, the fact that communication is not modeled explicitly does not mean that it is forbidden. Our point of view is that communication is not necessary to satisfy such basic requirements as Axiom 2. However, greater collective rationality, as in Axiom 3, may indeed require communication in some games. This line of argument is the same as in Farrell and Maskin (1989) and Bernheim and Ray (1989), which introduce the re-

[12] To see this, remove the underlined letters from $s = (ab\underline{ac}\,\underline{a}bac\,\underline{a}b\underline{ac}\,aba\ldots)$.

[13] Families of games are mentioned in the previous section and defined in the Supplement. The latter also contains a link to a program that outputs the family of any 2×2 stage game entered as input.

finement of renegotiation proofness into standard repeated games: players agree ex-ante to play a subgame-perfect equilibrium, but they are able to renegotiate the continuation after every period, a process which is not modeled explicitly.

5 Characterizations

The results in this section characterize, fully or partially, the set of conventions that satisfy the axioms. Existence will be studied in Section 6.2.

5.1 Stability and Efficient Simplicity

5.1.1 Main Results

Stability is equivalent to requiring that a convention pay every player i at least as much as she could get by playing her best mixed action in a worst-case scenario,

$$\pi_i(s) \geq \max_{\alpha_i \in \Sigma_i} \min_{a_j \in R_j(s)} u_i(\alpha_i, a_j) \quad \forall i. \tag{3}$$

The rhs is the minimal payoff that a player believes she can secure, if her beliefs give probability zero to the unobserved actions of the other player. This expression is always weakly larger than the standard maxmin in (1), hence stability implies individual rationality (formally defined as Axiom 4 in the next section).

A convention is *individually rational* if each player receives at least her maxmin payoff (1), and it is *almost constant* if some action profile is played with frequency one.

Theorem 1.
(1) The set of cyclic conventions that satisfy Axioms 1 and 3 contains only cyclic individually rational s that are either (i) constant such that a static Nash equilibrium is played in every period or (ii) such that players switch between two Pareto unranked action profiles, a and b, with $a_i \neq b_i \forall i$.

(2) The set of conventions satisfying Axioms 1 and 3 contains only individually rational s that are either (i) almost constant (some a^* is played with frequency 1) or (ii) such that players switch between two Pareto unranked action profiles, a and b, almost all the time. Moreover, if a^* is not a static Nash equilibrium (because player i strictly wants to deviate) in (i) or if $a_j = b_j$ in (ii), then for each $\hat{a}_i \in A_i$, there is $\hat{a}_j$ with $u_i(\hat{a}) \leq \pi_i(s)$ that player j plays occasionally.

Cyclic conventions are important in repeated games because this is the class generated by automata (Abreu and Rubinstein (1988)). In this case, the axioms imply that players must play the same static Nash equilibrium in all periods, or they must switch between two Pareto unranked profiles such that the convention is individually rational and both change actions across these profiles. This result places significant restrictions on the outcomes. Figure 2 in the Appendix illustrates it with well-known examples. For repeated two-player two-action (2×2) games, the most studied case in the literature, the first part of the theorem provides a full characterization (see Figure 2).

Corollary 1. *For 2×2 stage games, the set of all cyclic conventions satisfying Axioms 1 and 3 is the set of individually rational s that are either (i) constant such that a static Nash equilibrium is played in every period, or (ii) such that players switch between two Pareto unranked action profiles, a and b, with $a_i \neq b_i \forall i$.*

Proof. From part (1) of Theorem 1, we only need to show that all cyclic individually rational s that satisfy either (i) or (ii) also satisfy Axioms 1 and 3. Take a cyclic individually rational s that satisfies (i). Since s is constant, $R(s) = \{a\}$, and a is a Nash equilibrium, $u_i(a) = \max_{\alpha_i \in \Sigma_i} u_i(\alpha_i, a_j)$. Therefore, s satisfies Axiom 1. Any constant convention must satisfy Axiom 3, because it has no proper subsequence. We conclude s satisfies both axioms. Now, take a cyclic individually rational s that satisfies (ii). Since the game is 2×2 and players switch between two Pareto unranked action profiles, a and b, with $a_i \neq b_i \forall i$, we have $R_i(s) = A_i$ for all i. Therefore, Axiom 1 is equivalent to individual rationality. Moreover, the alternation between two Pareto unranked action profiles prevents the existence of a subsequence s' of s such that $\pi(s') \gg \pi(s)$. Hence, s satisfies Axioms 1 and 3. □

In the cyclic case, the tension between stability and efficient simplicity results in strong selection, because it forces punishments to be used and requires some (internal) efficiency. Beyond cyclic conventions, however—that is, if players are more sophisticated than the deterministic automaton model allows—the second part of the theorem allows for violations of the Nash or the bilateral switching requirement if, in response, occasional profiles play a strategic role of deterrence. When players play the same non-Nash profile in every period, or when one player only switches action in the Pareto alternation, there always exists a player who wants to deviate from the convention. The occasional profiles serve as a reminder of why he should not do so.

Let us illustrate the importance of the latter in an example. Consider Buchanan's (1977) Samaritan's Dilemma.

	W	$\overline{\text{W}}$
$\overline{\text{H}}$	2,2	1,1
H	5,3	3,5

Samaritan's Dilemma

A player chooses to help (H) or not to help ($\overline{H}$) another player accomplish a task, and the other player chooses to work (W) or not to work ($\overline{W}$). The dilemma is that the samaritan's help is crucial to both players' welfare, but if she helps, then the other player prefers not to work.

While it is straightforward to see that $(H, \overline{W})$ is the unique NE, experimentally it is observed that $(H, \overline{W})$ and (H, W) appear remarkably often. In light of Theorem 1, it is clear that a cyclic convention cannot give rise to such outcomes. The use of occasional profiles sheds light on why we might observe such a play. The row player prefers (H, W) but, the column player would prefer to play $\overline{W}$ if the row player chooses H. The way the row player can induce the column player to choose W is by choosing not to help ($\overline{H}$) occasionally to serve as a reminder. Since not helping too often would harm both players, it must be used only sparingly. Indeed, the second part of the theorem shows that outcomes besides $(H, \overline{W})$ are consistent with our axioms when we allow for occasional profiles.

5.1.2 Maximal Conventions

A simple characterization of Axioms 1 and 3 may not be available in general, but one is for maximal conventions. A convention s is *maximal* if there is no other convention s' such that $\pi(s') = \pi(s)$ and $R(s') \supseteq R(s)$, as defined in (2). That is, a convention is maximal if no other convention generates the same payoffs with additional action profiles. Clearly, maximal conventions must include all action profiles a, some with probability zero and some not. Let Co denote the convex hull.

Proposition 1. *The set of all maximal conventions that satisfy Axioms 1 and 3 is*

$$\left\{s : s \text{ is maximal, individually rational, and } \nexists \pi \in Co(\Gamma(s)) \text{ s.t. } \pi \gg \pi(s)\right\} \quad (4)$$

Proof. For a maximal convention s, $R_i(s) = A_i$ for all i, so that Axiom 1 is satisfied if and only if s is individually rational. Now, it remains to prove that a convention (in fact, a maximal one) satisfies Axiom 3 if and only if there is no $\pi \in \text{Co}(\Gamma(s))$ such that $\pi \gg \pi(s)$. This is shown in Lemma 2 in Section B.2 of the Appendix. □

This result is illustrated in Figure 2 in the Appendix.

5.2 Individual Rationality and Efficient Simplicity

In this section, we characterize solutions under individual rationality.

Axiom 4. *s is such that no i gets $\pi_i(s) < \underline{u}_i(G)$.*

Proposition 2. *The set of payoffs generated by all cyclic conventions that satisfy Axioms 2 and 4 is*

$$\left\{\pi : \exists s \text{ s.t. } \pi = (\pi_1(s), \pi_2(s)) \text{ and } \nexists u \in \Gamma(s) \text{ s.t. } u \gg \pi\right\}. \tag{5}$$

In the Prisoners' Dilemma (on p.22), any payoff in the white triangular area is eliminated, as it requires playing (C, C), but (C, C) is better for both players.

By all accounts, Axioms 2 and 4 are basic requirements, yet they are enough to dissipate most equilibrium multiplicity in many games. In Stag Hunt, for example, the only conventions that survive these two axioms are constant play of a static Nash equilibrium. There even are games for which Axioms 2 and 4 imply uniqueness: Proposition 5 (on p.23) identifies a class of common-interest games in which a *unique* convention satisfies the axioms.

Proposition 3. *The set of conventions that satisfy Axioms 3 and 4 only contains individually rational s that are almost constant or in which players spend almost all the time switching between two Pareto unranked action profiles.*

The proof is omitted because it is a consequence of Proposition 1. The comparison between Proposition 3 and the theorem points to the weaker strategic implications that follow from weakening stability. For example, any constant individually rational convention satisfies Axioms 3 and 4, whether or not it is a Nash equilibrium.

Both propositions are illustrated in Figure 3.

6 Further Properties

6.1 Payoff Invariance

The strategic aspects of a game remain essentially the same if players' utilities are subjected to a positive linear transformation. So, it is desirable that a solution be unaffected by such transformations (see p.70 in Harsanyi and Selten (1988)). A solution $\mathcal{S}$ (which is the set of all conventions satisfying some axioms) is said to be *payoff-invariant* if for any stage games G and G', where the utility functions in G' are positive linear transformations of those in G—i.e., $u_i' = \beta_i u_i + \delta_i$ with $\beta_i > 0$ and $\delta_i \in \mathbb{R}$—, we have $\mathcal{S}(G) = \mathcal{S}(G')$. Since the liminf is linear in the space of conventions, and since the min and the max operators are also linear, implying that the maxmin is too, a convention survives any one of our axioms in G, if and only if, it survives it in G'. Thus, the solution satisfying any combination of our axioms will be payoff-invariant.

6.2 Existence

It is straightforward to see that if there exists a pure strategy Nash equilibrium in the stage game, then playing it forever is a (cyclic) convention that satisfies Axioms 1 and 3. However, can we say something meaningful in games that may not have a pure Nash equilibrium? Proposition 1 answers this question positively if (4) is nonempty, and the next proposition proves that it is.

Proposition 4. *For any finite two-person stage game, the set of conventions that satisfy Axioms 1 and 3 is nonempty.*

Proof. Let $\pi^* = (\pi_1^*, \pi_2^*)$ be a feasible payoff on the Pareto frontier and such that $\pi_i^* \geq \underline{u}_i(G) \forall i$ (the existence of such π^* follows from the definition of $\underline{u}_i(G)$). Since π^* lies on the Pareto frontier, there exist two action profiles, a and b, and a convention s such that $R(s) = \{a, b\}$ and $\pi(s) = \pi^*$. Now consider an alternative convention, s', that plays a and b with the same frequencies as s, but also plays all the other profiles $A \backslash \{a, b\}$ infinitely often but with frequency zero. Then s' must survive Axiom 3, since there can be no Pareto improvements on $\pi(s)$, and it must also survive Axiom 1, since stability comes down to individual rationality (Axiom 4) when $R_i(s) = A_i$ for all i. □

7 Discussion

7.1 An Equilibrium Approach

In repeated situations, we could interpret conventions as real-time interactions between the players (learning perspective), or as dynamic norms that players attain through an unspecified process after playing for some time (equilibrium perspective). In this paper, we take an equilibrium perspective and interpret conventions as the long-run or stable outcome of the relationship.

But the equilibrium perspective prompts questions. How do conventions emerge? How does a player know in real time that the dynamic interaction will never satisfy some principle (e.g., individual rationality) and, thus, know how to adjust her behavior, unless she sees the entire future?

Standard equilibrium theory prompts similar questions. If a Nash equilibrium of a repeated game represents players' interaction in real time, then how does each player know all the future consequences of her move today? Similarly, how do players know their opponents' repeated game strategies? And if Nash equilibria of repeated games are not played in real time, then what process leads to them? Nachbar (2005) shows that there is no obvious answer.

Conventions are the starting point of our analysis. We do not explain their emergence or guarantee it, but if some conventions do emerge, our goal is to construct a solution that captures those and only those.

7.2 Time Structure

Unlike non-cyclic conventions, cyclic conventions have a time structure that axioms can exploit. Say that an axiom exploits the time structure of (cyclic) conventions when one can find (cyclic) s and s' such that s survives said axiom but s' does not, while all $a \in A$ are played with the same frequencies in s as in s'.

Although the main axioms in this paper ignore the time structure, they are particularly plausible in repeated (hence, temporal) environments. Take individual rationality. Although it seems hard to dispute, two players can easily miscoordinate in Battle of the Sexes if they play it once, but they would not make a habit of it if they played repeatedly.

Importantly, other axioms than those presented in this paper can be studied in our framework, including some that insist on the time structure. Here is a simple axiom that exploits the time structure, only provided for the sake of example:

Axiom 5. (Preference for shorter cycles) *s is such that there is no s' for which $\pi(s) = \pi(s')$, $R(s) = R(s')$ and $\ell(s') < \ell(s)$.*

In the spirit of Axiom 3(1), it is reasonable to believe that players would not increase the complexity of their conventions to obtain the exact same payoffs. This implies that players may play $s' = (ababab...)$, for example, instead of $s = (aaaabbbbaaaabbbb\ldots)$. Axiom 5, however, says nothing about the order of action profiles within the cycle.

7.3 Towards a Positive Theory of Repeated Games

Given an experimental data set, it is easy to draw basic conclusions about the performance of our theory, due to the observable nature of its predictions. Here, we consider the experimental data set of Mathevet and Romero (2012).

The data consist of 434 sequences of play generated by human subjects who played 8 different stage games for more than 100 periods on average. In each session, subjects were paired with a new partner and played a repeated game with a new stage game. The length of the interactions was chosen randomly (see Roth and Murnighan (1978)) with termination probability .01, which corresponds to discount factor $\delta = .99$.[14]

The games tested cover a wide range of 2×2 games, including symmetric, asymmetric, coordination, anti-coordination, prisoners' dilemma, threat-vulnerable, and common-interest games.

Two striking features emerge from the data. First, players often adopt a cyclic convention. About two thirds of all sequences exhibit a cyclic pattern, which we extracted by different procedures (e.g., pattern extraction algorithms) to get robust conclusions. Second, *among the observed cyclic conventions, more than 90% are constant or such that the players switch between two Pareto unranked action profiles* (not necessarily with 50-50 frequency).

Among the constant conventions, most are a static Nash equilibrium. Two important exceptions are (*i*) the Prisoners' Dilemma, where mutual cooperation is very common; and (*ii*) Chicken, where players massively play (C, C). It is also interesting to see that

[14] The authors considered another treatment in which each repeated game started with 30 rounds with certainty, after which $\delta = 0.9$. Since the differences between the two data sets were minor, they merged them.

	C	D
C	3,3	1,4
D	4,1	2,2

Prisoners' Dilemma (70 obs)

	C	D
C	1,1	2,4
D	4,2	1,1

Battle of the Sexes (70 obs)

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

Stag Hunt (50 obs)

	C	D
C	3,3	1,4
D	4,1	0,0

Chicken (70 obs)

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

Common Interest (18 obs)

	W	$\overline{\text{W}}$
$\overline{\text{H}}$	2,2	2,2
H	5,3	3,5

Samaritan's Dilemma (60 obs)

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

Ultimatum Game (60 obs)

	C	D
C	4,1	1,2
D	1,2	2,1

Mixed Game (36 obs)

Figure 1: 2×2 Games Tested and Number of Observations

players in the Ultimatum game play the static Nash equilibrium (C, D) and not the (non-equilibrium) profile (C, C), even though both give the same payoffs.

Pareto switching happens in Battle of the Sexes, Chicken (rarely), and the Samaritan's Dilemma. In the data, players always switch between *two* Pareto unranked profiles, and both players always switch actions across these profiles, except in (*iii*) the Samaritan's Dilemma, where the samaritan always helps and the other player works one time out of two.

We do not believe that instances (*i*), (*ii*) and (*iii*) necessarily contradict the theory. Take the Samaritan's Dilemma. Although many subjects end up helping the column player all the time, while column players work one time out of two, this convention usually appears in the data after many punishments by the samaritan. Compared to the other observed convention, which is the static Nash equilibrium, there are many more punishments in the periods that precede the alternations work/not-work (and this difference is statistically significant). This observation is in line with Theorem 1. Our conjecture is that samaritans establish a reputation that causes the other player to work, but this reputation is a stock that disappears if not replenished. This points to the role of the occasional profiles in our theory, but we have little to say empirically at this point. In infinitely repeated games, events such as 'punishing with vanishing frequency' are delicate to interpret in a finite horizon, as in the real world.

8 Conclusion

We have presented a new approach to determine which social norms should result from repeated play of a game. This approach gives perspective to situations traditionally dominated by multiplicity.

When some axiomatic solution seems to be valid, we may want to explain it by modeling players' behavior and describe explicitly how they can come to satisfy it. This paper does not address this issue, but it seems promising to first propose a set of reasonable axioms, and only then build (learning or equilibrium) models that satisfy them.

A Figures

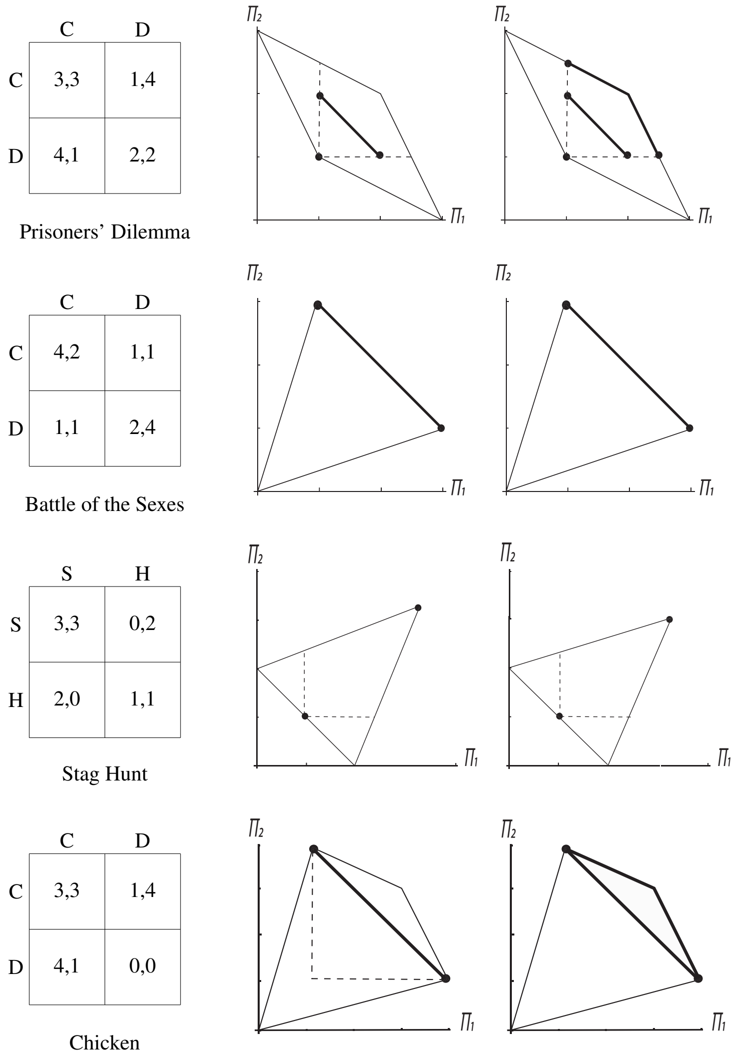

Figure 2: Payoff Sets under Stability and Efficient Simplicity (Axioms 1 and 3. Left: cyclic, Right: (maximal) general conventions).

Prisoners' Dilemma

	C	D
C	3,3	1,4
D	4,1	2,2

Battle of the Sexes

	C	D
C	4,2	1,1
D	1,1	2,4

Stag Hunt

	S	H
S	3,3	0,2
H	2,0	1,1

Chicken

	C	D
C	3,3	1,4
D	4,1	0,0

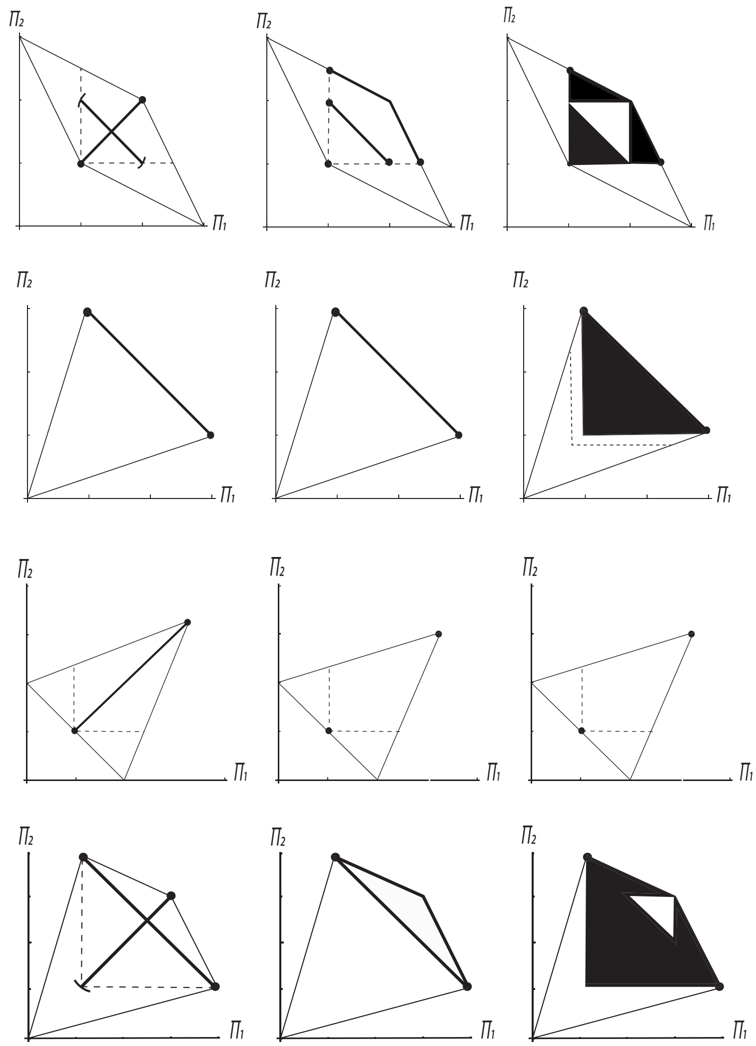

Figure 3: Payoff Sets under Individual Rationality and Efficient Simplicity (Left: Abreu and Rubinstein (1988)'s equilibrium payoffs; middle: Axioms 3 and 4; right: Axioms 2 and 4)

B Proofs

B.1 Uniqueness Result

Let $\underline{u}(G) = (\underline{u}_1(G), \underline{u}_2(G))$ be the maxmin payoff vector in stage game G (Section 3). The next proposition defines a family of common interest games—games where $u_i(a^*) \geq u_i(a)$ for all a implies $u_j(a^*) \geq u_j(a)$ for all a—for which a unique convention survives our two weakest axioms.

Proposition 5. *Suppose that game G has an action profile a^* such that $u(a^*) \gg u(a)$ and $u_1(a) + u_2(a) < \underline{u}_1(G) + \underline{u}_2(G)$ for all $a \neq a^*$. Then, the set of cyclic conventions that satisfy Axioms 2 and 4 is $\{(a^* \ldots)\}$.*

Proof. If players play any cyclic convention s involving a^* and another profile, then $\pi_i(s) < u_i(a^*)$ for all i, because $u_i(a^*) > u_i(a)$ for all i and hence Axiom 2 is violated. If players play any cyclic convention s that does not involve a^*, then $u_1(G) + u_2(G) > \pi_1(s) + \pi_2(s)$, because $\underline{u}_1(G) + \underline{u}_2(G) > u_1(a) + u_2(a)$ for all $a \neq a^*$. As a result, there must be some player i for whom $\pi_i(s) < \underline{u}_i(G)$, which contradicts Axiom 4. In conclusion, the only convention that survives both axioms is the constant play of profile a^*. □

B.2 Theorem 1

Define a convention s to be internally efficient if there is no $\pi \in \mathrm{Co}(\Gamma(s))$ such that $\pi \gg \pi(s)$.

Lemma 2. *A convention s satisfies Axiom 3 if and only if it is internally efficient.*

Proof. First, suppose that s satisfies Axiom 3. For every $\pi \in \mathrm{Co}(\Gamma(s))$, there is a subsequence s' of s such that $\pi(s') = \pi$. By Axiom 3, there is no subsequence s' such that $\pi(s') \gg \pi(s)$. Thus, there is no $\pi \in \mathrm{Co}(\Gamma(s))$ such that $\pi \gg \pi(s)$. By definition of $\mathcal{P}(\cdot)$, this implies that $\pi(s) \in \mathcal{P}(\mathrm{Co}(\Gamma(s)))$. Suppose now that s is internally efficient. Then $\pi(s) \in \mathcal{P}(\mathrm{Co}(\Gamma(s)))$. There cannot be a subsequence s' of s such that $\pi(s') \gg \pi(s)$, for otherwise there would be a convex combination (in particular, one using rational numbers as coefficients) of elements of $\Gamma(s)$ that Pareto dominates $\pi(s)$. Therefore, Axiom 3 holds. □

Proof of Theorem 1.

1. It follows from Lemma 2 that a convention s satisfies Axiom 3 if and only if it is internally efficient, that is, $\pi(s) \in \mathcal{P}(\mathrm{Co}(\Gamma(s)))$. Every point in $\mathcal{P}(\mathrm{Co}(\Gamma(s)))$ lies on the boundary of $\mathrm{Co}(\Gamma(s))$. This means that $\pi(s)$ lies on a line segment connecting two pure payoff profiles. Therefore, either s is almost constant (if $\pi(s)$ is an extremity of the line segment), or players switch almost all the time between two Pareto unranked action profiles (which are the extremities of the line segment on which $\pi(s)$ lies. As a result, the set of cyclic conventions that satisfy Axioms 1 and 3 only contains individually rational conventions that are either (*i*) constant or (*ii*) in which players switch between

two Pareto unranked action profiles, $a^1 = (a_1^1, a_2^1)$ and $a^2 = (a_1^2, a_2^2)$. In case (i), let a be the unique profile played under convention s, $R(s) = \{a\}$. Since $R_1(s) = a_1$ and $R_2(s) = a_2$, Axiom 1 immediately implies that a_1 and a_2 are mutual best-responses, and hence a must be a Nash equilibrium of the stage game. In case (ii), suppose by way of contradiction that $a_1^1 = a_1^2$ (the argument is similar for $a_2^1 = a_2^2$). Then there is $a_2 \in \{a_2^1, a_2^2\}$ for which $u_2(a_1, a_2) > \pi_2(s)$ for all $a_1 \in R_1(s) = \{a_1^1\}$ (because in the games we study, $a = a'$ whenever $u_i(a) = u_i(a')$ for some i). This violates Axiom 1.

2. The initial part follows from Lemma 2. Suppose now that convention s falls into one of these two cases: some a^* is played with frequency 1 but $u_i(a_i', a_j^*) > u_i(a^*)$ for some a_i', or player j does not switch action across the two Pareto unranked profiles a and b played almost all the time. In both cases, player j plays a unique action a_j^* with frequency 1 (in case (ii), denote $a_j^* = a_j = b_j$). If this were the only action that j plays, $|R_j(s)| = 1$, then s would violate Axiom 1 for the above reasons. Since s satisfies the axioms, for every action $\hat{a}_i \in A_i$, there must be some $\hat{a}_j \in R_j(s)$ such that $u_i(\hat{a}) \leq \pi_i(s)$. Because a_j^* is played with frequency 1, no $\hat{a}_j$ can be played with strictly positive frequency. Thus, actions $\hat{a}_j$ are played occasionally. □

References

Abreu, D. and A. Rubinstein. 1988. "The Structure of Nash Equilibrium in Repeated Games with Finite Automata." *Econometrica* 56 (6):1259–81.

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

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

Blonski, M., P. Ockenfels, and G. Spagnolo. 2011. "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence." *American Economic Journal: Microeconomics* 3 (3):164–192.

Buchanan, J. M. 1977. *Freedom in Constitutional Contract*. Texas A&M University Press.

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

Fudenberg, Drew and David K. Levine. 1998. Cambridge, Massachusetts: The Theory of Learning in Games.

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

Kalai, E. and M. Smorodinsky. 1975. "Other Solutions to Nash's Bargaining Problem." *Econometrica* 43 (3):513–18.

Mathevet, Laurent and Julian Romero. 2012. "Predictive Repeated Game Theory: Measures and Experiments."

Nachbar, John H. 2005. "Beliefs in Repeated Games." *Econometrica* 73 (2):459–480.

Nash, J. 1950. "The Bargaining Problem." *Econometrica* 18 (2):155–162.

Ray, D. 1994. "Internally Renegotiation-Proof Equilibrium Sets: Limit Behavior with Low Discounting." *Games and Economic Behavior* 6:162–177.

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

Rubinstein, A. 1986. "Finite automata play the repeated prisoner's dilemma." *Journal of Economic Theory* 39 (1):83–96.

Spiegler, Ran. 2005. "Testing threats in repeated games." *Journal of Economic Theory* 121 (2):214–235.