Procedural Rationality in Repeated Games*

Rajiv Sethi†

November 12, 2019

Abstract

This paper considers finitely repeated games played by procedurally rational players, who sample their available alternatives and use realized payoffs as a basis for strategy selection. The corresponding solution concept is that of (payoff) sampling equilibrium, which is a distribution over strategies that is self-replicating under the sampling procedure. Sampling equilibria are rest points of a disequilibrium dynamic process, and stability with respect to this process can be used as an equilibrium selection criterion. The structure of stable sampling equilibria in symmetric, finitely repeated games is characterized, and illustrated with applications to cooperation and coordination over time.

*For comments on an earlier version I thank Yuval Levin, Ariel Rubinstein, Karl Schlag, and conference participants at the Stockholm School of Economics.

†Department of Economics, Barnard College, Columbia University and the Santa Fe Institute.

1 Introduction

Nash equilibrium remains the dominant solution concept in the theory of games, but gives rise to paradoxical and counter-intuitive predictions in some environments. This is especially true in games with a long but finite sequence of moves, such as the centipede, the chain store game, and the finitely repeated prisoner's dilemma. The usual resolution of such paradoxes has relied on the introduction of behavioral types who mechanically choose particular strategies, and whose presence is taken into account by rational players. In some instances it makes sense for rational players to mimic behavioral types for extended periods of time, and this can give rise to sharply different predictions than would arise in the standard complete information setting.

For a different approach to the resolution of such paradoxes, one may turn to alternative conceptions of rationality and equilibrium. This paper explores a model of procedural rationality, originally introduced by Osborne and Rubinstein (1998), that is based on the idea that players sample their available alternatives and use realized payoffs as a basis for strategy selection. The corresponding solution concept is a distribution over strategies that is self-replicating under the sampling procedure. This may be interpreted as the the steady state of a dynamic process in which new entrants to a large population decide on a strategy by trying out each available alternative, and picking from among those that yield the highest realized payoff. Stability with respect to this dynamic process can then be used as an equilibrium selection criterion.

The focus in this paper is on stable sampling equilibria of symmetric finitely repeated games. The set of strategies in any such game can be partitioned into equivalence classes based on the following criterion: two strategies belong to the same class if and only if they give rise to the same sequence of outcomes when self-matched. For example, in the repeated prisoner's dilemma, all strategies that are never the first to defect belong to one class, while those that are never the first to cooperate belong to another. The equivalence classes obtained in this manner can be (weakly) ordered based on the payoffs generated in within-class matches. Again using the example of the prisoner's dilemma, the highest payoff class is composed of strategies that are never the first to defect, while the lowest payoff class is composed of those that are never the first to cooperate. Between these two extremes lie classes of strategies that generate outcomes involving varying levels of cooperation and defection in within-class matches.

Such a partition of strategies turns out to be useful in characterizing stable sampling equilibria in repeated games, provided that tie-breaking is regular in the following sense:

when multiple strategies are tied for best under the sampling procedure, each of these is selected with positive probability. This includes, but is not limited to uniform tie-breaking. Since ties arise generically in repeated games, the set of sampling equilibria is sensitive to the manner in which ties are assumed to be broken. A regular sampling equilibrium is defined as one in which all alternatives that are tied for best under sampling are selected with positive probability. At any such equilibrium, strategies that are played with positive probability are said to be present, as are classes that contain at least one present strategy. If all strategies in a class are present, the class is said to be fully present. A best present class at any sampling equilibrium is one that obtains the highest payoff in within-class matches among all classes that are present.

It is shown that the best present class at any regular sampling equilibrium must be fully present. This is then used to show that the sequence of outcomes that arise with positive probability in a regular sampling equilibrium cannot be Pareto-dominated when the number of repetitions is sufficiently large. These results have implications for cooperation and coordination over long periods of time. In particular cooperation must arise with positive probability in the repeated prisoners' dilemma, and a payoff-dominant strategy must be played with positive probability in a repeated coordination game. In repeated common interest games, there exists an efficient sampling equilibrium that is locally asymptotically stable under the sampling dynamics. And in the two-player case, this equilibrium is (almost) globally asymptotically stable, despite the possible existence of many other equilibria. Dynamic stability thus serves as a very powerful refinement criterion in common interest games.

Since the set of strategies in a repeated game rises very sharply with the number of repetitions, it often makes sense to restrict attention to strategies that have bounded complexity. Restricting attention to strategies that can be represented by finite automata with an exogenous bound on the number of states, sharper results can be obtained. In particular, in the finitely repeated prisoners' dilemma, when strategies are restricted to those that can be represented by finite automata with (at most) two states, all such strategies must be played with positive frequency at any sampling equilibrium. That is, there is a rich ecology of mutually reinforcing strategies present, despite the fact that the stage game has a strictly dominant strategy. This reveals and illustrates the starkly different predictions that arise when sampling equilibrium replaces subgame perfection as a solution concept in repeated games.

2 Related Literature

The concept of (payoff) sampling equilibrium used in this paper dates back to Osborne and Rubinstein (1998), who called it $S(k)$ equilibrium. The corresponding disequilibrium dynamics were introduced and used as an equilibrium selection device in Sethi (2000), where it was shown that strictly dominated strategies can be played with positive probability at stable sampling equilibria, while dominant strategy equilibria can be unstable. Mantilla et al. (2019) showed that even efficient dominant strategy equilibria can be unstable when the number of players is at least three. Cárdenas et al. (2015) examined experimental data on common pool resource games, arguing that stable sampling provides a better fit than competing models, including quantal response equilibrium.

Sandholm et al. (2019a) generalized the sampling dynamics to allow for different testing and tie-breaking rules, and showed that at stable rest points of the centipede game, play continues with high probability until the last few nodes. Sufficient conditions for the stability and instability of strict Nash equilibria under these generalized sampling dynamics are derived in Sandholm et al. (2019b). An alternative notion of (action) sampling equilibrium was developed in Osborne and Rubinstein (2003); see Salant and Cherry (2019) for further exploration of this idea.

The notion of a strict equilibrium class of strategies is closely connected to set-valued attractors examined in evolutionary game theory, in particular evolutionarily stable sets (Thomas, 1985; Swinkels, 1992; Balkenborg and Schlag, 2001), sets closed under rational behavior (Basu and Weibull, 1991) and sets closed under better-replies (Ritzberger and Weibull, 1995). Such sets have strong stability properties under evolutionary game dynamics, but (as shown below) they can be unstable under sampling dynamics.

The specific restriction of the strategy space considered in this paper relies on the length of histories on which actions can be made contingent. The use of finite automata to model repeated game strategies with bounded complexity in this sense dates back to Rubinstein (1986) and Abreu and Rubinstein (1988), who considered infinitely repeated games. Evolutionary stability with lexicographic costs of complexity were explored in Binmore and Samuelson (1992), again in the context of infinitely repeated games.

The classical approach to finitely repeated games (and extensive form games more generally) relies on backward induction. Only when the stage game has multiple equilibria can interesting and complex patterns of play arise in the finitely repeated game; otherwise the unique stage game equilibrium is simply played repeatedly (Benoit and Krishna, 1985). As noted above, this gives rise to some counterintuitive and indeed counterfactual predictions in many instances. The reputational approach to resolving these paradoxes has its origins in Kreps and Wilson (1982), Milgrom and Roberts (1982), and Kreps et al. (1982), with applications to the centipede game in McKelvey and Palfrey (1992), and bargaining in Abreu and Gul (2000). It is this reputational literature with which the approach taken here may be contrasted.

3 The Model

3.1 Preliminaries

Consider a symmetric, n-player game G with a finite action set A and payoff function $u : A^n \to \mathbb{R}$. Here $u(a_i, a_{-i})$ denotes the payoff to a player taking action a_i against opponents who choose $a_{-i} \in A^{n-1}$. An action profile $a \in A^n$ is said to be an outcome of G.

Let G_T denote the T-fold repetition of G, with strategy set S and payoff function $\pi : S^n \to \mathbb{R}$. Any strategy profile generates a sequence of outcomes and hence a sequence of stage game payoffs. Let $u_t(s_i, s_{-i})$ denote the payoff in period t to a player adopting strategy s_i against opponents adopting $s_{-i} \in S^{n-1}$. The repeated game payoff $\pi(s_i, s_{-i})$ is obtained by averaging the stage game payoffs across all the outcomes generated by the strategies:

$$\pi(s_i, s_{-i}) = \frac{1}{T} \sum_{t=1}^{T} u_t(s_i, s_{-i}).$$

Since the stage game is finite and finitely repeated, the set of strategies is also finite, though in some cases we may further restrict the set of available strategies to those with limited complexity in a manner discussed below. Let $m = |S|$ denote the (possibly restricted) number of pure strategies that players may use in the repeated game, and let Δ denote the unit simplex:

$$\Delta = \left\{ \sigma \in \mathbb{R}^m \quad \sigma_i \geq 0, \sum \sigma_i = 1 \quad \right. .$$

Any $\sigma \in \Delta$ may be interpreted as a distribution over pure strategies played in a large incumbent population. A sampling procedure involves the testing of various strategies against independent draws from such a population.

3.2 Payoff Sampling

Suppose that the repeated game is played by procedurally rational players in the sense of Osborne and Rubinstein (1998). That is, players choose a strategy from among those that result in the highest realized payoff after sampling each strategy against opponents drawn from an incumbent population.

Let $\sigma \in \Delta$ denote the frequencies with with the various strategies in S are played in the incumbent population. A sampling procedure involves m trials, where each trial involves a play of the repeated game G_T against an independently drawn opponent. These trials result in m realized payoffs for the sampling player, who then chooses a strategy from among those that yield the highest realized payoff. If several strategies are tied for best under this criterion, then an exogenously given tie-breaking rule is invoked.

Let $w_i(\sigma)$ denote the probability that strategy s_i is selected based on the sampling procedure, given that opponent strategies are independently drawn from the distribution σ. A *sampling equilibrium* is defined as a frequency distribution σ^* such that

$$w_i(\sigma^*) = \sigma_i^*$$

for each i. That is, a sampling equilibrium is a probability distribution over strategies that is self-replicating in the sense that the likelihood with which a strategy is selected under the sampling procedure matches that with which it is currently being played.

In repeated games the sampling procedure gives rise to ties with positive probability, so the set of sampling equilibria is sensitive to the manner in which ties are assumed to be broken. The results in this paper hold for any tie-breaking rule that places positive selection probability on each of the tied strategies that result in the highest payoff, including but not limited to uniform tie-breaking. Such a tie-breaking rule is said to be regular, and a *regular sampling equilibrium* is one that is obtained under any regular tie-breaking rule.

3.3 Dynamic Stability

One interpretation of a sampling equilibrium is as the rest point of a dynamic process under which strategies that are selected under sampling with higher probability than they are being played in an incumbent state increase in population frequency. Stability with respect to this dynamic process may then be used as a criteria for selection among sampling equilibria.

Let $\sigma(t)$ denote the frequency distribution over strategies in the incumbent population at time t, and suppose that the population evolves according to

$$\dot{\sigma}_i = w_i(\sigma) - \sigma_i. \tag{1}$$

That is, strategies increase in population frequency if they are selected with greater likelihood than they are being played in the incumbent population. Clearly a rest point of the differential equation system (1) is a sampling equilibrium, although rest points may or may not be stable.

The following stability definitions are standard. A rest point $\sigma^* \in \Delta$ is said to be:

- (Ljapunov) *stable* if, for every neighborhood U of σ^* there exists a neighborhood V such that $\sigma(t_0) \in V \Rightarrow \sigma(t) \in U$ for all t,
- *unstable* if it is not stable,
- *asymptotically stable* if it is stable and there exists a neighborhood U of σ^*such that if $\sigma(t_0) \in U$, then $\lim_{t\to\infty} \sigma(t) = \sigma^*$.
- *globally asymptotically stable* if $\lim_{t\to\infty} \sigma(t) = \sigma^*$ for all $\sigma(t_0) \in \Delta$, and
- *almost globally asymptotically stable* if $\lim_{t\to\infty} \sigma(t) = \sigma^*$ for all $\sigma(t_0) \in \text{int}(\Delta)$

The main solution concept used in the analysis to follow is stable sampling equilibrium. In repeated games, outcome-based partitions prove to be a useful device for this analysis.

3.4 Outcome-Based Partitions

Since G is symmetric, if all players adopt the same strategy they will generate a sequence of symmetric outcomes, and obtain the same payoff. This may also happen when players use different strategies. For instance, in the repeated prisoner's dilemma, any pair of strategies that are never the first to defect will generate the same sequence of outcomes when matched with each other than they do when matched with themselves: in this case a fully cooperative sequence.

The set of strategies S can be partitioned into equivalence classes based on this criterion. Specifically, two strategies s_i and s_j are placed in the same class if and only if they generate the same sequence of outcomes when matched against each other as they do when matched against themselves. This construction of equivalence classes is an *outcome-based partition* of S.

A unique outcome-based partition exists in any symmetric game. This can be seen as follows. Consider any set of strategies that generate the same symmetric sequence of outcomes $\{a^1, a^2, ..., a^T\}$ when self-matched, and consider the sequence of outcomes generated when each player uses a strategy from within this set. Clearly the first period outcome is precisely a^1. The second period outcome must therefore be a^2 since each player faces the same history that she would have if she were self-matched. Proceeding inductively, we see that all outcomes must be the same as they would be if all strategies were self-matched. And since any pair of strategies that generate identical outcomes when self-matched are placed in the same class, the outcome-based partition is unique.

When strategies belonging to the same equivalence class are matched with each other, the sequence of outcomes is indistinguishable from that which would arise if each strategy were self-matched. As a result, it is impossible for any player to deduce from the sequence of outcomes which strategy from within this set the opponents happen to be using. It is possible, of course, that each class consists of just a single strategy, though in repeated games it will generally be true that the set of classes is much smaller than the set of strategies. Some examples are provided in the sections to follow.

Suppose that the unique outcome-based partition of S results in K equivalence classes, denoted $S_1, ..., S_K$. Associated with each set S_k of strategies is a unique payoff π_k that is obtained by a member of class k when all opponents also use strategies in S_k. We call π_k the *within-class payoff* for strategies in S_k, and say that a class of strategies S_k is *better than* class S_l if $\pi_k > \pi_l$.

There may, of course, be classes that give rise to the same payoff in within-class matches but without giving rise to the same sequence of outcomes. For instance, in the repeated prisoners' dilemma with an even number of periods, strategies that alternate between cooperation and defection in within-class matches give rise to the same within-class payoff, but can belong to different classes depending on whether they cooperate or defect in the initial period. These strategies cannot be consolidated into a single class.

The concepts of sampling equilibrium, dynamic stability, and outcome-based partitions may be illustrated using a few examples before proceeding to more general results.

4 Examples

Consider the one-shot prisoner's dilemma, with payoffs (to the row player) given by:

	C	D
C	b	0
D	$b+1$	1

(2)

where $b > 1$. Here the two players each have a unit endowment that has value 1 to themselves, but value b to the other player. Cooperation C involves giving the endowment to one's opponent, defection D involves keeping it. The parameter b may then be interpreted as a benefit-cost ratio.

In the static game there are just two strategies. Let σ_1 denote the proportion of the incumbent population that plays cooperate. Then cooperation is selected under sampling if and only if the opponent cooperates when cooperation is sampled, and defects when defection is sampled. The probability of this event is

$$w_1(\sigma) = \sigma_1(1 - \sigma_1) \leq \sigma_1,$$

with strict inequality whenever $\sigma_1 > 0$. Hence, under the sampling dynamics (1),

$$\dot{\sigma}_1 = w_1(\sigma) - \sigma_1 < 0$$

whenever $\sigma_1 > 0$. The only sampling equilibrium involves defection with probability one, corresponding to the dominant strategy equilibrium of this game, and this sampling equilibrium is globally asymptotically stable.

More generally, any strict symmetric Nash equilibrium must also be a sampling equilibrium: when the entire population plays the corresponding strategy this is the only one that can possibly be selected under the sampling procedure. But such equilibria need not be stable, even if they involve strictly dominant strategies.

To illustrate, consider the two-player public goods game with payoffs given by

	H	M	L
H	6	3	0
M	7	4	1
L	8	5	2

The dominant strategy equilibrium in which L is played with probability one is clearly also a sampling equilibrium. However, as noted by Osborne and Rubinstein (1998), there is a second sampling equilibrium at $\sigma^* \approx (0.20, 0.28, 0.52)$. And as shown in Sethi (2000), the former equilibrium is unstable, while the latter is almost globally asymptotically stable. That is, the concepts of stable sampling equilibrium and Nash equilibrium give sharply different predictions in this game.

Now consider the repeated prisoners' dilemma with stage game payoffs as in (2), and suppose that $T = 2$. In this case there are eight possible strategies and the outcome-based partition consists of four classes:

Strategy	Behavior	Class	Within-class Payoff
s_1	C in both periods	S_1	b
s_2	C initially, C if opponent cooperates		
s_3	C initially, then D	S_2	$(b+1)/2$
s_4	C initially, D if opponent cooperates		
s_5	D initially, then C	S_3	$(b+1)/2$
s_6	D initially, C if opponent defects		
s_7	D initially, D if opponent defects	S_4	1
s_8	D in both periods		

Note that although classes S_2 and S_3 generate the same payoffs in within-class matches, they give rise to different outcome sequences in within-class matches, and hence cannot be consolidated. For example, if s_4 is matched with s_6, the result is cooperation by the former and defection by the latter in all periods, so the former gets payoff zero and the latter $b + 1$. These payoffs are quite different from those that arise when these strategies are self-matched, and the two strategies must therefore be placed in different classes.

Any regular sampling equilibrium in this game must involve some cooperation. To see this, consider any distribution σ^* such that $\sigma^*_i > 0$ only if $s_i \in S_4$. Then both s_7 and s_8 are matched within-class with probability 1, and both get payoff 1 when sampled. Under regular tie breaking, both strategies much be selected with positive probability if σ^* is a sampling equilibrium. In particular, we must have $\sigma^*_7 > 0$.

In this case, the following event has positive probability: when s_3 is sampled, the opponent plays s_7, and when all other strategies are sampled, the opponent plays s_8. Conditional on this event, the payoff obtained when s_3 is sampled is

$$\pi(s_3, s_7) = (b+1)/2 > 1,$$

while the payoffs obtained when all other strategies are sampled does not exceed 1. Hence s_3 is selected with positive probability under the sampling procedure at σ^*. That is, we have $w_3(\sigma^*) > 0$. Since $\sigma^*_3 = 0$ by hypothesis, σ^* cannot be a sampling equilibrium.

This is a special case of a more general phenomenon in repeated games.

5 Results

5.1 Best Present Classes

We say that a strategy s_i is *present* at any sampling equilibrium σ^* if $\sigma_i^* > 0$. A class S_k is present if there exists some strategy $s_i \in S_k$ that is present. A class is *fully present* if all strategies in the class are present. And a class S_k is the *best present class* if no better class is present.

The following result applies to any symmetric repeated game G_T, regardless of the number of repetitions.

Proposition 1. *At any regular sampling equilibrium, the best present class is fully present.*

The proof of this result relies on the observation that at any sampling equilibrium, all present strategies can be matched within-class with positive probability when sampled. In addition, all strategies that are not present, but belong to a class that is present, can also be matched within-class with positive probability when sampled. If this event occurs, then all strategies in the best present class will be tied for best, unless some other strategy that is not present is best. Since only present strategies can be tied for best in a sampling equilibrium, this means that the best present class must be fully present.

Proposition 1 can be used to narrow down the set of possibilities that can arise in a sampling equilibrium. We say that a symmetric action profile $a' \in A^n$ is dominated if there exists another symmetric action profile $a \in A^n$ such that

$$u(a) > u(a').$$

That is, all players obtain a higher payoff in the stage game if they all play the action associated with a than if they all play the action associated with a'. This is simply the usual notion of payoff-dominance, applied to symmetric action profiles. We say that a class of strategies in the repeated game G_T is dominated if the strategies in this class involve the play of dominated action profiles in each period whenever these strategies are matched within-class. A class is undominated if it is not dominated. The following result rules out sampling equilibria in which only dominated classes are present, as long as the number of repetitions is not too small.

Proposition 2. *At any regular sampling equilibrium, the best present class is undominated if T is sufficiently large.*

The proof of this result is based on the following reasoning. Suppose that the best present class is S_k is dominated. Then there exists an action $a_j \in A$, and a symmetric action profile $a = (a_j, ..., a_j)$, such that a dominates all outcomes that arise when a strategy in S_k is matched within-class. Let s_j denote the strategy that plays a_j unconditionally in all periods. Note that S_k contains a strategy s_k that switches permanently to action a_j whenever any opponent plays a_j in the initial period. Since S_k is fully present from Proposition 1, this strategy must be present. Now consider what happens when s_j is sampled. There is a positive probability that all opponents will adopt s_k, since this strategy is present. If this happens, then s_j will face the outcome a in all periods except the first. If T is large enough, the resulting payoff to s_j will exceed π_k. If all present strategies are matched within-class, then s_j will obtain a higher payoff than all present strategies, so must also be present. But then S_k cannot be the best present class, since s_j secures a higher payoff than π_k when matched within class.

Proposition 2 implies that in repeated symmetric public goods games a sampling equilibrium cannot involve the choice of an inefficient contribution level in each period when the number of repetitions is large. This is true even if the stage game has a strictly dominant strategy, so the repeated game has a unique subgame perfect equilibrium. The example of the iterated prisoners' dilemma is an instance of this phenomenon.

In addition, Proposition 2 implies that in any common interest game, the efficient action profile in the stage game (preferred by all players to all other action profiles) must be played with positive probability in any regular sampling equilibrium if the number of periods is sufficiently large. Much more than this can be said of common interest games, however, as shown below.

5.2 Strict Equilibrium Classes

In one-shot games, any symmetric strict Nash equilibrium must also be a sampling equilibrium: if the corresponding action is played by all individuals in the population, then no other action can ever be selected under sampling. However, repeated games cannot have strict Nash equilibria since some histories are not realized along the equilibrium path, and strategies that differ only on such histories will secure the same payoffs as equilibrium strategies.

The notion of strict equilibrium can be extended to apply to sets of strategies in such a manner as to recover the connection between strict equilibria and sampling equilibrium. A *strict equilibrium class* of strategies is an element S_k of the outcome based partition that has the following property: for any $s_i \notin S_k$ and any $s_{-i} \in S_k^{n-1}$, we have

$$\pi(s_i, s_{-i}) < \pi_k.$$

That is, if all other players choose strategies in S_k, then there is no best reply for a player that lies outside S_k.

We say that a sampling equilibrium is *single-class* if only one class is present at the equilibrium. By Proposition 1, this class must be fully present as long as the equilibrium is regular. The following result connects strict equilibrium classes to sampling equilibria:

Proposition 3. *A single-class regular sampling equilibrium has support S_k if and only if S_k is a strict equilibrium class.*

In fact, if S_k is a strict equilibrium class, then there exists a sampling equilibrium in which only strategies in this class are played, even if the tie-breaking rule is not regular. Regular tie-breaking ensures that all strategies in the class are present. And if S_k is not a strict equilibrium class, then there must exist a strategy outside S_k that has a positive probability of being best or tied for best when only strategies in S_k are present, so a single-class regular sampling equilibrium with support S_k is impossible, and Proposition 3 follows.

Proposition 3 is useful in identifying sampling equilibria of repeated coordination games, such as the following, with payoffs (to the row player) given by

	H	L
H	2	x
L	0	1

(3)

where $x < 1$. By Proposition 2, the action H must be played with positive probability in any sampling equilibrium of the T-fold repetition of this game, provided that the number of periods is sufficiently large. This is true for any $x < 1$, and hence true regardless of whether the payoff-dominant equilibrium is also risk-dominant. But Proposition 3 tells us much more: any strict equilibrium class of the repeated game must be the support of a sampling equilibrium. For instance, the class that results in outcomes $L\overline{H}$ (initially L followed by H for all remaining periods) in within-class matches is a strict equilibrium class. And other sampling equilibria are possible if strategies are restricted to those depending only on a one-period history. In this case, as shown in Section 6, there are sampling equilibria involving alternation between the two stage-game pure strategy equilibria. And if $x < 0$, even the highly inefficient class that involves outcomes $H\overline{L}$ in within-class matches is the support of a sampling equilibrium. However, it turns out that the stability refinement proves is very powerful in eliminating most of these sampling equilibria.

5.3 Common Interest Games

A symmetric common interest stage game is one in which there exists a symmetric action profile a that is preferred to all other action profiles by all players. That is, there exists $a \in A^n$ such that $u(a) > u(a')$ for all action profiles $a' \neq a$. The coordination game with payoffs (3) is an example.

If G is a common interest game, then G_T is itself a common interest game, in the sense that there exists a class of strategies in the outcome-based partition that is preferred by all players to any strategy profile outside the class. That is, there exists a class S_k such that $\pi(s) < \pi_k$ for all strategy profiles $s \notin S_k^n$. In this case S_k is a strict equilibrium class, and is the support of a regular sampling equilibrium. However, there may exist other strict equilibrium classes and sampling equilibria as we have seen.

In common interest games the efficient sampling equilibrium has strong stability properties, and in two player games the only stable sampling equilibrium involve the repeated play of the efficient action profile:

Proposition 4. *If G is a common interest game, then there exists an efficient regular sampling equilibrium of G_T that is locally asymptotically stable under the dynamics (1); if G is a two-player game then this equilibrium is almost globally asymptotically stable.*

The following example illustrates both why the number of players matters, and the crucial role played by repetition. Let G be a three player common interest game with payoffs given by

H

	H	L
H	$(7,7,7)$	$(0,6,0)$
L	$(6,0,0)$	$(5,5,0)$

L

	H	L
H	$(0,0,6)$	$(0,5,5)$
L	$(5,0,5)$	$(4,4,4)$

where player one chooses rows, player 2 columns, and player 3 matrices. In the absence of repetition there are just two strategies to be sampled. Note that H will be selected under the sampling procedure if and only if both opponents also choose H when H is sampled (otherwise the realized payoff will be zero when H is sampled, and at least 4 when L is sampled). If the population share of those choosing H is $\eta > 0$, the likelihood that H is selected under sampling will be $\eta^2 < \eta$. This share will decline under the sampling dynamics from all initial conditions, so the sampling equilibrium in which only L is chosen is globally asymptotically stable.

Now consider the repeated game with $T = 2$, and suppose that players can only condition second period actions on the number of opponents choosing H or L in the first period, and not their specific identities. According to Proposition 4, there is an efficient sampling equilibrium that is locally asymptotically stable. Note that the class of strategies that are never the first to play L has three elements: the strategy that plays H unconditionally, the one that switches to L if one opponent initially plays L, and the strategy that switches if both opponents play L. A strategy in this class will be selected under sampling if and only if at least one of these three strategies is matched within-class when sampled. The probability of this event $\mathcal{E}$ is

$$\Pr(\mathcal{E}) = 1 - (1 - \eta^2)^3.$$

It is easily verified that this exceeds η when η is close to 1. That is, there is a neighborhood of the sampling equilibrium in which only strategies in the efficient class are played, such that the aggregate frequency of these strategies is monotonically increasing. This efficient sampling equilibrium is therefore locally asymptotically stable.

In this example it is also true that there is an inefficient sampling equilibrium in which only L is chosen, and this too is locally asymptotically stable. According to Proposition 4, this cannot happen in two player games. Efficient equilibria of two-player common interest games have stronger stability properties because one does not need to simultaneously encounter multiple within-class opponents when sampling an efficient strategy.

To this point the set of available strategies has not been restricted in any way, but these sets can be large even with just a few repetitions, and it is worth considering natural truncations of these sets if sampling is to be a practicable procedure. Doing so also leads to some new insights and results.

6 Bounded Complexity

Since we are considering finite repetitions, the strategies in S can be represented by finite automata. That is, any strategy in the repeated game may be represented by a tuple

$\{Q, q_0, \lambda, \mu\}$, where Q is a finite set of states, q_0 is the initial state, $\lambda : Q \to A$ is an output function that determines which action is taken in state $q \in Q$, and $\mu : Q \times A \to Q$ is a transition function that determines the successor state, given any incumbent state and opponent action.

One approach to modeling bounded complexity is to restrict the length of the history on which a current action can be made contingent. This is equivalent to restricting the set of strategies S to contain only those that can be represented using automata with a limited number of states. When the stage game involves two actions and two players, a natural restriction is to consider automata with at most two states (thus allowing for conditioning on a one period history). This allows for a more detailed analysis of stable sampling equilibria in repeated games involving cooperation and coordination.

6.1 Cooperation

Consider strategies for the repeated prisoners' dilemma with payoffs (2) that can be implemented by (at most) two state finite automata. Any such strategy may be represented by a binary number with five bits, where the first identifies the initial state, the second and third respectively identify the successor states when the automaton is in the cooperative state and the opponent cooperates or defects, and the fourth and fifth bits respectively identify the successor states when the automaton is in the defecting state and the opponent cooperates or defects.

Letting 0 stand for cooperation and 1 for defection, the number 00101 corresponds to the tit-for-tat strategy, which cooperates initially, and regardless of its current state, transitions to the defect state if and only if the opponent defects, and to the cooperate state if and only if the opponent cooperates. Similarly, for any values of x and y, the strategy $000xy$ cooperates always, never leaving the initial cooperate state.

There are exactly 26 behaviorally distinct strategies of this kind; see Binmore and Samuelson (1992) for a complete enumeration. Of the 32 possible five-bit numbers, four correspond to a strategy that always cooperates, and four to a strategy that always defects, so six of the 32 possible five-bit numbers may be dropped without loss of generality. These can be partitioned into six equivalence classes depending on the payoffs they obtain in within-class matches. Let $S_1, ..., S_6$ denote these equivalence classes, in decreasing order of self-matched payoff. The following table enumerates the six equivalence classes, showing the binary representations, as well as the outcomes and payoffs that arise in within-class matches (assuming an even number of repetitions).

Class	Representation	In-Class Outcomes	In-Class Payoff
S_1	$00xyz$	$\overline{C}$	b
S_2	$10xy0$	$D\overline{C}$	$(1+(T-1)b)/T$
S_3	$01xy0$	$\overline{CD}$	$(b+1)/2$
S_4	$11xy0$	$\overline{DC}$	$(b+1)/2$
S_5	$01xy1$	$C\overline{D}$	$(b+(T-1))/T$
S_6	$1xyz1$	$\overline{D}$	1

There are four strategies representable by two-state automata in each class, as well as the two one-state strategies that unconditionally cooperate and unconditionally defect, for a total of 26. Within-class payoffs range from b to 1.

As we know from Proposition 2, any regular sampling equilibrium in the prisoners' dilemma involves some cooperation: the best present class cannot be the one that defects in all periods when matched within class. This still leaves open a broad range of possibilities. A sharper characterization can be obtained if we let the strategy set S contain only those strategies that can be represented by finite automata with at most two states, which we call the repeated prisoners' dilemma with bounded complexity. In this game, provided that the number of repetitions exceeds the benefit-cost ratio, all permitted strategies must be played with positive frequency in any sampling equilibrium:

Proposition 5. *If $T > b$ and σ^* is a regular sampling equilibrium of the repeated prisoners' dilemma with bounded complexity, then $\sigma_i^* > 0$ for all $s_i \in S$.*

Figure 1 shows the distribution of strategies at a computed stable sampling equilibrium, ordered on the horizontal axis based on their binary representations. While all strategies are played with positive frequency at any sampling equilibrium of this game, there are significant differences in the frequencies with which they are played, both across and within equivalence classes. The following table enumerates the eight strategies played with highest frequency at the stable sampling equilibrium shown in the figure:

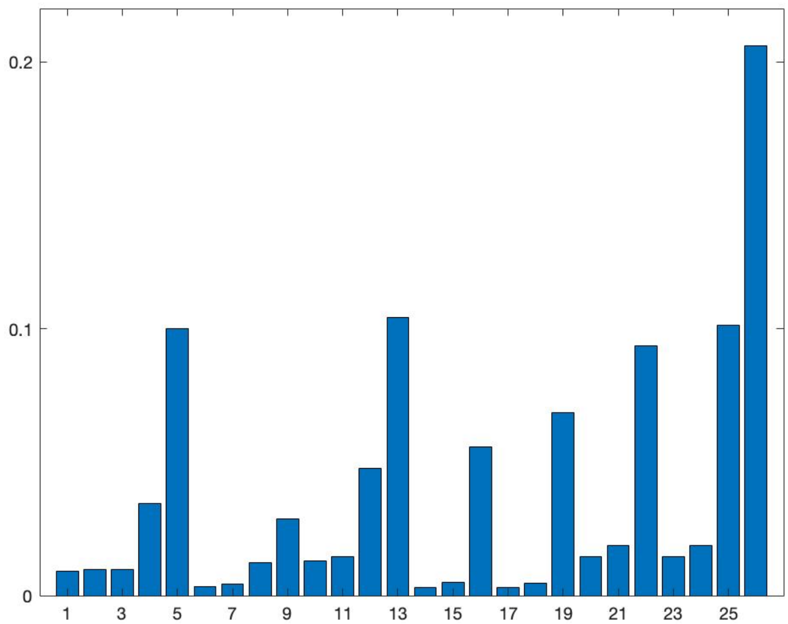

Figure 1: A stable sampling equilibrium in the repeated prisoner's dilemma, with strategies numbered in order of binary representation.

Frequency	Representation	Class	In-class Outcomes	In-class Payoffs
0.206	11111	S_6	$\overline{D}$	1
0.104	01111	S_5	$C\overline{D}$	$(b + (T-1))/T$
0.101	11110	S_4	$\overline{DC}$	$(b+1)/2$
0.100	00111	S_1	$\overline{C}$	b
0.094	11010	S_4	$\overline{DC}$	$(b+1)/2$
0.069	10110	S_2	$D\overline{C}$	$(1+(T-1)b)/T$
0.056	10010	S_2	$D\overline{C}$	$(1+(T-1)b)/T$
0.048	01110	S_3	$\overline{CD}$	$(b+1)/2$

Members of all classes are present, with the strategy of unconditional defection having the highest frequency, being used about one-fifth of the time. The second most frequent strategy is very similar, and switches to unconditional defection after a single period of cooperation, regardless of the initial opponent action. This is followed by a strategy that starts with defection and then alternates between the two actions in within-class matches, but stays in the defecting state if the opponent cooperates unconditionally. This strategy can exploit unconditional cooperation without sacrificing too much when encountering itself. Next is the grim trigger, which does extremely well in within-class matches. And so on.

The structure of stable sampling equilibrium in this game illustrates the sharp differences that arise in comparison with standard approaches. The key criterion for survival here is the ability to do better than other strategies when sampled, and this depends in complex ways on the frequencies with which these other strategies are encountered. Some strategies survive by doing well in within-class matches, while others thrive by doing well when matched with strategies in other classes. The result is an ecology of strategies that sustains itself and maintains considerable behavioral heterogeneity. This is not what one finds in either the standard complete information analysis or in the reputational approach.

6.2 Coordination

If the set of repeated game strategies is restricted to contain only those contingent on a one-period history, then the total number of strategies S depends only on the number of players and actions in the stage game, and not on the payoffs. Hence there are 26 behaviorally distinct strategies and six equivalence classes in the outcome-based partition in any 2×2 game, including the repeated coordination game with payoffs as in (3). Within-class payoffs range from 2 to 1, as shown in the following table:

Class	Within-Class Outcomes	Within-Class-Payoff
S_1	$\overline{H}$	2
S_2	$L\overline{H}$	$(1+2(T-1))/T$
S_3	$\overline{HL}$	$3/2$
S_4	$\overline{LH}$	$3/2$
S_5	$H\overline{L}$	$(T+1)/T$
S_6	$\overline{L}$	1

There are five strategies that are never the first to play L, and these are in class S_1. Class S_2 is composed of four strategies that initially choose L, and then choose H for all other periods when matched within-class.

We have already seen that there are at least two single-class sampling equilibria in this game, corresponding respectively to the strict equilibrium classes S_1 and S_2. We also know from Proposition 2 that there can be no sampling equilibrium in which only the dominated action is played, which rules out a single-class equilibrium involving S_6 when T is sufficiently large. It turns out that the repeated coordination game has at least four and possibly five single-class equilibria:

Proposition 6. *The repeated coordination game with bounded complexity has four single-class sampling equilibria, corresponding respectively to the classes $S_1, ..., S_4$. In addition, if $x < 0$, there is also a sampling equilibrium in which S_5 is the only present class.*

Despite the multiplicity of sampling equilibria in this repeated coordination game, we know from Proposition 4 that all the inefficient equilibria are dynamically unstable. This illustrates the power of the stability refinement in this game, and more generally in repeated common interest games with many strict equilibrium classes.

7 Discussion

The concept of (payoff) sampling equilibrium, augmented with a dynamic stability refinement, provides a sharp contrast with Nash equilibrium and other solution concepts. As has been shown in previous work, strictly dominant strategies can be played with positive probability at unique stable sampling equilibria of one-shot games, and stable sampling equilibria in extensive form games can be far less counterintuitive and paradoxical than subgame perfect equilibria.

This paper has considered symmetric finitely repeated games, for which quite general results can be obtained using the device of outcome-based partitions. Stable sampling equilibrium can give rise to significant levels of cooperation in such environments, although considerable heterogeneity in strategies persists in the iterated prisoners' dilemma. In repeated coordination games one finds both behavioral homogeneity and a tendency to efficiency.

A natural direction for future research would be to consider asymmetric games, with strategies defined in a manner that is contingent on player position. That is, a strategy would encode multiple contingent plans of actions depending on the player position to which an individual is assigned. In this case it is meaningful to consider outcomes that arise when strategies are self-matched, and one could proceed to explore the consequences of outcome-based partitions. Contrasting the results with more standard equilibrium approaches, especially in light of experimental and empirical regularities, would seem to be a fruitful exercise.

Appendix

Proof of Proposition 1. Let σ^* denote a regular sampling equilibrium at which S_k is the best present class. Let P denote the set of strategies present at σ^*:

$$P = \{s_i \,|\, \sigma_i^* > 0\}.$$

Suppose, by way of contradiction, that S_k is not fully present. Then there exist $s_i \in S_k$ and $s_j \in S_k$ such that $s_i \notin P$ and $s_j \in P$. Suppose that each strategy in $s_i \cup P$ is matched within-class when sampled. This event has positive probability, since s_i can be matched with s_j and all strategies in P can be self-matched. Since S_k is the best present class, the resulting payoff to s_i will be at least as great as that to any strategy in P conditional on this event. Hence, either s_i will be tied for the best realized payoff among sampled strategies, or there will be some other strategy $s_l \notin P$ that is (at least) tied for the best realized payoff among sampled strategies. Under regular tie-breaking, either $w_i(\sigma^*) > 0$ or $w_l(\sigma^*) > 0$. But neither s_i nor s_l are in P, so $\sigma_i^* = \sigma_l^* = 0$. Hence either $w_i(\sigma^*) \neq \sigma_i^*$ or $w_l(\sigma^*) \neq \sigma_l^*$, so σ^* cannot be a regular sampling equilibrium. □

Proof of Proposition 2. Let σ^* denote a regular sampling equilibrium at which S_k is the best present class, and suppose that S_k is dominated. Let $\{a^t\}_{t=1}^T$ denote the (symmetric) sequence of outcomes generated by strategies in S_k when matched within-class. Since S_k is dominated, there exists an action $a_i \in A$ such that

$$u(a_i, ..., a_i) > u(a^t) \tag{4}$$

for all t. There exists as strategy s_j in S_k that generates the outcomes $\{a^t\}$ when matched within-class, but chooses action a_i in all periods after the first when matched with a strategy that plays a_i unconditionally. Since S_k is fully present from Proposition 1, this strategy is present. Now consider the following positive probability event: all present strategies are matched within-class when sampled, and the strategy that plays a_i unconditionally is matched with s_j when sampled. If T is sufficiently large, then from (4), the strategy that plays a_i unconditionally will obtain a higher payoff under this event that any present strategy. Hence it must also be present. But this strategy belongs to a better class than S_k, which contradicts that hypothesis that S_k is the best present class at σ^*. □

Proof of Proposition 3. Suppose S_k is a strict equilibrium class. Then $\pi(s_i, s_{-i}) < \pi_k$ for all $s_i \notin S_k$ and $s_{-i} \in S_k^{n-1}$. At any population state σ whose support is a subset of S_k, only strategies in S_k can be selected under sampling. Furthermore, all strategies in S_k get the same payoff against σ, so all must be selected with positive probability under regular tie-breaking. Hence there exists a single-class sampling equilibrium with support S_k.

To prove the converse, suppose that there exists a single-class sampling equilibrium σ^* with support S_k. If S_k is not a strict equilibrium class, then there exists $s_i \notin S_k$ and $s_{-i} \in S_k^{n-1}$ such that $\pi(s_i, s_{-i}) \geq \pi_k$. Consider the following positive probability event: when s_i is sampled the opponents play s_{-i}, resulting in a payoff of at least π_k. Not that since all present strategies belong to S_k, they must be matched within-class when sampled, and thus obtain payoff π_k. Hence s_i has a positive probability of selection, contradicting the hypothesis that it is played with zero probability at the sampling equilibrium σ^*. □

Proof of Proposition 4. Let σ^* denote the efficient regular sampling equilibrium of a repeated common interest game G_T, S_1 the highest payoff class that is its support, and $m_1 = |S_1|$ the number of strategies present at this equilibrium. Consider any state $\sigma \in \text{int}(\Delta)$. Let

$$\eta = \sum_{s_i \in S_1} \sigma_i$$

denote the frequency with with strategies in S_1 are played at σ. Since $\sigma \in \text{int}(\Delta)$, we have $\eta > 0$. Any given strategy $s_i \in S_1$ is matched within class with probability η^{n-1}, and matched outside class with probability $1 - \eta^{n-1}$. The probability that all m_1 strategies in S_1 are matched outside class is therefore $(1 - \eta^{n-1})^{m_1}$. Let $\mathcal{E}$ denote the following event: when all strategies are sampled against σ, at least one strategy $s_i \in S_1$ is matched within-class. The likelihood of this event is

$$\Pr(\mathcal{E}) = 1 - (1 - \eta^{n-1})^{m_1}. \tag{5}$$

First consider the case $n = 2$. Since $m_1 > 2$, we have

$$\Pr(\mathcal{E} \,|n = 2) = 1 - (1 - \eta)^{m_1} > 1 - (1 - \eta)^2 = 2\eta - \eta^2 > \eta$$

for all $\eta \in (0, 1)$. If $\mathcal{E}$ occurs, then at least one strategy in S_1 yields the highest attainable payoff, since G is a common interest game. In this case no strategy $s_j \notin S_1$ can be selected under sampling. Hence

$$\sum_{s_i \in S_1} w_i(\sigma) > \eta = \sum_{s_i \in S_1} \sigma_i.$$

Under the sampling dynamics (1), we therefore have

$$\sum_{s_i \in S_1} \dot{\sigma}_i = \sum_{s_i \in S_1} (w_i(\sigma) - \sigma_i) > 0.$$

The above reasoning holds not only for $\sigma \in \text{int}(\Delta)$, but for all σ such that $\eta > 0$. That is, along any trajectory that is initially at $\sigma(0) \in \text{int}(\Delta)$, we have η increasing for all

$\sigma(t)$, regardless of whether such trajectories remain in the interior of Δ, as long as $\eta < 1$ continues to hold. Hence $\lim_{t\to\infty} \eta = 1$ and G_T has an almost globally stable sampling equilibrium in which all present strategies are members of S_1.

Now consider the case of $n \geq 3$, and note that $m_1 > n$ (the number of strategies in this class must exceed the number of players). Using (5), define

$$\phi(\eta) = \Pr(\mathcal{E}) - \eta = 1 - (1 - \eta^{n-1})^{m_1} - \eta$$

so that

$$\phi'(\eta) = m_1(n-1)(1-\eta^{n-1})^{m_1-1}\eta^{n-2} - 1.$$

Note that $\phi(1) = 0$ and $\phi'(1) = -1$. Since ϕ is continuous, we therefore have $\phi(\eta) > 0$ for η sufficiently close to 1. As in the two player case, if $\mathcal{E}$ occurs, then at least one strategy in S_1 yields the highest attainable payoff and no strategy $s_j \notin S_1$ can be selected under sampling. That is, there exists a neighborhood of the efficient sampling equilibrium within which

$$\sum_{s_i \in S_1} w_i(\sigma) > \eta = \sum_{s_i \in S_1} \sigma_i$$

whenever $\eta \neq 1$. Hence

$$\sum_{s_i \in S_1} \dot{\sigma}_i = \sum_{s_i \in S_1} (w_i(\sigma) - \sigma_i) > 0$$

in this neighborhood, and the efficient sampling equilibrium is locally asymptotically stable under the dynamics (1). □

Proof of Proposition 5. Let σ^* denote a regular sampling equilibrium of the repeated prisoners' dilemma with bounded complexity, and suppose that $T > b$. The proof is completed in four steps, with each claim following from the previous ones. Specifically, at σ^*, (1) unconditional defection is present, (2) the class S_6 is fully present, (3) unconditional cooperation is present, and (4) all strategies are present.

1. Unconditional defection is present at σ^*. To see why, let S_k denote the best present class at σ^*. If this class is S_6, then S_6 must be fully present from Proposition 1, and since unconditional defection is a member of S_6, it must also be present. Now suppose that S_6 is not the best present class at σ^*, so $k \leq 5$. Note that each class S_k with $k \leq 5$ contains a strategy that cooperates in all periods, except possibly the first, when matched with unconditional defection. In class S_1 such a strategy is unconditional cooperation. In classes S_2–S_5 respectively, such strategies are shown in Figure 2. In the figure, the node on the left in each case is the initial state, and the

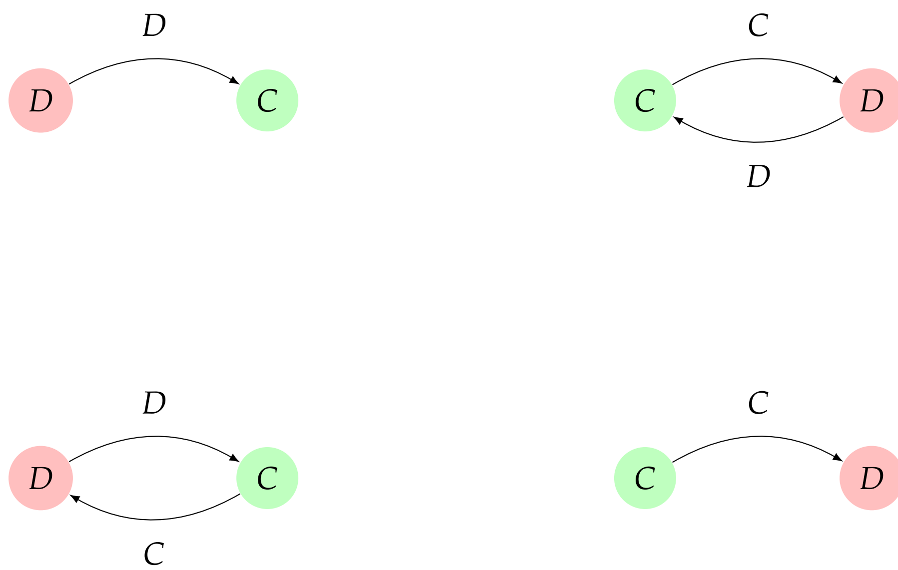

Figure 2: Strategies that cooperate for all or almost all periods when matched with unconditional defection, in classes S_2–S_5 respectively.

arrows and associated actions describe the conditions for transitions between states. Regardless of which class is the best present, therefore, there must exist at least one strategy that is present and that generates a payoff to unconditional defection that is at least

$$\frac{1+(T-1)(b+1)}{T} > b$$

given $T > b$. If all present strategies are matched within-class when sampled, the highest attainable payoff is b. If unconditional defection is matched with the most favorable present strategy, then it obtains a payoff greater than b, and hence greater than that of any present strategy, unless it is itself present. Hence unconditional defection must be present at σ^*.

2. The class S_6 is fully present at σ^*. To see why, suppose that $s_i \in S_6$ is not present. Since unconditional defection is present, there is a positive probability that all present strategies are matched with unconditional defection when sampled. This results in payoff at most 1, which is also the payoff obtained by s_i when sampled against unconditional defection. So s_i has a positive probability of being at least tied for best when sampled, and must therefore be present.

3. Unconditional cooperation is present at σ^*. Let s_j denote the strategy of unconditional cooperation. Since S_6 is fully present, there exists a strategy that cooperates in almost all periods when matched with s_j; see Figure 3 for an illustration. If this strategy is encountered when s_j is sampled, the payoff to s_j is $(T-1)b/T$, which

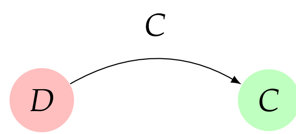

Figure 3: A strategy in S_6 that cooperates in almost all periods when matched with unconditional cooperation.

exceeds 1 when $T \geq 2$. All present strategies get at most 1 when sampled against unconditional defection, which also happens with positive probability since S_6 is fully present. Hence there is a positive probability that s_j is best under sampling, and must therefore be present at σ^*.

4. All strategies are present at σ^*. Suppose some strategy s_l is not present, but is matched with unconditional cooperation when sampled, getting at least b, while all others are matched with unconditional defection when sampled, getting at most 1. In this case s_l must be selected with positive probability at σ^*, contradicting the hypothesis that it is not present. □

Proof of Proposition 6. Given Proposition 3, to show that a single-class equilibrium with support S_k is present, we need only show that S_k is a strict equilibrium class.

First consider class S_1 with within-class outcomes $\overline{H}$ and payoff $\pi_1 = 2$. This is clearly a strict equilibrium class, so there exists a sampling equilibrium with support S_1.

Next consider S_2, with within class outcomes $L\overline{H}$ and payoff $\pi_2 = (1 + 2(T-1))/T$. For any strategy $s_i \notin S_2$ and $s_j \in S_2$ there must be some period t in which the outcome deviates from the symmetric play of $L\overline{H}$ for the first time. In this period the outcome is asymmetric, so $u_t(s_i, s_j) < 1 \leq u_t(s_j, s_j)$. In all periods after t the payoff to s_i is at most 2, while in periods before t the payoff is equal to that which s_j obtains when matched within class. Hence $\pi(s_i, s_j) < \pi_2$. Since this holds for all $s_i \notin S_2$, S_2 is a strict equilibrium class, and there exists a sampling equilibrium with support S_2.

Next consider S_3, with within-class outcomes $\overline{HL}$. Suppose that a strategy $s_i \notin S_3$ is matched with some strategy $s_j \in S_3$. Note that since $s_j \in S_3$, this strategy will choose action L following outcome (H, H). Hence the outcome (H, H) can only arise in at most half the periods (if T is even) or at most $(T+1)/2$ periods (if T is odd). In remaining periods the highest payoff is obtained by coordination on L. This is precisely the sequence of outcomes that arises when s_j is matched within-class. Since $s_i \notin S_3$, the resulting outcomes must result in a strictly lower payoff to s_i than is obtained when s_j is matched within-class. That is, S_3 is a strict equilibrium class, and there exists a sampling equilibrium with support S_3.

The case of S_4, with within-class outcomes $\overline{LH}$, is virtually identical. Suppose that a strategy $s_i \notin S_4$ is matched with some strategy $s_j \in S_4$. Since $s_j \in S_4$, this strategy will choose action L following outcome (H, H). Hence the outcome (H, H) can only arise in at most half the periods (if T is even) or at most $(T-1)/2$ periods (if T is odd). In remaining periods the highest payoff is obtained by coordination on L. This is precisely the sequence of outcomes that arises when s_j is matched within-class. Since $s_i \notin S_4$, the resulting outcomes must result in a strictly lower payoff to s_i than is obtained when s_j is matched within-class. That is, S_4 is a strict equilibrium class, and there exists a sampling equilibrium with support S_4.

Finally consider S_5, with within class outcomes $H\overline{L}$ and payoff $\pi_5 = (T+1)/T > 1$. Suppose $x < 0$. To show that S_5 is a strict equilibrium class, it suffices to show that $\pi(s_i, s_j) < \pi_5$ whenever $s_i \notin S_5$ and $s_j \in S_5$. Suppose first that s_i involves the play of H initially, generating outcome (H, H). This outcome must occur again if $\pi(s_i, s_j) \geq \pi_5$ is to hold; otherwise all remaining periods involve either miscoordination or coordination on L, with miscoordination in at least one period, since $s_i \notin S_5$. There are just four possible outcome paths that result in a return to (H, H), given that $s_j \in S_5$, and these are as follows:

$$\begin{array}{ccccccccc} (H,H) & \rightarrow & (H,L) & \rightarrow & (H,H) & & & & \\ (H,H) & \rightarrow & (H,L) & \rightarrow & (L,H) & \rightarrow & (H,H) & & \\ (H,H) & \rightarrow & (L,L) & \rightarrow & (H,L) & \rightarrow & (H,H) & & \\ (H,H) & \rightarrow & (L,L) & \rightarrow & (H,L) & \rightarrow & (L,H) & \rightarrow & (H,H) \end{array}$$

These are the only possibilities because s_j plays L after either symmetric outcome, and all other feasible paths never return to (H, H). It is easily verified that in all four cases, $\pi(s_i, s_j) < \pi_5$. The case where s_i initially plays L can be dealt with similarly. In either case, S_5 is a strict equilibrium class, and there exists a sampling equilibrium with support S_5.

If $x > 0$, however, S_5 is not a strict equilibrium class, and there exists no sampling equilibrium with support S_5 as long as T is sufficiently large. To see this, note that there exists a strategy $s_j \in S_5$ that chooses H initially and after any asymmetric outcome, and chooses L after any symmetric outcome; see Figure 4 for an illustration. Consider strategy s_i matched with s_j, where s_i chooses H unconditionally. The resulting sequence of stage game payoffs to s_i is $\{2, x, 2, x, ...\}$. Suppose $x \geq 0$ and T is odd. Then

$$\pi(s_i, s_j) = \frac{2 + (x+2)(T-1)/2}{T} \geq \frac{T+1}{T} = \pi_5.$$

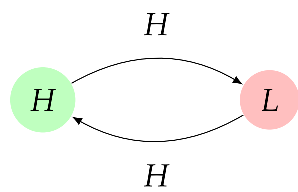

Figure 4: A strategy that plays H initially and after asymmetric outcomes, and plays L otherwise.

Hence there cannot be a sampling equilibrium with S_5 as the best present class as long as $x \geq 0$, and T is odd and at least equal to 3. If T is even, then

$$\pi(s_i, s_j) = \frac{2+x}{2} \geq \frac{T+1}{T} = \pi_5$$

as long as $T > \overline{T} = 2/x$. As long as $x > 0$ there exists such a $\overline{T}$, and hence S_5 is not a strict equilibrium class when T is sufficiently large. □

References

Dilip Abreu and Faruk Gul. Bargaining and reputation. Econometrica, 68(1):85–117, 2000.

Dilip Abreu and Ariel Rubinstein. The structure of Nash equilibrium in repeated games with finite automata. Econometrica, pages 1259–1281, 1988.

Dieter Balkenborg and Karl H. Schlag. Evolutionarily stable sets. International Journal of Game Theory, 29(4):571–595, 2001.

Kaushik Basu and Jörgen W. Weibull. Strategy subsets closed under rational behavior. Economics Letters, 36(2):141–146, 1991.

Jean-Pierre Benoit and Vijay Krishna. Finely repeated games. Econometrica, 53:905–922, 1985.

Kenneth G. Binmore and Larry Samuelson. Evolutionary stability in repeated games played by finite automata. Journal of Economic Theory, 57(2):278–305, 1992.

Juan Camilo Cárdenas, César Mantilla, and Rajiv Sethi. Stable sampling equilibrium in common pool resource games. Games, 6(3):299–317, 2015.

David M. Kreps and Robert Wilson. Reputation and imperfect information. Journal of Economic Theory, 27(2):253–279, 1982.

David M. Kreps, Paul Milgrom, John Roberts, and Robert Wilson. Rational cooperation in the finitely repeated prisoners' dilemma. Journal of Economic Theory, 27(2):245–252, 1982.

César Mantilla, Rajiv Sethi, and Juan Camilo Cárdenas. Efficiency and stability of sampling equilibrium in public goods games. Journal of Public Economic Theory, forthcoming, 2019.

Richard D. McKelvey and Thomas R. Palfrey. An experimental study of the centipede game. Econometrica, 60(4):803–836, 1992.

Paul Milgrom and John Roberts. Predation, reputation, and entry deterrence. Journal of Economic Theory, 27(2):280–312, 1982.

Martin J. Osborne and Ariel Rubinstein. Games with procedurally rational players. American Economic Review, 88(4):834–847, 1998.

Martin J. Osborne and Ariel Rubinstein. Sampling equilibrium, with an application to strategic voting. Games and Economic Behavior, 45(2):434–441, 2003.

Klaus Ritzberger and Jörgen W. Weibull. Evolutionary selection in normal-form games. Econometrica, 63(6):1371–1399, 1995.

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

Yuval Salant and Josh Cherry. Statistical inference in games. Unpublished Manuscript, Northwestern University, 2019.

William H. Sandholm, Segismundo S. Izquierdo, and Luis R. Izquierdo. Best experienced payoff dynamics and cooperation in the centipede game. Theoretical Economics, forthcoming, 2019a.

William H. Sandholm, Segismundo S. Izquierdo, and Luis R. Izquierdo. Stability for best experienced payoff dynamics. Journal of Economic Theory, forthcoming, 2019b.

Rajiv Sethi. Stability of equilibria in games with procedurally rational players. Games and Economic Behavior, 32(1):85–104, 2000.

Jeroen M. Swinkels. Evolutionary stability with equilibrium entrants. Journal of Economic Theory, 57(2):306–332, 1992.

Bernhard Thomas. On evolutionarily stable sets. Journal of Mathematical Biology, 22(1): 105–115, 1985.