Robert Samuel Simon

The challenge of non-zero-sum stochastic games

Article (Accepted version) (Refereed)

Original citation:
Simon, Robert Samuel (2016) *The challenge of non-zero-sum stochastic games.* International Journal of Game Theory. ISSN 0020-7276

DOI: 10.1007/s00182-015-0497-3

The Challenge of Non-Zero-Sum Stochastic Games

Robert Samuel Simon

August 17, 2015

London School of Economics
Department of Mathematics
Houghton Street
London WC2A 2AE
R.S.Simon@lse.ac.uk
+44-207-955-6753
fax +44-207-955-6877

Abstract

For a broad definition of time-discrete stochastic games, their zero-sum varieties have values. But the existence of ϵ-equilibrium for their corresponding non-zero-sum games has proven elusive. We present the problems associated with ϵ-equilibria in non-zero-sum stochastic games, from both the perspectives of proving existence and demonstrating a counter-example.

Key words: Stochastic Games, Topological Dynamics, Markov decision processes, equilibrium existence

1 Introduction

A time-discrete stochastic game is played in stages. At every stage the game is in some state of the world. The players are informed of the whole history, including the current state, and they choose simultaneously a pair of actions. The current state and actions of the players chosen determine a probability distribution according to which a new state for the next stage is chosen. The payoffs for the players are determined by the history of play, which may be defined with the help of payoffs at each stage.

Beyond the above description, there is no commonly accepted definition of what is a stochastic game. Are there finitely many states or infinitely many? Should the ultimate payoffs be determined primarily by stage payoffs, and if so, then how? Should the actions at every stage be finite, infinite, or in some sense compact?

The first model of stochastic games is due to L. Shapley (1953), introduced in the context of two-player zero-sum games with finitely many states, finitely many actions, and discounted sums of stage payoffs. Shapley proved that these stochastic games have values. A value of a game is a real number r such that for every $\epsilon > 0$ the first player has a strategy that guarantees the expected payoff of $r - \epsilon$ and the second player has a strategy that holds down the expected payoff of the first player to no more than $r + \epsilon$.

Because our interest in stochastic games has been in the difference between the zero-sum and non-zero-sum games, in this paper we restrict ourselves to a class of stochastic games for which their corresponding zero-sum varieties do have values.

A stochastic game is *normal* if
(1) there are finitely many players and at any state the action sets for all players are finite,
(2) there are countably many states,
(3) the strategy of a player is a determination, for each state and finite history leading to that state, of a probability distribution on the actions that are available at that state,
(4) the payoff for each player is a function on the infinite histories of play that is measurable with respect to the smallest σ-algebra that includes, for every finite history truncation, the set of all its infinite stage continuations,

(5) the payoff functions for all players are bounded, meaning that there is an $M > 0$ such that no matter what happens in the game the payoffs to all players are within the set $[-M, M]$.

Both discounted payoffs and un-discounted limit average payoffs (either lim-inf or lim-sup) are Borel measurable functions on the infinite histories of play. By limit average payoffs we mean that at every stage there are stage payoffs and the ultimate payoff for a player is a limit of the averages of these payoffs. Indeed it is difficult to formulate a stochastic game satisfying all the conditions except for (4) that is not normal. Using an axiom of choice, D. Gale and F. Steward (1953) found a (not normal) zero-sum stochastic game that does not have a value.

For any non-negative ϵ, an *ϵ-equilibrium* in a game is a profile of strategies, one for each player, such that no player can gain in expected payoff by more than ϵ by choosing a different strategy, given that all the other players do not change their strategies. An equilibrium is a 0-equilibrium. We say that approximate equilibria exist if for every positive ϵ there exists an ϵ-equilibrium.

The combined work of D. Martin (1975, 1998) and its application by A. Maitra and W. Sudderth (1998, 2003) proved that zero-sum normal stochastic games do have values. The result for zero-sum stochastic games with finitely many states and actions and undiscounted limit average payoffs was proven by J.-F. Mertens and A. Neyman (1981). These discoveries pose the natural open question: do all non-zero-sum normal stochastic games have approximate equilibria?

In this context, we should mention that when a normal stochastic game has perfect information, an application of the above mentioned work of Martin, Maitra and Sudderth by J.-F. Mertens and A. Neyman (J.-F. Mertens 1986) showed that it has approximate equilibria. Because the challenges of stochastic games of this paper concern mostly the problem of interpreting a player's intentions from her behaviour through statistical analysis, this result is a side issue here. Nevertheless it is interesting in its own right. Alter the game of chess so that (1) the winner gets 1 point, the loser -1 point, (2) the rules allow for endless play, and (3) whenever there is a draw or the game never ends both players receive -2 points. Players heading toward what used to be a draw from perpetual play are now torn between pressing toward a win and preferring to lose the match just to avoid the perpetual play. It is not obvious that such a non-zero-sum game with an infinite game tree should have approximate equilibria, let alone other such games with more players.

There are not many major results showing the existence of approximate equilibria for non-zero-sum normal stochastic games when the payoffs are not defined by a discount factor. The most important such result is that of N. Vieille (2000 a,b,c), for the special case of two-person games with finitely many states and limit average payoffs.

We have identified at least three problems with the existence of approximate equilibria for normal stochastic games, problems which need to be overcome if there is to be a general proof of their existence. It is our work with limit average payoffs (in particular recursive games, described below) that has highlighted these three problems; however these problems are present with more general types of normal stochastic games.

(A) The first problem is that of converting δ-equilibria of induced one stage games into ϵ-equilibria of the infinite stage stochastic game. Let $\sigma_1, \ldots, \sigma_n$ be the strategies of the n players of an infinite stage stochastic game. By an induced one stage game we mean the game where at some stage (after some finite history) the players are allowed to choose their actions independently of the strategies $\sigma_1, \ldots, \sigma_n$, however after this stage they are required to adhere to $\sigma_1, \ldots, \sigma_n$.

(B) The second problem involves the existence of an orbit of a correspondence that has a topological structure. A basic result relevant here is the Structure Theorem of E. Kohlberg and J.-F. Mertens (1986), which gives a topological structure to equilibrium correspondences.

(C) The third problem concerns the coordination of punishment.

Although one could attempt to tackle the problem of approximate equilibrium existence in all generality, it would be wiser to work on these three types of stochastic games that lie on the frontier of the question.

Before describing these three types of games, a minor technical definition is necessary. A state in a stochastic game is *absorbing* if once the play reaches this state it never leaves it, and furthermore by reaching this state the game is over, meaning that the payoffs for all players are determined.

The three types of normal stochastic games considered here are:

(1) two-player games with countably many non-absorbing states,

(2) games with many players and only one non-absorbing state,

(3) games with three players and at least two non-absorbing states.

These three types of games highlight the three main problems in establishing the existence of approximate equilibria. Roughly speaking, Problem (A) is a problem of both game types (1) and (3), though it is easiest to identify with game type (1); Problem (B) is very relevant to game type (2) and may be relevant also to the other two game types; and Problem (C) occurs with game type (3) but not with the other two game types.

A stochastic game is *recursive* if the payoff at all non-absorbing states is zero, no matter what the players do. A recursive stochastic game is *positive* recursive if there is some player whose payoffs at all absorbing states are positive. Call the game *uniformly* positive recursive if there is a positive value r and a player who receives at least r at all absorbing states. A positive recursive stochastic game has the absorbing property if some player who receives only positive payoffs at all absorbing states can force the play to absorption.

Recursive games are an important class of stochastic games. First, their payoffs can be interpreted as limit averages of stage payoffs; once an absorbing state is reached one could say that the payoffs corresponding to this state are the payoffs the players receive again and again at each following stage to the end of time. And if no absorbing state is reached then 0 is the limit average payoff for each player. Second, by expanding the state space of any normal stochastic game so that there is a one-to-one relationship between the finite histories of play and the states, any state corresponds to a clopen (open and closed) subset of the infinite histories of play and every open subset of the infinite histories of play will correspond to some collection of states. A stochastic game where all non-zero payoffs are determined by membership in an open set of the infinite histories of play becomes in this way equivalent to a recursive game. Notice that if all absorbing payoffs are positive then the payoffs are lower-semi-continuous and if all absorbing payoffs are negative then the payoffs are upper-semi-continous (as functions on the infinite histories of play).

The rest of this article is organised as follows. We present the three types of games and their special problems for approximate equilibria. In conclusion, we reveal one way one might find a normal stochastic game that lacks approximate equilibria, were one to exist. This way applies to all three types of games handled in this paper.

2 Two Players, Infinitely many States

N. Vieille's proof for the existence of approximate equilibria for two-player games with finitely many states and limit average payoffs is rather complex, and we cannot explain here in detail how it works. However we highlight a few aspects of his proof that pertain to the problem of generalising the result to countably many states.

After a reduction argument, done by Vieille, it suffices to show that there are approximate equilibria for (uniformly) positive recursive games with the absorbing property. Given any probability $p > 0$, from the above assumptions, there will be an $\epsilon > 0$ small enough so that with any ϵ-equilibrium the probability that the game will stay forever in non-absorbing states will be no more than p. The question remains, however, how and in what time frame will the absorbing states be reached?

This is an important question because a first approach to the problem suggests a counter-example by embedding into some stochastic game the famous game (without approximate equilibria) where two children state natural numbers (the fictitious income of their parents) and whoever states the larger number is the winner. The idea is that the numbers could reflect the stage on which a player starts to bring the game toward absorption and in general we assume it is more advantageous to let the other player bring the game toward absorption than to do it one's self. However experience with two-player stochastic games with such a structure suggests that there is an equilibrium where one player is primarily responsible to bring the game toward absorption and another where this is true of the other player. Indeed, Vieille's proof uses this asymmetry between the roles of the players, albeit with the complexity of many non-absorbing states. If there is only one non-absorbing state, the proof is much simpler, see F. Thuijsman and O.J. Vrieze (1989). This asymmetry of the players is puzzling, because with finitely defined bi-matrix games, generically the number of equilibria is odd. This asymmetry suggests both a proof and a counter-example to more general classes of stochastic games, as one could imagine both that Vieille's approach is unique to the class of games for which it applies or that the player asymmetry is part of a more general and successful approach.

With the strategies utilised in Vieille's proof, there is a finite number k such that the approximate equilibria are defined by breaking each non-absorbing state into k historical classes such that for each non-absorbing state and each historical class each player is asked to perform a stationary strategy. Punishment strategies may be set off if some player engages in behaviour that is statistically deviant. We call such a strategy profile (minus the possibility of punishment) "situationally" stationary, where the critical property is the finitely many historical classes to every state. To support this approach, a δ - ϵ relationship was established by Vieille between the one-stage induced games and the original game; for every $\epsilon > 0$ there is a $\delta > 0$ small enough so that if the strategies define δ-equilibria of the induced one stage games then these strategies, potentially with the help of punishment, will define an ϵ-equilibrium. The idea is that each player is requested to act according to the situationally stationary strategies, and whenever cumulatively a player obtains an advantage of at least ϵ, that player is punished. To make this work, punishment must occur with sufficiently low probability so as to not alter significantly the expected payoffs.

This approach follows a more general principle, illustrated by R.S. Simon (2007): if a decision maker is asked to choose actions in a stationary way such that at each stage the induced future expected differences in payoff from these actions is never more than δ, then the probability that this player will gain cumulatively more than ϵ in payoff is no more than ϵ for the relationship $\delta = \frac{\epsilon^3}{nM}$, where n is the number of different states and M is the maximal difference in payoffs. The proof follows an application of the Doob Martingale inequality. The result was generalised further (R.S. Simon 2007), with the role of M and n replaced by the expectation of the total variation of changes in the future expected payoff. This result allows one to punish the decision maker for gaining cumulative advantage without punishing the honest decision maker unduly. A similar result was used also by E. Solan and N. Vieille (2001) in their analysis of quitting games, but where the strategies are Markovian rather than stationary (an issue addressed below).

There are two limitations to this approach: first that the behaviour of the decision maker must be stationary, and second that the δ-ϵ relationship depends on n, the number of states.

Can the δ-ϵ relationship (with finitely many states) be generalised from stationary to Markovian processes? This question appears to be critical to understanding stochastic games, and it is posed in R.S. Simon (2007) as a conjecture. If there is only one state where the decision maker can have any influence, the conjecture is easy to prove (R.S. Simon 2007), and this is the context of quitting games.

The second limitation is the severe one. Consider the example where a decision maker is asked to move left or right with equal probability, and whenever the difference between the number of left and right moves equals some positive integer N or its negative $-N$ then the process stops. Assume also that the decision maker receives a payoff of 1 if the process stops with N more right moves than left moves, and vice-versa a payoff of -1 if there are N more left moves than right moves. Given that the choice of left and right are given equal probability throughout, the expected value for the payoff at any given history is $\frac{k}{N}$, where k is the difference between the number of right moves and left moves so far with that history. A single choice of left will decrease this value by $\frac{1}{N}$ and a single choice of right will increase this value by $\frac{1}{N}$. With $\epsilon > 0$ given and N any number larger than $\frac{2}{\epsilon}$, the decision maker will not increase her expected payoff by more than ϵ at any stage and yet the honest decision maker will receive 1 with a probability of $\frac{1}{2}$ and -1 with the same probability, thus guaranteeing a payoff of 0 when the initial state is 0. This dependency of the $\delta - \epsilon$ relationship on the number of states inspires the following approach to stochastic games with countably many states.

Let us keep the assumption that the game is uniformly positive recursive with the absorbing property and include the simplifying condition that from any state only finitely many states can be reached on the next stage. Assume further that the number of non-absorbing states that can be reached at stage n is represented by a function $s(n)$ such that $\lim_{n\to\infty} s(n) = \infty$. How fast should $s(n)$ grow so that the non-existence of approximate equilibria becomes plausible? If it grows too slowly, the distinction with finitely many states may disappear and Vielle's proof may be relevant for these games as well. On the other hand, if it grows too quickly, either the property that some player can force absorption (to a positive payoff for herself) may be broken or that this player may be too easy to control through statistical analysis.

A key question is whether or not $s(n)$ should be bounded by a polynomial function. In particular, $\sqrt{n}$ could play a critical role in defining the non-absorbing states on the nth stage. As with games with finitely many states, we assume that in general the approximate equilibria will have to be supported by punishment strategies. What should set off the punishment of a player? The easy answer is that a player should be punished if it obtains cumulatively through its actions too much of an advantage. One way of reaching an absorbing state could be the consequence of an independent random process, for example the repeated choice of left or right by some player X such that absorption occurs if more left than right (or more right than left) is chosen by a difference of at least $C\sqrt{n}$, where n represents the number of stages so far played and C is a uniform constant. The other player Y may not be able to determine with any sufficient confidence through statistical testing whether this player X chose to reach an absorbing state through intention or through honest random behaviour. Without any basis for trust, to keep player X "honest", e.g., unable to obtain significant payoff gains through cheating, it may be necessary to resort from the start with punishment strategies defined by player X's zero-sum game. Such ways of playing by Y could result in bad payoffs for both players and not define approximate equilibria. A "conflict of interest" may arise between pursuing one's payoff and keeping the other player honest, a conflict that is overcome in finite stage games through an application of a fixed point theorem to obtain a Nash equilibrium (and in the Vieille proof through an auxiliary game which is also finite dimensional in character).

And if Player X should choose one of three actions with a $\frac{1}{3}$ probability for each, or both players should be making $\frac{1}{2}$- $\frac{1}{2}$ choices of two actions, Cn states would be necessary (hence the question whether the function $s(n)$ should be bounded by a polynomial).

When moving from finitely many to infinitely many states, consider the following variation of the pigeon hole principle: if S is a finite subset of a set T and $(t_1, t_2, \ldots)$ is an infinite sequence of elements of T such that for all $s \in S$ there are only finitely many i such that $t_i = s$, then there is an N such that $n \geq N$ implies that $t_n \notin S$. This fact has its probabilistic variation: if a random process visits the elements of T such that for every $s \in S$ the expected number of visits to s is finite, then with probability one the process will eventually leave the finite set S permanently. With countably many non-absorbing states one can visit every one of them only finitely many times and still never reach an absorbing state. The break between potential and actual payoffs (illustrated best in the theorem in the last section of this paper) essential to our analysis of stochastic games becomes more difficult to bridge when there are infinitely many states. With infinitely many states it is much easier for play to appear to be moving toward certain absorbing states and their corresponding payoffs without this actually ever happening.

Before one concludes that there must be a counter-example with countably many states, there is, however, a mystery to two-person stochastic games which leaves some hope for the existence of approximate equilibria.

A quasi-stationary strategy profile is one that usually remains stationary, however in exceptional circumstances resorts to punishment strategies. If the limit average payoff is used as an evaluation, such approximate equilibria exist in zero-sum games with finitely many actions and states (Flesch, J., Thuijsman, F. and Vrieze, O.J. 2000). N. Vieille asked whether there exists "quasi-stationary" approximate equilibria for all two player games with finitely many states. The question was answered in the negative by R.S. Simon (2006). This question is closely related to the following question for finite state games: at what rate will absorption occur with an ϵ-equilibria? If generically approximate equilibria in finite state games involves complex and very slowly absorbing processes, as suggested by the Vieille proof, then there would appear to be no hope to extend this result to infinite state games where perhaps the absorbing states are running away from the players faster than they can get to them through equilibrium-like play. Although this example of R.S. Simon (2006) does not have quasi-stationary approximate equilibria, it has an additional equilibrium involving a rapid rate of absorption which is neither stationary in character nor based on the Vieille proof. All attempts to remove such an equilibrium from this game, while preserving the non-existence of quasi-stationary approximate equilibria, have failed. It is possible that there lurks a significantly different proof for two-person games with finitely many states, one that calls for faster absorption rates and also can be generalised to infinitely many states. If so, we suspect that this is related to the topological structure of stochastic games, (handled in the next section). The reason is that this topological structure regulates the changes in future expected payoffs, and this rapidly absorbing equilibrium in R.S. Simon (2006) also involves rapidly changing future expected payoffs.

3 One Non-absorbing State, Many Players

In this section, we consider quitting games, though much written will be relevant to the wider class of stochastic games.

A quitting game has only one non-absorbing state. Each player at the only non-absorbing state has only two moves, c for continue and q for quit. As soon as one or more of the players at any stage chooses q, the game stops (enters an absorbing state) and the players receive their payoffs, which are determined by the subset of players that choose simultaneously the move q. As long as no player has stopped the game, all players receive a payoff of zero.

In a quitting game the histories of play are of no use to the players because for any given stage of play there is only one history that counts, the history for which all players had chosen c on all stages. The only strategies for the players are Markovian, meaning that behaviour is determined only by the position and the stage (and in this case only by the stage since there is only one position that is non-absorbing).

The formal definition of quitting games was introduced in abstraction by E. Solan and N. Vieille (2001) but studied first on explicit examples by J. Flesch, F. Thuijsman, and O.J. Vrieze (1997). Interest in quitting games is to a large extent due to the latter authors' discovery of a game with cyclic symmetry with respect to the players such that for all sufficiently small $\epsilon > 0$ and ϵ-equilibria the future expected payoffs conditioned by the event that nobody has ended the game change dramatically with the progression of stages. A more radical example was presented by E. Solan and N. Vieille (2003), where the future expected payoffs conditioned on non-absorption involve a alternation between two distinct vectors.

There may be approximate equilibria in a quitting game without having any 0-equilibrium (E. Solan 2001), and the same even holds for zero-sum games (R.S. Simon 2012). In a more complex setting of stochastic games, there is an example by H. Everett (1957).

E. Solan (1999) proved that all three player quitting games have approximate equilibria (and this proof concerns stochastic games with only one non-absorbing state). Approximate equilibria for a larger class of two-player stochastic games was proven by F. Thuijsman and O.J. Vrieze (1989). The ultimate complexity of quitting games lies in the potentially large number of players, yet even for four players the question of approximate equilibrium existence is still open.

E. Shmaya and E. Solan (2004) proved the existence of approximate equilibria for a larger class of two-player stopping games, where the transitions to absorption are dependent on the stage of play. Their approach used a stochastic version of Ramsey Theory; Z. Xu (2011) generalised this to a profound result in stochastic Ramsey Theory.

There is a proof of approximate equilibria by E. Solan and N. Vieille (2001) for a subclass of quitting games involving very restricted conditions on the payoffs, and this result was extended to "escape" games by R.S. Simon (2007).

Quitting games benefit from the above mentioned δ-ϵ relationship between δ-equilibria of the induced one stage games and the ϵ-equilibria of the infinite stage game, proven first by E. Solan and N. Vieille (2001). The above mentioned generalisation to Markovian processes is unproblematic when there is only one non-absorbing state.

The key to understanding quitting games is the reduction of the game into one-stage games, as described above; strategies for all players at all stages are given and the players are allowed to choose their actions, potentially differently, in only one stage of play. Let n be the number of players. We consider the following simple n-dimensional matrices. In all positions where at least someone has chosen q the corresponding absorbing payoff vector is placed. Where all players choose the move c we place a variable vector payoff $x \in \mathbf{R}^n$ that represents a future expected payoff x on the next stage conditioned on non-absorption. We call this one-stage game $\Gamma(x)$. The equilibria lying over the space $\mathbf{R}^n$ of such vectors x has a topological structure, as is suggested by the Structure Theorem of E. Kohlberg and J.-F. Mertens (1986). In order to understand the dynamics of future expected payoffs, (by the going from one stage to the next), one must understand better what structure theorems say about this dynamic process.

A quitting game has *stationary approximate equilibria* if for every $\epsilon > 0$ there is a $p \in [0,1]^N$ such that $(p, p, p, \ldots)$ is an ϵ-equilibrium. A quitting game has *instant approximate equilibria* if for every $\epsilon > 0$ there is a $p \in [0,1]^N$ with $p^j = 1$ for some player $j \in N$ such that an ϵ-equilibrium is described by the behaviour p on the first stage followed by punishment of player j by the other players from the second stage and on, given that player j has failed to quit on the first stage.

Even without a structure theorem, the existence of approximate equilibria is implied by the existence of infinite sequences of one-stage approximate equilibria that involve a certainty of absorption (E. Solan and N. Vieille 2001); furthermore they can be assumed to be cyclic in nature, meaning that these one-stage equilibria are repeating with a finite and fixed period of repetition. Let F_ϵ be the correspondence from $\mathbf{R}^n$ to the subsets of $\mathbf{R}^n$ such that $F_\epsilon(x)$ is the set of y such that y represents the expected payoffs of an ϵ-equilibrium of $\Gamma(x)$. Once the special cases of quitting games with stationary or instant approximate equilibria are removed, the question of approximate equilibria becomes equivalent (R.S. Simon 2012) to the existence, for every $\epsilon > 0$, of orbits of the correspondence F_ϵ in $\mathbf{R}^n$ with infinite total variation (meaning that $\sum_{i=1}^{\infty} ||x_i - x_{i-1}|| = \infty$ when $x_0, x_1, \ldots$ is the orbit). In other words, lacking such an orbit of F_ϵ for some fixed $\epsilon > 0$, we can confirm that a quitting game is a counter-example.

When we discovered a relationship between the topological structure of one stage approximate equilibria and quitting games, we posed a question on dynamics defined by a correspondence with certain topological properties whose affirmation would prove the existence of approximate equilibria in quitting games. Reducing to the special case of functions, we had the following question:

Let E be a Euclidean space and C a compact subset in E of the same dimension. Let $f : C \to E$ be a continuous function such that $f(\delta C)$ is contained in C. **Question:** Does there exists an infinite orbit of f, meaning a sequence $x_0, x_1, \ldots$ in C such that $x_{i+1} = f(x_i)$ for all $i = 0, 1, \ldots$?

Unfortunately the answer is no; the counter-example to the above functional question is in M. Gobbino and R.S. Simon (2013) and the counter-example to the question on correspondences is a slight variation of an example presented in the same paper.

What is the connection between topological dynamics and quitting games? Assume for a quitting game that every player quitting alone gets a positive payoff by doing so. Without loss of generality, we can normalise this payoff to 1 for every player. With n the number of players, let W be the subset of $\mathbf{R}^n$ defined by $W := \mathbf{R}^n \backslash \{r \mid \forall j \quad r^j > 1\}$, equivalently W is the set of vectors where some player gets no more than 1. A structure theorem proven in R.S. Simon (2012) shows that the part of the equilibrium correspondence where some player quits with positive probability is a homotopic image of the interior of the set W such that a boundary point x of W is mapped to the game $\Gamma(x)$ and its equilibrium where all players choose to continue with certainty. By slightly perturbing the equilibria so defined at the boundary of W, by letting some player (whose continuation payoff is close to 1) quit with very small probability, one creates a correspondence that represents a subset of the one-stage ϵ-equlibria.

Looking at the above question on functions, notice that the iterations of the function f can appear to be continuous, but in reality they are not. If $f(r) = s$ and s is in δC, the iteration f^2 may be not properly defined on a neighbourhood of r: for one point r_1 near to r the point $f(r_1)$ may fall inside of C with $f^2(r_1)$ potentially far away from $f(r_1)$ and for another point r_2 near to r, the point $f(r_2)$ may be outside of C with no continuation to any $f^2(r_2)$. However if we require that near to the boundary of C the function f causes only a very slight change in position, we are led to different questions of dynamics. But because it is difficult to quantify what we mean by only a very slight change in position, one is led to questions that ask not only for the existence of an orbit but of an orbit with infinite total variation. In R.S. Simon (2012), a revised question (still open) along these lines was posed whose affirmation also would confirm the existence of approximate equilibria in quitting games. Because this question leaves much less room for discontinuous behaviour through iterations, a positive answer is very plausible.

Without implying that there must be a counter-example, we relate what would be necessary for a quitting game to be a counter-example, if one exists. Every orbit of an ϵ-equilibrium correspondence (of the one-stage games) must have finite total variation, meaning that the orbits must eventually converge to a point y such that $F_\epsilon(y) = \{y\}$. The set of such vectors we call D, (for the dead vectors). The set D will have a positive distance from $W = \mathbf{R}^n \backslash \{r \mid \forall j \quad r^j > 1\}$; that distance may depend on the chosen ϵ. Such orbits cannot represent approximate equilibria because they would imply that after some stage reached with significant probability the players never choose to quit with any positive probability, contradicting the positive payoff that some player would obtain by quitting alone (R.S. Simon 2012).

Lets look at a subclass of quitting games for which there are approximate equilibria, proven in R.S. Simon (2007): if x is not in $W = \mathbf{R}^n \backslash \{r \mid \forall j \quad r^j > 1\}$ then the only equilibrium of $\Gamma(x)$ is the one where all players choose to continue with certainty. The proof involves looking at any pathway P of vectors in $\mathbf{R}^n$ such that one end of P lies slightly inside of W, the other end lies far outside of W, and the intersection of P with the boundary of W is only one point. Let y be the end point of P inside of W and assume it represents what happens after one player quits with some small probability. Let x be the intersection of P with the boundary of W. If some iteration of y ever involves motion exclusively to points of distance greater than ϵ from W, the topological structure of the equilibrium correspondence (in particular the connectivity that it implies) requires that there is some point between y and x such that the same iteration send this point back to a point exactly a distance of ϵ from W (and from which the process can continue with some player quitting with very small probability). In this way an orbit with infinite total variation can be found.

How does one frustrate the above structural argument? The answer is simple; at some such paths P the equilibrium correspondence over P can curve around so that there is some point in $P \backslash W$ that maps to a point inside of W. Naturally a lot of motion back into the set W from outside of W suggests a diminished set D and therefore potentially the existence of orbits with infinite total variation. However with higher dimension, e.g. a large number of players, the dynamics in the space $\mathbf{R}^n \backslash W$ could allow for very complex motions involving travelling several times away from and back again into W before final motion to the set D.

Lets look more closely at a one-stage game $\Gamma(r)$ where $r^j > 1$ for all players j and yet there is some equilibrium of $\Gamma(r)$ where some players choose to quit with positive probability and the result is an expected payoff that lies inside of W, (meaning to a point $s \in \mathbf{R}^n$ such that $s^j \leq 1$ for some j). Looking at any player j who quits with positive probability, it does so because at least one other player also quits with positive probability. The same is true of that other player. The most likely induced structure is a cycle of players who inspire each other to quit with positive probability. As we would not want such a cycle to include all the players (as that would suggest either a serious difficulty ever landing into the subset D or a tendency for vectors outside of W to stay outside of W), we should look at quitting games with proper subsets of such cycles of players. This gives some clue why it is difficult to find quitting game counter-examples with only four players, as any directed graph on four vertices has a very limited structure of proper cycles. Only with five vertices can we find a good structure of proper cycles. However due to the ways one can move from the use of one cycle of players to the next cycle of players, we would expect a counter-example, should it exist, to have at least six players. We think that eight player quitting games are much more promising, but because of the extreme number of cases generated by higher dimensions we expect a confirmation of a counter-example to be possible only with the help of computers.

4 Three Players and Many States

This class of games introduces a unique problem, that of effective punishment. The minimax theorem does not apply to games with at least three players: if Player Two and Player Three are trying to punish Player One, there can be a big difference between what they can accomplish if they choose their punishment strategy first (the min-max value) and what happens if they respond to a fixed strategy of Player One (the max-min value). In an equilibrium, approximate or exact, when considering a deviant strategy that may inspire punishment, one must consider how the punished player will respond to a predetermined punishment strategy profile of the other players. In other words, it is the player's min-max value that matters, which may be higher than the max-min value. In some sense, the problem is swept under the carpet when dealing with games of a finite structure (finitely many stages, actions, and players) or with discounted payoffs through a Nash equilibrium obtained from a fixed point argument; the Nash equilibrium ensures that players get at least their min-max values without that being stated explicitly as a goal. Even a "folk theorem" type result, an approximate equilibrium supported by punishment, may become problematic if there are three players. With two player stochastic games, the players could be requested to engage in deterministic behaviour; if some player violates the agreement, she will be punished according to her zero-sum value at the first stage after the violation. Because zero-sum normal stochastic games have values, those zero-sum values are also the max-min values, a fact which helps support such a process. With three players, their min-max values may be relatively high and make the stability of such a process problematic. At present we cannot see why there cannot be a three-person normal stochastic game in which the payoffs for all infinite histories always sum to zero and yet the min-max value for each player at the start is positive. Such a game would be a counter-example to the existence of approximate equilibrium. In the continuous time context, such a counter-example does exist (R. Laraki, E. Solan, and N. Vieille 2005) for stopping games.

When there are many states, especially infinitely many, there is another complication with three player games introduced by punishment. When a player is required to play in a non-deterministic way, it is difficult to detect deviant behaviour. Such detection involves this player crossing a threshold of statistical deviation. In the context of supporting an equilibrium, the natural threshold is that a player is punished for gaining cumulatively too much advantage in payoff. (Other thresholds for punishment can apply, however they may not eliminate the fundamental problem of identifying who has deviated.) Let us assume that this happens with very small probability according to some prescribed strategies for the players. Nevertheless, with three or more players, there are other ways to obtain undue advantage than this direct approach. Player One could engage in behaviour that does not generate any cumulative advantage to herself *directly* through the prescribed strategies, yet brings an innocent Player Two into situations where he suffers a high risk of obtaining too much cumulative advantage and getting punished. Player One could benefit greatly from the joint punishment of Player Two by Player One and Player Three! This problem does not happen with two player games, as then Player Two could punish Player One at the same time that Player One is punishing Player Two. But the lack of a way for all three players to be punished simultaneously could sabotage many otherwise plausible approximate equilibria.

5 A general result

The following general approach to finding a counter-example with positive recursive games comes from our reflections on stock market bubbles. We look for an example where all future expected payoffs in "equilibrium" are fictitious and not derived from the reality of any future behaviour of the players.

Let $\mathcal{H}_m$ stand for the non-absorbing histories ending at the mth stage of play at a non-absorbing state. For any assignment $P : \mathcal{H}_m \to \mathbf{R}^n$ of payoffs for the n players at the mth stage, define the P truncation of the game to be the m stage game that defines all m stage histories to be absorbing with final payoffs defined by P.

Theorem: Let G be an n player normal stochastic game that is uniformly positive recursive with the absorbing property. Assume that all payoffs for the players lie within $[-M, M]$ for some $M > 0$. Let $f : \mathbf{N} \to \mathbf{N}$ be a function such that $f(k) > k$ for all k. The game G does not have approximate equilibria if there exists a $\rho > 0$, an $\epsilon > 0$ and an $N > 0$ such that for every $k > N$ and for every assignment P of payoffs in $[-M, M]$ for the $f(k)$th stage truncations, with every ϵ-equilibrium of the P truncated game the probability of absorption before the kth stage is no more than $1 - \rho$.

Proof: Let $r > 0$ be the positive value such that player j, who can force absorption, gets at least r at all absorbing states. Let us assume that there exists a $\min(\epsilon, \rho r/2)$ -equilibrium of the game, yet the above condition is met. As player j can force absorption for a payoff of at least r, there must be a stage k such that with this $\min(\epsilon, \rho r/2)$-equilibrium the probability of absorption before reaching the stage k must be at least $1 - \rho/2$. This $\min(\epsilon, \rho r/2)$-equilibrium establishes expected payoffs on the $f(k)$th stage of play, and these payoffs define a truncation P. However by the assumed condition this $\min(\epsilon, \rho r/2)$-equilibrium involves a probability of at least ρ of not absorbing before stage k, a contradiction. q.e.d.

The converse statement is problematic for games with at least three players because a player could deviate in a way that prevents absorption. If the game had only two players, they could punish each other if they fail over a sufficiently long time to reach an absorbing state. But with at least three players, we could have a situation similar to that described above, where it may be difficult for the players to identify who has deviated, as the detection of deviation may be statistical in character.

The above theorem could be improved by establishing which assignments P could correspond to the payoffs of approximate equilibria. However we fail to see the significance of such an improvement, given the large difference that could exist between k and $f(k)$.

6 References

Everett, H. (1957) "Recursive Games", in Dresher, M., Tucker, A., and Wolfe, P. (Eds), *Contributions to the Theory of Games III*, (*Annals of Mathematical Studies*), **39**, pp. 47-78.

FLESCH, J., THUIJSMAN, F., and VRIEZE, O.J. (1997). "Cyclic Markov Equilibria in a Cubic Game", *International Journal of Game Theory*, **26**, pp. 303-314.

FLESCH, J., THUIJSMAN, F., and VRIEZE, O.J. (2000). "Almost Stationary Epsilon-equilibria in Zero-sum Stochastic Games", *Journal of Optimization Theory and Applications*, **105 (2)**, pp. 371-389.

GALE, D. and STEWART, F. (1953). "Infinite Games with Perfect Information", in: *Contributions to the Theory of Games II*, Princeton University Press.

GOBBINO, M. and SIMON, R. S. (2013). "How Many Times can a Function be Iterated?", Journal of Difference Equations and Applications, **19 (2)**, pp. 332-354.

KOHLBERG, E. and MERTENS, J.-F. (1986). "On the Strategic Stability of Equilibria", *Econometrica*, **54 (5)**, pp. 1003-1037.

LARAKI, R., SOLAN, E and VIELLE, N. (2005) "Continuous-time Games of Timing," *Journal of Economic Theory*, **120 (2),** pp. 206-238.

MAITRA, A. and SUDDERTH, W. (1998). "Finitely Additive Stochastic Games with Borel Measurable Payoffs", *International Journal of Game Theory*, **27 (2)**, pp. 257-267.

MAITRA, A. and SUDDERTH, W. (2003). "Stochastic Games with Borel Payoffs", in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, **570**, Chapter 24, pp. 367-373.

MARTIN, D. (1975). "Borel Determinacy", *Annals of Mathematics*, **102**, pp. 363-371.

MARTIN, D. (1998). "The Determinacy of Blackwell Games", *The Journal of Symbolic Logic,* **63**, pp. 1565-1581.

MERTENS, J.-F. (1986) . "Repeated Games", In: Proceedings of the International Congress of Mathematicians, Berkeley, California, USA. American Mathematical Society, Providence. pp. 1528-1577.

MERTENS, J.F. and NEYMAN, A. (1981) . "Stochastic Games," *International Journal of Game Theory,* **10**, pp. 53 - 66.

SHAPLEY, L.S. (1953). "Stochastic Games", *Proceedings of the National Academy of Sciences, U.S.A.*, **39**, pp. 1095-1100.

SIMON, R.S. (2012). "A Topological Approach to Quitting Games", *Mathematics of Operations Research*, **37 (1)**, pp. 180 - 195.

SIMON, R.S. (2007). "The Structure of Non-Zero-Sum Stochastic Games", *Advances in Applied Mathematics*, **38 (1)**, pp. 1-26.

SIMON, R.S. (2006). "Value and Perfection in Stochastic Games", *Israel Journal of Mathematics*, **156 (1)**, pp. 285-310.

SHMAYA, E. and SOLAN, E (2004) "Two-player Non Zerosum Stopping Games in Discrete Time", *Annals of Probability*, **32 (3B)**, pp. 2733-2764.

SOLAN, E. (1999). "Three-Player Absorbing Games", *Mathematics of Operations Research*, **24**, pp. 669-698.

SOLAN, E. (2001). "The Dynamics of the Nash Correspondence and n-Player Stochastic Games", *International Game Theory Review* **3(4)**, pp. 291-299.

SOLAN, E. and VIEILLE, N. (2001). "Quitting Games", *Mathematics of Operations Research*, **26**, pp. 265-285.

SOLAN, E. and VIEILLE, N. (2003). "Quitting Games- An Example", *International Journal of Game Theory*, **31** , pp. 365-381.

THUIJSMAN F. and VRIEZE (1989). "On Equilibria in Repeated Games with Absorbing States", *International Journal of Game Theory*, **18 (3)**, pp. 293-310.

VIEILLE, N. (2000a). "Two-Player Stochastic Games I: A Reduction", *Israel Journal of Mathematics*, **119**, pp. 55-91.

VIEILLE, N. (2000b). "Two-Player Stochastic Games II: The Case of Recursive Games", *Israel Journal of Mathematics*, **119**, pp. 92-126.

VIEILLE, N. (2000c). "Small Perturbations and Stochastic Games", *Israel Journal of Mathematics*, **119**, pp. 127-142.

XU, Z. (2011). "A Stochastic Ramsey Theorem", *Journal of Combinatorial Theory, Series A*, **118(4)**, pp. 1392-1409.