TI 2002-111/1
Tinbergen Institute Discussion Paper

Altruism, Fairness and Evolution: the Case for Repeated Stochastic Games

Matthijs van Veelen

Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, and Tinbergen Institute

Altruism, fairness and evolution: the case for repeated stochastic games.

Matthijs van Veelen
Tinbergen Instituut
Keizersgracht 482
1017 EG Amsterdam

September 11, 2002

Abstract

Abstract

This paper is an effort to convince the reader that using a stochastic stage game in a repeated setting - rather than a deterministic one - comes with many advantages. The first is that as a game it is more realistic to assume that payoffs in future games are uncertain. The second is that it allows for strategies that make an evolutionary approach possible, while folk theorem strategies do not allow for such an analysis. But the most important feature is that such a setting allows for equilibrium strategies that look very much like human behaviour; altruism and fairness will be shown to feature in a natural way in equilibrium.

Journal of Economic Literature Classification Number: C70.

1 Introduction

How selfish soever man may be supposed, there are evidently some priciples in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. [...] That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment is by no means confined to the virtuous and humane, though they may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it.

Adam Smith.

Thus begins *the Theory of Moral Sentiments* by someone we mostly refer to as an economist now, but then held the Chair of Moral Philosophy at the University of Glasgow. One can voice all kinds of modern times critique on this book - that it is overly lengthy in its effort to be all-embracing and obviously written before both Nash and Darwin - but not being inhibited by contemporary scientific mores, it is a forceful attempt to discuss all of our everyday moral sentiments. Now, on the other hand, we do have Darwin and Nash, and the aim of this paper will be to pick up on only a small question from the wide scope of moral-related issues with the insights from evolutionary game theory. We will wonder whether game theory and evolution can explain this evidently present sentiment we call altruism, and the slightly more complicated but unavoidable notion of fairness. In this explanation repeated stochastic games will play a central role.

Of course there are models and opinions concerning the evolutionary history of fairness and altruism that are more up to date than Adam Smith's 1759 book. Some of those more recent insights feature in Section 3 and it will be clear that not all current researchers hold the same opinion on the evolutionary stability of altruism. After that, we will pay some attention to the role of repeated games in general and the repeated prisoners dilemma in particular as the focal point of the discussion on cooperation. This will be done in Section 4, where we will also discuss the difficulties that come with an evolutionary approach to this particular kind of games. In Section 5 we come to why I think the explaining power of (evolutionary) game theory can be taken a bit further if we allow for the stage game to be stochastic. I hope to show that this generalization improves on the deterministic version in three ways. The first is that I think it is more realistic to presume that if we are engaged in a repeated game, we do not know exactly what game we will play the next time we meet. A second advantage is that such a model opens the door for a well defined evolutionary approach, which is inhibited in a standard folk theorem setting. But the most important improvement is that the human tendency to think about behaviour in terms of fairness could, at least partially, be explained by identifying 'fair' preferences with strategies in a repeated stochastic game. In this respect, Theorem 8 is more or less central; for a large class of repeated stochastic prisoners dilemma's, it states that if players would initially be restricted to strategies that have an equivalent in terms of a fair preference, then an equilibrium of a natural type (satisfying *Independence of Hypothetical Situations)* remains an equilibrium if we allow for all strategies, including those that do not have a fair preference equivalent. In section 6, this result is put back into the perspective of the debate on altruism, fairness and evolution.

But first of all we will have to properly define altruism and fairness.

2 What is altruism and what is fairness ?

In this article both the notions of altruism and fairness will be related to sentiments that make us choose one action over another, or as we call them nowadays: preferences. We first confine ourselves to altruism. If we assume a world with two people and one commodity, we can imagine both of them having a preference over the possible quantities and divisions of that commodity. If we define x_1 as how much person 1 gets and x_2 as the quantity allocated to person 2, we can imagine the first person having a utility function $u_1(x_1, x_2) = x_1 x_2$. Such a utility function is a clearcut case of perfect altruism; both quantities enter the function symmetrically. We could also imagine person 2 having a utility function that completely ignores how much the other one gets; $u_2(x_1, x_2) = x_2$. Such preferences and the person who has them we usually call selfish. This may all seem too obvious to mention, but it should be noted that other definitions of altruism do appear in the literature every once in a while. For instance in a model by Eshel, Samuelson and Shaked (1998), players who fail to understand the consequences of their actions are somewhat misleadingly referred to as altruists, when the actions they choose to play in a misguided effort to get the maximal payoff, happen to benefit someone else. In this paper, rather than modelling altruism as behaviour under ignorance, I would like to focus on players that are fully aware of what payoffs go with what action, and yet consciously choose for an altruistic alternative.

Another point that should be mentioned is that there is no need to worry about the impossibilities that arise if we would want to formulate altruism in such a way that it allows for my preference to depend on yours and vice versa. We will see the payoffs as nothing more or less than evolutionary succes, the blunt probability to survive and procreate, however indirect that may be. Our (altruistic) preferences simply judge how much of our own success we are willing to give up for the success of someone else.

To define fairness is a bit less straightforward. There is an abundance of definitions and some of those are no less fitting to describe what we mean when we use the word in real life than the one that I chose. I will mention two alternative definitions, but before that I will try to defend the choice that I made, which is to see fairness as *reciprocal altruism*.[1]

The reciprocity part of fairness is, I think, no less evident than the altruism. If our neighbour has - again - been listening to Slayer at maximum volume untill three at night, we will not be very much inclined to help him the next morning with his ill-starting engine, or at least much less than if he always turnes off Classic FM at ten o'clock. If it wasn't for his inconsiderate musical habits, we would be happy to help him out, but given his behaviour we think it only fair that he reaps what he sowed. Our altruism apparently is not unconditional. Actually, I will argue that altruism between genetically unrelated people only survives because it is not unconditional, but we will save that for later.

Andreoni and Miller (2002) chose another definition. In their paper fairness, like altruism, is a characteristic of a preference over how much the players involved get. Fairness then is a preference for equal distributions. In our setting, this is subsumed under altruism and what they call a preference for fairness would in our case coincide with a curved or kinked altruistic iso-utility curve.

Karni and Safra (2002) take an axiomatic approach to describe choice behaviour of what they call *self-interested moral individuals.* Their setting is slightly different, since their subjects have random allocation pocedures of an indivisible good to choose from, which makes the set on which the preferences are defined a simplex in $\mathbb{R}^n$ rather than the whole of $\mathbb{R}^n$. Futhermore they endow

[1] In this paper, we will formalise and extend Trivers' (1971) notion of reciprocal altruism.

people with two sets of preferences: one representing actual choice behaviour and one that represents a person's notion of fairness. The assumptions they make, rule out what one could call odd combinations of these two different types of preferences, thereby formalising the position they take, which is that ethical value judgements do affect individual behaviour. In this paper we will only consider actual choice behaviour. That does not imply that I disagree with using two sets of preferences in an axiomatic setting, but since we will be looking for equilibria in games, it is only the actual behaviour that matters here.

A last possibility that I would like to mention, is that one could also include fairness as a social norm. Of course in real life there are actions we do take if no-one is looking, but refrain from before an audience - or vice versa. However decisive the judgement of a third party may be in reality, this paper's model will be too limited to include it. Here we will only look at the evolutionary stability of people 'doing the right thing' just because they want to and not because they anticipate the reaction or sanctions of third parties. Any preference that someone not directly involved could have over the division between two people other than him- or herself is therefore not considered or explained here. This leaves us with fairness being nothing more than reciprocal altruism. I admit that that is a limitation, but it is one that leaves more than enough to be explained.

Formally, we restrict ourselves to a model with only two players. Both of them are endowed with a *set* of preference relations; one for every possible history of the game. Suppose H is the set of all possible histories and assume that (x_1, x_2) is a vector that indicates current stage payoffs of both players - that is: possible combinations of incremental evolutionary succes. If we then let $\mathcal{R}$ denote the set of all possible preference relations over $\mathbb{R}^2$, we can define what we are interested in: functions $CP : H \to \mathcal{R}$ that are to reflect the conditional (altruistic) preferences of the two players.

3 Altruism and evolution

Like Adam Smith, I don't think we really need experiments to find out whether or not altruism exists. It may not always be there with an intensity we would wish for, but I would say it is indeed perfectly clear that none of us is completely indifferent to the fate of everyone else in the world.[2] Now given that we do not appear to behave a hundred percent self centered, another question would be: how did our species develop such behaviour. Before trying to answer this question, there is one thing we can check off first, and that is altruism towards those to whom we are genetically related. Genes that make us care for others that carry the same genes do better than genes that don't. The Selfish Gene by Dawkins (1976) contains an entertaining description of how lots of the behaviour that we usually (and rightfully) think of in terms of love and affection is perfectly congruent with payoff- or fitness maximizing behaviour from the gene's point of view. Since this is a perfectly adequate explanation for our altruistic behaviour towards our relatives, we will assume that the individuals we consider are not genetically related. After all, we do not only love our siblings and our children, but we also care about our friends and we even prefer not be too unfair towards people we hardly know. Our aim is to find out if it is possible that evolution has made us do so and we would like to see how mutation and selection might have shaped our notion of fairness.

In their article *"Is altruism evolutionary stable?"* Bester and Güth (1998) answer the question

[2] Of course experiments have provided us with useful insights in the peculiarities of our deviations from money maximizing equilibrium behaviour; I only want to mention Henrich et al. (2001) and Andreoni and Miller (2002), but there is an endless list of papers that would qualify for a reference here.

in their title with a straightforward yes. Their argument is more or less that nice people will be treated nicely. Crucial to their model is the assumption that (altruistic) preferences are common knowledge; when I meet another player, I can read his or her 'affective condition' from the face. Given this assumption they correctly argue that for a certain set of one-shot games that exhibit strategic complementarities, egoists will always be treated so much worse than altruists - who their opponents know will choose a more cooperative strategy - that their being selfish becomes a disadvantage rather than an advantage.

Binmore's answer to the same question is an equally staightforward no. He disagrees with the model of Bester and Güth, criticizing the assumption that altruism can credibly be signalled. In his book (1994) he remarks for a similar case that such a population can always be invaded by a mutant that looks like an altruist but behaves like an egoist. This mutant would succesfully exploit the cooperative attitude towards players that look like altruists and wipe out the true altruists. The new population in turn would not be evolutionary stable either, since it can be invaded by a mutant that is an altruist and bears a new sign of its altruism, but that does not really weaken the statement that the first population is not evolutionary stable. On top of that, Binmore actually states that altruism by definition is evolutionary unstable; after all, giving up fitness reduces your fitness.[3]

Now I tend to think that Binmore's argument against the evolutionary stability of altruism in the one shot game is rather convincing. Nonetheless, I am not sure that this should make us think that the altruism we observe is bound to disappear (ceteris paribus). In section 5 I show that altruism, together with reciprocity, can be a building block of an equilibrium strategy in a stochastic repeated game. Thereby it is of crucial importance that we are aware of the exact definitions we use. With Binmore's argument in mind, we should not think of altruism as a preference over the discounted outcome of the whole repeated game. By definition evolution selects winners and any such preference is indeed tautologically unstable. But while the selection takes place at a meta-player level, our players might just play their everyday games in a short sighted fashion. My idea of altruism is that it is to be seen as a myopic preference that is updated from time to time, depending on the actions being played. I do believe that we can be sincerely altruistic, but also that we do not always see beyond the ends of our noses. These two weaknesses however can make one strong strategy. Our altruism and the appropriate degree of reciprocity, that is, the right dependency of our level of altruism on the actions of the other players, might just add up to a winning strategy in the repeated stochastic game.

This myopic definition of altruism is not only interesting because it allows for an explanation of altruism. The possibility of a level of altruism that can be changed during the course of the game also makes it more realistic. After all, we do grow attached to some people and we do get disappointed in others. Updating opinions on and attitudes towards others is something we do from time to time and if this reciprocity can be taken into account, that, I think, is a good thing.

There are other angles though that should not be ignored. Both the models of Bester and Güth (1998) and Binmore (1994) as well as the model I will present here only consider natural selection, but this may not be the only type of selection at work. Since survival alone is not enough to pass on one's genes, Miller (2001) suggests not to overlook sexual selection as an explanation for

[3]Binmore does not deny that there is some true altruism, but in his opinion this is to be seen as a relic from the time that we lived in groups in which we were genetically related to all members. Nowadays we mainly meet people we are not genetically related to, but this situation, on an evolutionary time scale, has emerged only recently. He suggests that time simply has been too short to (fully) adapt. I do not want to reject this thought, but I do have some doubts since I can not think of a model of genetic competition that goes with the story, and emperical evidence is not very hopeful either; rather than supporting Binmores claim, experiments in small scale societies by Henrich et al (2001) seem to point in the opposite direction.

altruism and morality. Actually, he states that it is mate choice rather than survival rates that determined our notion of fairness. This line of reasoning certainly deserves attention, but it is more easily discussed after the possibilities that repeated stochastic games allow for are disclosed a little. Therefore the assessment of the different explanations is best postponed to Section 6. The same goes for Sober and Wilson (1998), who argue that group selection arguments are wrongfully dismissed in the discussion on the evolution of altruism.

Finally one could draw a parallel between the evolution of altruism and the evolution of the mouth. The mouth of our earliest ancestors that had one, served a single purpose: eating. Nowadays we -and other animals- do all kinds of other things with it as well, like speaking and singing and smiling and kissing. Since evolution tends to take the line of the least resistance, one can imagine that adapting an eating instrument for speach is easier than developing a seperate device to make us talk. The same might just go for altruism. Primarily it may have been there towards relatives only, optimizing the gene's fitness. Suitably adapted however it might very well serve as an optimal strategy to play repeated stochastic games with. In this light one can imagine that the growing complexity of human interactions and the growing possibilities of labour division can be characterised in terms of an increase in the numbers of non-zero sum games we play or an increase in the gains from cooperation. This, I think, would nurture a (complex) system of reciprocal altruism rather than weed out what altruism we have left as we drift away from simple society.

4 Repeated games: folk theorems and finite automata

The focal point of the game theoretic discussion on cooperation is the repeated prisoners dilemma. Generally speaking, one could say that there are two ways to get at such a game. One is to prove and use a variety of folk theorems, that show which outcomes can be supported by equilibria. This is done both for cases with and without discounting and for different types of equilibria, such as equilibria that are subgame perfect, renegotiation proof or equilibria that only use 'simple strategies'. (For a good overview of the different folk theorems, I would like to refer to Chapter 8 in van Damme (1987)[4]). The other possibility is to think of players as finite automata, that is, binary stimulus responders that react on the history but have neither perfect recall nor unlimited computational abilities. This restricts the set of strategies to those that can be represented by those finite automata. Since these finita automata are also called (Moore) machines, the game with this restricted strategy set is also referred to as a machine game, to emphasize that players choose machines to play the repeated game for them rather than chosing actions every period. This approach was initiated by Rubenstein (1986) and Abreu & Rubinstein (1988) while Binmore and Samuelson (1992) and Probst (1995) investigated the evolutionary stability of those finite automata. It is important to note that this type of analysis uses averaged, undiscounted payoffs only and that as from Abreu & Rubinstein (1988), the standard is to use lexicographic preferences, where complexity matters only when payoffs are equal. Apart from looking for equilibria in this restricted set of strategies, one is also tempted to put some of these machines together and see which one gets the upper hand. This is more or less what is done in the justly famous tournament called by Axelrod (1984). The experiment was followed by what could be seen as an evolutionary case study, where different mixtures of strategies were chosen as starting points.

There are three reasons why formal definitions from evolutionary game theory are not very helpful if one wants to pick an equilibrium that natural selection would be likely to pick (or have picked) as well. One applies to folk theorem strategies, one to finite automata and one to both.

[4]See also Fudenberg & Maskin (1986), Aumann & Shapley (1976), Abreu (1988), van Damme (1989).

Possibly superfluous, I will give the definition of evolutionary stability as stated in Taylor and Jonker (1978).[5] The strategies x and y that feature in this definition are not necessarily pure.

Definition 1 *A strategy x is an evolutionary stable strategy (ESS) if there exists an invasion barrier ϵ_y for every strategy $y \neq x$, where ϵ_y is an invasion barrier for $y \neq x$ if $u[x,(1-\epsilon)x+\epsilon y] > u[y,(1-\epsilon)x+\epsilon y]$ for all $\epsilon \in (0,\epsilon_y)$*

An equivalent definition of evolutionary stability is that x is an evolutionary stable strategy if both $u[x,x] \geq u[y,x]$ for all x and $u[x,y] > u[y,y]$ if $u[x,x] = u[y,x]$ for all $y \neq x$, which is how it was originally defined by Maynard Smith and Price (1973). I would like to stress that if an *ESS* indeed exists, then this static notion can serve as a useful shortcut to get at stable points in the replicator dynamics of a population that is playing such a game. We should keep in mind though that, more than finding an *ESS*, it is those rest points in a natural type of population dynamics that we are really after.

The most important feature of these definitions is that for a strategy x to be evolutionary stable, it is not enough to do at least as good as other strategies in a nearby population. Strategies need to do better in order to drive out mutant strategies (where the driving out, in the case of a mixed equilibrium strategy, is to be seen as a return to the equilibrium mixture of pure strategies). It is this strictness that poses a problem for both the approaches if one wants to apply the definition of evolutionary stability to (deterministic) repeated games and it is on the off-equilibrium paths where it goes wrong. To clarify that, I'd like to give two examples in a folk theorem setting. That is to say that even though some of the featuring strategies are finite automata as well, we will first stick to the case that allows for discounting and in which complexity does not play a role.

Please first consider the strategy *always defect*. This strategy clearly is in equilibrium with itself. Now think of a mutant strategy that defects in the first round and for all other rounds imitates the move of the other player in the previous round. (This strategy is sometimes called *tat-for-tit*, since it retaliates like *tit-for-tat*, but is its mirror image concerning the initial move). The strategy *always defect* always defects, so the condition for *tat-for-tit* to play cooperate is never met and therefore this mutation in inconsequential. Also when playing against itself, the mutated strategy leads to a strain of mutual defections, and therefore both strategies do equally well against each other and they also do just as well against the other as they do against themselves. The possibility of such a mutation therefore shows that the strategy *always defect* does not qualify for evolutionary stability as it is defined above. The situation however is a little worse than just that. To see why, assume a population of pure defectors. As we just saw, there is no evolutionary pressure against *tat-for-tit*, so there is no reason why it could not enter the population and attain a share that is sufficient to open the door for yet another mutant. Think for instance of a second mutant that plays cooperate in the first two rounds and from then on imitates the move of the other player in the previous round. With sufficiently high discount rates and a sufficient share of the first mutant, this second mutant does strictly better than *always defect* and its arrival also makes the first mutation improve its performance by dragging it into perpetual cooperation.

A similar strain of reasoning can be held for any (symmetric) strategy from the folk theorem, but then with falling incidence of cooperation. As an example we can take the simplest trigger strategy, where a player cooperates and only switches to forever defect if the other player deviates. (In a finita automata setting this strategy is sometimes referred to as *grim*). Such a strategy is vulnerable, not to the strategy *always cooperate* itself, but to the predators on pure cooperators. Indeed, if we take a strategy that prescribes cooperation and switches to a punishment phase if

[5] Note that these equilibria are always symmetric; we focus on populations with random matching.

the other deviates, then the equilibrium path is eternal cooperation. Invaders that simply always cooperate are therefore not detected until a second mutation arrives that always defects and thereby feeds on the first mutation lowering its guard.

For the finita automata case Binmore and Samuelson (1992) clearly show that no single machine can be an evolutionary stable strategy for the repeated prisoner's dilemma. Probst (1995) uses a setwise concept called an evolutionary stable collection (*ESC*) and shows that for the repeated prisoners dilemma that does not help either (proposition 4). To resolve the non-existence of an *ESS*, Binmore and Samuelson (1992) define a modified evolutionary stable strategy (*MESS*) and, among other things, show that such a strategy does exist.[6] In appendix A the exact definitions of a *MESS* and an *ESC* are given, along with examples why I think these two concepts are not the most appropriate or the most fruitful way to proceed once the notion of an *ESS* turns out not to be helpful.

Now that we failed to find one single strategy that is evolutionary stable, we could try for a mixture, but there we run into a second problem. This time it is a problem that only applies to the folk theorem case. Please recall that this type of theorem is meant to prove, with different assumptions, that there can be a whole set of expected payoffs that can be supported by equilibria, if only the players manage to coordinate on them. Typically, there are two ways to reach such an averaged or discounted expected payoff. One is to use a public randomization device, which, roughly speaking, is an external party that at every stage flips a coin in order to determine what will be the next step on the equilibrium path that the players are to adhere to. A strategy then prescribes actions to go both with the outcome of the public randomization and with whether or not the other player played according to the equilibrium path. The other possibility is to have a prespecified sequence of actions that amounts to a certain averaged or discounted payoff. Now for both methods, strategies are identified as equilibria by evaluating what happens if players make one step deviations from the equilibrium strategy, while it is assumed that players do coordinate because they can correctly anticipate their opponents strategy. However, if we want to have a mixture of strategies as a candidate for an *ESS*, we would also have to evaluate what happens if a player, thinking to coordinate on an outcome a, meets another player that assumes that he or she is coordinating on an outcome b that differs from a. In the public randomization case, there is no consistent definition for the payoffs that come with such an encounter, since the randomization to attain an expected payoff if there is coordination is not even unique. The same goes for the other case; in general there will be different sequences that lead to one and the same discounted or averaged payoff, so if there is miscoordination in terms of the outcome to be supported, the payoffs that those two players get when they play against each other could basically still be anything. Mixtures of folk theorem type strategies therefore do not have well defined payoffs.[7]

Perhaps one can formulate the problem even a little bit stronger. An evolutionary approach more or less comes down to theorizing about what happens if you would have a population in which there are a few different strategies present, or a population that faces mutations every once in a while. In such a setting, it is not very natural to make strong assumptions concerning the players' ability to anticipate. Yet the folk theorems rely heavily on the assumption that the players correctly anticipate the others strategy, which makes folk theorem strategies not compatible with

[6]They also remark that a weak point of a *MESS* is that it can still be vulnerable against mutants opening doors for other mutants.

[7]With finite automata this problem does not arise; on the contrary, it is very clear how they do against each other when they meet. Yet Binmore and Samuelson (1992) do not stick to the standard definition of a mixed strategy equilibrium, nor does Probst (1995) use the standard setwise generalization of evolutionary stability. Again, for details I refer to appendix A.

an evolutionary approach.

But suppose that payoffs in mixtures can be defined, then mixed equilibria may still not help out. In that case, it might be possible to fall back on a setwise generalization of an *ESS*. Thomas (1985) has two equivalent definitions, of which we will state the shorter one:

Definition 2 *A set X of strategies is an evolutionary stable (ES) set if it is nonempty and closed and each $x \in X$ has some neighbourhood U such that $u[x,y] \geq u[y,y]$ for all $y \in U$, and $u[x,y] > u[y,y]$ for all $y \in U, y \notin X$*

The existence of such a set is not guaranteed, but if there is one, then this set is also asymptotically stable in the replicator dynamics. (See also Weibull (1995)). Unlike *MESS* or *NSS*[8], this setwise concept is not vulnerable to at first harmless mutants opening doors for other mutants. (To see why this is indeed the case, assume that a given strategy is an element of such an ES set. This implies that all strategies with any share of initially inconsequential mutants would have to be elements of this ES set as well. But then there can not be a second mutant type that can enter the population for large enough a share of the first mutant, because that would be a violation of the condition of an ES set, if we take x to be the strategy with this large enough share of the first mutant.)

A problem for the finite automata case - not the biggest, but still not one to ignore - is that there it is more or less standard to have lexicographic preferences. Abreu & Rubinstein (1988) chose to have preferences that first evaluate game-payoffs and only if these are the same, look at the complexity of the strategies. From Binmore and Samuelson (1992) it is clear that with this type of preference one can still sensibly define an *ESS*, but it should be noted that then the link between ESS and stability in the replicator dynamics is lost.

In the next section we will consider repeated *stochastic* games, which can help eliminate at least the second problem mentioned; as we will see there, one can easily define strategies that need no randomization device nor coordination on a sequence of actions. Thereby it overcomes what I think is the most serious problem and it paves the way to apply the concepts of an *ESS* or an ES set.

I would like to make two remarks before we go to the core of the paper. The first is that even if a static evolutionary approach is facilitated, its outcome may, for the first reason mentioned, still be unsatisfactory. If it is, however, I once more would like to mention that these things are only shortcuts: they are useful if they work, but if they don't, that does not mean that there is no other way to get where we want to go. After all, along with evolutionary stability, the population dynamics are also defined as soon as payoffs of strategies playing each other are known.[9] The second remark is that I would like to emphasize that this paper is not a report on a complete evolutionary analysis of all such games; it is exploratory in nature and only contains a first few steps in looking at some interesting equilibria.

[8] Neutrally stable strategy; the definition can be found in appendix A.

[9] Allowing for an evolutionary analysis of course in itself is not a reason to introduce stochasticity. If the repeated prisoners dilemma is the game we play, then that is what we should analyse. But I think that if a setting allows for (equilibrium) strategies that we can evaluate when playing against themselves and against other strategies without randomization or assumptions about coordination, then this is a feature that is extra realistic in the sense that a (small) difference between people (read: their strategies) does not neccessarily mean that they do not play each other at all or that any cooperation between them immediately breaks down.

5 Repeated stochastic games

In this paper, a repeated stochastic game will be defined as follows. Assume that there is a collection G of (stage) games. These games are alike in the sense that both the set of players I and the action spaces $A_i,\ i \in I$, are the same for all games g in this collection G, so they only differ in their payoffs. Assume furthermore that there is a probability measure μ over this G. Now at every stage, first a game g is drawn and once the payoff function π^g of the game g is known, the players decide which action they play. When playing they therefore know what game they play now, but not what games they will play in the future. Furthermore we assume that there is a constant probability of termination $(1-\delta)$, which can also be seen as a discountfactor δ, but we should be careful not to think of this as representing the time-preferences of the players. The repeated stochastic game with probability of termination $(1-\delta)$ will be called $\Gamma(\delta)$.[10]

We will restrict ourselves to the situation where $\Gamma(\delta)$ is a symmetric 2 player game. In order to properly define symmetry in 2 player games, we will use a function $\pi^{\leftarrow}$ that is nothing but the payoff function π where the players have swopped their roles. Formally, for a 2 player game with $A_1 = A_2 = A$ and a given π, we define $\pi^{\leftarrow}$ as follows: $\pi_1^{\leftarrow}(a_1, a_2) = \pi_2(a_2, a_1)$ and $\pi_2^{\leftarrow}(a_1, a_2) = \pi_1(a_2, a_1)\ \forall\ a_1, a_2 \in A$. The game $g^{\leftarrow}$ is the corresponding game with the roles being reversed. Symmetry then will be defined as follows:

Definition 3 *A 2 player repeated stochastic game $\Gamma(\delta)$ is symmetric if the action sets of the two players coincide and if the following is true for the probability measure μ:*

$$\text{If} \quad g \in E \quad \Leftrightarrow \quad g^{\leftarrow} \in F \quad \text{holds for } E, F \subset G, \text{ then } \mu(E) = \mu(F).$$

Please note that this does not mean that every stage game has to be symmetric, it only guarantees that the odds are equally distributed.

5.1 An example

First we will focus on a very simple example of such a game. In fact, we will look at the most simple stochastic extention of the deterministic repeated prisoners dilemma that I can think of. For this case, we take G to be the set of symmetric prisoners dilemma's with $\pi_1(C, D) = 0$, $\pi_1(D, C) = 2$ and $\pi_1(C, C) + \pi_1(D, D) = 2$, where C denotes the action *cooperate* and D the action *defect*. Furthermore we let X be uniformly distributed on $[0, 1]$ and define the payoffs as follows:

$$\begin{aligned}
S &= \pi_1(C, D) = 0 \\
P &= \pi_1(D, D) = 1 - X \\
R &= \pi_1(C, C) = 1 + X \\
T &= \pi_1(D, C) = 2.
\end{aligned}$$

This gives us a well-defined distribution over G. Let the history up to time t be represented by $h_t = ((g_1, a_{11}, a_{21}), ..., (g_{t-1}, a_{t-1,1}, a_{t-1,2}))$ where g_n represents the game drawn at stage n and

[10] Our type of game is in fact a special case of what Mertens, Sorin and Zamir (1994) call a stochastic game. In this broader class of games, the distribution of stage games in subsequent stages can depend on the actions that the players chose and on the game that has been played (formally: the distribution of states at $t + 1$ depends on the state at time t and the actions at time t). Our repeated stochastic game can be seen as a degenerate case where the distribution of stage games is the same every period. This however is not a special case that seems to have drawn attention; they focus more on games that have interesting transition probabilities. In the games they focus on, it is the influence that present actions have on future games that matter, rather than the influence it might have on future behaviour of other players, which is what we are particularly interested in.

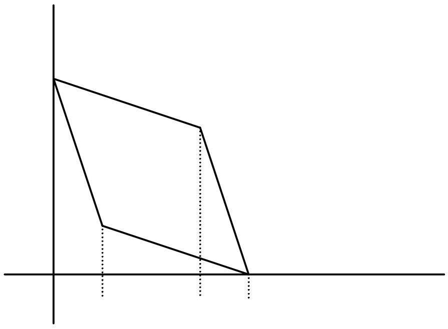


Figure 1:

$a_{n,1}$ and $a_{n,2}$ the actions that were played at the same stage. A strategy S assigns an action to every combination of a history h_t and game g_t; $S : H \times G \to A$. Given a history h_t, the set of games G is therefore divided up into subsets in which a player will play a certain action. In our example, this amounts to the following: for a given history, a strategy gives a subset of $[0, 1]$ such that if a game is drawn that corresponds with an X in this subset, then the player cooperates, and if not, then the player defects.

To be able to properly contrast some of the equilibrium strategies this game allows for with the equilibrium strategies from folk theorems, we will first have to do a little refrasing in Friedman's "Nash-threats" folk theorem (1971) in order to produce a stochastic equivalent. Therefore we also need to adapt the public randomization to the new, broader setting. We will consider two options how to publicly randomise in order to reach a given expected payoff and a general description that embraces both of them. The first is simply to condition actions on the draw by nature from G, the second possibility is having a seperate randomization device. Just to get an idea of how that works, we can look at the example at hand.

For the first way of randomizing, consider a match of strategies that leads to the following actions being played:

$$\begin{array}{ll} (D,D) & \text{if } X \in [0, z_1) \\ (D,C) & \text{if } X \in [z_1, z_2) \\ (C,D) & \text{if } X \in [z_2, z_3) \\ (C,C) & \text{if } X \in [z_3, 1] \end{array}$$

of course with $0 \leq z_1 \leq z_2 \leq z_3 \leq 1$. Then play is conditioned on the game being drawn. Solving

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = z_1 \begin{pmatrix} 1 - \frac{1}{2}z_1 \\ 1 - \frac{1}{2}z_1 \end{pmatrix} + (z_2 - z_1) \begin{pmatrix} 2 \\ 0 \end{pmatrix} + (z_3 - z_2) \begin{pmatrix} 0 \\ 2 \end{pmatrix} + (1 - z_3) \begin{pmatrix} 1\frac{1}{2} + \frac{1}{2}z_3 \\ 1\frac{1}{2} + \frac{1}{2}z_3 \end{pmatrix}.$$

with $0 \leq z_1 \leq z_2 \leq z_3 \leq 1$ gives us a range of vectors (z_1, z_2, z_3) for which equilibrium play along these lines leads to expected payoffs e.

The second possibility is to draw an $\omega \in \Omega$, where Ω is a set of public signals, independent from the draw of the stage game g from G. This brings us more or less back in the standard public randomizing situation, where ω determines which actions will be played. In our case - and in any other case where the games in G are identical up to a monotone transformation of the payoffs - both ways of randomizing lead to the same set of feasible payoffs, but in general that need not be the case. An expected payoff that can be reached with separate randomizing can always be constructed with conditioning on the game g as well, but the converse is not true. We can overcome this by using a more general definition, where equilibrium play is characterized by a function $a : G \times \Omega \to \prod_{i \in I} A_i$. Please note that both the game and the public signal are arguments of this function. Let $\mathcal{A}$ be the set of all such functions. The set of feasible payoffs then becomes $\left\{v \in \mathbb{R}^{|I|} \mid \exists\ a \in \mathcal{A} \text{ such that } \mathbb{E}\left[\pi^g\left(a\left(g, \omega\right)\right)\right] = v\right\}$.[11] For most elements of this set there will be a variety of functions a that do the trick and dependence on g resp. ω only are extreme cases.

With the randomization sorted out, we can state the appropriate version of Friedman's Nash-threats folk theorem (1971):

Theorem 4 *Let $a^* \in \prod_{i \in I} A_i$ be an equilibrium of the stage game for all games $g \in G$ and let its expected payoffs be a vector e. Suppose that payoff differences in the set G are bounded, meaning that there is an upper bound M such that $\pi_i^g\left(\widehat{a}_i, a_{-i}\right) - \pi_i^g\left(a\right) < M$ for all $g \in G$, all $a \in \prod_{i \in I} A_i$, all $\widehat{a}_i \in A_i$ and all $i \in I$. Then for every feasible individually rational payoff v, that is, every feasible v with $v_i > e_i$, there is a δ^* such that for all $\delta \in (\delta^*, 1)$ there is a subgame perfect equilibrium of $\Gamma\left(\delta\right)$ with expected payoffs v.*

Proof. In order for the strategy to play according to $a \in \mathcal{A}$ with $\mathbb{E}\left[\pi^g\left(a\left(g, \omega\right)\right)\right] = v$ and play a^* after any deviation to be in subgame perfect equilibrium with itself, the following is sufficient:

$$\max_{\widehat{a}_i \in A_i} \pi_i^g\left(\widehat{a}_i, a_{-i}\left(\pi, \omega\right)\right) - \pi_i^g\left(a\left(\pi, \omega\right)\right) < \tfrac{\delta}{1-\delta}\left[v_i - e_i\right] \qquad \forall\ g \in G, \forall\ \omega \in \Omega \text{ and } \forall\ i \in I$$

Since payoff differences in G are bounded, one can always find a δ^* such that this indeed holds for all $\delta \in (\delta^*, 1)$. ■

In our example $\pi_i^g\left(\widehat{a}_i, a_{-i}\left(\pi, \omega\right)\right) - \pi_i\left(a\left(\pi, \omega\right)\right)$ is at most $1 - X$, which in turn never exceeds 1 and that reduces the inequality in the proof to $1 < \frac{\delta}{1-\delta}\left[v_i - e_i\right]$.

This version of Friedman's Nash threats folk theorem has given us a set of subgame perfect equilibria of the game. This is not surprising nor very interesting; the equilibrium strategies we found are much like those of its deterministic counterpart and it only takes more effort and notation to describe them. No gains so far from introducing a stochastic stage game. However, the stochastic version of the repeated prisoner's dilemma also allows for equilibria that are not merely a reformulation of the folk theorem type of strategies. Below we will introduce a few strategies that fall into this more interesting category and I hope to stage them as steps in an effort to approximate human behaviour more closely than the equilibrium strategies we found so far in either one of the two deterministic approaches. We will begin with two particular equilibrium strategies and then give a more general description.

[11] It may not be very elegant to have π appear as part of its own argument, but I hope the interpretation is clear: a game g with payoff function π^g is drawn, along with a public signal ω. The function $a\left(g, \omega\right)$ then prescribes actions and $\pi^g\left(a\left(g, \omega\right)\right)$ are the corresponding payoffs.

Strategy 1

Consider the following strategy for some $0 \leq \overline{x} \leq 1$:

Play C if both players played C so far and $X \in [\overline{x}, 1]$. Play D otherwise.

In appendix B there are some calculations to show that such a strategy is indeed in equilibrium with itself for $\delta > 0.5$. Please note that even though this is a trigger strategy, it really differs from the ones we described before. While the gun never goes off in the folk theorem strategies, it definately fires here; with probability 1 we get to the punishment fase in finite time.

A nice thing about this equilibrium is that is does not need randomization nor a prespecified sequence in order to coordinate on an outcome, while yet all the symmetric expected discounted payoffs can be supported by such an equilibrium, if only we chose the right $\overline{x}$. Moreover, if we have two different strategies of this type - that is, two such strategies but with a different $\overline{x}$ - then their expected payoffs when playing against each other are perfectly well defined. Another nice thing is that one can easily find a conditional altruistic preference relation (see Section 2) and even a conditional altruistic utility function, such that myopically maximizing every periods utility is equivalent to following this strategy. The strategy is indeed perfectly consistent with what happens if player 1 maximizes

$$u_1 (\ x_1, x_2 \mid h_t \ ) = \begin{cases} x_1 + \frac{1-\overline{x}}{1+\overline{x}} \ x_2 & \text{if } h_t = \left(g_1, (C,C)_1, ..., g_{t-1}, (C,C)_{t-1}\right) \text{ for any } g_1, ..., g_{t-1} \\ x_1 & \text{otherwise} \end{cases}$$

and player 2 maximizes the mirror image of this conditional altruistic utility function. This may not be hard to see, but since this observation is more or less crucial, I will go into it briefly.

For any history that does not consist of a sequence of mutual cooperation only, it is obvious that myopically maximizing the current stage utility leads to playing D, since that is a dominant strategy for every possible stage game. For a history of mutual cooperation on the other hand, we need to check for the two obvious possibilities; one where the opponent plays C and one where he or she plays D.

In case the opponent plays C, the options are either to play C as well, in which case the utility is $\pi_1(C,C) + \frac{1-\overline{x}}{1+\overline{x}} \cdot \pi_2(C,C) = (1+X) + \frac{1-\overline{x}}{1+\overline{x}} \cdot (1+X)$, or to play D, in which case the utility is $\pi_1(D,C) + \frac{1-\overline{x}}{1+\overline{x}} \cdot \pi_2(D,C) = 2 + \frac{1-\overline{x}}{1+\overline{x}} \cdot 0 = 2$. The first is larger than the second if $X > \overline{x}$, in which case playing cooperate yields a higher utility.

If the other would play D, a similar story holds; $\pi_1(C,D) + \frac{1-\overline{x}}{1+\overline{x}} \cdot \pi_2(C,D) = 0 + \frac{1-\overline{x}}{1+\overline{x}} \cdot 2$ is to be compared to $\pi_1(D,D) + \frac{1-\overline{x}}{1+\overline{x}} \cdot \pi_2(D,D) = (1-X) + \frac{1-\overline{x}}{1+\overline{x}} \cdot (1-X)$ and a, player maximizing one shot utility, will choose to cooperate if $X > \overline{x}$. With an altruism parameter $\frac{1-\overline{x}}{1+\overline{x}}$, playing C is therefore dominant if $X \in (\overline{x}, 1]$ and playing D is dominant if $[0, \overline{x})$.[12]

This little exposition shows that if players are not aware of the fact that they are engaged in a repeated game and if they both myopically maximize their history dependent altruistic utilities, then neither of them would do any better in the 'real' game by deviating from their myopic altruistic behaviour. That, I think, is the main attraction of such an equilibrium strategy and its equivalent in conditional altruistic preferences.

This equilibrium strategy also has its less attractive sides though and one of them is that, apart from not being evolutionary stable, it is also vulnerable to mutants opening doors for other mutants that then do strictly better than the incumbent strategy. Consider for instance the strategy that

[12] We could introduce a tie-breaking rule here, but the probability that the players will actually need it is 0.

only checks its own side of the history; it cooperates if it has cooperated itself so far and $X \in [\overline{x}, 1]$ and it defects otherwise. This mutation can at first be inconsequential - it does not alter the actions being played against the incumbent strategy or against itself - but it gives pure defectors an edge.[13]

Strategy 2

Now for $0 \leq \overline{x} \leq 1$ consider the strategy:

Play C if both players played C the last round and $X \in [\overline{x}, 1]$.
Play C if not both players played C in the last round and $X \in [\widetilde{x}, 1]$.
Play D otherwise,

If $\widetilde{x}$ is sufficiently much larger than $\overline{x}$, then this strategy is also in equilibrium with itself. Again appendix B contains the algebra needed to show that this is indeed an equilibrium strategy for well chosen combinations of δ, $\overline{x}$ and $\widetilde{x}$. This strategy also has an equivalent in terms of myopic history dependent utilities:

$$u_1\ (\ x_1, x_2 \mid h_t\ ) = \begin{cases} x_1 + \frac{1-\overline{x}}{1+\overline{x}}\ x_2 & \text{if } h_t = \left(..., g_{t-1}, (C,C)_{t-1}\right) \text{ for any } g_t \text{ and any history up to } t-2 \\ x_1 + \frac{1-\widetilde{x}}{1+\widetilde{x}}\ x_2 & \text{otherwise} \end{cases}$$

In contrast to the first strategy, this one is not relentless and although its form may still not be very subtle, this one does know forgiveness (or forgetfulness). Something else worth noting is that when this strategy plays against itself, we have an equilibrium in which almost all paths will show both periods of war and times of peace, which could be qualified as an interesting feature. A disadvantage - one that this strategy shares with strategy 1 - is that it is also vulnerable to mutants that only check up on their own past actions and thereby give way to pure defectors.

A general formulation of a set of equilibrium strategies for this example

To allow for more sophisticated equilibrium strategies, we can make a more general formulation of which strategies 1 and 2 are special cases. While strategies 1 and 2 both only have two cut-off levels ($\overline{x}$ and 1 for strategy 1 and $\overline{x}$ and $\widetilde{x}$ for strategy 2), we could basically want to allow for as many cooperation thresholds as one likes. To that end, take functions f_1 and f_2 as follows:

$$f_i : \{C, D\} \times [0,1] \times [0,1] \rightarrow [0,1]$$

where the variables indicate the other players action at time $t-1$, the cooperation threshold $\overline{x}_{i,t-1}$ of player i at time $t-1$, the realization of X at time $t-1$ and the cooperation threshold $\overline{x}_{i,t}$ of player i at time t. Players are understood to use these functions to update their willingness to cooperate (or their level of altruism) and this type of updating could be seen as Markovian since different histories leading to the same cooperation threshold $\overline{x}_{i,t}$ are not distinguished at time t. The action of player i at time t is then determined by his or her cooperation threshold $\overline{x}_{i,t}$ and the game being drawn; C if $X_t \in [\overline{x}_{i,t}, 1]$ and D if $X_t \in [0, \overline{x}_{i,t})$.

[13] Another, less negative example is a mutant strategy that plays this equilibrium strategy, with only one alteration: after a number of rounds of mutual defection (s)he would switch back to the cooperative phase as soon as it finds the other player play cooperate. Such a forgetting (or forgiving) mutation would not show since no other player would ever initiate the return to cooperation. But if enough such mutations accumulate, it paves the way for another mutation that does initiate the making up and therefore does better than the relentless strategy and increases the payoff of the first mutant

For given functions f_1 and f_2 we could compute the expected payoff and for a given function f we could also compute a set of best responses $BR(f)$ (not necessarily all fitting the given formulation). An equilibrium would then be a strategy f such that $f \in BR(f)$. Some good and bad aspects of strategies 1 and 2 carry over to all equilibria in this more general setting. A good thing is that all such equilibrium strategies again have a counterpart in the set of conditional altruistic utility functions; for player 1 that would be

$$u_1 (x_1, x_2 \mid \overline{x}_{1,t}) = x_1 + \tfrac{1-\overline{x}_{1,t}}{1+\overline{x}_{1,t}} x_2,$$

where the updating of $\overline{x}_{i,t}$ is done according to f_1, while player 2 myopically maximizes the mirror image of this conditional utility function. A bad thing is that equilibrium play is always symmetric, which makes updating on the behaviour of the other interchangeable with updating using ones own past actions only, which in turn makes it vulnerable to cheaters.

It should be noted that for this particular example, with its very particular set of stage games from which the stage game is drawn, it is close to trivial that an equivalent of a strategy in the form of a conditional altruistic preference relation can be found. With a larger set of possible stage games this is not neccesarily so; the set of strategies can be much richer than the set of strategies for which there is an equivalent in conditional altruistic preferences. However, for a rather broad class of stochastic repeated prisoners dilemmas, we can prove something interesting. In the next subsection we will look at equilibria of repeated stochastic games where players are restricted to strategies that do have such a conditional preference equivalent. For what I think is a reasonable subset of equilibria (that is strategies that satisfy *Independence of Hypothetical Situations*) we will show that such an equilibrium in the game with a restricted strategy space is also an equilibrium in the game without the restriction that a strategy should be the result of following a conditional altruistic preference. To get there, we will have to make an effort to get the notation and the definitions right. This may be a little tedious, but once it is done, the proof of the result becomes relatively easy.

5.2 Conditional altruism in equilibrium

In the definition of a history, we will write π_t instead of π^{g_t} to denote the payoff function of the stage game drawn at time t. Since players and possible actions are assumed to be all the same, the payoff function is the only variable that actually allows variation and therefore the payoff function completely determines the stage game. If we agree to let an empty pair of brackets denote the history 'no history', then the following definition makes sense:

$$h_1 = ()$$
$$h_t = \left((g_1, a_{1,1}, a_{1,2}), ..., (g_{t-1}, a_{t-1,1}, a_{t-1,2})\right), \qquad t = 2, 3, ...$$

Sometimes we will write $(h_t, (g_t, a_{t,1}, a_{t,2}))$ for a history h_{t+1}. Furthermore we will need a way to write down a history with the roles of the players reversed. Given a history h_t, its mirror image $h_t^{\leftarrow}$ is found by simply renumbering the players:

$$h_t^{\leftarrow} = \left((g_1^{\leftarrow}, a_{1,2}, a_{1,1}), ..., (g_{t-1}^{\leftarrow}, a_{t-1,2}, a_{t-1,1})\right).$$

The set of possible histories at time t is denoted by H_t and the set of all possible histories is $H = \bigcup\limits_{t=1}^{\infty} H_t$.

In evaluating the past, a player can conceive different histories as equivalent. The idea behind this equivalence is that later on we will assume that players react the same to those equivalent histories, because the choice behaviour of their opponents in those different histories is the same. For two histories to be equivalent to a player, two things need to hold. First, the two sequences of actions need to coincide. The second requirement is, loosely formulated, that his or her opponent would have faced the same options to choose from at every stage in either of the two histories. Formally, we will think of histories h_t and h'_t as *equivalent to player 1* if

$$\begin{array}{lll} a_{\tau,i} = a'_{\tau,i} & \text{for } \tau = 1,...,t-1 \text{ and } i = 1,2 & \text{and} \\ \pi_\tau\left(a_{\tau,1}, \widehat{a}\right) = \pi'_\tau\left(a_{\tau,1}, \widehat{a}\right) & \text{for } \tau = 1,...,t-1 \text{ and for all } \widehat{a} \in A, & \end{array}$$

where A is the action space. Similarly, h_t and h'_t are *equivalent to player 2* if

$$\begin{array}{lll} a_{\tau,i} = a'_{\tau,i} & \text{for } \tau = 1,...,t-1 \text{ and } i = 1,2 & \text{and} \\ \pi_\tau\left(\widehat{a}, a_{\tau,2}\right) = \pi'_\tau\left(\widehat{a}, a_{\tau,2}\right) & \text{for } \tau = 1,...,t-1 \text{ and for all } \widehat{a} \in A. & \end{array}$$

In the upcoming theorem, we will go back and forth between conditional preferences and strategies. A conditional preference of player i, as mentioned in Section 2, is a function $CP : H \to \mathcal{R}$ where $\mathcal{R}$ is the set of all preference relations over $\mathbb{R}^2$. A strategy is a fuction $S : H \times G \to A$, where G is the set of games. The set of all conditional preferences is denoted by $\mathcal{CP}$ and the set of all strategies is $\mathcal{S}$.

Now both these functions are defined from the perspective of the player who holds them. That is to say, if players with strategies S_a and S_b meet, they both call themselves number 1 and the other number 2. Choosing one of the two ways of numbering the players - for instance, if we choose to say that the player with strategy S_a is number 1 and that the other is number 2 - then in order to get the proper result of these two strategies meeting, we will have to evaluate $S_a\left(h, g\right)$ and $S_b\left(h^{\leftarrow}, g^{\leftarrow}\right)$ for the histories h and the games g that occur. A player that accidently gets the label 'two' will therefore look at mirror images of histories and games that are represented in the chosen numbering. The same goes for conditional preferences; without detours they define preference relations on $\mathbb{R}^2$ of a player that got numbered 1, but if a player happens to get number 2, this function has to be evaluated in the point $h^{\leftarrow}$, where the resulting preference is over points in $\mathbb{R}^2$ where the original axes are interchanged.

To get from a conditional preference relation to a strategy, we need to know what a player expects the other to play. An expectation pattern of a player is a function $e : H \times G \to A$. Together with a conditional preference, an expectation pattern defines a choice set $\widetilde{C}$ for a history h and a game draw g in the following straightforward way:

$$\widetilde{C}\left(g, e\left(h, g\right), CP\left(h\right)\right) = \left\{\ a_1 \in A \mid \ \left(\pi^g\left(a_1, e\left(h, g\right)\right), \pi^g\left(a, e\left(h, g\right)\right)\right) \in CP\left(h\right)\ \ \forall\ a \in A\ \right\}.$$ [14]

This set can, in principle, contain more than one element. To get a well defined choice function C, we introduce a tie-breaking rule, that is a function $T : \mathcal{A} \to A$ such that $T\left(B\right) \in B$ where $\mathcal{A}$ is the set of all subsets of A.

$$C_T\left(g, e\left(h, g\right), CP\left(h\right)\right) = T\left(\widetilde{C}\left(g, e\left(h, g\right), CP\left(h\right)\right)\right).$$

[14] Alternatively, one can write:
$\widetilde{C}\left(g, e\left(h, g\right), CP\left(h\right)\right) = \left\{\ a_1 \in A \mid \pi^g\left(a_1, e\left(h, g\right)\right) \succeq_{CP(h)} \pi^g\left(a, e\left(h, g\right)\right)\ \forall\ a \in A\ \right\}$

For the interesting cases however, the choice set already is a singleton almost everywhere, that is, only for a set of stage games that occurs with probability zero the rule actually does any tie breaking. In those cases the choice of a tie-breaking rule is inconsequential in the sense that with probability one a history unfolds that would be unaltered by another choice of a tie-breaking rule. Consequentially, expected payoffs in those cases are independent of the tie-breaking rule.

The expectation pattern of a player is called *correct* if the other player always behaves as is expected and it is called *empathy consistent* if a player expects the other to behave as he or she would behave herself. Speaking in formulas, we say that the expectation pattern e_1 of a player is correct when it meets a player with expectation pattern e_2, conditional preference CP_2 and tie-breaking rule T_2 such that for all $h \in H$ and all $g \in G$

$$e_1(h,g) = C_{T_2}\left(g^{\leftarrow}, e_2\left(h^{\leftarrow}, g^{\leftarrow}\right), CP_2\left(h^{\leftarrow}\right)\right).$$

An expectation pattern e_1 is empathy consistent if it is combined with a conditional preference CP_1 and tie-breaking rule T_1 such that for all $h \in H$ and all $g \in G$

$$e_1(h,g) = C_{T_1}\left(g^{\leftarrow}, e_1\left(h^{\leftarrow}, g^{\leftarrow}\right), CP_1\left(h^{\leftarrow}\right)\right).$$

Introducing the final three definitions, one can imagine that strategies, expectation patterns and conditional preferences do not discern equivalent histories. The options the other player has, given what actions you played, are obviously relevant when evaluating his or her behaviour (or formally: to condition your reaction on). However, it seems a little unnatural to distinguish histories because one of them would have offered the other player a different set of payoffs to chose from in a situation that did not arise since I myself did not play the action required to give the opponent those options. Therefore we define *Independence of Hypothetical Situations (IHS)* for conditional preferences, expectation patterns and for strategies.

Definition 5 *A conditional preference CP satisfies Independence of Hypothetical Situations if $CP(h_t) = CP(h'_t)$ for all t and for all h_t and h'_t that are equivalent to player 1.*

Definition 6 *An expectation pattern e satisfies Independence of Hypothetical Situations if $e(h_t, g) = e(h'_t, g)$ for all games g in G, for all t and for all h_t and h'_t that are equivalent to player 1.*

Definition 7 *A strategy S satisfies Independence of Hypothetical Situations if $S(h_t, g) = S(h'_t, g)$ for all games g in G, for all t and for all h_t and h'_t that are equivalent to player 1.*

Before we can state and prove the main result (theorem 8), there is one conceptual issue that we will have to settle. In the theorem there will be two games, a stochastic repeated game $\Gamma(\delta)$ as we have defined in the beginning of this section, and a game $\Gamma^*(\delta)$ that only differs in the size of the strategy space. In the latter game, only those strategies are permitted that can be the result of having a certain conditional preference in combination with an expectation pattern and a tie-breaking rule. Formally: a strategy S is an element of the strategy space of $\Gamma^*(\delta)$ if and only if there is a conditional preference CP, an expectation pattern e and a tie-breaking rule T such that

$$S(h,g) = C(g, e(h,g), CP(h)) = T\left(\widetilde{C}(g, e(h,g), CP(h))\right) \qquad \text{for all } h \in H \text{ and all } g \in G.$$

It is quite crucial to note that here the expectation pattern is a fixed thing; in this setting it is not possible, given the same history, to expect one action from one player and something else from another. It is obvious that then we can not demand these expectations to be correct for any two strategies meeting one another. We could have made a different choice, but I hope to convince the reader that the setting we chose might lead to results that are more significant.

The obvious other choice would be to assume common knowledge at every stage about each others conditional preferences and then restrict the possible expectations that players can have to those that are correct. This could be seen as a model in the line of Bester and Güth (1998), where players are assumed to carry true signals of their altruism, while our model more or less meets Binmore's (1994) criticism. The main reason why I did not chose this other model with true signals, is that it would not really be much of an explanation. Consider for instance a strategy that always cooperates with itself and always defects on all other strategies. This is obviously an equilibrium in such a setting - it is even evolutionary stable - but I do not think there is a reason to exclude a strategy that, carrying the signal of this first strategy, cheats and always plays defect. Furthermore it would by no means be a generalization of what we have for the deterministic repeated game; neither the folk theorem setting nor the finite automata approach allows for (conditioning on) such signals.

Now that the stage is set to state the theorem, we will go through the assumptions. The game $\Gamma(\delta)$ is specified by a set of stage games G and the measure μ that determines the distribution from which these stage games are drawn. Our first assumption is that all stage games are prisoners dilemmas for which each of the four payoff vectors lie in a different quadrant. Formally, that is:

- **Assumption 1**

  For all $g \in G$ the following holds:

  g is a prisoners dilemma,

  $\pi_1(C,C), \pi_1(D,C), \pi_2(C,C), \pi_2(C,D) > 0$ and

  $\pi_1(D,D), \pi_1(C,D), \pi_2(D,D), \pi_2(D,C) < 0$.

An alternative assumption under which the theorem would still hold is that every game $g \in G$ must be a prisoners dilemma for which $\pi_1(D,D) = \pi_2(D,D) = 0$.

The second assumption is that the game is symmetric (see definition 3). Since we already assumed that the stage games are prisoners dilemma's, this reduces to the symmetry of μ.

- **Assumption 2**

  $\Gamma(\delta)$ is symmetric.

Finally we have a second restriction on the possible measures μ and this one modestly demands that the expectations of the payoffs exist

- **Assumption 3**

  Both $\mathbb{E}[\ \pi_1(D,C)\ ]$ and $\mathbb{E}[\ \pi_1(C,D)\ ]$ exist.

This assumption makes sure that we can, for $\delta \in (0,1)$ and for any history h_t, compute the expected payoff to a player using strategy S_1 against a player with strategy S_2:

$$\begin{aligned}\mathbb{E}_\delta[\ S_1, S_2 \mid h_t\ ] = {} & \mathbb{E}[\ \pi_t(S_1(h_t, g_t), S_2(h_t^{\leftarrow}, g_t^{\leftarrow}))\ ] + \\ & \delta\ \mathbb{E}[\ \mathbb{E}_\delta[\ S_1, S_2 \mid (h_t, (g_t, S_1(h_t, g_t), S_2(h_t^{\leftarrow}, g_t^{\leftarrow})))\ ]\ ]\end{aligned}$$

In the last expression $(h_t, (g_t, S_1(h_t, \pi_t), S_2(h_t^{\leftarrow}, \pi_t^{\leftarrow})))$ is a history h_{t+1} and since g_t (and therefore h_{t+1}) is stochastic, we have to take an expectation over $\mathbb{E}_\delta[\, S_1, S_2 \mid h_{t+1}\,]$. When we feed this recursive definition back into itself, we get a discounted infinite sequence of stochasts and each of them has an expectation with an upper limit of $\mathbb{E}[\,\pi_1(D, C)\,]$. This makes $\mathbb{E}_\delta[\, S_1, S_2 \mid h_t\,]$ exist for all h_t.

Theorem 8 *Let $\Gamma(\delta)$ satisfy assumptions 1, 2 and 3. Then a symmetric equilibrium strategy S of $\Gamma^*(\delta)$ for which the conditional preference CP satisfies IHS, is also an equilibrium of $\Gamma(\delta)$.*

Proof. Assume S is an equilibrium of $\Gamma^*(\delta)$ but not of $\Gamma(\delta)$. Then there is a strategy S' that does better against S than S itself. The idea is that we will turn that strategy S' into a strategy S'' that does at least as good as S' against S, but fits in the strategy space that $\Gamma^*(\delta)$ allows for, that is: it comes with a conditional preference CP and an expectation pattern e such that $S''(h_t, g) = C_T(\pi^g, e(h_t, g), CP(h_t))$ for any tie-breaking rule T. This would contradict that S is an equilibrium of $\Gamma^*(\delta)$.

The first step is to group histories in equivalence classes. The equivalence class of a history h_t is defined by:

$$E(h_t) = \{\, h_t' \mid h_t' \text{ and } h_t \text{ are equivalent to player 2}\,\}.$$

For all histories h_t and games g we define the expectation pattern by $e(h_t, g) = S(h_t^{\leftarrow}, g^{\leftarrow})$. Please note that the fact that S is *IHS* makes $S(h_t'^{\leftarrow}, g^{\leftarrow})$ as a function of g equal for all $h_t' \in E(h_t)$. Then, for a given history h_t and a given strategy S we can also define equivalence classes of games:

$$E_{S,h_t}(g) = \{\, g' \mid e(h_t, g') = e(h_t, g) \text{ and } \pi^{g'}(a, e(h_t, g)) = \pi^g(a, e(h_t, g))\ \forall\, a \in A\,\}.$$

Since $S(h_t'^{\leftarrow}, g^{\leftarrow})$ does not show variation as long as $h_t' \in E(h_t)$, we know that $E_{S,h_t'}(g) = E_{S,h_t}(g)$ for all $h_t' \in E(h_t)$.

Now for given strategies $S_1$and S_2, we define a function V_{S_1,S_2} that gives the continuation payoff of S_1 against S_2 for history h_t and game g:

$$V_{S_1,S_2}(h_t, g) = \pi^g(S_1(h_t, g), S_2(h_t^{\leftarrow}, g^{\leftarrow})) + \delta\mathbb{E}_\delta[\, S_1, S_2 \mid (h_t, (g, S_1(h_t, g), S_2(h_t^{\leftarrow}, g^{\leftarrow})))\,].$$

Then we fix a tie breaking rule T and define

$$S''(h_t, g) = T(B_{S',S}(h_t, g)) \text{ with}$$
$$B_{S',S}(h_t, g) = \left\{\, a \in A \mid a = S'\left(\widetilde{h}_t, \widetilde{g}\right) \text{ where } \left(\widetilde{h}_t, \widetilde{g}\right) \in \arg \max_{\left(h_t' \in E(h_t), g' \in E_{S,h_t}(g)\right)} V_{S',S}(h_t, g) \,\right\}.$$

Now that we have defined the strategy S'', we will have to show that it does at least as good as S' against S and that it comes with a conditional preference. The intuition behind the answer to the question why S'' cannot be beaten by S' in its performance against S is the following. Their common opponent S will, by definition of *IHS,* not distinguish one history from another as long as they are elements of one and the same equivalence class. The definition of e implies that the same goes for different games that are elements of one equivalence class. Yet S' might still prescribe different actions for different histories or games in such sets. Therefore, if there is any difference between the value of $V_{S',S}(h_t, g)$ for different histories in $E(h_t)$ or for different games in $E_{S,h_t}(g)$, then this is caused by the different action that S' takes. In that case, one might just as well choose the best continuation out what different suggestions S' apparently has for different elements of the same equivalence class, and that is exactly what S''does.

The formal argument uses induction, where the induction step uses a second induction argument. First we define a sequence of strategies $\{S''_n\}_{n=0,1,2,...}$ as follows:

$$S''_n(h_t, g) = \begin{cases} S''(h_t, g) & t \leq n \\ S'(h_t, g) & t > n \end{cases}$$

The induction starts with assuming that for a given n the following holds:

$$(1) \qquad \mathbb{E}_\delta[\, S''_n, S \mid h_t \,] \geq \mathbb{E}_\delta[\, S''_i, S \mid h_t \,] \qquad \text{for } i = 0, ..., n \text{ and for all } h_t \in H_t,\ t = 1, ..., n.$$

Now by construction, $V_{S''_{n+1},S}(h_{n+1}, g) \geq V_{S''_n,S}(h_{n+1}, g)$ for all $g \in G$ and $h_{n+1} \in H_{n+1}$.[15] If we let g be drawn according to μ and take expectations on both sides, we get:

$$(2) \qquad \mathbb{E}_\delta\left[\, S''_{n+1}, S \mid h_{n+1} \,\right] \geq \mathbb{E}_\delta[\, S''_n, S \mid h_{n+1} \,].$$

This, however, is not enough to prove our complete induction step; not only do we need this to hold for histories at time $t = n + 1$, but also for histories with time indexes between 1 and n. To get there, we employ a second induction: assume that for a given $j \in \{2, ..., n + 1\}$

$$(3) \qquad \mathbb{E}_\delta\left[\, S''_{n+1}, S \mid h_j \,\right] \geq \mathbb{E}_\delta[\, S''_n, S \mid h_j \,].$$

Writing down the definition of $V_{S''_{n+1},S}$ gives

$$V_{S''_{n+1},S}(h_{j-1}, g) = \\ \pi^g\left(S''_{n+1}(h_{j-1}, g), S\left(h^{\leftarrow}_{j-1}, g^{\leftarrow}\right)\right) + \delta\, \mathbb{E}_\delta\left[\, S''_{n+1}, S \mid \left(h_{j-1}, \left(g, S''_{n+1}(h_{j-1}, g), S\left(h^{\leftarrow}_{j-1}, g^{\leftarrow}\right)\right)\right) \,\right].$$

Since $S''_{n+1}(h_{j-1}, g) = S''_n(h_{j-1}, g)$ for given $j = 1, ..., n + 1$ we can use (3) to find that

$$V_{S''_{n+1},S}(h_{j-1}, g) \geq \\ \pi^g\left(S''_n(h_{j-1}, g), S\left(h^{\leftarrow}_{j-1}, g^{\leftarrow}\right)\right) + \delta\, \mathbb{E}_\delta\left[\, S''_n, S \mid \left(h_{j-1}, \left(g, S''_n(h_{j-1}, g), S\left(h^{\leftarrow}_{j-1}, g^{\leftarrow}\right)\right)\right) \,\right].$$

The right hand side however is nothing but $V_{S''_n,S}(h_{j-1}, g)$. If we let g again be drawn according to μ and take the expectation on both sides, we get

$$(4) \qquad \mathbb{E}_\delta\left[\, S''_{n+1}, S \mid h_{j-1} \,\right] \geq \mathbb{E}_\delta[\, S''_n, S \mid h_{j-1} \,].$$

Since (2) gave us a start for $j = n + 1$, this second induction indeed establishes that

$$(5) \qquad \mathbb{E}_\delta\left[\, S''_{n+1}, S \mid h_t \,\right] \geq \mathbb{E}_\delta[\, S''_n, S \mid h_t \,] \qquad \text{for all } h_t \in H_t,\ t = 1, ..., n + 1.$$

We now combine the result of our second induction with (1), the assumption of the first, to conclude that $\mathbb{E}_\delta\left[\, S''_{n+1}, S \mid h_t \,\right] \geq \mathbb{E}_\delta[\, S''_n, S \mid h_t \,] \geq \mathbb{E}_\delta[\, S''_i, S \mid h_t \,]$ for $i = 0, ..., n$ and for all $h_t \in H_t, t =$

[15] In fact $V_{S''_{n+1},S}(h_{n+1}, \pi) \geq V_{S''_i,S}(h_{n+1}, \pi)$ for all $\pi \in \mathcal{G}$ and $h_{n+1} \in H_{n+1}$ holds for all $i = 0, ..., n + 1$. We will only use this for $i = n$.

$1, ..., n$. Since $\mathbb{E}_\delta\left[\ S''_n, S \mid h_{n+1}\ \right]$ obviously equals $\mathbb{E}_\delta\left[\ S''_i, S \mid h_{n+1}\ \right]$ for $i = 1, ..., n$, we can do the same for histories with index $n+1$; $\mathbb{E}_\delta\left[\ S''_{n+1}, S \mid h_{n+1}\ \right] \geq \mathbb{E}_\delta\left[\ S''_n, S \mid h_{n+1}\ \right] = \mathbb{E}_\delta\left[\ S''_i, S \mid h_{n+1}\ \right]$ for $i = 0, ..., n$. For $i = n+1$ the whole thing is tautological, and therefore we can summarize this as:

$$(6) \qquad \mathbb{E}_\delta\left[\ S''_{n+1}, S \mid h_t\ \right] \geq \mathbb{E}_\delta\left[\ S''_i, S \mid h_t\ \right] \quad \text{for } i = 0, ..., n+1 \text{ and for all } h_t \in H_t, t = 1, ..., n+1.$$

But then we have actually proved the whole induction step of the first induction. The induction can be initialized by simply observing that (1) obviously holds for $n = 1$. By induction, it now holds for all $n \geq 0$.

While $S''_0(h_t, \pi) = S'(h_t, \pi)$ and $\lim_{n\to\infty} S''_n = S''$, we have proved that $\mathbb{E}_\delta\left[\ S'', S \mid h_1\ \right] \geq \mathbb{E}_\delta\left[\ S', S \mid h_1\ \right]$ which by assumption is strictly larger than $\mathbb{E}_\delta\left[S, S \mid h_1\right]$.

The only thing that remains to be proved, is that S'' indeed comes with a straightforward conditional preference. To that end, define $CP(h_t)$ on all relevant pairs $(x_1, x_2), (y_1, y_2)$, that is pairs for which there is a game $g \in G$ such that $(x_1, x_2) = \pi^g(C, e(h_t, g))$ and $(y_1, y_2) = \pi^g(D, e(h_t, g))$ as follows:

$(x_1, x_2) \succeq_{CP} (y_1, y_2)$ if there is a game g such that
$S'(h_t, g) = C$ and $V_{S',S}(h_t, g) \geq V_{S',S}(h'_t, g')$ for all $h'_t \in E(h_t)$ and all $g' \in E_{S,h_t}(g)$.

$(x_1, x_2) \preceq_{CP} (y_1, y_2)$ if there is a game g such that
$S'(h_t, \pi) = D$ and $V_{S',S}(h_t, g) \geq V_{S',S}(h'_t, g')$ for all $h'_t \in E(h_t)$ and all $g' \in E_{S,h_t}(g)$.

Assumption 1 makes sure that (y_1, y_2) is in quadrant IV if (x_1, x_2) is in I and in III if (x_1, x_2) is in II. This preference is well defined, since for given history h and given strategy S all games that lead to the same pair $(x_1, x_2), (y_1, y_2)$ fall within the same equivalence class. Furthermore CP, together with e and T as they are defined resp. chosen above, make up strategy S'', that is,

$$S''(h, g) = C(g, e(h, g), CP(h)) = T\left(\widetilde{C}(g, e(h, g), CP(h))\right) \text{ for all } h \in H \text{ and all } g \in G.$$

Strategy S'' is therefore an element of the action space of the game $\Gamma^*(\delta)$. Since we have established that $\mathbb{E}_\delta\left[\ S'', S \mid h_1\ \right] \geq \mathbb{E}_\delta\left[\ S', S \mid h_1\ \right]$ which is by definition larger than $\mathbb{E}_\delta\left[\ S, S \mid h_1\ \right]$ for some S' if S is not an equilibrium of $\Gamma(\delta)$, this contradicts our assumption that S is an equilibrium of $\Gamma^*(\delta)$ but not of $\Gamma(\delta)$. S must therefore also be an equilibrium of $\Gamma(\delta)$ if it is an equilibrium of $\Gamma^*(\delta)$. *Q.E.D.* ■

5.3 Interpretation

Now what exactly have we achieved with this theorem? Summarizing, one could say that for a repeated stochastic game, we have chosen first to restrict strategies to those that fit a given structure, where degrees of altruism or willingness to cooperate are updated depending on the behaviour of the opponent. The theorem states that if we, for whatever reason, find ourselves in an equilibrium that furthermore satisfies *Independence of Hypothetical Situations*, then lifting the restriction on the strategy set will not upset the equilibrium. A scenario in which this knowledge could be useful is one where there is reason to assume that dynamics in a population of strategies that do fit the restriction gets us to an equilibrium before mutants with an altogether different structure can get a foothold. That would be the case if we for instance assume that it is much more likely for a person to beget offspring that has a conditional preference different from his or her own, than it is to have descendants whose behaviour does not follow any such structure. This is not at all a wild assumption; sexual reproduction implies that it does not even take mutations to get children that are not exact copies of oneself. After all, they are only drawn from the distribution over the possible combination of both parents genes. Alternatively, one could say that the probability of succes with a mutation to a different conditional preference is much higher than the odds to hit a winning strategy if a mutation loses all conditional preference structure and must be chosen from the unstructured remainder of the strategy set, that will mainly consists of strategies that do much worse than a strategy with a closeby conditional preference. In both these settings, the theorem suggests that the set of strategies with a conditional preference equivalent is rich enough to contain all the relevant strategies; if we get to an equilibrium with people that behave according to a conditional preference that satisfies *IHS*, then the remainder of the strategy set does not contain strategies that can keep it from being in equilibrium.

One thing to note though is that this does not explain why these preferences would satisfy revealed preference axioms and come with continuous, convex and monotonic utility functions, as Andreoni & Miller (2002) find. It would take a more elaborate evolutionary setup with assumptions on mutation probabilities to see if these findings can be reconciled with evolution of strategies.

6 Trying to find the right game

The typical task of a game theorist could be seen like this: he or she is given a game and with the toolbox of equilibria, equilibrium refinements and their motivations, the game theorist is expected to predict the behaviour of the players. But with the same ingredients, one could also make another puzzle. For given behaviour and using the same toolbox, one could also try to find the game or type of game that is being played. Now I tend to think that a study of behaviour ought to be a bit of both, going back and forth in these opposite directions. An explanation then consists of three things: a game, behaviour and a link between the two. The quality of the model is then to be judged on whether the game is convincingly realistic, the behaviour is anything like we observe and the link between the two fits a consistent idea or equilibrium concept.

The deterministic repeated prisoners dilemma is pretty much the standard for explaining cooperative behaviour. This is definitely not without reason since it does give us very good insights, but the temptation is strong to mistake conclusions from the model for conclusions about reality. In this paper, I have presented a stochastic version of the same game and I hope that I have convincingly argued that it does better than the deterministic one on these three points. Yet I would like to emphasize that I do not want to say that because it improves on the standard, it should also replace it. In this section I would like to mention a few games or ideas that one first of all could see as competing models. Another possibility however is that one could also think of some of them as games that are being played simultaneously with overlapping action sets. To rephrase that, one can imagine that the 'real' game, the payoffs of which drive the selection process, is a composition of a few of those different games and a parameter for fairness or altruism may be a simultaneous move on a few boards at a time.

In or out of equilibrium?

The first thing to do is to get back to Binmore's (1994) opinion on this. His idea is that the genuine preferences for fairness or altruism that are there towards our non-relatives are simply out of equilibrium. Although experiments do not seem to support this particular type of being out of equilibrium, I do agree that we can not always be sure that we are in equilibrium if indeed the games and the stakes have been changing fast relative to the generations timescale of the selection process.

Signaling arms race

The game of Bester and Guth (1998), where altruism can be recognised on forehand, may indeed be vulnerable to mutant phoneys, but one could image that such a game becomes the subject of an arms race, where altruists develop better signals, egoists make better fakes and all players get more sensitive signal testing.

Judgement from third parties

As mentioned in Section 2, we have imposed a rather severe restriction, and that is that we have defined fairness as nothing more than reciprocal altruism, whereas in real life punishment and reward may also come from others than those we hurt or help. Binmore (1994, 1998) actually even argues that our system of fairness considerations is nothing but a coordinating device, accomodating a system of everybody guarding everybody to behave according to the equilibrium coordinated upon. Now I do think that this paper (as well as for instance Miller's (2001) book) shows that altruism and fairness can, to a certain extend, be explained without stripping it of its candour. Yet I do agree that including judgements from players that are not directly involved would very much enhance the realistic content of the model, but it might very well complicate the analysis even more. Nonetheless it is important not to overlook how restricting our definition of fairness is.

Prisoners dilemma or gift exchange?

One could argue that as a model, a stochastic gift exchange is no less realistic than a repeated stochastic prisoners dilemma. The situation that Trivers (1971) describes in order to introduce reciprocal altruism is in fact a (deterministic) gift exchange and not a prisoners dilemma. The idea of a stochastic gift exchange would be that today I may have the opportunity to help someone out big time at low cost, while tomorrow I might be able to do the same at much bigger expenses and the day after tomorrow the roles could be reversed. To show that it does make a difference, Appendix C looks at an example of such a game. This one could be seen as the equivalent of our simple example of a repeated stochastic prisoner's dilemma, but then with moves that are sequential instead of simultaneous. We look at whether strategies of type 1 from the one example can also be an equilibrium strategies of the second. Unfortunately - but not surprisingly - it turns out that for most discount factors only one of the extremes in the set of cooperation thresholds (or altruism parameters) correspond with an equilibrium in the stochastic gift exchange game.

Natural or sexual selection?

Geoffrey Miller (2001) quite forcefully argued that there has been too much of a focus on natural selection, whereas sexual selection can answer questions that are destined to remains unanswered if we would stick to natural selection only. I will not try to repeat all of the books content, but it is indeed a line up of puzzles that go with sexual selection wonderfully well. I will only pick up on one point from the book that I think is important in our context.

However much we refine our natural selection game, we will always have a hard time explaining why we tip in restaurants that we will never return to, or why we help people that will pretty sure never be able to repay us. His explanation is that there is a benefit that remains hidden if we look at natural selection only and that is that being selfish and unfair is extremely unattractive. My favourite illustration from the chapter on morality is Scrooge. By being stingy he indeed did very well materially and he will definitely not die of starvation, but his genes nonetheless will not make it to the next generation. After all, who wants to date a miser. If we therefore observe behaviour that seems out of equilibrium at first sight, we may simply be looking at the wrong game. Indeed giving away can look stupid, but only if we overlook the fact that the selection mechanism that counts is sexual.

Although Miller does emphasize that he does not claim that sexual selection explains everything, I think it is good to mention explicitly what the limits are in the case of altruism and morality. He argues that traits that evolve as a result of sexual selection must be fitness indicators. One of the neccessary conditions for something to serve as a fitness indicator is that it must be costly. He thereby refers to Zahavi's (1997) *handicap principle*: cheap talk does not work and any fitness indicator is as credible as it is wasteful. This I think hints at the possibility that there are also limits to the explanatory power of sexual selection in this respect. In the arms race of ever costlier fitness indicators, one would expect altruism to remain unconditional; if it is exploitable, so much the better. Yet we have preferences for reciprocity and equality and although we do get warm glows from giving, we hardly ever feel all right when we think we are being exploited.

Again, we should be aware that we are not forced to choose only one of these selection types. We can have both and we can expect that they are interacting in more intricate ways than have been described so far.

Group selection?

In their book, Sober and Wilson (1998) convincingly argue that the heretic status of group selection arguments is unjust. Yet, I would hesitate to use their argument in this specific case, since it hinges very much on an assumption for which I have heard no reason why it would hold. I will shortly describe their argument in order to bring out what they, half tacitly, assumed. They think of a population that is divived up into groups, where the share of altruists and egoists differ. Assuming that altruists can raise the fitness of the group at their own expense, they correctly argue that a decline of the share of altruists within each group not necessarily implies that the share of altruists in the total population also goes down, and they give examples to show how the greater share of altruists in the faster growing groups can outweigh overall within group decline. Furthermore they assume that, after a number of generations, the population is regrouped and that this circle is repeated over and over again. With the proper regrouping, this process would then lead to a stable positive share of altruists in the population.

Now a simple consideration can show where the rub is. If we assume that the total population is large and that the groups are randomly formed, then the odds of ending up in a group which, apart from yourslf, consists of a certain number of altruists and a certain number of egoists are the same for altruists and egoists. (If the total population is not large, the odds are even a little less advantageous for altruists). Since egoists outperform altruists for every possible composition of the rest of the group, the egoists will outperform the altruist in the total population. (Please note that this is not a case of what they call the averaging fallacy; the argument is not that egoists outperform altruists within groups, but that, facing the same probability distribution of ending up in different groups, one outperforms the other in expectation.) Therefore, the only way in which such a group selection argument could hold, is if there would be reasons to assume that the regrouping is not random; altruist would on average have to find themselves in groups with more altruists - apart, of course, from themselves - than egoists, and they give no reason why this would be the case.[16]

A MESS and ESC

To get the same notation in both definitions, we will use M and N for finita automata (machines) and π for the payoff function. The first two definitions are from Binmore & Samuelson (1992)

Definition 9 *A (Moore) machine M is a modified evolutionary stable strategy (MESS) if, for all possible mutants N,*

1) $\pi(M, M) > \pi(N, M)$ *or*
2) $\pi(M, M) = \pi(N, M)$ *and* $\pi(M, N) > \pi(N, N)$ *or*
3) $\pi(M, M) = \pi(N, M)$ *and* $\pi(M, N) = \pi(N, N)$ *and* $|M| \leq |N|$

where $|M|$ denotes the number of states of a machine M.

In their mixed extention, F is a distribution over machines $M_1, ...M_n$ where M_i occurs with frequency f_i. The expected payoff of a machine N against such a distribution (population) is then $\pi(N, F) = \sum_{i=1}^{n} f_i \pi(N, M_i)$

Definition 10 *A distribution F is a polymorphous MESS if, for all possible mutants N and for all M_i in the support of F,*

1) $\pi(M_i, F) > \pi(N, F)$ *or*
2) $\pi(M_i, F) = \pi(N, F)$ *and* $\pi(M_i, N) > \pi(N, N)$ *or*
3) $\pi(M_i, F) = \pi(N, F)$ *and* $\pi(M_i, N) = \pi(N, N)$ *and* $|M_i| \leq |N|$

The modifications for the pure strategy MESS can be seen as demanding only neutral stability - that is, it requires $|M| \leq |N|$ rather than $|M| < |N|$ - combined with lexicographic preferences over complexity after payoffs. The definition of a polymorphous MESS however is not the equivalent of a mixed *NSS*,[17] nor would the definition turn into equivalent of a mixed ESS if $|M_i| \leq |N|$ would be replaced by $|M_i| < |N|$ (which I then would call a strict MESS). I would like to give two examples to show that. Suppose all possible machines are M_1, M_2 and N respectively and we gather the payoffs they earn against each other in a matrix.

$$\begin{bmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 5 \\ 0 & 1 & 0 \end{bmatrix}$$

To see that $F = \left[\frac{1}{2}, \frac{1}{2}, 0\right]$ satisfies *2)* in Definition 10, we evaluate $\pi(M_i, F) = \pi(N, F) = \frac{1}{2}$ and both $\pi(M_1, N) = \frac{1}{2} > 0 = \pi(N, N)$ and $\pi(M_2, N) = 5 > 0 = \pi(N, N)$. F is therefore a MESS

[16]Some examples they give, may suggest otherwise, but that is because they do not mention that not alle groups are equally likely; with the proper combinatorics, the sum of all possible groups amounts to the same as the expected performance of the different types.

[17]Neutrally stable strategy; $u[x, x] \geq u[y, x]$ for all x and $u[x, y] \geq u[y, y]$ if $u[x, x] = u[y, x]$.

and since it is so by *2)*, it is also a strict MESS. Yet it is not an ESS nor an NSS since for the mutant $G = \left[0, \frac{3}{4}, \frac{1}{4}\right]$, $\pi(F,F) = \frac{1}{2} = \pi(G,F)$ and $\pi(F,G) = \frac{17}{16} < \frac{9}{8} = \pi(G,G)$.

The opposite inclusion does not hold either, which we can see if we look at a second payoff matrix

$$\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

and at the same $F = \left[\frac{1}{2}, \frac{1}{2}, 0\right]$. For all possible distributions G over the three machines we know that $\pi(F,G) > \pi(G,G)$[18], which makes F globally and therefore also locally superior. (This is equivalent to evolutionary stability; see Weibull (1995) or Hofbauer, Shuster and Sigmund (1979)). F is not a MESS though; for M_1 we see that $\pi(M_1,F) = \frac{1}{2} = \pi(N,F)$ and $\pi(M_1,N) = -1 < 0 = \pi(N,N)$.

It is important to see that the paper suggests that the modification to allow for an equal number of states, that is to require $|M| \leq |N|$ instead of $|M| < |N|$, is the only deviation from the standard definition. For the first definition, where a pure strategy ESS is modified, that indeed is true. These two examples however show that its polymorphous extention deviates from the mixed equilibrium ESS in more ways. While the attraction of evolutionary stability and neutral stability is that they imply asymptotic stability and Lyapounov stability respectively in the replicator dynamics, the first example shows that such stability implications are lost in the definition of a polymorphous MESS. This is not due to what is presented and defended as the modification - complexity does not even show up in any of the examples - but to the other differences between the standard definitions and the definition used here.

Probst (1995) extends a strict polymorphous MESS to a setwise equivalent called evolutionary stable collection (ESC), which suffers from the same weaknesses as the notion it generalizes; the counterexamples still apply and an ESC is therefore not the same as an ES-set, which I again think is a preferable concept.

B Equilibrium strategies in the simple example

B.1 Strategy 1

The strategy we look at is:

> Play C if both players played C so far and $X \in [\overline{x}, 1]$. Play D otherwise.

and the claim is that this is an equilibrium strategy for $\delta > 0.5$.

It is obvious that for any history other than mutual cooperation, it is optimal to defect when playing against this strategy, since for all of those histories the opponent will always defect, whatever happens. So we only have to consider different actions for a history that consist of cooperation only. For those histories, it is clear that a strategy that cooperates in games where the opponent defects ($X \in [0, \overline{x})$) can always be improved by making it defect for those games as well. This means that the only serious candidates for outperforming the given strategy must play cooperate only for a subset of games for which the given strategy cooperates. We can further restict the strategies to be checked to those with a cooperation threshold higher than or equal to $\overline{x}$. The

[18] $\pi(F,G) - \pi(G,G)$ with of course $g_1, g_2, g_3 \geq 0$ and $g_1 + g_2 + g_3 = 1$ attains a strict maximum at F.

reason for this is that any other strategy among the candidates we had left, could be improved upon by a chosing a strategy with a threshold level; a strategy that cooperates for a subset of games that occur with probability p can never do better than a strategy that cooperates for games for which $X \in [1-p, 1]$ since future losses for not cooperating are the same in both cases, but the one shot gains of the strategy with cooperation threshold $1-p$ are an upper bound to the one shot gains of all strategies that cooperate with probability p.

Therefore we compute the expected payoff to a player with swiching thresholds $\overline{x}_a$ against a player with switching threshold $\overline{x}_b \leq \overline{x}_a$.

$$\mathbb{E}[\text{ Payoff }] = \mathbb{P}(X_1 \leq \overline{x}_b)\left(\mathbb{E}[\ 1 - X_1 \mid X_1 \leq \overline{x}_b\ ] + \mathbb{E}\textstyle\sum_{i=2}^{\infty} \delta^{i-1}[1 - X_i]\right) +$$

$$\mathbb{P}(\overline{x}_b < X_1 \leq \overline{x}_a)\left(\mathbb{E}[\ 2 \mid \overline{x}_a < X_1 \leq \overline{x}_b\ ] + \mathbb{E}\textstyle\sum_{i=2}^{\infty} \delta^{i-1}[1 - X_i]\right) +$$

$$\mathbb{P}(X_1 > \overline{x}_a)\left(\mathbb{E}[\ 1 + X_1 \mid X_1 > \overline{x}_a\ ] + \delta\ \mathbb{E}[\text{ Payoff }]\right)$$

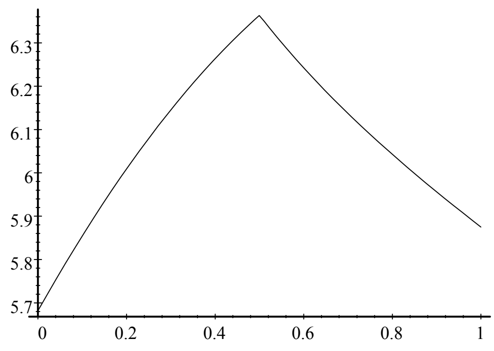
$$\Longleftrightarrow$$

$$\mathbb{E}[\text{ Payoff }] = \left((\overline{x}_b)\left(1 - \tfrac{1}{2}\overline{x}_b + \tfrac{\delta}{1-\delta}\tfrac{1}{2}\right) + (\overline{x}_a - \overline{x}_b)\left(2 + \tfrac{\delta}{1-\delta}\tfrac{1}{2}\right) + (1 - \overline{x}_a)\left(\tfrac{1\frac{1}{2} - \overline{x}_a - \frac{1}{2}\overline{x}_a^2}{1 - \overline{x}_a} + \delta\ \mathbb{E}[\text{ Payoff }]\right)\right)$$

$$\Longleftrightarrow$$

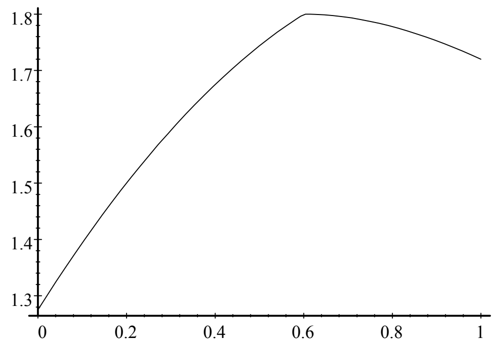
$$\mathbb{E}[\text{ Payoff }] = -\tfrac{1}{2}\tfrac{-2\overline{x}_b + 2\overline{x}_b\delta - \overline{x}_b^2 + \overline{x}_b^2\delta + 2\overline{x}_a - \delta\overline{x}_a + 3 - 3\delta - \overline{x}_a^2 + \overline{x}_a^2\delta}{(-1+\delta)(1-\delta+\delta\overline{x}_a)}.$$

Drawing the expected payoff as a function of the threshold level $\overline{x}_a$, there are two possible types of pictures we can get:



$(\delta = 0.9, \overline{x}_b = 0.5)$

and



$(\delta = 0.5, \overline{x}_b = 0.6)$

On forehand, we already knew that choosing a $\overline{x}_a$ lower than $\overline{x}_b$ would lead to a lower expected payoff, so we have focused on the part between $\overline{x}_b$ and 1. Now the right derivative of the payoff function in $x_a = x_b$ equals

$$\left.\frac{d\left(-\frac{1}{2}\frac{-2x_b+2x_b\delta-x_b^2+x_b^2\delta+2x_a-\delta x_a+3-3\delta-x_a^2+x_a^2\delta}{(-1+\delta)(1-\delta+\delta x_a)}\right)}{dx_a}\right|_{x_a\downarrow x_b} = \frac{1}{2}\left.\frac{-4\delta-x_a^2\delta+2x_a\delta+2\delta x_b+x_b^2\delta+2-2x_a}{(1-\delta+x_a\delta)^2}\right|_{x_a\downarrow x_b} =$$

$$= \frac{-2\delta+2x_b\delta+1-x_b}{(1-\delta+x_b\delta)^2} = \frac{(1-2\delta)(1-x_b)}{(1-\delta+x_b\delta)^2}.$$

Which is smaller than 0 as long as $\delta > 0.5$.

Then we compute a second derivative for all x_a:

$$\frac{d^2\left(-\frac{1}{2}\frac{-2x_b+2x_b\delta-x_b^2+x_b^2\delta+2x_a-\delta x_a+3-3\delta-x_a^2+x_a^2\delta}{(-1+\delta)(1-\delta+\delta x_a)}\right)}{dx_a^2} = \frac{d\left(\frac{1}{2}\frac{-4\delta-x_a^2\delta+2x_a\delta+2\delta x_b+x_b^2\delta+2-2x_a}{(1-\delta+x_a\delta)^2}\right)}{dx_a} =$$
$$-\frac{1-3\delta^2+2\delta^2x_b+x_b^2\delta^2}{(1-\delta+x_a\delta)^3} = -\frac{1-\delta^2\left(3-2x_b-x_b^2\right)}{(1-\delta+x_a\delta)^3}.$$

$0 \leq x_a \leq 1$ implies that we are in the first picture case if $1-\delta^2\left(3-2x_b-x_b^2\right) < 0$ and in the second picture case if not. The only possible threat to optimality therefore is that in this first case where $\frac{1}{\delta^2} < \left(3-2x_b-x_b^2\right)$ the expected payoff in $\overline{x}_a = 1$ exceeds the expected payoff in $\overline{x}_a = \overline{x}_b$

$$-\frac{1}{2}\frac{-2x_b+2x_b\delta-x_b^2+x_b^2\delta+2-\delta+3-3\delta-1+\delta}{(-1+\delta)(1-\delta+\delta)} > -\frac{1}{2}\frac{-2x_b+2x_b\delta-x_b^2+x_b^2\delta+2x_b-\delta x_b+3-3\delta-x_b^2+x_b^2\delta}{(-1+\delta)(1-\delta+\delta x_b)}+$$

$$\Longleftrightarrow$$

$$\delta < \frac{1}{x_b+3}.$$

This however is not possible if $\delta > 0.5$.

For all $\delta > 0.5$ the best response to the given strategy is therefore to switch to defect exactly when the other player does so.

Please note that there are more conditions for continued cooperation that would form an equilibrium strategy together with defecting after any deviation from mutual cooperation. Part of the reason why such a strategy would be an equilibrium is that both players switch at the same time to the punishment phase. This strategy 1 however naturally comes with a preference that is increasing in both the shares.

B.2 Strategy 2

The second strategy under consideration is:

Play C if both players played C the last round and $X \in [\overline{x}, 1]$.
Play C if not both players played C in the last round and $X \in [\widetilde{x}, 1]$.
Play D otherwise.

We will not determine exactly for which combinations of $\overline{x}$, $\widetilde{x}$ and δ this is an equilibrium; it can be done following the same procedure as with strategy 1, but I do not think it is very interesting to see the whole exercise repeated for a different strategy. After the essential calculations, we therefore just pick three values and show that it at least works for those values. A combination of $\overline{x}$, $\widetilde{x}$ and δ that makes strategy 2 an equilibrium consists of a δ that is large enough and an $\widetilde{x}$ that is sufficiently larger than $\overline{x}$.

For the same reasons as with strategy 1, we only need to look at strategies with cooperation thresholds larger than or equal to $\overline{x}$ and $\widetilde{x}$. First we compute the expected payoff for a players with swiching thresholds $\overline{x}_a, \widetilde{x}_{ab}$ against a player with switching thresholds $\overline{x}_b, \widetilde{x}_{ab}$, $\overline{x}_a \geq \overline{x}_b$.

$$\begin{aligned}\mathbb{E}[\text{ Payoff} \mid C\ ] = {} & \mathbb{P}(X_1 \leq \overline{x}_b)\,(\mathbb{E}[\ 1 - X_1 \mid X_1 \leq \overline{x}_b\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]) + \\ & \mathbb{P}(\overline{x}_b < X_1 \leq \overline{x}_a)\,(\mathbb{E}[\ 2 \mid \overline{x}_b < X_1 \leq \overline{x}_a\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]) + \\ & \mathbb{P}(X_1 > \overline{x}_a)\,(\mathbb{E}[\ 1 + X_1 \mid X_1 > \overline{x}_a\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ])\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\text{ Payoff} \mid D\ ] = {} & \mathbb{P}(X_1 \leq \widetilde{x}_{ab})\,(\mathbb{E}[\ 1 - X_1 \mid X_1 \leq \widetilde{x}_{ab}\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]) + \\ & \mathbb{P}(X_1 > \widetilde{x}_{ab})\,(\mathbb{E}[\ 1 + X_1 \mid X_1 > \widetilde{x}_{ab}\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]) + \\ & \Longleftrightarrow\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\text{ Payoff} \mid C\ ] = {} & (\overline{x}_b)\left(1 - \tfrac{1}{2}\overline{x}_b + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) + (\overline{x}_a - \overline{x}_b)\,(2 + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]) + \\ & (1 - \overline{x}_a)\left(\tfrac{1\frac{1}{2} - \overline{x}_a - \frac{1}{2}\overline{x}_a^2}{1 - \overline{x}_a} + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right)\end{aligned}$$

$$\mathbb{E}[\text{ Payoff} \mid D\ ] = (\widetilde{x}_{ab})\left(1 - \tfrac{1}{2}\widetilde{x}_{ab} + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) + (1 - \widetilde{x}_{ab})\left(\tfrac{1\frac{1}{2} - \widetilde{x}_{ab} - \frac{1}{2}\widetilde{x}_{ab}^2}{1 - \widetilde{x}_{ab}} + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right)$$

.

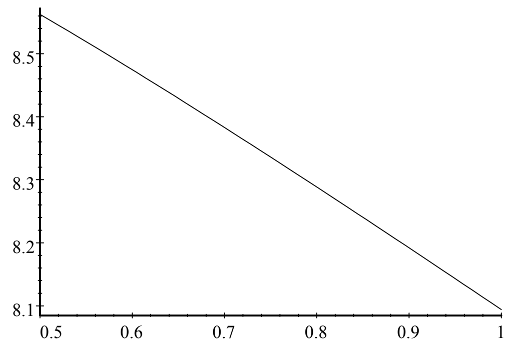
Rewriting the second equation gives:

$$\mathbb{E}[\text{ Payoff} \mid D\ ] = \left(\frac{1}{2}\frac{2\widetilde{x}_{ab}{}^2-3-2\delta\mathbb{E}[\text{Payoff} \mid C]+2\delta\mathbb{E}[\text{Payoff} \mid C]\ \widetilde{x}_{ab}}{-1+\widetilde{x}_{ab}\delta}\right).$$

Filling in in the first and rewriting gives:

$$\mathbb{E}[\text{ Payoff} \mid C\ ] = \frac{1}{2}\frac{-2\overline{x}_b+2\overline{x}_b\delta\widetilde{x}_{ab}-\overline{x}_b^2+\overline{x}_b^2\delta\widetilde{x}_{ab}+2\overline{x}_a-2\delta\overline{x}_a\widetilde{x}_{ab}-2\overline{x}_a\delta\widetilde{x}_{ab}{}^2+3\delta\overline{x}_a+3-3\widetilde{x}_{ab}\delta-\overline{x}_a^2+\delta\overline{x}_a^2\widetilde{x}_{ab}}{1-\widetilde{x}_{ab}\delta-\delta^2\overline{x}_a-\delta+\delta^2\widetilde{x}_{ab}+\delta\overline{x}_a}.$$

For $\delta = 0.9, \widetilde{x}_{ab} = 0.9, \overline{x}_b = 0.5$ we get:



which is decreasing for all $\overline{x}_a \geq 0.5$.

Then we compute the expected payoff for a players with swiching thresholds $\overline{x}_{ab}, \widetilde{x}_a$ against a player with switching thresholds $\overline{x}_{ab}, \widetilde{x}_b$, $\widetilde{x}_a \geq \widetilde{x}_b$.

$$\mathbb{E}[\text{ Payoff} \mid C\ ] = \mathbb{P}(X_1 \leq \overline{x}_{ab})\left(\mathbb{E}[\ 1 - X_1 \mid X_1 \leq \overline{x}_{ab}\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) +$$

$$\mathbb{P}(X_1 > \overline{x}_{ab})\left(\mathbb{E}[\ 1 + X_1 \mid X_1 > \overline{x}_{ab}\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right)$$

$$\mathbb{E}[\text{ Payoff} \mid D\ ] = \mathbb{P}(X_1 \leq \widetilde{x}_b)\left(\mathbb{E}[\ 1 - X_1 \mid X_1 \leq \widetilde{x}_b\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) +$$

$$\mathbb{P}(\widetilde{x}_b < X_1 \leq \widetilde{x}_a)\left(\mathbb{E}[\ 2 \mid \widetilde{x}_b < X_1 \leq \widetilde{x}_a\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) +$$

$$\mathbb{P}(X_1 > \widetilde{x}_a)\left(\mathbb{E}[\ 1 + X_1 \mid X_1 > \widetilde{x}_a\ ] + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right) +$$

$$\Longleftrightarrow$$

$$\mathbb{E}[\text{ Payoff} \mid C\ ] = (\overline{x}_{ab})\left(1 - \tfrac{1}{2}\overline{x}_{ab} + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) + (1 - \overline{x}_{ab})\left(\frac{1\frac{1}{2}-\overline{x}_{ab}-\frac{1}{2}\overline{x}_{ab}^2}{1-\overline{x}_{ab}} + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right)$$

$$\mathbb{E}[\text{ Payoff} \mid D\ ] = (\widetilde{x}_b)\left(1 - \tfrac{1}{2}\widetilde{x}_b + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) + (\widetilde{x}_a - \widetilde{x}_b)\left(2 + \delta\ \mathbb{E}[\text{ Payoff} \mid D\ ]\right) +$$

$$(1 - \widetilde{x}_a)\left(\frac{1\frac{1}{2}-\widetilde{x}_a-\frac{1}{2}\widetilde{x}_a^2}{1-\widetilde{x}_a} + \delta\ \mathbb{E}[\text{ Payoff} \mid C\ ]\right).$$

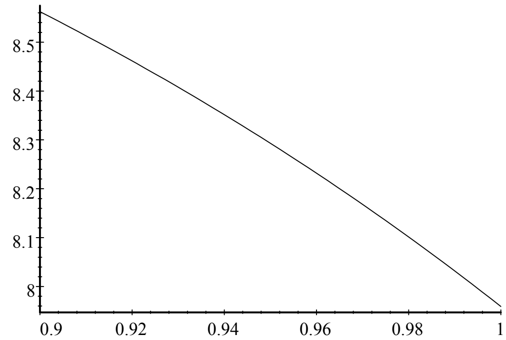
Rewriting the second equation gives:

$$\mathbb{E}[\text{ Payoff } \mid D\ ] = \frac{1}{2}\frac{2\widetilde{x}_b + \widetilde{x}_b^2 - 2\widetilde{x}_a - 3 + \widetilde{x}_a^2 - 2\delta\mathbb{E}[\text{Payoff} \mid C] + 2\delta\mathbb{E}[\text{Payoff} \mid C]\widetilde{x}_a}{-1+\delta\widetilde{x}_a}$$

Filling in in the first and rewriting gives:

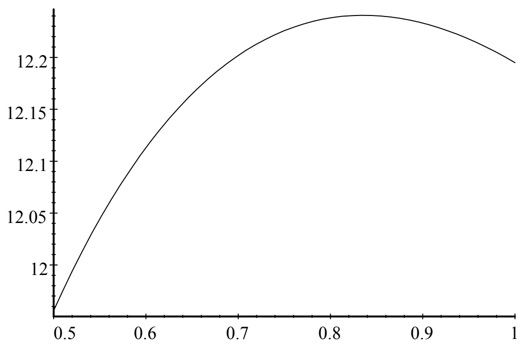
$$\mathbb{E}[\text{ Payoff } \mid C\ ] = -\frac{1}{2}\frac{2\overline{x}_{ab}{}^2 - 2\overline{x}_{ab}{}^2\delta\widetilde{x}_a + 2\overline{x}_{ab}\widetilde{x}_b\delta + \overline{x}_{ab}\delta\widetilde{x}_b{}^2 - 3\overline{x}_{ab}\delta + \overline{x}_{ab}\delta\widetilde{x}_a{}^2 - 3 + 3\delta\widetilde{x}_a - 2\overline{x}_{ab}\delta\widetilde{x}_a}{1 - \delta\widetilde{x}_a - \overline{x}_{ab}\delta^2 - \delta + \delta^2\widetilde{x}_a + \overline{x}_{ab}\delta}$$

For $\delta = 0.9, \widetilde{x}_b = 0.9, \overline{x}_{ab} = 0.5$ we get:



which is decreasing for all $\widetilde{x}_a \geq 0.9$.

Please note that equilibrium can not be reached for any randomly chosen combination of $\widetilde{x}$ and $\overline{x}$: with $\delta = 0.9, \widetilde{x}_{ab} = 0.6$ and $\overline{x}_b = 0.5$ the first formula yields



which is increasing at $\overline{x}_a = 0.5$.

C Stochastic gift exchange

Let X be uniformly distributed on $[0, 1]$. At every period an x is drawn from this distribution. In odd periods player 1 can choose between cooperate and defect. Cooperate results in payoffs $x - 1$ for him- or herself and $x + 1$ for the other player, and defect leads to 0 for both. In even periods the roles are reversed. If we look at a strategy like strategy 1 from the other example, that is:

play C if both players played C so far and $X \in [\overline{x}, 1]$. Play D otherwise.

then we can compute expected payoffs for a cooperation threshold $\overline{x}_a$ against $\overline{x}_b$

$$\mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] = \int_{\overline{x}_a}^{1} (x-1)\, dx + \delta \left(1-\overline{x}_a\right) \int_{\overline{x}_b}^{1} (x+1)\, dx + \delta^2 \left(1-\overline{x}_a\right)\left(1-\overline{x}_b\right) \mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] \Leftrightarrow$$

$$\mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] = \frac{-\frac{1}{2}(1-\overline{x}_a)^2 + \delta(1-\overline{x}_a)\left(\frac{1}{2}(1-\overline{x}_b)(3+\overline{x}_b)\right)}{1-\delta^2(1-\overline{x}_a)(1-\overline{x}_b)}$$

For $\delta > \frac{9}{16}$ we can see that $\arg\max_{\overline{x}_a \in [0,1]} \mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] < \overline{x}_b$ for any $\overline{x}_b \in (0,1)$. Moreover, for $\overline{x}_b = 0$ we find that $\arg\max_{\overline{x}_a \in [0,1]} \mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] = 0 = \overline{x}_b$. This implies that for $\delta > \frac{9}{16}$ there is only one value of $\overline{x}$ for which this strategy is an equilibrium; $\overline{x} = 0$, which reduces the strategy to *grim*.

If we take $\delta < \frac{1}{2}$ then we find that $\arg\max_{\overline{x}_a \in [0,1]} \mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] > \overline{x}_b$ for any $\overline{x}_b \in (0,1)$. Then for $\overline{x}_b = 1$ we find that $\arg\max_{\overline{x}_a \in [0,1]} \mathbb{E}\left[\text{ Payoff } \mid \overline{x}_a \right] = 1 = \overline{x}_b$. Thus for $\delta < \frac{1}{2}$ this strategy can only be an equilibrium if $\overline{x} = 1$ and that makes the strategy no different from *always defect*.

Only for a small window of deltas, that is for δ between $\frac{1}{2}$ and $\frac{9}{16}$, there is an equilibrium strategy with cooperation threshold $\overline{x}$ that equals $\frac{-\frac{1}{2}+\delta+\frac{1}{2}\sqrt{(9-16\delta)}}{\delta}$. Such an equilibrium could be called stable in the sense that this $\overline{x}$ lies in an open interval such that the best response to any strategy with an x in this interval is a strategy with a cooperation threshold that lies between x and $\overline{x}$. For an even smaller window, $\frac{1}{2}\sqrt{17} - \frac{3}{2} < \delta < \frac{9}{16}$, there is a second equilibrium value for $\overline{x} = \frac{-\frac{1}{2}+\delta-\frac{1}{2}\sqrt{(9-16\delta)}}{\delta}$. The ad hoc stability condition does not hold for those equilibria; slightly smaller $\overline{x}$ run away to 0.

References

[1] Abreu, D. (1988): "On the theory of infinitely repeated games with discounting," *Econometrica,* 56, 383-396

[2] Abreu, D. and A. Rubinstein (1988): "The structure of Nash equilibria in repeated games with finite automata," *Econometrica*, 56, 1259-1281.

[3] Andreoni, J., and J. Miller (2002): "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70, 737-753.

[4] Aumann, R., and L. Shapley (1976): "Long term competition: a game theoretic analysis," mimeo, Hebrew University.

[5] Axelrod, R., (1984): *The evolution of cooperation.* New York: Basic Books.

[6] Bester, H. and W. Guth (1998): "Is altruism evolutionary stable?" *Journal of Economic Behaviour and Organization*, 34, 193-209.

[7] Binmore, K.G. (1994): *Game Theory and the Social Contract, Volume I: Playing Fair.* Cambridge, MA: MIT Press.

[8] Binmore, K.G. (1998): *Game Theory and the Social Contract, Volume II: Just Playing.* Cambridge, MA: MIT Press.

[9] Binmore, K.G., and L. Samuelson (1992): "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory*, 57, 278-305.

[10] Binmore, K.G., and L. Samuelson (1997): "Muddling Through: Noisy Equilibrium Selection," *Journal of Economic Theory*, 74, 235-265.

[11] van Damme, E. (1987): *Stability and Perfection of Nash Equilibria*. Berlin: Springer Verlag. (2nd ed. 1991)

[12] van Damme, E. (1989): "Renegiotion-proof equilibria in repeated prisoner's dilemma," *Journal of Economic Theory*, 47, 206-217

[13] Dawkins, R. (1976): *The selfish gene*. Oxford: Oxford University Press.

[14] Eshel, I., L. Samuelson and A. Shaked, 1998 Altruists, egoists and hooligans in a local interaction model, *American Economic Review*, 88, 157-179.

[15] Friedman, J. (1971): "A noncooperative equilibrium for supergames," *Review of Economic Studies*, 38, 1-12.

[16] Fudenberg, D. and E. Maskin (1986): "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica*, 54, 533-554.

[17] J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr (2001): "In Search of Homo economicus: Experiments in 15 Small-Scale Societies," *American Economic Review*, 91, 73-78..

[18] Hofbauer, J. and K. Sigmund (1998): *Evolutionary Games and Population Dynamics*, Cambridge UK: Cambridge University Press.

[19] Karni, E. and Z. Safra (2002): "Individual sense of justice: a utility representation," *Econometrica*, 70, 263-284.

[20] Maynard Smith, J. and G.R. Price (1973): "The logic of animal conflict," *Nature*, 246, 15-18.

[21] Mertens, J-F., S. Sorin and S. Zamir (1994): *Repeated games*, Louvain-La-Neuve: Univ. Catholique de Louvain, Center for Operations Research & Econometrics.

[22] Miller, G. (2001): *The Mating Mind; How Sexual Choice Shaped the Evolution of Human Nature*, New York: Anchor Books.

[23] Probst, D., 1996 On evolution and learning in games, PhD thesis, University of Bonn

[24] Samuelson, (2001): "Analogies, Adaptations, and Anomalies," *Journal of Economic Theory*, 97, 320-367.

[25] Smith, A. (1759): *The Theory of Moral Sentiments*, Indianapolis: Liberty Fund. (New ed.: 1984)

[26] Sober, E. and D. S. Wilson (1998): *Unto Others; the evolution and psychology of unselfish behavior,* Cambridge MA: Harvard University Press.

[27] Taylor, P. and L. Jonker (1978): "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, 40, 145-156.

[28] Trivers, R.L. (1971): "The evolution of reciprocal altruism," *Quarterly Review of Biology,* 46, 35-57.

[29] Weibull, J.W. (1996): *Evolutionary Game Theory*, Cambridge MA: MIT Press.

[30] Zahavi, A. (1997): *The handicap principle: A missing piece of Darwin's puzzle*, Oxford: Oxford University Press.