Learning to be fair

Luciano Andreozzi
Università degli Studi di Trento
Facoltà di Economia
Via Inama, 5
38100 Trento Italy
E.mail luciano.andreozzi@economia.unitn.it

April 21, 2009

Abstract

We study the process of equilibrium selection in games when players have "sophisticated" preferences of the type discussed, among others, by Rabin [16] and Segal and Sobel [17]. To this end, we employ standard noisy version of the best response dynamics. We obtain several results concerning some popular games such as the Prisoner's Dilemma, the Battle of the Sexes and the Dictator Game. For example we show that with the preferences for reciprocity introduced by Rabin, the cooperative Nash equilibrium in the Prisoner's Dilemma is never stochastically stable.

> You really shouldn't leave a bar or restaurant without leaving a tip of at least fifteen percent and about the same should be added to taxi fares. A hotel porter should get roughly $1 for each bag carried to your room; a coatcheck clerk should receive the same per coat. (The Rough Guide to California)

1 Introduction

Although the notion of Nash equilibrium is a cornerstone of non cooperative game theory, few scholars believe that Nash equilibria are good predictors of the way games are played by ordinary people facing a game for the first time. There are many sources for this skepticism. For example, players can only compute their Nash equilibrium strategies if their payoff functions are common knowledge. Young [22] notices that this assumption "strain[s] credulity: if anything is common knowledge, it is that utility functions are almost never common knowledge." (117)

One might reply that such a claim is too strong, at least when a game's outcomes are simple monetary amounts as it is usually the case in experiments. If one could assume that players were only concerned with their own monetary payoffs, then the only source of uncertainty about other players' preferences would be their degree of risk aversion. A large amount of experimental evidence gathered in the last two decades, however, shows that such an assumption is unwarranted. Players trust other players (and repay their trust) in Trust Games, give money in Dictator Games and reject unfair offers in Ultimatum Games. (See Camerer [6] for an overview) None of these choices can be rationalized assuming that players are only concerned with the money they receive at the end of the experiment.

To make sense of these experimental findings, several models have been proposed in which human beings are assumed to be concerned with other things other than their material payoffs. This literature is customarily divided into two main strands. There is a first class of models (whose most representative articles are Bolton and Ockenfels [5] and Fehr and Schmidt [11]) in which players take into consideration the amount of money obtained by other players beside themselves. They might be altruists (sacrificing their own well being to improve the wellbeing of others), inequality adverse (being willing to pay to reduce inequality between themselves and other players) and so on. Despite the success of these models in accounting for many experimental findings, there is a large agreement on the fact that a pure concern for the distribution of the outcomes is unable to account for many important anomalies observed in experiments. (See Sobel [18])

The second strand of research, pioneered by Rabin [16] and developed further by Charness and Rabin [8], Falk and Fischbacher [10] and Dufwenberg and Kirchsteiger [9], is based on a somewhat stronger set of assumptions. Besides a concern for the outcome of the game, these models assume that people's utilities are influenced by their mutual expectations concerning the way in which the game will be played. A typical (but not unique) way in which such a concern manifests itself is a taste for reciprocity. A player who expects to be treated fairly by other players will be willing to sacrifice part of his payoff in order to be fair to them. The same player would be willing to forego part of his monetary payoff in order to "punish" other player's unfairness. Since players' preferences depend not only upon the *outcome* of the game, but also upon their mutual expectations, this approach requires a more complex analytical toolset, which is usually taken from (a simplified version of) psychological game theory introduced by Geanakoplos, Pearce, and Stacchetti [12].

Making more sophisticated assumptions on people's preferences has two effects. On the one hand, models become more flexible, which allows for the existence of equilibria that are more in line with what one observes in experiments. However, and this is the second effect, in the presence of more sophisticated preferences equilibrium predictions become even more doubtful. If common knowledge of payoff functions looks suspicious when preferences include only material payoffs, it looks even more so if other aspects of human motivation like altruism and reciprocity are taken into the picture. Furthermore, more sophisticated preferences usually do not reduce the number of equilibria, so that the equilibrium selection problem presents itself just like in ordinary game theoretical models.

There is a well established tradition in game theory that looks suspiciously at NE as predictors of individual's behavior in one-shot interactions, but considers it a plausible approximation for situations in which players are allowed to learn and experiment. It is customary to refer to this approach as "evolutionary", although it is by no means confined to situations in which genetic selection takes place. Rather than trying to answer the question of what is the rational course of action in a given game, evolutionary models assume that players are boundedly rational and adjust their behavior over time to the behavior of other players. In this context, Nash equilibria are interpreted as *conventions*, that is as beliefs concerning how a certain game is to be played, which are shared within a population of individuals.

This paper presents an evolutionary model in which agents' preferences are of the kind first studied by Rabin [16]. The model has thus two main ingredients. First, a "sophisticated" theory of player's preferences. I take this part from Segal and Sobel [17], which is the most general approach to interdependent preferences proposed so far for simultaneous games. Some popular models such as those proposed by Rabin [16] and Dufwenberg and Kirchsteiger [9], are particular cases of their axiomatic model. The second ingredient is a dynamic process of adjustment, which is taken from the best response models introduced by Kandori, Mailath and Rob [13] and Young [20], and popularized in many other papers afterwards. In particular, I shall follow the recent comprehensive treatment proposed in Binmore, Samuelson and Young [4]. To show the relevance and flexibility of the model I will discuss a few games the Prisoner's Dilemma, the Battle of the Sexes and the Dictator Game the approach to reciprocity pioneered by Rabin [8].

The paper proceeds as follows: Section 2 introduces the necessary technicalities concerning player's preferences. Section 3 discusses the dynamical adjustment process and proves some fairly general lemmata. Section 5 contains the applications. Section 6 concludes.

2 The game

We shall present here the axiomatic approach to (simultaneous) psychological games pioneered by Segal and Sobel [17]. Our presentation will necessarily be cursory. For ease of comparison, we shall follow them as closely as possible. The interested reader is to look at the original paper for the details and the main motivation.

We shall deal with simultaneous two-players games. The approach we propose can be easily extended to n players games, although it would become more cumbersome and less straightforward. It can be extended to sequential games, but in this case some more substantial work needs to be done. We leave this task for future research.

Players will be denoted as 1 and 2. Let X_i be the set of outcomes for player i. Player i has a preference relation $\succcurlyeq_i^{out}$ over the set of lotteries on X_i we shall indicate as $\Delta(X_i)$. These preferences satisfy the assumptions of expected utility. They can thus be represented by a pair of von Neumann-Morgernstern functions $\pi_i : \Delta(X_i) \to \mathcal{R}$. We shall refer to π_i as player i's *material payoff function.*

Let A_1 and A_2 be the set of *pure* strategies for player 1 and 2 resp., with generic element a_i^k. S_1 and S_2 are the sets of mixed strategies derived from A_1 and A_2, with generic element s_i. $s_i(a_i^k)$ is the probability with which a_i^k is played in s_i. $A = A_1 \times A_2$ is the set of pure strategy combinations and $S = S_1 \times S_2$ is the set of mixed strategy combinations. The *outcome function* for player i is $O_i : A \to X_i$. It assigns an element in X_i to any pure strategy combination chosen by the two agents. We shall indicate with $\pi_i(a_1^k, a_2^h)$ (rather than $\pi_i(O_i(a_1^k, a_2^h)))$ player i's utility for the outcome he gets when the strategy combination chosen by the players is (a_1^k, a_2^h).

Since preferences over outcomes obey the axioms of expected utility, the extension of material payoffs to mixed strategies is immediate:

$$\pi_i(s_1, s_2) = \sum_{a_1^k \in A_1} \sum_{a_2^h \in A_2} s_1(a_1^k) s_2(a_2^h) \pi_i(a_1^k, a_2^h)$$

In standard game theory, preferences over the strategies of the game are directly derived from those for outcomes. If player i expects player j's to use his (possibly mixed) strategy s_j, he prefers strategy s_i to s_i' iff $\pi_i(s_i, s_j) > \pi_i(s_i', s_j)$. A generic game G is a quadruple $G = < A_1, A_2, \pi_1, \pi_2 >$, where A_i is the set of pure strategies for player i, and π_i is payoff function.

S&S assume that player i's preferences over strategies depend, beside the strategy chosen by the other player (which determines the outcome of the game) upon the expected pattern of play. Suppose that players' expectations concerning the way in which the game will be played is represented by a (possibly mixed) of strategies $s^* = (s_1^*, s_2^*)$. Then, player i's preferences over strategies will be represented by a complete and transitive preference relation $\succcurlyeq_{i,s^*}$.

The relation $\succcurlyeq_{i,s^*}$ is assumed to satisfy the following axioms:

1. Continuity (C): For every $s_i \in S_i$ and $s^* \in S$, the sets $\{(s_i', s^*) : s_i' \succcurlyeq_{i,s^*} s_i\}$ and $\{(s_i', s^*) : s_i \succcurlyeq_{i,s^*} s_i'\}$ are closed subsets of $S_i \times S$.

2. Independence (I): For every $s_i, s_i', s_i'' \in S_i$, for every $s^* \in S$ and for every $\alpha \in (0, 1]$, $s_i \succcurlyeq_{i,s^*} s_i'$ if and only if $\alpha\ s_i + (1-\alpha)s_i'' \succcurlyeq_{i,s^*} \alpha\ s_i' + (1-\alpha)s_i''$.

3. Self-interest (SI) Suppose $\pi_j(s_i, s_j^*) = \pi_j(s_i', s_j^*)$. Then $s_i \succcurlyeq_{i,s^*} s_i'$ if and only if $\pi_i(s_i, s_j^*) \geq \pi_i(s_i', s_j^*)$.

The self-interest assumption (SI) has the following intuitive content. Suppose player i expects player j to play s_j^* and can choose two different strategies s_i and s_i' which (given s_j^*) yield player j the same utility over outcomes. Then he will prefer the strategy that yields the highest utility (over outcomes) to himself.

S&S main result is the following theorem.

Theorem 1 *(Segal and Sobel [17] Theorem 1) If preferences $\succcurlyeq_{i,s^*}$ satisfy axioms (C) (I) and (SI) they can be represented by a utility function with the following form:*

$$V_i(s_i, s^*) = \pi_i(s_i, s_j^*) + \alpha_i(s^*)\pi_j(s_i, s_j^*) \tag{1}$$

where s^* is the "context" of the game.

This theorem states that each player i's utility function is the sum of two terms:

1. his own utility over outcomes when he plays the strategy s_i and expects j to play her strategy s_j^*.

2. the utility i "gives" to j when he chooses s_i and expects j to play s_j^*, weighted by the coefficient $\alpha_i(s^*)$.

$\alpha_i(s^*)$ represents i's attitude to j (either altruistic of spiteful) and depends upon the expected pattern of play s^*.[1] Standard game theoretical models assume that $\alpha^j(s^*) = 0$ for any s^*. Psychological game theory allows for the possibility that an agent is altruistic in context s^* ($\alpha_i(s^*) > 0$) and spiteful

[1]Notice that in evaluating strategy a_1^k, player 1 considers $s^* = (s_1^*, s_2^*)$ as the context of the game rather than (a_i^k, s_2^*). The idea is that s_1^* is the strategy that player 1 is expected to play in that game when the context is s^*. If he picks a different pure strategy a_1^k, he gets a different payoff, but he cannot alter the context of the game. In a simultaneous move game, player i cannot assume that j observes his unilateral deviation from the expected pattern of play.

in a different context s^{**} ($\alpha_i(s^{**}) < 0$). For example, a player might behave altruistically in a Prisoner's Dilemma if he believes that the other player will cooperate, and selfishly if he believes that the other player will defect.

We shall indicate with $G^V =< G, V_1, V_2 >$ the psychological game obtained by an ordinary game G, when player's preferences over strategies are represented by V_1 and V_2.

Let us now turn to the notion of equilibrium appropriate for a psychological game G^V. The choice of the notion of equilibrium hinges upon the interpretation of mixed strategies. Mixed strategies raise problems in psychological games that are even more delicate than the problems they raise in standard games. The reason is that it makes a difference whether one interprets them as conscious randomization or uncertainty about other player's behavior (see Segal and Sobel [17] and Dufwenberg and Kirchsteiger [9]). We have no space here to discuss this matter. We shall limit ourselves to assume that players do not randomize, and believe that other players do not randomize too. Mixed strategies are thus interpreted as uncertainty concerning other player's behavior. This is the natural interpretation in view of the model we shall present in Section 3. In that model, mixed strategies will be explicitly modelled as distributions of strategies within populations of identical agents, each of whom only uses a pure strategy.

We shall indicate with $BR_i(s^*)$ player i's set of best responses to context s^*, calculated with utility functions as in Equation 1. Formally, $BR_i(s^*) = \{a_i^k \in A_i : V_i(a_i^k, s^*) \geq V_i(s_i, s^*) \ \forall s_i \in S_i\}$.

Since players do no randomize, the relevant notion of equilibrium is equilibrium in beliefs as defined in Aumann and Brandenburger [1].

Definition 1 *A belief profile $s^* \in S$ is an equilibrium in beliefs if $s_i^*(a_i^k) > 0$ implies that $a_i^k \in BR_i(s^*)$.*

Notice that since $s_i^*(a_i^k)$ is the probability with which j expects i to play his pure strategy a_i^k, s^* is an equilibrium if each player expects the other to play a strategy with positive probability only when that strategy is a best response to s^*. Notice also that when $\alpha_i(s^*) = 0$ for all i, s^* can only be an equilibrium iff s^* is a Nash equilibrium for the game G.

3 The Learning Model

Suppose there are two large populations which (with a slight abuse of notation) we shall denote as 1 and 2. Each population numbers N individuals. Time is discrete. The psychological game G^V is played in each round by pairs of agents drawn at random from the two populations. The agent drawn from population $i = 1, 2$ always occupies the corresponding role in the game G^V.

Each agent adopts a *pure* strategy to play the game at every match. Let N_i^k be the number of individuals playing strategy a_i^k in population i. $\sigma_i(a_i^k) = \frac{N_i^k}{N}$ is the fraction of population i playing a_i^k. The *state* of the two populations is $\sigma = (\sigma_1, \sigma_2)$. Σ_i is the set of possible states for population i and $\Sigma = \Sigma_1 \times \Sigma_2$

is the set of possible states for the entire process. Notice that the set of states Σ_i and the set of mixed strategies S_i are both probability distributions over the set of pure strategies A_i.

At the end of each period, each agent receives, with probability λ, the possibility to revise his strategy. (With probability $(1-\lambda)$ his strategy remains unchanged.) In this case we say that an agent receives the learning draw. When an agent receives the learning draw, he first observes the current state of the two populations σ and updates his expectations concerning the way in which people behave (and expect him to behave) in the game. In the terminology introduced in Section 2, updating agents take the current state of the two populations σ as the conventional way of playing game G^V.

Given σ, a revising agent will choose a pure strategy a_i^k which maximizes $V_i(a_i^k, \sigma)$, that is a strategy $a_i^k \in BR(\sigma)$. This process defines a Markov chain over the possible states of the population, which we shall denote as $\Gamma(0, N)$. At the beginning of each period, depending on the current state of the population σ, there is a well defined probability $\gamma_{\sigma\sigma'}$ that the system will enter a new state σ'. As it is customary in this kind of literature, we shall assume that with a strictly positive probability ε each agent is then selected to make a "mistake". Such an agent will take a pure strategy at random.

The process derived from the combination of the learning process $\Gamma(0, N)$ plus the mistakes will be denoted as $\Gamma(\varepsilon, N)$. For any $\varepsilon > 0$, this process will have a strictly positive transition matrix and will therefore be unreducible. As a consequence, there will be a unique stationary distribution $\mu_{\varepsilon,N}$. The elements of $\mu_{\varepsilon,N}$ represent the average time the process spends in each state as t gets large. We are interested in the stationary distribution $\mu_{\varepsilon,N}$ for small values of ε. We say that a state σ is *stochastically stable* if $lim_{\varepsilon\to 0}\mu_{\varepsilon,N}(\sigma) > 0$. Stochastically stable states are those states we expect to observe in the long run when mistakes become arbitrarily rare.

3.1 Absorbing states in the unperturbed process

We say that a set of states $Q \subseteq \Sigma$ is *absorbing* in the process $\Gamma(0, N)$ if: (i) it is impossible from a state $\sigma \in Q$ to reach a state $\sigma' \notin Q$ and (ii) for any two states $\sigma, \sigma' \in Q$ it is possible to reach σ from σ'. When an absorbing set Q is a singleton, we say that it is an *equilibrium.* We will indicate with $\mathcal{S}$ the set of equilibria, and with $\breve{\mathcal{S}} \subseteq \mathcal{S}$ the set of stochastically stable equilibria.

We shall deal with games that have the following property.

Definition 2 *No-cycling condition: the process* $\Gamma(0, N)$ *satisfies the no-cycling condition if all its absorbing states are singletons.*

A state $\sigma^* \in \Sigma$ can be an equilibrium only provided that none of the agents, when receiving the learning draw, will be induced to adopt a different strategy. This requires that there is a single best response to σ^* for both players, and in both populations all agents adopt their (unique) best response. Formally, σ^* is an equilibrium if for any agent in population $i = 1, 2$ there is a strategy a_i^k such that $\sigma_i^*(a_i^k) = 1$ and $V_i(a_i^k, \sigma^*) > V_i(a_i^{k'}, \sigma^*)$ for any $k' \neq k$.

We shall denote as σ_{kh} the state in which all agents in population 1 play a_1^k and all agents in population 2 play a_2^h. If σ_{kh} corresponds to an equilibrium we shall indicate it as σ_{kh}^*. Clearly, σ_{kh}^* can be an equilibrium only provided that two conditions are met (i) (a_1^k, a_2^h) is an equilibrium in beliefs as in Definition 1, and (ii) $BR_i(\sigma^*)$ is a singleton for $i = 1, 2$.

Young [20] proved that, in standard game theory, the no-cycling condition holds for a large class of games, known as weakly acyclical games. An analogous definition can be made with minimal modifications in our context as well.

Let $a = (a_1^k, a_2^h)$ be a pair of pure strategies for the two players. We shall define a *best-reply graph* of G^V in the same way it is defined in Young [20]. Each vertex of the graph is a pair of pure strategies $a = (a_1^k, a_2^h)$. Let a and a' be two vertexes of the graph. There is a directed edge $a \to a'$ if and only if $a \neq a'$ and either $a_1 = a_1'$ and $a_2' \in BR_2(a)$, or $a_2 = a_2'$ and $a_1' \in BR_1(a)$.

In other words, two edges a and a' are connected only if one of the two players is playing the same strategy both in a and a', while the other player plays a best reply to that strategy in a'. We say that a game G^V is *weakly acyclical* if for each vertex a of its best-reply graph, there is a directed path to a vertex a^* from which there is no outgoing arrow. Clearly, a vertex without outgoing arrows corresponds to an equilibrium in pure strategies of the game G^V.

The following theorem is a counterpart of Young [20], Theorem 1.

Theorem 2 *Let G^V be a weakly acyclical game. The unperturbed process $\Gamma(0, N)$ converges almost surely to an equilibrium σ_{kh}^*.*

Proof. Consider any state σ and suppose that $a_i^k \in BR_i(\sigma)$. Let all agents in population i receive the learning draw (this event has clearly a positive probability). Since they will all observe σ, there is a strict probability that they will all choose a_i^k. We have thus reached a new state σ' in which all agents in population i play the same strategy a_i^k. Suppose in the next period all (and only) agents in population j receive the learning draw (again, this event has a non-zero probability under the process $\Gamma(0, N)$). Since in σ' all agents in population i play a_i^k, there is a positive probability that all agents in j will choose the same strategy $a_j^h \in BR(\sigma')$. We have thus reached a state σ'' in which all agents in population i play a_i^k and all agents in population j play a_j^k. If these strategies form a strict equilibrium (that is if $a_i^k = BR_i(\sigma'')$ and $a_j^h = BR_j(\sigma'')$) then we are done: the process has reached an absorbing state.

Suppose instead that σ'' is not an equilibrium. Since the game is weakly acycical, there is a path $\sigma'' \to \sigma^1 \to \sigma^2 ... \to \sigma^n = \sigma^*$ with the following characteristics. First, for any $t = 1, 2, .., n$, $\sigma^t = \sigma_{kh}$ for some h and k. In other words, all agents in population 1 play the same pure strategy a_1^k and all agents in population 2 play the same pure strategy a_2^h. Second, in σ^t either a_1^k or a_2^h are best responses to σ^{t-1} for any $t > 1$. That is to say, σ^t can be reached from σ^{t-1} by having all agents in one of the two populations to receive simultaneously the learning draw, another event which occurs with strictly positive probabiliy under the unperturbed process $\Gamma(0, N)$. As a consequence, there is a non-zero probability that from any state like σ'' in which all agents in both populations play the same pure strategy one reaches an absorbing state σ^*. ■

This proposition shows that when playing a weakly acyclical game G^V, two populations of myopic players will sooner or later find their way to an equilibrium. In games with multiple equilibria, this theorem is of little use, however, as one would like to know which equilibrium is most likely to be selected. To answer this question, we have to turn to the perturbed process $\Gamma(\varepsilon, N)$ and to the notion of stochastic stability.

3.2 Stochastic stability

The idea behind stochastic stability has a fairly intuitive content, although the technical apparatus to make it rigorous is somewhat demanding. In a weakly acyclical game G^V the only absorbing states for the unperturbed process $\Gamma(0, N)$ are equilibria. In the perturbed process $\Gamma(\varepsilon, N)$ agents occasionally make mistakes. These mistakes produce occasional jumps from one equilibrium to another. The intuitive idea is that the larger the number of mistakes needed for a transition from one equilibrium to another, the less likely that transition will be. Some equilibria will be easier to reach than others, because they require fewer mistakes. Intuitively, an equilibrium σ^* will be stochastically stable if it is the easiest equilibrium to reach from all other equilibria.

We shall now make this idea precise. Consider two equilibria $\sigma^*, \sigma^{**} \in \mathcal{S}$. The *resistance* between σ^* and σ^{**}, $r(\sigma^*, \sigma^{**})$, is the minimum number of mistakes required to go from σ to σ^{**}. We say that a collection of directed edges over the set of equilibria $\mathcal{S}$ is a tree rooted at σ^* if for any other equilibrium $\sigma^{**} \in \mathcal{S}$ there is a single path from σ^{**} to σ^*. Let $T(\sigma^*)$ be a tree rooted at σ^* and $\mathcal{T}_{\sigma^*}$ the set of such trees. We say that the resistance of a tree is the sum of the resistances over its edges. The *stochastic potential* of an equilibrium is the minimum resistance over all trees rooted at σ^*. The following well known theorem, due to Young [20], characterizes the set of stochastically stable equilibria in terms of their stochastic potential.

Theorem 3 *Young [20], Theorem 4. An equilibrium σ^* is stochastically stable if and only if there is no other equilibrium σ^{**} with a lower stochastic potential.*

The following piece of notation will be useful in the rest of the paper. Let σ^* be an equilibrium. The *basin of attraction* of σ^*, $B(\sigma^*)$, is the set of states σ such that the unperturbed process $\Gamma(0, N)$ converges to σ^* with probability one. The *basin of potential attraction* of σ^*, $B^*(\sigma^*)$, is the set of states σ such that the unperturbed process $\Gamma(0, N)$ converges to σ^* with positive probability. The resistance between two equilibria $r(\sigma^*, \sigma^{**})$ is the minimum number of mistakes required to enter the basin of *potential* attaction of equilibrium σ^{**} starting from σ^*. To see this, consider that once the system has entered the basin of potential attraction of a new equilibrium σ^{**}, it can reach the equilibrium σ^{**} with positive probability even if no more mistakes intervene.

4 Reciprocal altruism

The applications we shall present in the next Section concern the particular specifications of the model presented above due to Rabin [16]. We shall briefly deal with Rabin's model, entering in some details only to show that it has a natural interpretation in terms of the two populations model presented above.

Let G^V be a psychological game played by pairs of agents drawn from population 1 and 2. Imagine that an agent in population i receives the learning draw, and observes the state of population $\sigma = (\sigma_1, \sigma_2)$. He then comes to believe that a fraction $\sigma_j(a_j^h)$ of population j chooses strategy a_j^h. He also assumes that each agent in population j will expect a fraction $\sigma_i(a_i^k)$ of population i to choose strategy a_i^k.

The first step in Rabin's analysis is to determine how kind it is to choose a strategy a_i^k after having observed σ. To this end, let $\Pi_j(\sigma_j)$ be the set of expected payoffs an agent in population i can "give" (on average) to agents in population j, when the state of that population is σ_j. In choosing his pure strategy a_i^k, the agent in population i "chooses" an expected payoff (for the agents in population j) within $\Pi_j(\sigma_j)$. Let π_j^h be the highest payoff in $\Pi_j(\sigma_j)$ and π_j^l the lowest payoff in $\Pi_j(\sigma_j)$, among those that are Pareto efficient.[2]

The *equitable payoff* for player j is simply the average between these two payoffs $\pi_j^e(\sigma_j) = \frac{1}{2}(\pi_j^h(\sigma_j) + \pi_j^l(\sigma_j))$. Player i's fairness in choosing pure strategy a_i^k is defined as the difference between the (expected) payoff he gives to members of population j and the equitable payoff: $f_i(a_i^k, \sigma_j) = \pi_j(a_i^k, \sigma_j) - \pi_j^e(\sigma_j)$. A similar definition holds for a player in the j population using a_j^h when the state of population i is σ_i.

In line with the assumption that players do not randomize, a player's fairness is only defined for pure, rather than mixed, strategies. However, when choosing a pure strategy a_i^k, an i agent knows that he will meet a j agent playing a_j^h with probability $\sigma_j(a_j^h)$. He also believes that a j agent who choose a_j^h knew that the state of the i population was σ_i and therefore that agent's fairness is $f_j(a_j^h, \sigma_i)$.

Rabin assumes that player's overall utility function is:

$$\begin{aligned} V_i(a_i^k, \sigma) &= \sum_{a_j^h \in A_j} \sigma_j(a_j^k)[\pi_i(a_i^k, a_j^h) + \alpha_i \ f_j(a_j^h, \sigma_i) \ f_i(a_i^k, \sigma_j)] \qquad (2) \\ &= \pi_i(a_i^k, \sigma_j) + \sum_{a_j^h \in A_j} \sigma_j(a_j^k) \ \alpha_i \ f_j(a_j^h, \sigma_i) \ (\pi_j(a_i^k, \sigma_j) - \pi_j^e(\sigma_j)) \\ &= \pi_i(a_i^k, \sigma_j) + \alpha_i \ f_j(\sigma_j, \sigma_i) \ (\pi_j(a_i^k, \sigma_j) - \pi_j^e(\sigma_j)) \end{aligned}$$

where $f_j(\sigma_j, \sigma_i) = \sum_{a_j^h \in A_j} \sigma_j(a_j^k) f_j(a_j^h, \sigma_i)$.

[2]Dufwenberg and Kirchsteiger [9] choose a different formulation, in which π_j^l is simply the minimum payoff in $\Pi_j(\sigma_j)$. This difference is immaterial in all the applications we propose below, and therefore we shall not discuss it further.

$V_i(a_i^k, \sigma)$ has the following interpretation. Consider the term inside the square bracket in the first equation. This is the payoff a player using a_i^k obtains in a match with an agent using a_j^h, when the state of the populations is σ. The first term, $\pi_i(a_i^k, a_j^h)$, is simply i's material payoff and therefore it does not depend on σ. The second term introduces fairness. $f_j(a_j^h, \sigma_i)$ is i's assessment of the fairness of an agent playing a_j^h. $f_i(a_i^k, \sigma_j)$ is i's own fairness, again under the hypothesis that he cannot observe a_j^h, but only σ_j. Finally, α_i is a parameter which weights fairness considerations with respect to material payoff.

The utility function $V_i(a_i^k, \sigma)$ models the idea that players are willing to reciprocate other people's nice and nasty behavior. This is shown clearly in the second equation above. Suppose that choosing a_j^h when the state is σ is a fair choice, so that $f_j(a_j^h, \sigma_i) > 0$. Then the material payoff of a j agent who plays a_j^h enters i's utility function with a positive weight. Otherwise, such a weight is negative. This means that i is willing to sacrifice part of his material payoff to increase the material payoffs of those j agents who employ nice strategies, and to reduce the material payoff of agents who choose nasty strategies.

5 Applications

This section presents three applications of the learning model discussed in Section 3.

5.1 The Battle of the Sexes

I consider the following parametrized version of the Battle of the Sexes game, with $k > 1$. S_2 is a player's most preferred strategy, and k measures how much it is preferred to S_1. This matrix represents player's "material payoffs". Players' extended preferences are assumed to be represented by equation 2. We assume that all agents have the same concern for fairness, represented by a parameter $\alpha = \alpha_1 = \alpha_2$.

	S_1	S_2
S_1	$0,0$	$1,k$
S_2	$k,1$	$0,0$

Consider the following open intervals: $I_1 = (0, \frac{2}{k^2})$, $I_2 = (\frac{2}{k^2}, \frac{2(2+k+k^2)}{k(k^3-1)})$, $I_3 = (\frac{2(2+k+k^2)}{k(k^3-1)}, 2k), I_4 = (2k, \infty)$. (To obtain less cumbersome results, we avoid dealing with border cases.)

Proposition 1 *If* $\alpha \in I_1$, $\mathcal{S} = \{\sigma_{12}^*, \sigma_{21}^*\}$. *If* $\alpha \in I_2 \cup I_3$, *then* $\mathcal{S} = \{\sigma_{12}^*, \sigma_{21}^*, \sigma_{22}^*\}$. *If* $\alpha \in I_4$ *then* $\mathcal{S} = \{\sigma_{12}^*, \sigma_{21}^*, \sigma_{22}^*, \sigma_{11}^*\}$.

This proposition states that when fairness considerations are sufficiently unimportant ($\alpha < \frac{2}{k^2}$), the only absorbing states of the process generated by the game G^V coincide with the strict Nash equilibria of the underlying Battle of the Sexes game. When α becomes sufficiently large ($\alpha \in (\frac{2}{k^2}, 2k)$), then a third absorbing state emerges, in which all agents in both populations use their favorite strategy S_2. Finally, if α becomes still larger ($\alpha > 2k$) a fourth absorbing state emerges, in which all agents play their least favorite strategy S_1.

This proposition is not new, as it is a simple consequence of Rabin's treatment of the Battle of the Sexes Game. The next proposition is more interesting, as it deals with stochastic stability.

Proposition 2 *If* $\alpha \in I_1 \cup I_2$*, then* $\breve{\mathcal{S}} = \{\sigma^*_{12}, \sigma^*_{21}\}$*. If* $\alpha \in I_3 \cup I_4$*, then* $\breve{\mathcal{S}} = \{\sigma^*_{22}\}$*.*

This proposition states that if players are sufficiently concerned with fairness ($\alpha > \frac{2(2+k+k^2)}{k(k^3-1)}$), the only stochastically stable equilibrium is the one in which all agents in both populations choose their most preferred strategy. The expected outcome is thus (S_2, S_2), which wouldn't be a NE if players were exclusively concerned with material payoffs.

Picture 1 illustrates the content of this proposition (it assumes that $k = 2$ and $\alpha_1 = \alpha_2 = \alpha = 1$). Every state of the two populations σ corresponds to a point in the square. σ^2_i represents the fraction of population i who employs strategy S_2. The shading of the different areas indicates the combination of strategies that are best responses to points in that area. For example, in the darkest area S_1 is a best response for agents in population 1 and S_2 is a best response for agents in population 2. In the lightest area the reverse is true. There are three equilibria σ^*_{12}, σ^*_{21} and σ^*_{22}. To see this, consider for example that the point in which all agents play S_2 ($\sigma^2_1 = \sigma^2_2 = 1$) belongs to the area with intermediate shading, in which playing S_2 is in fact a best response for agents belonging to both populations.

The gist of the proof is to show that the easiest transition out from the basin of attraction of each equilibrium involves only mistakes in one of the to populations. In particular, the segment r_{22} represents the easiest way out from the basin of attraction of equilibrium σ^*_{22}, while r_{12} is the simplest way out from the basin of attraction of equilibrium σ^*_{21}. The symmetrical segment, not shown in the picture, is the easiest way out from equilibrium σ^*_{12}. Proposition 2 establishes the conditions under which $r_{22} > r_{12}$ (in which case σ^*_{22} is stochastically stable) and $r_{12} > r_{22}$ (in which case both σ^*_{12} and σ^*_{21} are stochastically stable).

5.2 The Prisoner's Dilemma

I shall consider the following, normalized version of the PD. The gains for mutual cooperation are set equal to one, while mutual non cooperation yields zero. T is the extra gain for unilateral defection, while S is the cost of being the only cooperator.

	C	D
C	$1, 1$	$-S, 1+T$
D	$1+T, -S$	$0, 0$

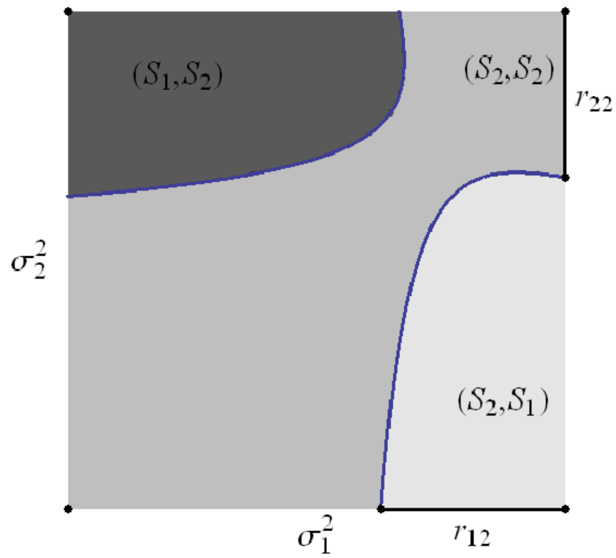

Figure 1: Basins of attraction for the Battle of the Sexes. $\alpha = 1$.

As in the previous section, agents are assumed to have preferences à la Rabin, and their concern for fairness is represented by a single parameter α. Let σ_i^C be the fraction of the population $i = 1, 2$ which chooses C. σ_{DD} is the state in which all agents in both populations play D, and σ_{CC} is the state in which all agents play C.

Proposition 3 *If $\alpha \leq \frac{2T}{(1+S)^2}$, then $\mathcal{S} = \{\sigma_{DD}\}$. If $\alpha > \frac{2T}{(1+S)^2}$, then $\mathcal{S} = \{\sigma_{DD}, \sigma_{CC}\}$. For any value of $\alpha \geq 0$ $\check{\mathcal{S}} = \{\sigma_{DD}\}$.*

The characterization of absorbing states is just a consequence of Rabin's treatment of the PD. If players are sufficiently motivated by reciprocity ($\alpha > \frac{2T}{(1+S)^2}$) there is an equilibrium in which all agents cooperate. The second part of this proposition deals with stochastic stability and it is original. It shows that, for any degree of concern for reciprocity, the cooperative equilibrium fails to be stochastically stable. The two populations will always spend longer periods of time in the equilibrium in which no cooperation is observed.

Figure 2 illustrates this proposition. r_C is the minimum number of mistakes needed in either population 1 or 2 to exit the basin of attraction of equilibrium σ_C and to enter the basins of potential attraction of σ_{DD}. r_D is the number of mistakes needed to leave the basin of attraction of the non-cooperative equilibrium σ_{DD} to enter the basin of potential attraction of σ_{CC}. The proof in the appendix shows that for any value of α, r_C is always shorter than r_D.

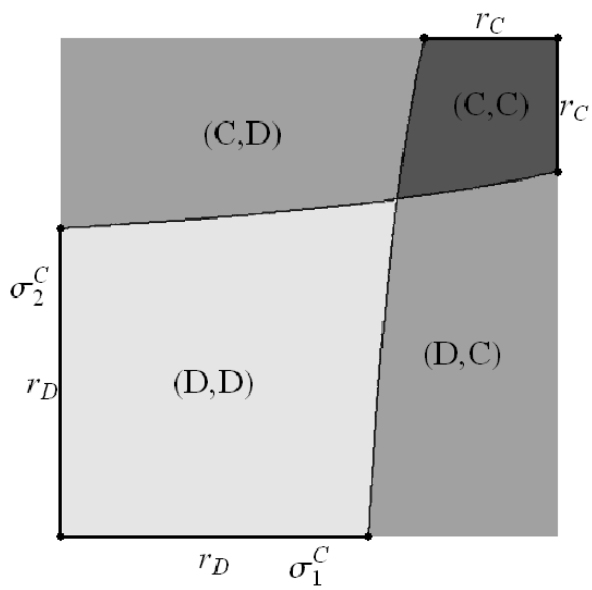

Figure 2: Basins of (potential) attraction for equilibria in the PD. $\alpha = 2$.

6 Conclusions

The model presented here lies at the interface between two areas of research that have rarely meet so far: "sophisticated" social preferences embedding psychologically plausible assumptions such as fairness and guilt, and social learning. It aspires to give a contribution to both. First, the model presented above might be a source of inspiration for scholars working on the empirical validation of the various hypothesis concerning "social" preferences. Rather than assuming that an equilibrium will be played in experiments as if individuals where fully rational, it would be more promising to investigate the stability properties of various equilibria. Although admittedly stochastic stability is not very accurate as a predictor of the way in which games are played in experiments, because it requires very long time spans before it becomes realistic, a knowledge of the size of the various basins of attraction of the existing

A Appendix

The following Lemma holds for any specification of the utility functions and greatly simplifies proofs.

Lemma 1 *Ler σ^*_{kh} be an equilibrium. For any σ such that either $a_1^k \in BR_1(\sigma)$ or $a_2^h \in BR_2(\sigma)$, $\sigma \in B^*(\sigma^*_{kh})$.*

Proof. Suppose that $a_1^k \in BR_1(\sigma)$ (the case in which $a_2^h \in BR_2(\sigma)$ can be treated symmetrically). If the process is in state σ and all agents in population 1 receive the learning draw, there is a positive probability that the in the following state σ' all agents in population 1 play a_i^k. (It requires that all agents who receive the learning draw will adopt the same best response). Since by hypothesis σ_{kh}^* is an equilibrium, $BR_2(\sigma') = a_2^h$ (because a_2^h is the only best response to a_1^k). If all players in population 2 receive the learning draw, the process will reach σ_{kh}^*. ■

A.1 Stochastic Stability in the Prisoner's Dilemma

We prove Proposition 3 by proving a certain number of claims. We prove the claims for agents in population 1, but symmetrical arguments also hold for population 2.

Claim 1 *For any value of $\alpha > 0$, and $p \in [0,1]$, $V_1(C, p, \frac{1}{2}) < V_1(D, p, \frac{1}{2})$.*

Proof. When a fraction p of Population 1 plays C, the equitable payoff would be $\pi_1^e(p) = \frac{1}{2}(\pi_1(p, C) + \pi_2(p, D))$. On the other hand, for any p, when population 2 is split equally between cooperators and not cooperators one has that $\pi_1(p, \frac{1}{2}) = \frac{1}{2}\pi_1(p, C) + \frac{1}{2}\pi_1(p, D)$, so that $\pi_1(p, \frac{1}{2}) = \pi_1^e(p)$, and hence $f_2(p, \frac{1}{2}) = 0$. It follows that $V_1(C, p, \frac{1}{2}) = \pi_1(C, \frac{1}{2}) < V_2(D, p, \frac{1}{2}) = \pi_2(D, \frac{1}{2})$. ■

Claim 2 *For $\alpha > \frac{1}{4}$, for any $p \in [0,1]$ $V_1(C, p, 1) > V_1(D, p, 1)$.*

Claim 3 *If $BR_i(\sigma) = \{D\}$, then $(p, q) \in B^*(\sigma_{DD})$*

Part 1 of Proposition is a consequence of Claim 2. In fact, by continuity of the utility functions $V_i(., ., .)$, if $\alpha > \frac{1}{4}$ there is an $\varepsilon > 0$ such that $V_1(C, p, 1-\varepsilon) > V_1(D, p, 1-\varepsilon)$ and (symmetrically) $V_2(C, 1-\varepsilon, q) > V_2(D, 1-\varepsilon, q)$. It follows that for any state σ such that $p > 1-\varepsilon$ and $q > 1-\varepsilon$ C is a strict best response. It follows that, in the absence of mutations, from all these states the system converges with probability one to the equilibrium σ_{CC}.

Part 2 of Proposition 3 is obtained as a combination of the three claims. From Claim 1 it follows that for any σ such that $p \leq \frac{1}{2}$ and $q \leq \frac{1}{2}$ D is a best response in both populations. As a consequence, if r_{DD} is the minimum number of mistakes that are necessary to leave the basin of attraction of equilibrium σ_{DD}, one has that $r_{DD} > \frac{1}{2}$ for any value of α. If $\alpha > \frac{1}{4}$ σ_{CC} is also an equilibrium. If a sufficiently large number of agents in one of the two populations switch from C to D, a state σ is reached such that D becomes a best response for agents in the other population. From Claim 3, it follows that such a stat σ belongs to $B^*(\sigma_{DD})$. Because of Claim 1, we know that if half of a population switches from C to D, D is in fact a best response for the other population. As a consequence, if r_{CC} it the minimum number of mistakes needed to leave the basin of attraction of σ_{CC}, $r_{CC} < \frac{1}{2}$ for any value of α. Therefore, $r_{DD} > \frac{1}{2} > r_{CC}$ which proves our proposition.

References

[1] Aumann, R. and A. Brandenburger, Epistemic conditions for Nash equilibrium, *Econometrica*, (1995) 63, 1161–1180

[2] P. Battigalli and M. Dufwenberg, Guilt in Games, *Am. Econ. Rev.*, 97 (2007) 170-176

[3] K. Binmore and L. Samuelson, Evolutionary drift and equilibrium selection. *Rev. Econ. Stud.* 66 (1999), 363–394

[4] K. Binmore, L. Samuelson and P. Young, Equilibrium selection in bargaining models, *Games and Economic Behavior*, 45 (2003), 296-328

[5] G. Bolton and A. Ockenfels, ERC: a theory of equity, reciprocity and competition, *Amer. Econ. Rev.* 90 (2000) 166–193.

[6] C. Camerer, *Behavioral Game Theory*, Princeton University Press, Princeton, NJ, 2003.

[7] Charness, G. and M. Dufwenberg, "Promises and Partnership." *Econometrica*, 74(6) (2006) 1579–1601.

[8] G. Charness and M. Rabin, Understanding social preferences with simple tests, Quart. J. Econ. 117 (3) (2002) 817–869.

[9] Dufwenberg, Martin and Georg Kirchsteiger, A theory of sequential reciprocity, *Games Econ. Behav.* 47 (2) (2004) 268–298.

[10] A. Falk, U. Fischbacher, A theory of reciprocity, *Games Econ. Behav.* 54 (2) (2006) 293–315.

[11] E. Fehr, K. Schmidt, A theory of fairness, competition, and cooperation, *Quart. J. Econ.* 114 (1999) 817–868.

[12] J. Geanakoplos, D. Pearce, E. Stacchetti, Psychological games and sequential rationality, *Games and Econ. Behav.* 1 (1989) 60–79.

[13] M. Kandori, G.J. Mailath and R. Rob, Learning, mutation, and long run equilibria in games. *Econometrica* 61 (1993) 29–56.

[14] Lewis, D. K. (1969) *Convention: a philosophical study*, Cambridge (MA) : Harvard University Press.

[15] D. Levine, Modelling altruism and spitefulness in game experiments, *Rev. Econ. Dyn.* 1 (1998) 593–622.

[16] M. Rabin, Incorporating fairness into game theory, *Amer. Econ. Rev.* 83 (5) (1993) 1281–1302.

[17] U. Segal and J. Sobel, Tit for tat: Foundations of preferences for reciprocity in strategic settings, *J. Econ. Theory* 136 (2007) 197 – 216.

[18] J. Sobel, Interdependent preferences and reciprocity, *J. Econ. Lit.* 43 (2) (2005) 396–440.

[19] Sugden, R. (1986) *The Economics of Rights, Cooperation and Welfare*, Oxford, Blackwell.

[20] P. Young, The evolution of conventions. *Econometrica* 61 (1993), pp. 57–84.

[21] P. Young, An evolutionary model of bargaining. *J. Econ. Theory* 59 (1993), pp. 145–168

[22] P. Young, *Individual Strategy and Social Structure*, Princeton Univ. Press, Princeton (1998).

Notation

G normal form game

A_i set of *pure* strategies for player $i = 1, 2$

m_i number of pure strategies for player i.

a_i^k generic element of A_i.

S_i mixed strategies set derived from A_i

s_i generic element of S_i.

$s_i(a_i^k)$ the probability with which a_i^k is played in s_i.

$\pi_i : S \to R$ player i's *material payoff function*.

$\pi_i(s_1, s_2)$ i's payoff when (s_1, s_2) is the outcome.

$s^* = (s_1^*, s_2^*) \in S$ the "context" of a game.

$V_i : S \times S \to R$ player's i preference over strategies

$V_i(a_i^k, s^*) = \pi_i(a_i^k, s_j^*) + \alpha_i(s^*)\pi_j(a_i^k, s_j^*)$

$\alpha_i(s^*)$ player i's "reciprocity coefficient"

$f_i(a_i^k, \sigma_j)$ player i's fairness coefficient

$\pi_i^e(\sigma_i)$ equitable payoff for player i

N number of individuals in each populations.

Σ_i is the set of possible states in which population i can be.

$\Sigma = \Sigma_1 \times \Sigma_2$ is the set of possible states

$\sigma = (\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2 = \Sigma$ state of the process

σ_i state of population i

$\sigma_i(a_i^k)$ fraction of population i that adopts strategy a_i^k in state σ_i.

λ probability of learning

$BR(\sigma) = (BR_1(\sigma), BR_2(\sigma))$ best responses to state σ.

ε probability of mutation

$\Gamma(\varepsilon, N)$ Markov chain over the set of states Σ

$\mu(\varepsilon)$ invariant distribution

$\mathcal{S}$ set of equilibria

$\check{\mathcal{S}}$ the set of stochastically stable equilibria.

σ_{hk} state in which all agents in population 1 adopt a_1^k and all agents in population 2 adopt a_2^h.

σ_{kh}^* σ_{kh} if it is an equilibrium

$B(\sigma^*)$ The *basin of attraction* of σ^*

$B^*(\sigma^*)$ The *basin of potential attraction* of σ^*