Learning in the Trust Game

Mark Howard

February 27, 2004

1 Introduction

This paper models the dynamics of a heterogeneous population of artificial agents playing the Trust Game. Individuals learn which strategies yield the higher payoff according to one of a number of different learning rules as suggested by Young (1998) . We include Fictitious Play, Reinforcement, Pavlovian and Imitation learning rules in the simulation. The main result is that the most widely used learning models in economics are unlikely to lead to any trust at all except by mistake given homogeneous agents. While heterogeneous learning is capable of sustaining significant levels of trust with minimal errors.

The purpose of the accompanying simulation is to model the intermediate term dynamics of a heterogeneous population of artificial agents playing the repeated Trust Game anonymously. We extend previous work on imitation and pavlovian learning in 2×2 games to $n \times n$ games. The expectation is that in small groups over the short run trust and trustworthiness can be achieved with sufficient numbers of conformist and pavlovian learners offsetting the effects of the statistical learners. In the long run we expect this population composition to favor statistical learners who can optimize their level of defection given the number of trustworthy types in the population.

Support for this approach has recently been given by Carpenter (2002) who compares the analytical approach that employs a standard learning dynamic and computes equilibria numerically to the agent based approach which simulates an environment with a finite population of interacting players. The simulation results suggest that the agent-based method for computing equilibria is substantially equivalent to the analytical/numerical approach.

Equilibrium theories and best response learning are relevant since they show the limit point of a learning process facilitate comparative statics. Learning would be trivial if all the strategic situations we faced had unique nash equilibria and if we learned quickly by whatever rule .[1] Where agents play a number of different games, some with multiple equilibria other with incomplete information so that learning the nature of the game and or the distribution of opponents types takes time, a behavioral theory of learning is needed. (Ho, Camerer and Chong 2001, pg.28)

1.1 The Trust Game

Trust has been widely herald as the most basic form of social capital and a precondition of the market itself. Trust is an crucial factor of economic exchanges since it lowers transaction costs. Arrow states that "there is an element of trust in every transaction" (1973, p. 23) and that trust is both necessary and sufficient for economic activity (1974). If agents trust each other they don't have to rely on costly enforcement in the form of contracting or monitoring to ensure that arrangements are adhered to. From a game theoretic perspective trust is surprising despite evidence in the lab and on main street to the contrary.

[1] For an introduction to learning models see Appendix One

The standard form of the Trust Game supposes that in each period a population of n agents are randomly paired and assigned the role of player one or player two so that each pair has one of each. Then each pair plays the following game: Player One is given $10, and instructed to pass any portion of this sum (whole dollars only) to the second player. This portion is tripled by the experimenter, before being passed on to the second player, who must then decide how much, if anything, to send back to the first player.

Various treatments specify the degree of reputation building or anonymity of the encounter. The Partners treatment calls for the game to be repeated with the same pairs for a given number of times with a certain probability. This allows for reputation effects. Under an the Perfect Strangers treatment the game is never repeated with the same pairs of players , so reputation-building and strategic play are ruled out. Here we use the strangers treatment with random matching where given a population of n agents the probability of playing the same player next round as this round is $1/(n\text{-}1)$.

The quantity sent by the first mover is evidence of trust, while the amount returned by the second mover is an sign of reciprocity, or trustworthiness. If agents are purely self interested, the second player will give back nothing, and realizing this, the first player will send nothing to the second player. So we expect there will be neither trust nor trustworthiness. Experimental results on human subjects indicate (Boyd and Gintis 2002) that both first and second movers propose positive amounts. Subjects are willing to trust when they expect a positive return. In a repeated Trust Game with changing partners Anderhub, Engelman and W.Guth (2002)have shown that if the population is composed of a sufficient portion of unconditional trust/trustworthy types then the remaining best responders optimal choice will be to reciprocate.

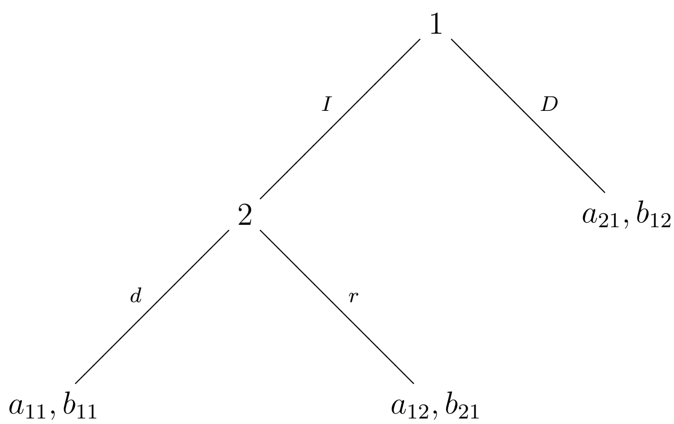

Figure 1: The extensive form of the trust game.

2 Literature Review

Evolutionary game theory has been widely used in an attempt to refine the equilibrium concept static stability with the ESS or dynamically with asymptotic stability. Samuelson and Zhang (1992) study the stability of replicator dynamics in asymmetric games and find that stability ensures only Nash equilibria will result. However they are not able to provide general conditions under which the backward induction equilibrium prevails. Recent work on certain classes of games has allowed for such condition to be specified for the reciprocity game as we show below.

Here we discuss the recent literature regarding the dynamics of various learning models and apply the results to the reciprocity game with a single type of learner - homogeneous learning. We begin with the replicator dynamic which has been widely studied and linked to three of the learning models presented here, fictitious play, cumulative reinforcement and directly bias imitation. Following Benaïm and Hirsch (1999) the average behavior of agents learning via stochastic algorithms in discrete time can be understood by looking at the related deterministic system in continuous time. So we start with continuous time a deterministic model of evolution in 2×2 games and move to discrete time stochastic learning models in two player games with more than two strategies.

The simplified reciprocity/investment/trust game has extensive and normal forms as in Figure 1 and Table 1 respectively where

$$\begin{gathered} a_{12} > a_{21} > a_{11} \\ \text{and} \\ b_{11} > b_{21} > b_{12} \end{gathered}$$

		Player 2	
		d	r
Player 1	I	a_{11}, b_{11}	a_{12}, b_{21}
	D	a_{21}, b_{12}	a_{21}, b_{12}

Table 1: The normal form of the reciprocity game

The only Nash equilibrium in pure strategies $\{D,d\}$ is also subgame perfect. In mixed strategies supposing player one plays I with probability β and player 2 plays d with probability α then then all α, β pairs such that $\beta = 0, \alpha \geq \frac{a_{12}-a_{21}}{a_{12}-a_{11}}$ are Nash.

Cressman (2003) provides analysis of continuous time replicator and best response dynamics for all 2×2 asymmetric(bimatrix) games. He shows that for the centipede game of length two (which is equivalent to the reciprocity game) the ordinary replicator dynamics converge to the Nash set mentioned above and that all interior trajectories of the best response dynamic converge to $\{D,d\}$ as shown below.

2.1 Replicator Dynamics

The two player two strategy replicator dynamic presented here follows Hofbauer and Sigmund (1998) and is based on the biological model where two large populations exist. In each round/generation each member of each population is randomly paired with a member of the other population and they play the reciprocity game above. Members of population one always play player ones role and likewise for population two. In biological terms the payoffs are the number of offspring. In both populations under this dynamic the frequency of a strategy increases when it has above average payoffs within it's own population. Therefor we associate a differential equation with each strategy in the game assuming the rate of increase over time $\dot{x}_1 / x_1$ in strategy one is equal to the difference between it's own payoff against y $(A\mathbf{y})_1$ and the average payoff against y $\mathbf{x} \cdot A\mathbf{y}$ in population one. The same is true for strategy two and both similarly for population two giving us

$$\dot{x}_i = x_i((A\mathbf{y})_i - \mathbf{x} \cdot A\mathbf{y})$$

$$\dot{y}_j = y_j((B\mathbf{x})_j - \mathbf{y} \cdot B\mathbf{x})$$

for i,j = 1,2

We can without losing anything add a constant to each column of A and B so we get zeros on the diagonals. And the resulting matrixes look like

$$A^{'} = \begin{bmatrix} 0 & a^{'}_{12} \\ a^{'}_{21} & 0 \end{bmatrix}, B^{'} = \begin{bmatrix} 0 & b^{'}_{12} \\ b^{'}_{21} & 0 \end{bmatrix}$$

$$a^{'}_{12} = a_{12} - a_{21} > 0$$

$$a^{'}_{21} = a_{21} - a_{11} > 0$$
$$b^{'}_{12} = b_{12} - b_{12} = 0$$
$$b^{'}_{21} = b_{21} - b_{11} < 0$$

In 2×2 games since $x_2 = 1 - x_1$ and $y_2 = 1 - y_1$ we only need to look at x_1 and y_1 which we simplify to x and y so that x is the percent that play I and y is the percent that play d and the replicator equations become

$$\overset{\bullet}{x} = x(1-x)(a^{'}_{12} - (a^{'}_{12} + a^{'}_{21})y)$$

$$\overset{\bullet}{y} = y(1-y)(-b^{'}_{21})x$$

If $b^{'}_{21}b^{'}_{12} \leq 0$ (which it is in this game since $b^{'}_{12} = 0$) one of the two strategies for player two dominates the other [2]and y is either constant or converges monotonically to 1 or 0. In this case y converges monotonically to 1 on the interior since all terms on the right side of population 2s replicator dynamic are positive. So all arrows in the interior of the phase space point

[2]See Hofbauer and Sigmund (1998) , Cressman (2000) and Cressman (2003)

north. Again y may not converge to 1 if $x \to 0$. This means if the proportion of games which reach player twos decision node approaches zero, there is less and less selection at work on population two and no reason to expect r to be driven to extinction. We also check on the interior for

1. $\dot{x} = 0$ which is true where

$$0 = (a_{12}^{'} - (a_{12}^{'} + a_{21}^{'})y)$$

or

$$y = \frac{a_{12}^{'}}{a_{12}^{'} + a_{21}^{'}}$$

2. $\dot{x} < 0$ where

$$y > \frac{a_{12}^{'}}{a_{12}^{'} + a_{21}^{'}}$$

so all arrows point west in this region.

3. $\dot{x} > 0$ where

$$y < \frac{a_{12}^{'}}{a_{12}^{'} + a_{21}^{'}}$$

so all arrows point east in this region.

On the boundary it is easy to check that when $x = 1$, $\dot{y} > 0$ and as noted above when $x = 0$, $\dot{y} = 0$ while when $y = 1$, $\dot{x} < 0$ and when $x = 0$, $\dot{y} > 0$.

2.2 Imitation Dynamics

According to Schlag (1998) his justification of imitation was the first microeconomic foundation for the replicator dynamic. Directly Biased imitation is

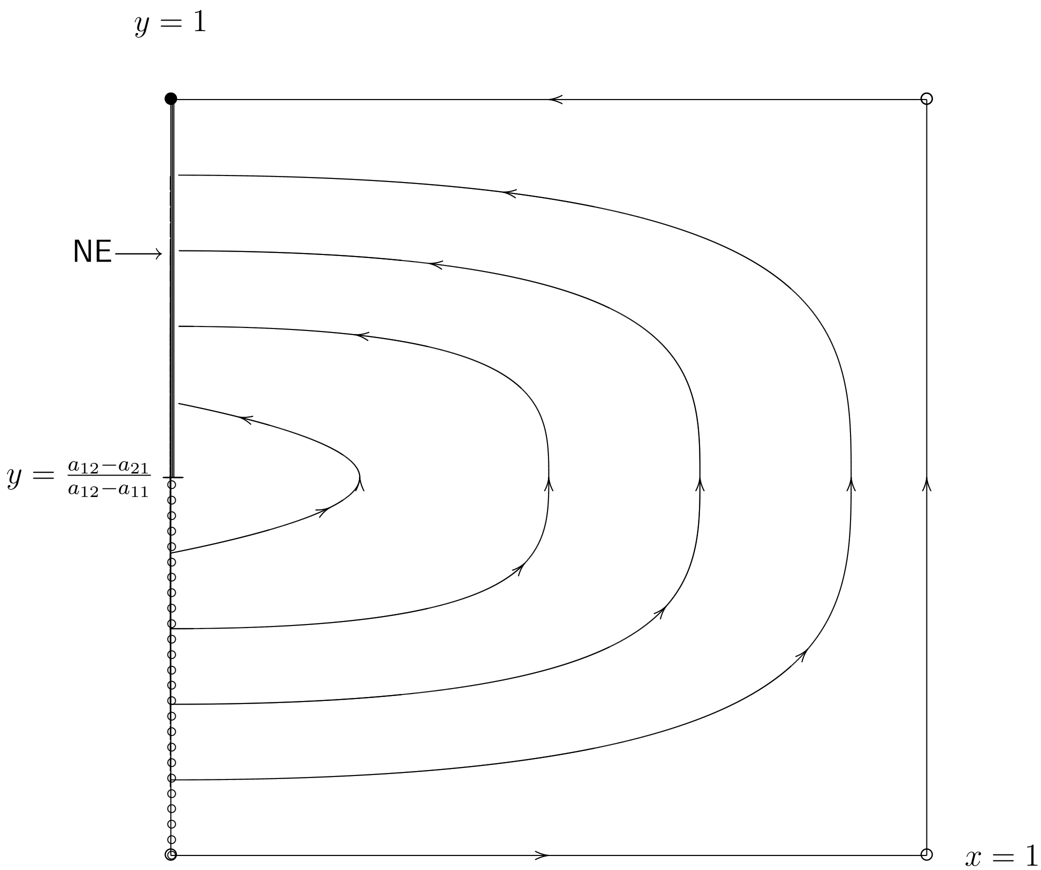

Figure 2: Phase diagram of the replicator dynamic for the reciprocity game.

imitation based on payoffs differences. Here we analyze Schlag's SPOR rule[3] which sequentially and proportionally imitates the strategy with the higher payoff in pairwise comparisons through a given sample. When all players use the SPOR rule actions yielding above average payoffs increase in frequency and those with below average payoffs decrease. The increase in play of a strategy in this case is a function of the product of it's frequency and of the

[3]Schlag(1999) analyzes imitation in the context of the multi armed bandit problem. He looks at various implementations of direct bias including imitate the best , imitate the best average and his own SPOR. He finds the SPOR rule to be best.

difference between it's own and the average expected payoff in the population. The resulting learning adjustment path is a discrete-time version of the replicator dynamic.[4] Following Benaïm and Hirsch (1999) the average behavior of agents using learning algorithms in discrete time can be understood by looking at the related system in continuous time. Therefore we can predict that the unperturbed SPOR imitation learning process will converge to the Nash set as shown in the phase diagram of the replicator dynamic for the reciprocity game in figure 2.

2.3 Simple Fictitious Play

The most widely study learning process among economist is that of fictitious play. In simple fictitious play (Fudenberg and Levine 1998, p29) players believe they are playing against a fixed but unknown distribution of opponent player types. What is is learned is the distribution of opponents types. Players then play the best response to this empirical distribution. Fictitious play (Brown 1951) was introduced as a method of calculating equilibria in finite zero-sum games. The idea is to fictitiously

> visualize two statisticians perhaps ignorant of min-max theory, playing many plays of the same discrete zero-sum game. One might naturally expect a statistician to keep track of the opponents past plays and, in the absence of a more sophisticated calculation, perhaps to choose at each play the optimum pure

[4] Increasing the sampling size in the SPOR method speeds up the dynamic. As the sample size goes to infinity the learning dynamic turns into the adjusted replicator dynamic of Maynard Smith (1982).

strategy against the mixture represented by all the opponents past plays (Brown 1951, pg.374).

Brown argued that no matter where they began, these two fictitious players would iterate towards the value (nash equilibrium) of the game.[5] Brown's iterative method utilizing fictitious play was subsequently argued as a theory of learning. As in Brown's method the simple fictitious play learning model adds one to the weight of a strategy each time it is used by an opponent. As a learning model simple fictitious play makes the prediction that in games with a unique subgame perfect nash equilibrium such as the reciprocity game, play will converge in to this nash.

The continuous time counterpart to fictitious play is the bimatrix best response dynamic which is shown in figure 3. We analyze the dynamics as follows. If a player two estimates any of the player ones will invest then his best response will be to return zero (play d) so all arrows on the interior point north. While if player one estimates the number of player twos who don't reciprocate is less than $\frac{a_{12}-a_{21}}{a_{12}-a_{11}}$ then his best response is to invest while if the number of player twos who do reciprocate is less than $\frac{a_{12}-a_{21}}{a_{12}-a_{11}}$ then his best response is to defect.

2.4 Cumulative Reinforcement

Following the simple form of the reinforcement model (Erev and Roth 1995), assume each player has some initial propensity $R_i^k(0)$ to play each of his m

[5](Robinson 1951) Robinson (1951) proved convergence in the limit - as the number of periods goes to infinity.

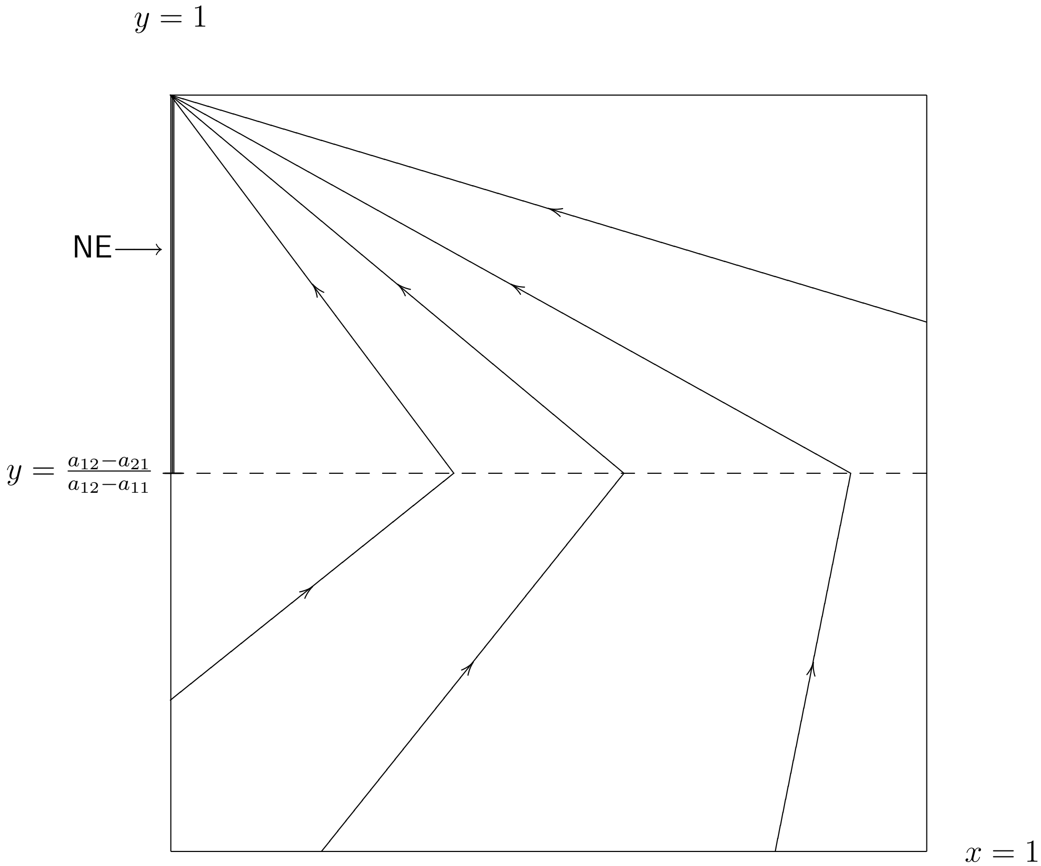

Figure 3: Phase diagram of best response dynamic for the reciprocity game.

strategies before any experience has occurred . Player i's propensity to play j is updated so that the propensity to play j in the next period is,

$R_i^j(t) = R_i^j(t-1) + \pi_i^i(s_i(t), s_{-i}(t))$

Let the probability that player i uses his jth pure strategy in period$t$ be,

$P_i^j(t) = \frac{R_i^j(t)}{\sum_{k=1}^{m} R_i^k(t)}$

where the sum is over all of player i's m strategies.

Hopkins and Posch(2002) argue that cumulative reinforcement learning in 2×2 games converges to the nash equilibrium set. They prove this for

$$
\begin{array}{cccc}
a_{11}, b_{11} & a_{12}, b_{21} & \bullet & a_{1m}, b_{m1} \\
a_{21}, b_{12} & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
a_{n1}, b_{1n} & \bullet & \bullet & a_{nm}, b_{mn}
\end{array}
$$

rescaled partnership games and claim that Beggs (2002) does so for rescaled zero-sum games.[6] Beggs (2002) shows that the Ever and Roth (1998) cumulative reinforcement learning (CRL) model has long run behavior described by adjusted replicator dynamics of Smith (1982). In 2×2 games this yields the same behavior as the ordinary replicator equations above for this game. [7]He shows that if all players use cumulative reinforcement then the average payoff approaches the value of the game. He points out that convergence of strategies in such games can be very slow under CRL while convergence of average payoffs is much faster. He proves that if the game has either a unique pure strategy equilibrium of a unique mixed strategy equilibrium then if both players use CRL then play converges to it. He also shows that in 2×2 zero-sum games for any positive level of forgetting in CRL converges to a pure strategy combination.[8]

2.5 Stochastic Learning Dynamics

Hofbauer and Sigmund (1998) pp 128 assuming the payoff matrix above shows that such bimatrix games are rescaled zero-sum games if $c < 0$ and there exist suitable c_{ij}, c_j and d_i and such that

$$a_{ij} = c_{ij} + c_j, b_{ji} = cc_{ij} + d_i$$

for all i= 1 to n and $j = 1$ to m

for the trust game above let

$$c_1 = (-a_{11} - b_{11}), c_2 = (-a_{12} - b_{21})$$

$$d_i = 0 \text{ for all } i \text{ and } c = -1$$

so the reciprocity game above is a rescaled zero-sum game [9].

Following Benaïm and Hirsch (1999) , Hofbauer and Hopkins (2001) predict the behavior of smooth fictitious play by analyzing the associated continuous time perturbed best response dynamics. They find that for any rescaled zero-sum game like the reciprocity game, stochastic fictitious play converges to the perturbed best response equilibria[10]. They show that in a zero sum game with two players each player has a unique equilibrium strategy or is

[6] 2×2 games are either rescaled zero-sum or rescaled partnership games (Hofbauer and Sigmund 1998).

[7] In particular the set of rest points is the same but their stability isn't always. Hopkins and Posch(2002) show that in 2×2 games the system is unstable under the ordinary replicator dynamic if and only if it is unstable under the adjusted version. Therefor we can use the simpler version presented above.

[8] He points out that while the rescaling process leaves rational agents choices unchanged this is not necessarily true for CRL unless initial reinforcements are rescaled accordingly.

[9] This result is easily extended to the $N \times N$ reciprocity game.

[10] This is equivalent the quantal response equilibria of of McKelvey and Palfrey (1995).

indifferent between strategies which guarantee a payoff equal to the value of the game. Adding noise breaks this indifference. They show that the expected motion of SFP is a discrete time representation of the perturbed best response dynamics. In other words the mean valued ODE corresponding to smooth fictitious play is the perturbed best response dynamics. They prove in Theorem 3.2 that under the perturbed best response dynamic in a two person rescaled zero-sum game there is a unique rest point which is globally asymptotically stable. In Theorem 5.1 they show for any two person rescaled zero-sum game smooth fictitious play converges to the unique perturbed equilibrium. [11]

This result along with Hopkins(2001) results for 2×2 games allows us to say both smooth fictitious play and cumulative reinforcement learning with experimentation will converge to the unique perturbed equilibrium associated with a given perturbation function in 2×2 games.

2.6 Perturbed Replicator Dynamics

Hart (2002) investigates the connection between replicator dynamics and backward induction in extensive form games and finds that the subgame perfect equilibrium becomes the only stable outcome if the mutation rate is low enough and the populations large enough as long as the expected number of mutations per generation is not zero). Hart has added the large population part. This suggests that those learning models whose dynamics

[11] The precise equilibrium selected depends on the perturbation function.

have been linked to replicator dynamics such as fictitious play, cumulative reinforcement and direct bias imitation will converge in homogeneous populations given low but positive error rates and large populations? The key in this game is that there be sufficient mutation/error by the player one's so that positive amounts are sent thereby allowing selection/learning in population 2. This is no problem as we move to our stochastic learning models where we get convergence to the perturbed equilibrium associated with a given perturbation function.

Hopkins (2001) looks at smooth fictitious play and cumulative reinforcement learning with experimentation and shows these stochastic discrete time models can be analyzed with their corresponding continuous time model. The same approach is taken by (Fudenberg and Levine (1998) and in analyzing stochastic fictitious play) He argues both models can be seen as noisy versions of the replicator dynamic with identical local stability. This means if they converge they will converge to the same point.

This equilibrium can be calculated following Hopkins (2001) by solving the following perturbed replicator dynamics,

$$\dot{x} = R(x)Ay + \frac{1}{\lambda}g(x)$$

given the perturbation function

$$g(x_n) = R(x_n)\phi'(x_n) = \begin{vmatrix} x(1-x) & -x(1-x) \\ -x(1-x) & x(1-x) \end{vmatrix} \cdot \begin{vmatrix} \frac{1}{x} \\ \frac{1}{1-x} \end{vmatrix} = \begin{vmatrix} 1-2x \\ 2x-1 \end{vmatrix}$$

and similarly for $\dot{y}$

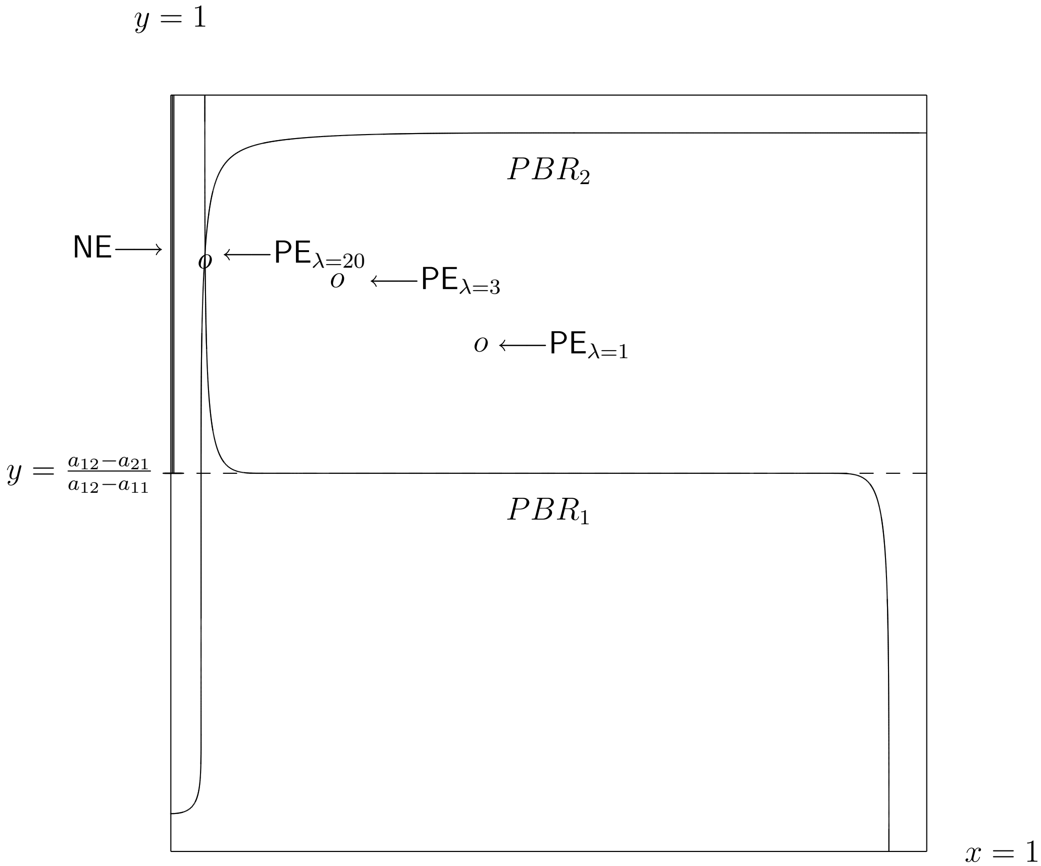

Figure 4: Perturbed Best Response Functions

$$\dot{x} = x(1-x)(a'_{12} - (a'_{12} + a'_{21})y) + \frac{1-2x}{\lambda}$$

$$\dot{y} = y(1-y)(-b'_{21})x + \frac{1-2y}{\lambda}$$

we can solve for $\dot{y} = \dot{x} = 0$ given the payoff matrix in table 2 resulting in

$$y = \frac{1}{2} - \frac{2x-1}{4\lambda x(1-x)}$$

and

$$x = \frac{2y-1}{4\lambda y(1-y)}$$

solving these simultaneously for different λ results in the various perturbed equilibrium in table 3

		P2 d	P2 r
P1	I	$0, 8$	$4, 4$
	D	$2, 2$	$2, 2$

Table 2: Payoffs associated with table 3

λ	x	y
20	.042	.785
10	.0815	.78
5	.152	.7705
4	.183	.764
3	.223	.755
2	.302	.735
1	.413	.686

Table 3: Perturbed equilibria for different λ

What these results say is that unless we turn the error rate way up all learning representable via the replicator dynamic will converge in homogeneous populations to no trust at which time selection no longer operates on the second population. Our simulations reveal that allowing two type of learners can result in trust becoming profitable.

3 Simulating the Trust Game

We use the agent based approach. Axelrod (1997) has used this approach to explore the foundations of cooperation in a wide variety of contexts (for a recent review of Axelrods contribution see Hoffman (2000).Support for this approach has recently been given by Carpenter (2002) who compares the analytical approach that employs a standard learning dynamic and computes equilibria numerically to the agent based approach which simulates an environment with a finite population of interacting players. The simulation results suggest that the agent-based method for computing equilibria is substantially equivalent to the analytical/numerical approach.

We simulate the Trust Game in it's normal form one issue in doing so is that while player one has 11 pure strategies player two has way to many . Therefore we will used an expanded simplified version of the game (Gunnthorsdottir, McCabe and Smith 2002) where player 1 is given 2 dollars with the game otherwise being the same. Player one now has 3 pure strategy, player two 28.

3.0.1 Simulating Heterogeneous Learning

To represent statistical learners we use the Experience Weighted Attractors model (Ho et al. 2001)[12]. This model has a number of advantages including the fact that it incorporates both fictitious play and reinforcement learning, two of the most widely used learning models in economics. This allows for

[12] see Appendix One

		Player 1		
		0	1	2
	00	2,2	5,1	8,0
	01	2,2	5,1	7,1
	02	2,2	5,1	6,2
	03	2,2	5,1	5,3
	04	2,2	5,1	4,4
	05	2,2	5,1	3,5
	06	2,2	5,1	2,6
Player 2	10	2,2	4,2	8,0
	11	2,2	4,2	7,1
	12	2,2	4,2	6,2
	13	2,2	4,2	5,3
	14	2,2	4,2	4,4
	15	2,2	4,2	3,5
	16	2,2	4,2	2,6
	20	2,2	3,3	8,0
	21	2,2	3,3	7,1
	22	2,2	3,3	6,2
	23	2,2	3,3	5,3
	24	2,2	3,3	4,4
	25	2,2	3,3	3,5
	26	2,2	3,3	2,6
	30	2,2	2,4	8,0
	31	2,2	2,4	7,1
	32	2,2	2,4	6,2
	33	2,2	2,4	5,3
	34	2,2	2,4	4,4
	35	2,2	2,4	3,5
	36	2,2	2,4	2,6

Table 4: Trust Game, Note:Payoffs = (Player2,Player1)

subsets of agents, some purely belief learners, others purely rote learners, as well as types who employ a mixture of the two, to be represented by the EWA model with different parameters specifications. Additionally The program provides for two types of imitators. One follows Schlag (1999) which sequentially and proportionally imitates the strategy with the higher payoff. The other following Boyd and Richerson, (Boyd and Richerson 1985) where the probability of imitating a given strategy is a (increasing or decreasing) function of the frequency of that type in the population. Finally we also permit agents who follow one of two pavlov types. One we term Simpleton follows Malawski and the other the win-stay lose-shift (WSLS) (Posch 1997, Kraines and Kraines 1995) learning rule. To win in this case is to achieve a given aspiration level. The simulation allows for the case where the aspiration level is fixed and where it updates according to payoffs in the previous rounds. The generalizations of conformist and pavlovian learning rules from the two strategy (dichotomous trait) case to the n strategy case are provided in appendix one.

3.0.2 Program Specification

We are interested in both the intermediate term dynamics as well as the long term properties of repeated games. Within a generation an agent may alter her strategy based on her learning rule but that rule does not change. The frequency of each distinct type of learner will not change over a single generation. The distinct learning rules will be the target of selection across generations.

This simulation is written in Borland C++ Builder and opens to a window where the user sets the initial parameters of the simulation. The user chooses the types of learners to include and their initial distribution. The program then graphically display both the intermediate term dynamics and the long term evolution of learning strategies. To this end one window displays the average and distribution of the amount sent, the amount returned and payoff to each of the types of learners in each round. Another window will show the proportion of each of the types in the population over generations. The program creates an output file to which to which it writes the initial parameters of the simulation, amount each individual sent or returned, and total payoff as well as each groups statistics for each round and the proportion of each type in the population in every generation.

3.0.3 Top Level Algorithm

1. Initialize
    (a) Prompt for and Error Check User Input to set Global Variables and the output file name
    (b) Set global Variables - random seed, number of generations, number of rounds per generation, and output file name
2. Create agents
    (a) Prompt for and Error Check User Input to Add Agents
    (b) Create each new type and add to a list of types
    (c) Create and add agents of each type to the population
3. Initialize output
4. Play game

(a) For G = 1 to Number of Generations Do

   i. For R = 1 to Number of Rounds per Generation Do

      A. Pair Players
      B. Update Payoffs (Play Game)
      C. Update Strategies
      D. Update Output

3.1 Early Results-Homogeneous Populations

We see a substantial amount of trust and trustworthiness in both homogeneous and heterogeneous population.

3.1.1 Simple Fictitious Play

As we would expect a homogeneous population of stochastic simple fictitious play types moves to the subgame perfect nash of this game in about 10 periods. Interestingly once this is achieved player two's are increasingly nearly indifferent between all their strategies so the average sent back climbs to above one third and the best response for the first player one's whose estimates pick this up becomes to send something. So they send positive amounts and those player two's who receive these positive amount's while initially nearly indifferent(and sending something back) learn to send nothing back and we get a little cycle producing low but positive levels of trust/trustworthiness. You could squeeze this out by cranking up lambda the sensitivity parameter.

3.1.2 Cournot Best Response

Given an uniform initial distribution of attractors we see continuous switching as the best response to the last opponents move creates cycles. In the two player case if player one sends 2 then player twos best response is to play A - 00,10,20 or 30 with equal probability. The best response to A will be 0 with prob .375 and 1 with probability .625. If it (the best response to A) is 1 the best response from two is to play 01-06 with equal probability. The best response to which will be 0 with probability 5/12 and 2 with probability 7/12, If the best response to A is 0 the best response from two is to randomly select a strategy. The best player one response to a randomly selected strategy will be to send 2 with probability 0.529762 and 1 with probability 0.333333 and therefore 0 with probability .137 so it's unlikely we get stuck at 0 for long. And so we get almost constant switching from both types. As the group size increases individuals are still switching while the average sent shows a cycle slightly above one.

3.1.3 Reinforcement

3.1.4 Pavlov

3.1.5 Simpleton

3.1.6 SPOR

3.1.7 FDB

References

Anderhub, V., D. Engelman, and W.Guth, "An Experimental Study of the Repeated Trust Game with Incomplete Information," *Journal of Economic Behavior and Organization*, 2002, *48*, 197–216.

Arrow, Kenneth J, *Information and Economic Behavior*, Federation of Swedish Industries, 1973.

Arrow, Kenneth J., *The Limits of Organization*, New York, NY: W.W. Norton., 1974.

Axelrod, Robert, *The Complexity of Cooperation*, Princeton University Press., 1997.

Beggs, A.W., "On the Convergence of Reinforcement Learning," March 2002. University of Oxford Dept. of Economics Discussion Paper.

Benaïm and Hirsch(1999), "Mixed Equilibria and Dynamical Systems Arising from Fictitious Play," *Games and Economic Behavior*, 1999, *29*, 36–72.

Boyd, Robert and Herbert Gintis, "Final Funding Request for the Preferences Network Plans for Future Research," 2002. Final Funding Request for the Preferences Network Plans for Future Research.

——— **and Peter Richerson**, *Culture and the Evolutionary Process*, Chicago: The University of Chicago Press, 1985.

Brown, G. W., "Iterative Solution of Games by Fictitious Play," in T. J. Koopmans, ed., *Activity Analysis of Production and Allocation*, New York: Wiley, 1951, pp. 374–376.

Carpenter, Jeffery P., "Evolutionary Models of Bargaining: Comparing Agent-Based Computational and Analytical Approaches to Understanding Convention Evolution," *Computational Economics*, 2002, *19*, 25–49,.

Cressman, Ross, "Subgame Monotonicity in Extensive Form Games," *Games and Economic Behaviour*, 2000, *32*, 183–205.

——, *Evolutionary Dynamics and Extensive Form Games*, Cambridge Mass.: MIT Press, 2003.

Erev, Ido and Alvin E. Roth, "On the need for low rationality, cognitive game theory: Reinforcement learning in experimental games with unique, mixed strategy equilibria," 1995. mimeo.

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

Gunnthorsdottir, Anna, Kevin McCabe, and Vernon Smith, "Using the Machiavellianism Instrument to Predict Trustworthiness in a Bargaining Game," *Journal of Economic Psychology*, 2002, *23*, 4966.

Hart, Sergiu, "Evolutionary Dynamics and Backward Induction," *Games and Economic Behavior*, 2002, *41*, 227–264.

Ho, Teck-Hua, Colin F. Camerer, and Juin-Kuan Chong, "Economic Value of EWA Lite: A Functional Theory," *Cal Tech Social Science Working Paper*, 2001, *1122.*

Hofbauer, Josef and Ed Hopkins, "Learning in Perturbed Asymmetric Games," October 2001. Working Paper.

Hofbauer, Joseph and Karl Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge, UK: Cambridge University Press, 1998.

Hoffmann, Robert, "Twenty Years on: The Evolution of Cooperation Revisited," *Journal of Artificial Societies and Social Simulation*, 2000, *3* (2).

Hopkins, Ed, "Two Competing Models of How People Learn in Games," *NAJ Economics Peer Reviews of Economics Publications*, 2001, *1.*

—— **and Martin Posch**, "Attainability of Boundry Points under Reinforcement Learning," 2002. Working Paper.

Kraines, David and Vivian Kraines, "Evolution of Learning Among Pavlov Strategies in a Competitive Environment with Noise," *The Journal of Conflict Resolution*, 1995, *39* (3), 439–466.

McKelvey, R.D. and T.R. Palfrey, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behaviour*, 1995, *10*, 6–38.

Posch, Martin, "Win Stay - Lose Shift an Elementary Learning Rule for Normal Form Games," 1997. Working Paper 97-06-056 E, Santa Fe Institute.

Robinson, Julia, "An Iterative Method of Solving a Game," *The Annals of Mathematics*, 1951, *54*, 296–301.

Samuelson, Larry and Jianbo Zhang, "Evolutionary Stability in Asymmetric Games," *Journal of Economic Theory*, 1992, *57*, 363–391.

Schlag, Karl, "Justifying Imitation," 1998. Working Paper.

—— , "Which One Should I Imitate?," *Journal of Mathematical Economics*, 1999, *31*, 493–522.

Young, H. Peyton, *Individual Strategy and Social Structure*, Princeton University Press, 1998.