Bonn Econ Discussion Papers

Discussion Paper 8/2001

Cooperation as a Result of Learning with Aspiration Levels

by

Jörg Oechssler

February 2001

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Cooperation as a Result of Learning with Aspiration Levels

Jörg Oechssler*
Department of Economics
University of Bonn

February 15, 2001

Abstract

It is shown that a win–stay, lose–shift behavior rule with endogenous aspiration levels yields cooperation in a certain class of games. The aspiration level in each round equals the current population average. The class of games includes the prisoner's dilemma and Cournot oligopoly and thus yields an explanation for cooperation and collusion.

JEL– classification numbers: C72, L13.

*Adenaueralle 24, 53113 Bonn, Germany, E–mail: oechssler@uni-bonn.de

1 Introduction

Psychological research as well as introspection strongly suggest that individuals' learning behavior depends on comparisons with the experiences of other individuals. It seems odd, therefore, that one of the most studied type of learning theories, reinforcement learning, depends only on one's own accrued payoffs. One way of reconciling reinforcement learning with social learning is to include an aspiration level relative to which strategies are reinforced and to make this aspiration level dependent on the payoff experiences of other individuals. The idea that aspiration levels should be endogenous is not new, however most authors considered endogenous aspiration levels that depend on individual payoff histories only (e.g. Karandikar et al., 1998; Posch et al., 1999; or Börgers and Sarin, 2000).[1]

In an interesting recent paper Dixon (2000) considers a continuum of duopoly markets in which firms base their aspiration levels on the industry's average profit level. He shows that in a certain class of games that include the prisoner's dilemma and standard Cournot duopolies, collusion will be the inevitable outcome of the following simple win–stay, lose–shift learning process.[2] When firms make profits that are above the average industry profit, i.e. above their aspiration level, they are satisfied and choose the same action again in the following period. If, however, the firm's profit is below its aspiration level, the firm will experiment with arbitrary actions.

In this note I simplify and generalize this insightful result by Dixon (2000) using standard stochastic stability analysis (see e.g. Kandori, Mailath and Rob, 1993). There are several differences to Dixon's (2000) paper. For example, I use a different methodology, namely stochastic stability analysis, which assumes that all players experiment with small probability, whereas Dixon assumes that only players below their aspiration level experiment

[1] However, in their discussion, Posch et al. (1999) mention a process YESTERMAX, which assumes an aspiration level equal to the maximum of both players' payoffs, and argue that this process yields efficiency in most 2×2 games.

[2] Win–stay, lose–shift learning processes were originally introduced already by Thorndike (1911).

with different actions. An advantage of stochastic stability analysis is that it allows (in this case) to give a very simple proof and in particular, it allows to generalize Dixon's result to the case of symmetric n–person games that satisfy a similar condition as the one used by Dixon. The method also shows that Dixon's assumption of an infinite number of markets is not necessary for the results to hold. Most importantly, it allows to dispense with Dixon's assumption that there exists already at the beginning a strictly positive mass of markets in which the efficient outcome is being played.

Results like these show how crucial it is to be precise about *who* is being imitated in learning processes. In the context of a Cournot oligopoly imitation results range from perfectly competitive outcomes to collusive ones depending on the type of imitation assumed. When the most successful action of direct opponents is imitated, competitive outcomes may result as shown by Vega–Redondo (1997). When the most successful actions of firms, who are in the same role but in a different market, are imitated (with a certain probability), then an imitation process behaves like the evolutionary replicator dynamics and converges to the Cournot–Nash equilibrium (see Schlag, 1998). Finally, Dixon (2000) and the current paper show that imitation may result in collusion if the population average is imitated in the form of an aspiration level.

2 Aspiration learning

Consider a symmetric normal form game with player roles $i = 1, ..., n$, a finite action set A for each player, and payoff functions $\Pi_i(a)$, with $a \in A^n$. This game is played in a finite number ($k \geq 2$) of groups, locations, or markets by the same fixed set of players in each time period $t = 0, 1, 2, ...$.

I assume that players use the following rule to determine their choice of action. In period 0 players choose some arbitrary action. In all subsequent periods players form an aspiration level which is given by the average payoff in the population at time t, $\bar{\Pi}^t$. If a player's payoff in period t is at least as high as this average payoff, he is satisfied and chooses with probability $1 - \varepsilon$

the same action again in $t + 1$. With some small probability ε he trembles and chooses an arbitrary action. If a player's profit falls short of $\bar{\Pi}^t$, he is unsatisfied and experiments with an action from A according to a fixed probability distribution with full support.

Below I will consider the class of games whose joint payoff functions have a unique, symmetric, local maximum. The joint payoff function is simply the sum of the payoffs of all n players, $\Pi(a) := \sum_{i=1}^{n} \Pi_i(a)$. I will say that the joint payoff function has a unique, symmetric, local maximum a^c, with $a_i^c = a_j^c, \forall i, j$, if

$$\left[\Pi(a) \geq \Pi(a_i', a_{-i}), \forall i \text{ and } a_i' \neq a_i\right] \Leftrightarrow a = a^c. \tag{1}$$

Thus, the modifier "local" refers to the requirement that the joint payoff at each symmetric action vector (other than a^c) can be improved upon through a *unilateral* deviation by some player. Note that the unique local maximum a^c must also be a global maximum. Clearly, a sufficient condition for (1) is for $\Pi(\cdot)$ to be symmetric and strictly concave in a. In bi–matrix games condition (1) simply requires that there exists a unique cell on the diagonal in which the sum of payoffs is higher than in any other cell of its respective row and column. For example in 2×2 games

	a_1	a_2
a_1	α	β
a_2	γ	δ

two types of game belong to this class. (i) Prisoner's dilemma type games ($\beta > \delta > \alpha > \gamma$) if $2\delta > \gamma + \beta > 2\alpha$. And (ii) games with a strictly dominant and efficient equilibrium, i.e. if $\alpha > \gamma, \beta > \delta$, $2\alpha > \beta + \gamma$. That is, to the class belong games in which the cooperative, joint payoff maximizing outcome is a Nash equilibrium outcome but also games in which this is not the case.

Another fitting example is a symmetric oligopoly with profit functions that are strictly concave in the own action a_i and concave in a. Many related games like public good games, resource extraction games etc. also satisfy condition (1).

Proposition 1 *Consider a game whose joint payoff function has a unique local maximum a^c. Then, the limit distribution for $\varepsilon \to 0$ puts probability one on the state s^c, in which the joint payoff maximizing action is chosen by all players.*

Proof. The assumptions define a Markov process on a finite state space. For $\varepsilon > 0$ the process is irreducible and aperiodic and, therefore, has a unique stationary distribution. Formally, we consider the limit distribution for $\varepsilon \to 0$. For $\varepsilon = 0$ the process may have several absorbing sets.[3] By standard arguments (see e.g. Samuelson, 1994) only the members of absorbing sets of the unperturbed process can appear in the support of the limit distribution.

I will show first that all absorbing sets other than s^c can be destabilized by a single tremble, which puts the process in the basin of attraction of s^c. Then, I will show that one tremble is not sufficient to leave the basin of attraction of s^c, which implies that s^c is the unique stochastically stable state.

Clearly, a state is absorbing if and only if all players receive the same payoff, and no state in which players receive different payoffs can be part of an absorbing set. Consider any absorbing state other than s^c with a corresponding joint payoff $\Pi(a)$. Suppose there is a tremble by one player j in some group to an action a'_j such that

$$\Pi(a'_j, a_{-j}) > n\bar{\Pi} > \Pi(a).$$

Such an action a'_j exists due to the fact that there is a unique local joint payoff maximum. Thus all players in the remaining groups are below the population average payoff. Hence, with positive probability an entire group will switch to playing a^c and subsequently all other groups must follow suit. To leave the basin of attraction of s^c, however, a tremble is required in *every* group because otherwise the process would return to s^c. ■

[3] A set of states is called absorbing if there is zero probability to exit the set and a positive probability of moving from any state in the set to any other state in the set in finite time.

References

[1] Börgers, T. and Sarin, R. (2000), "Naive Reinforcement Learning with Endogenous Aspirations", *International Economic Review,* 41, 921-950.

[2] Dixon, H. (2000), "Keeping up with the Joneses: Competition and the Evolution of Collusion", *Journal of Economic Behavior & Organization,* 43, 223-238.

[3] Kandori, M., Mailath, G., and Rob, R. (1993). "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica*, 61, 29-56.

[4] Karandikar, R., Mookerherjee, D., Ray, D., and Vega–Redondo, F. (1998), "Evolving Aspirations and Cooperation", *Journal of Economic Theory*, 80, 292-331.

[5] Palomino, F., and Vega–Redondo, F. (1999), "Convergence of Aspirations and (Partial) Cooperation in the Prisoner's Dilemma, *International Journal of Game Theory*, 28, 465-488.

[6] Posch, M., Pichler, A. and Sigmund, K. (1999), "The Efficiency of Adapting Aspiration Levels", *Proceedings of the Royal Society of London: Series B, Biological Sciences,* 266, 1427-1435.

[7] Samuelson, L. (1994). "Stochastic Stability with Alternative Best Replies", *Journal of Economic Theory*, 64, 35-65.

[8] Schlag, K. H. (1998). "Why Imitate, and if so, How? A Bounded Rational Approach to Multi–Armed Bandits", *Journal of Economic Theory,* 78, 130-156.

[9] Thorndike, E. (1911), *Animal Intelligence,* New York: Macmillan.

[10] Vega–Redondo, F. (1997), "The Evolution of Walrasian Behavior", *Econometrica,* 65, 375-84.