

# Learning in a local interaction hawk-dove game

Do not learn as your neighbours do

Jurjen Kamphorst<sup>1</sup> and Gerard van der Laan<sup>2</sup>

## Abstract

We study how players in a local interaction hawk dove game will learn, if they can either imitate the most succesful player in the neighborhood or play a best reply versus the opponent's previous action. From simulations it appears that each learning strategy will be used, because each performs better when it is less popular. Despite that, clustering may occur if players choose their learning strategy on the basis of largely similar information. Finally, on average players will play Hawk with a probability larger than in the mixed Nash equilibrium of the stage game.

*Keywords:* Learning, Local Interaction, Hawk-Dove game

*Subject Classification:* C73

---

<sup>1</sup>Department of Economics, Faculty of Law, Leiden University, Steenschuur 25, Leiden, The Netherlands. E-mail: j.j.a.kamphorst@law.leidenuniv.nl

<sup>2</sup>Department of Econometrics and Tinbergen Institute, Free University, De Boelelaan 1105, Amsterdam, The Netherlands. E-mail: glaan@feweb.vu.nl

# 1 Introduction

In the economic literature agents are typically rational. To be specific, they are fully aware of the options (formally, the *strategies*<sup>3</sup>) available to them, as well as of their own preference ordering over the possible outcomes. In addition they are capable of making costlessly and flawlessly any computations necessary to determine their optimal strategy. Moreover, it is typically assumed that the rationality of agents is common knowledge, that the agents have full information on the strategies available to other agents, as well as some idea on the preferences of the other agents over the different possible outcomes. Under these assumptions economists have been able to study many economic phenomena and to derive many interesting and insightful conclusions. However, often agents do not use all the information available to them, do not make all necessary computations and/or do not trust that other agents act rationally. Nevertheless, it is not *a priori* clear whether irrationality in the decision making of real-life agents leads to qualitatively different outcomes. Agents may still behave as if these typical assumptions are satisfied. They may have learned in some way which action is optimal, for instance from their own experience in similar situations in the past or the experience of other agents, or they may act as rational by imitating others who apparently followed a good strategy themselves. Of course, it also may happen that agents never learn the optimal decisions. What agents will learn, given the way in which they learn, is studied in the field of evolutionary game theory.

Since the 1960s biologists have applied game theory in evolutionary biology. Evolutionary game theory as such was developed by biologists in the early 1970s, see e.g. Maynard Smith and Price (1973) and Maynard Smith (1982). In this literature agents start with a predetermined strategy (where different agents may have different strategies) and the rewards are measured in terms of Darwinian fitness instead of utility, that is, the payoffs of a player determine how much offspring he has (all with the same strategy as their parent). This gives rise to population dynamics and the criterion of rational behavior by agents was replaced by (stability conditions on) these population dynamics to generate the solutions. Such dynamics are commonly named *replicator dynamics*.

Economists have picked up these developments by recognizing that these models can also be used to study the effects of a learning process of agents by allowing players to occasionally choose another strategy, see e.g. Samuelson (1997), Vega-Redondo (1996), Weibull (1995), Young (1998) for textbooks on Evolutionary Game Theory. Where biologists assume that agents with a better strategy have on average more offspring, economists assume that agents are more likely to adopt a strategy that they perceive as more successful than one that is less successful. Since payoff is measured in offspring, the dynamics in biology are clear. There is an important strand in the economic evolutionary game theoretic literature which adopts the replicator dynamics. It assumes

---

<sup>3</sup>An action (decision) is a behaviour performed in a particular situation. A strategy is a specification of which action an agent will choose (or with what probability which action will be chosen) in any situation in which the player may find itself.

that the share of the population using each strategy grows at a rate proportional to how well the agents using that strategy are doing relative to the whole population (see for example Fudenberg and Levine (1998)). However, learning in economics, meaning that there is some *learning rule* (also: *update rule*) by which agents convert information from the past into current strategies, is more complicated. Learning rules exist in a huge variety, including best response to previous strategies of other players (e.g. Anderlini and Ianni (1996), Goyal and Janssen (1997) and Kandori, Mailath and Rob (1993)), imitation of successful strategies or agents (e.g. Janssen (2000), Kirchkamp (2000), Noailly, Van den Bergh and Withagen (2004) and Possajennikov (2000)), or rules based on reciprocity like tit-for-tat or 'Win Cooperate, Lose Defect' (e.g. Tieman, Houba and Van der Laan (2000)). Unfortunately, although we can ignore many 'unreasonable' learning rules (such as 'use the strategy that performs worst'), the remaining learning rules can lead to different results, see e.g. Eshel, Samuelson and Shaked (1998). Hence it is important to know which learning rules are most applicable under what circumstances, and, if we can expect that some players use a different learning rule than others, how these learning rules interact.

In this paper we address these issues by letting players choose between two learning rules, and allowing them to change their choice every once in a while, according to some long run rule. This is similar to regarding the capacity choice of firms which is fixed in the short run, but variable in the long run. A learning rule is thus interpreted as a long term strategic variable. The main aim of the paper is to study the consequences of having multiple learning rules in the population. Moreover, we investigate whether one learning rule will drive out the other learning rule, so whether it is likely that all players will learn the same way. Since many papers have shown the significance of the interaction structure, see e.g. Ellison (1993) and Goyal (1996), we study these questions with two different interaction structures. Our focus is on a local interaction structure, where players are distributed on a torus. In this set-up, two players are matched with positive probability to play the stage game if and only if they are neighbors. By also considering a model with global interaction (any pair of players can be matched with positive probability) we gain insight into the effects of the interaction structure on the choice of players between the two learning rules. The stage game is the *Hawk-Dove* game. When two players are called to play the stage game, each can choose one of the two actions Hawk and Dove. When both play 'Dove' both get the so-called selection benefit payoff, when one player plays 'Dove' and the other 'Hawk', the 'Dove' player gets a coordination benefit payoff and other receives both benefits, when both play 'Hawk', both get nothing.

Recent similar approaches can be found in Kirchkamp (1999), Josephson (2001) and Droste, Hommes and Tuinstra (2002). Kirchkamp (1999) looks at a local interaction model where players cannot reselect their action every time they play. When a player chooses his action, he compares the performance of his own action with that of a randomly selected neighbor. His strategy specifies the general willingness to switch to the neighbor's rule as well as the weights attached to the information on the own rule and on that of the neighbor. Joseph-

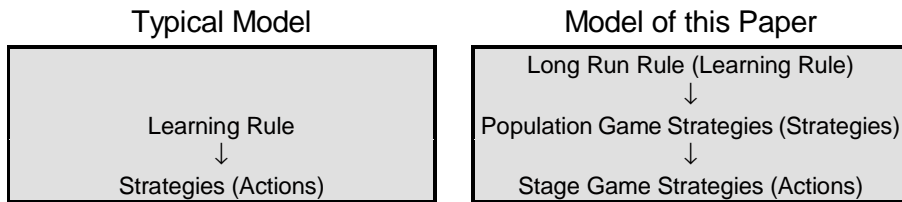


Figure 1: Terminology in this paper

son (2001) looks at which rules out of a class of learning rules are evolutionarily stable in a setting with uniform interaction. In a uniform interaction model Droste *et al.* (2002) look at a Cournot stage game on which the best-reply dynamics are unstable and where some rules may be more costly to use than other rules.

In Section 2 we describe the framing of the learning situation. Section 3 presents the analysis and the results. We find that this stage game provides much insight into the strengths and weaknesses of the learning rules, as well as into the interaction between these rules. Section 4 concludes.

## 2 The learning situation

In this paper we consider a multiperiod model in which in any period two players from a population of players are selected to play the stage game. In a typical model the learning rule of a player prescribes which strategy he has to play in the stage game. In this paper we consider a situation in which there are two learning rules within the population. Occasionally a player can switch from one learning rule to the other. This decision is based on a long run learning rule. In this way the long rule learning rule determines the strategic learning rule and the strategic learning rule determines which strategy (action) a player has to play in the stage game. To avoid confusion, in the remaining we use the following terminology. The stage game strategy of a player is called an *action*. The learning rule that determines which action a player has to play in the stage game is called the *strategy* of a player in the *population game*, being the game in which the sequence of stage games is played. The long run learning rule that determines which strategy a player plays in the population game is called the *long run rule* (see Figure 1).

The population in our game with local interaction contain  $n \times n$  players spread over a torus consisting of  $n \times n$  cells. In Figure 2 we show a  $4 \times 4$  torus. In each cell there is exactly one player. Every player has a neighborhood of the surrounding 8 players and himself. A torus has no borders, which means that a player on row (column)  $n$  has also surrounding players on row (column) 1. For instance, the neighborhood of player X in the top right cell consists of the nine

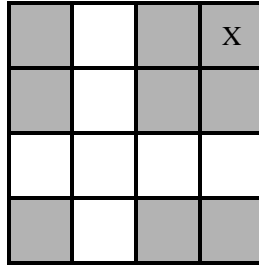


Figure 2: A  $4 \times 4$  torus

grey cells. Players observe only their own neighborhood, and whenever they are selected to play, they play against one of their neighbors. In the game with global interaction the neighborhood of a player consists of all other players (on the torus) in the population.

In the *population game* a sequence of stage games is played. The stage game is a 2-player normal form game in which both players play an action from the action space  $\mathcal{A} = \{\text{Dove}, \text{Hawk}\}$ . The payoff matrix of a player is given by

$$A^{HD} = \begin{bmatrix} \alpha & \varepsilon \\ \alpha + \varepsilon & 0 \end{bmatrix}, \quad (1)$$

where  $\alpha > \varepsilon > 0$ , where the top row (left column) corresponds to the action 'Dove' and the bottom row (right column) to the action 'Hawk'. The player itself chooses a row and his opponent a column. We denote the payoff of a player who plays  $s \in \mathcal{A}$  against some opponent playing  $s'$  by  $\pi(s, s')$ . Since  $\alpha > \varepsilon > 0$  the stage game has two asymmetric equilibria (in pure actions), namely (Dove,Hawk) and (Hawk,Dove). In a Nash equilibrium, so if players manage to coordinate to play different actions, both players get the payoff  $\varepsilon$ . In addition the player choosing 'Hawk' in a Nash equilibrium gets the additional payoff  $\alpha$ . We call  $\varepsilon$  the *coordination benefit* and  $\alpha$  the *selection benefit*, as it is the benefit that a player receives if he would be allowed to select the most preferred action in the Nash equilibrium. When both players play 'Dove' both receive the selection benefit, when both play 'Hawk', both receive zero payoff.

In the population game the stage game is played for a number of periods. In each period any player is assumed to know from all players in his neighborhood the chosen actions and the obtained payoffs in the last stage game they played. In every period one randomly selected pair of neighbors plays the Hawk-Dove stage game. Each player in the selected pair of neighbors chooses a stage game action according to his population game strategy. In the population game each player plays a strategy from the strategy space  $\mathcal{S} = \{\text{BR}, \text{IBA}\}$ . A player that uses the BR (IBA) strategy is called a BR (IBA) player.

A BR player plays in each stage game a (myopic) best reply versus his opponent. This means that he plays a best reply against the last chosen action

by his opponent. So, a BR player acts as if he regards this last chosen action as the best forecast of his opponent's action in the current period. Note that this particular variant of a best reply strategy is consistent with short memory (only the last played action of the players are known) and knowledge of the identity of the opponent. In other variants of best reply strategies players may expect a mixed action instead of a pure action, for instance in case of *fictitious play*, see Fudenberg and Levine (1998), where players have information on the complete history of their opponent, or *adaptive play*, see *e.g.* Young (1998), where players have information on a random sample of their opponent's strategy. In these variants players are assumed to recall more off the past. In for instance Ellison (1993)) it is assumed that players do not know the identity of their opponent and therefore base their beliefs on the actions of more than one player.

An IBA player tries to learn from the best performing player. To do so, he considers all the payoffs that the players in his neighborhood obtained in the last stage game they played and then he plays the action that was chosen by the player with the highest payoff. This behavior is consistent with people who read success stories and try to copy them, businesses that imitate business models of successful firms, etc.

The strategy of a player prescribes the action that he chooses. As such it corresponds to the learning rule in many other papers, where strategies and actions coincide. In this paper, on top of the strategy (learning rule) there is an exogeneously determined, long run (learning) rule. In each period, the first selected player of the pair gets the opportunity to evaluate his strategy. The decision to stay at the current strategy or the switch to the alternative strategy is determined by the long run rule of the population game. In this paper we consider two long run rules: the *Best Reply Rule* (BRR) or the *Imitate Best Behavior Rule* (IBBR). When IBBR is applied, the players are assumed to gather information from their neighborhoods on the average payoffs of the two strategies BR and IBA. When a player has the opportunity to reconsider his strategy, he adopts the strategy with the highest average payoff. If BRR is applied, a player chooses the strategy that would have given him the highest payoff in his last matching. So if the last time that he played, the IBA strategy would have given him a higher payoff than the BR strategy, then he selects IBA as his strategy, and vice versa. In both cases players stay at their old strategy in case of a tie. We would like to emphasize that all players have the same long run rule, and that this rule does not change during the population game. In our analysis we mainly consider the IBBR. As a check on the robustness of our choice of IBBR as the long run rule, we also compare some of the outcomes with the outcomes when using BRR.

Within the population game the timing is as follows. At the start of a new period one player is selected randomly from the population. This player gets the opportunity to change his current strategy according to the evaluation of his strategy by the long run rule. Then one of his neighbors (in the game with global interaction one of the other players) is selected and the selected pair plays the stage game. In the stage game both players choose the action (Dove or Hawk) prescribed by their strategy. Then both players receive their payoffs

and a new period starts.

In the analysis we mainly focus our attention to questions such as: how will the play of the stage game evolve, and what does this depend on? Which of the two strategies (BR and IBA) will survive under what circumstances? If both strategies survive, what will determine the relative popularity of each strategy (we measure this popularity as the *share of BR players* in the population)? If both strategies survive, will there be large areas with only IBA or only BR players? What is the impact of placing the population on a torus, instead of using global interaction and which results will qualitatively change if we impose the BRR as the long run rule instead of IBBR?

Most of the questions cannot be answered by analytical treatment of the model. Therefore the majority of the analysis is based on simulations. Each simulation is run for a population of 900 players, in case of local interaction spread over a  $30 \times 30$  torus. Each simulation is of a certain type, namely the parameters  $\alpha$  and  $\varepsilon$  in payoff matrix  $A^{HD}$  and the other settings of the model (*e.g.* the initial expected share of BR players, the long run rule, etc.). All simulations with local interaction are run 1000 times. All simulations with global interaction are run a 100 times. The data reported for each simulation is the average of all runs of the simulation. The initial situation is generated independently for each of the runs: each player is assigned to have Dove as their (virtually) last played action with 50 percent probability and each player is assigned the same payoff from the (virtually) last played stage game, which is lower than the minimal payoff of the stage game. In most simulations the probability that a player initially is a BR player is 50 percent. In some simulations we look at the influence of starting out with many (or few) BR players. Then either one, respectively eight, out of nine players are BR players. Observe that the variety in initial situations of the runs of any one simulation is very large.

### 3 Results and Analysis

In Table 1 (see the Appendix for the meaning of the columns in the table) we present the simulation results for a variety of stage game payoff matrices of the form (1). We summarize the results in Facts 7.1 to 7.10. We first present some facts regarding the selection of actions by the different strategies. Then we can explain the facts related to the popularity of the two strategies. From the simulations it appears that only the relative size of  $\alpha$  to  $\varepsilon$  is important, not the absolute value of  $\alpha$ . Therefore, in Table 1 only the ratio is given.

**Fact 7.1** *BR players play Dove roughly half of the time if players can only use the BR strategy. If also the IBA strategy can be used, then BR players will tend to play more Dove than Hawk.*

**Fact 7.2** *IBA players will tend to play Hawk. The more BR neighbors an IBA player has, the stronger this tendency is.*

**Fact 7.3** *Under global interaction and with BR players in the population, IBA players play Hawk with probability one.*

ID	Rel. Sel. Ben. ( $\alpha/\epsilon$ )	Descr.	Share BR	St. Dev. Share BR	p(D)	p(D B)	p(D I)	Share BR next to BR
HD1	100		0.356	0.028	0.286	0.669	0.075	0.451
HD2	60		0.356	0.028	0.288	0.671	0.075	0.451
HD3	30		0.354	0.029	0.287	0.671	0.076	0.449
HD4	20		0.353	0.029	0.286	0.672	0.076	0.448
HD5	$16 \frac{2}{3}$		0.356	0.028	0.288	0.671	0.076	0.452
HD6	15		0.355	0.028	0.286	0.669	0.074	0.450
HD7	$12 \frac{1}{2}$		0.360	0.030	0.288	0.668	0.074	0.455
HD8	$6 \frac{1}{4}$		0.394	0.026	0.298	0.670	0.056	0.466
HD9	5		0.413	0.024	0.305	0.670	0.048	0.471
HD10	1		0.541	0.020	0.354	0.636	0.022	0.564
HD11	0.5		0.590	0.019	0.372	0.619	0.018	0.605
HD12	0.01		0.633	0.018	0.388	0.603	0.017	0.641
HD13	15	IBA only	0.000	0.000	0.311	n.a.	0.311	n.a.
HD14	15	BR only	1.000	0.000	0.500	0.500	n.a.	n.a.
HD15	15	B:I = 1:8	0.356	0.029	0.287	0.670	0.075	0.450
HD16	15	B:I = 8:1	0.356	0.028	0.285	0.669	0.074	0.451
HD17	1	B:I = 1:8	0.541	0.020	0.354	0.635	0.023	0.564
HD18	1	B:I = 8:1	0.540	0.021	0.353	0.635	0.022	0.563
HD19	30	BRR	0.619	0.012	0.402	0.630	0.030	0.585
HD20	20	BRR	0.620	0.013	0.401	0.629	0.030	0.585
HD21	15	BRR	0.619	0.012	0.403	0.631	0.031	0.585
HD22	0.01	BRR	0.624	0.012	0.403	0.629	0.027	0.591
HD23	15	BRR, B:I = 1:8	0.620	0.012	0.401	0.629	0.030	0.585
HD24	15	BRR, B:I = 8:1	0.619	0.013	0.401	0.629	0.030	0.585
HD25	15	GI	0.122	0.046	0.108	0.897	0.000	n.a.
HD26	5	GI	0.287	0.048	0.221	0.775	0.000	n.a.
HD27	15	GI, BRR	0.643	0.012	0.392	0.609	0.000	n.a.

Table 1: Simulation results.

**Fact 7.4** *In the simulations with IBA players, the outcomes are inefficient because too many players play Hawk. Moreover the dynamics do not converge to a Nash equilibrium.*

**Evidence for Fact 7.1.** Simulation HD14 shows that in the absence of the IBA strategy players play Dove 50 percent of the time. However, all other simulations in Table 1 show that when also the IBA strategy is available and the IBA players tend to select Hawk, then the BR players favor Dove (*i.e.* play Dove more than 50 percent of the time). The intuition for this result is the following. In the Hawk-Dove game, a BR player plays the action that is opposite to the last action that his opponent played. Hence, if the whole population uses only the BR strategy then in the long run (*i*) BR players play Dove 50 percent of the time and (*ii*) in every period roughly 50 percent of the population plays Dove. However, when some players use an alternative strategy and this strategy prescribes Dove with a probability higher than 50 percent, then BR players are more likely to meet Dove than Hawk, and therefore they are more likely to play Hawk. Likewise BR players tend to select Dove if the alternative strategy favors Hawk. In our model the alternative strategy is IBA. Since IBA favors Hawk, our simulations indeed show that the BR players play more Dove than Hawk.

**Evidence for Fact 7.2.** In case that players can only use the IBA strategy in the population game, Simulation HD13 shows that IBA players tend to select Hawk. That IBA players tend to play Dove with even smaller probability if the probability of being matched with a BR player increases can be seen in Table 1 (HD6).

To explain why IBA players tend to select Hawk, notice that IBA players select the stage game action which yields the maximal payoff observed in their neighborhood. Hence, after the first periods, they select Dove only if no player in their neighborhood has payoffs equal to  $\pi(\text{Hawk}, \text{Dove})$ . However, because there are only four different payoff levels, and each player observes the payoffs of nine players, it is very likely that at least one of these nine players has this payoff. For example, ignoring any effect due to spatial distribution, if 50 percent of the players selected Dove last period, then IBA players select Dove with an average probability of roughly  $(1 - \frac{1}{2}(1 - \frac{1}{2}))^9 = 0.075$ . In fact the equality

$$(1 - p_D(1 - p_D))^9 = p_D$$

has one stable solution: approximately  $p_D = 0,2$ . This is even lower than the 31 percent that Simulation HD13 shows in case only the IBA strategy is available, due to the fact that we ignored the effects of the spatial distribution, which counts for correlation between the choices of neighbors. Neighbors will tend to choose the same actions because they base their decision on related data sets.

Figure 3 shows a scatterplot of the share of BR players versus the share of players that select Top. The data are taken from the initial periods of 5 runs of Simulation HD6. Of each run the data was stored at the end of period  $900k$ ,

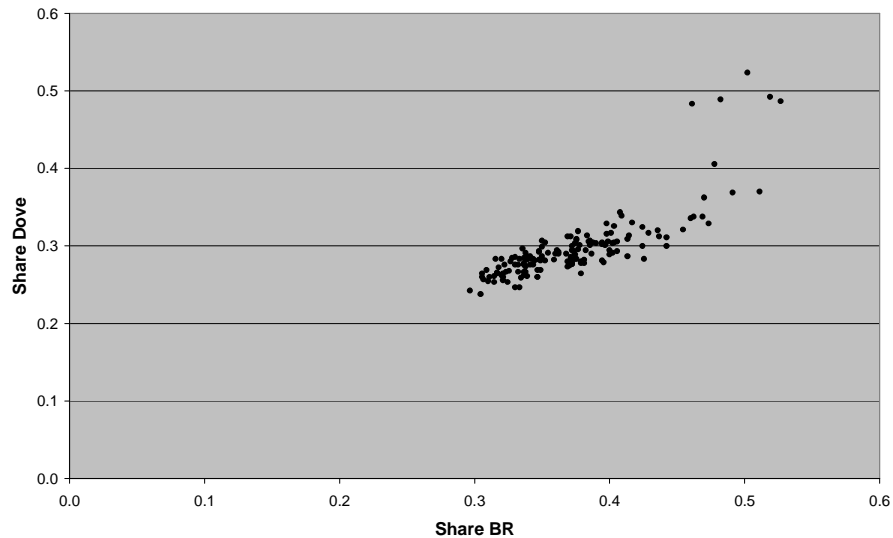


Figure 3: Scatterplot of Share of BR Players versus the Share of Players who select Dove. Data is generated by runs of Simulation HD6

for every  $k = 0, 1, 2, \dots, 30$ . The same data on the share of BR players is also presented in five time series in Figure 5. The Figure shows that the probability with which a player plays Dove increases if the share of BR players increases, although both BR players and IBA players are less likely to play Dove.

**Evidence for Fact 7.3.** This fact is observed in the Simulations HD25 to HD27. This is intuitive, since, as we observed in the evidence for Fact 7.2, the probability that the maximal payoff  $\pi$  (Hawk,Dove) is realized by some player in a neighborhood is quite high. The population is a hundred times larger than a neighborhood. Thus a player observes the maximal payoffs almost surely under global interaction.

**Evidence for Fact 7.4.** The maximal joint payoff in a period is either  $2\alpha$  or  $\alpha + 2\varepsilon$ . In both cases at least one player has to select Dove. So if the outcomes are efficient, the population needs to play Dove at least 50 percent of the time. Clearly this is not the case (Table 1). Note that in the pure Nash equilibria precisely one of the players plays top, and that in the mixed Nash equilibrium both players play Dove with 50 percent probability. Hence it is clear that the dynamics do not converge to any Nash equilibrium.

We now consider which strategy is more likely to obtain which benefit. This provides some intuition for the results regarding the popularity of the different strategies.

**Fact 7.5** *In the long run, BR players are more likely to obtain the coordination benefit  $\varepsilon$ , while IBA players are more likely to receive the selection benefit  $\alpha$ .*

**Evidence for Fact 7.5.** Simulations HD1 to HD12 show that there are more BR players if the coordination benefit is larger relative to the selection benefit. Since the long run rule considers only the average payoffs of each strategy for some neighborhood, it follows that BR players obtain a relatively large share of their payoffs from the coordination benefit, while IBA players obtain a larger share of their total payoffs via selection benefits. However, also in absolute sense BR players are more likely to obtain coordination benefits and less likely to obtain the selection benefit than IBA players, as is shown by the following, rough, numerical estimation.

Using data from Simulation HD6, we estimate several probabilities in Table 2. With these probabilities we estimate that the average BR player receives the coordination benefit with probability<sup>4</sup>

$$p_{co}^{BR} = 2p_{BB}p_{BB}^H p_{BB}^D + p_{BI} (p_{BI}^H p_{IB}^D + p_{BI}^D p_{IB}^H),$$

and the selection benefit with probability

$$p_{sel}^{BR} = p_{BB}p_{BB}^D + p_{BI}p_{IB}^D.$$

For the average IBA player these probabilities are

$$p_{co}^{IBA} = 2p_{II}p_{II}^H p_{II}^D + p_{IB} (p_{BI}^H p_{IB}^D + p_{BI}^D p_{IB}^H)$$

and

$$p_{sel}^{IBA} = p_{II}p_{II}^D + p_{IB}p_{BI}^D.$$

According to these estimates, the average BR player receives the coordination benefit with 0.68 probability and the selection benefit with 0.20 probability. The average IBA player receives the coordination benefit with 0.35 probability and the selection benefit with 0.33 probability. Thus these simulation results suggest that BR players are indeed more likely to receive the coordination benefit, while the IBA players are more likely to receive the selection benefit. According to our estimates an IBA player selects Hawk in about 93 percent of the cases. Hence in roughly 93 percent of the cases the maximal payoffs in the neighborhood of an IBA player are equal to the maximal payoffs in payoff matrix  $A^{HD}$ . The same analysis for HD9 gives us that the coordination benefit and the selection benefit are obtained by BR players with a probability of 0.69 and 0.19 respectively, and by IBA player with a probability of 0.39 and 0.38 respectively.

To see the intuition behind Fact 7.5, let us consider how these two strategies select the action and how they interact. BR players act as if they assume that the opponents behavior is given, namely the opponent's last action. They respond

---

<sup>4</sup>The descriptions of the different variables, and how their value is estimated is explained in Table 2.

Description	Sym- bol	Esti- mate HD6	Esti- mate HD9	Calculation
Prob(BR matched with BR)	$p_{BB}$	0.45	0.47	Share BR nghb BR, Tabel 1
Prob(BR matched with IBA)	$p_{BI}$	0.55	0.53	$1 - p_{BB}$
Prob(IBA matched with BR)	$p_{IB}$	0.30	0.37	Share BR nghb IBA, Tabel 1
Prob(IBA matched with IBA)	$p_{II}$	0.70	0.63	$1 - p_{IB}$
Prob(BR player plays Dove)	$p^D_B$	0.67	0.67	Pr(D B) , Tabel 1
Prob(IBA player play Dove)	$p^D_I$	0.07	0.05	Pr(D I) , Tabel 1
Prob(IBA plays Dove vs IBA player)	$p^D_{II}$	0.07	0.05	$p^D_I$ , because IBA players ignore their opponent, we assume that the probability with which an IBA player plays Dove does not depend on the strategy of his opponent.
Prob(IBA plays Hawk vs IBA player)	$p^H_{II}$	0.93	0.95	$1 - p^D_{II}$
Prob(IBA plays Dove vs BR player)	$p^D_{IB}$	0.07	0.05	$p^D_I$
Prob(IBA plays Hawk vs BR player)	$p^H_{IB}$	0.93	0.95	$1 - p^H_{IB}$
Prob(BR plays Dove vs IBA player)	$p^D_{BI}$	0.93	0.95	$p^H_{IB}$ , BR players adapt to the action of the opponent.
Prob(BR plays Hawk vs IBA player)	$p^H_{BI}$	0.07	0.05	$p^D_{IB}$ , BR players adapt to the action of the opponent.
Prob(BR plays Dove vs BR player)	$p^D_{BB}$	0.35	0.35	Solution to equation $p^D_B = p_{BB} p^D_{BB} + p_{BI} p^D_{BI}$
Prob(BR plays Hawk vs BR player)	$p^H_{BB}$	0.65	0.65	$1 - p^D_{BB}$

Table 2: Numerical approximation based on Simulations HD6 and HD9.

optimally to that belief in order to get the coordination benefits. Observe that because their opponent's strategy is believed to be predetermined, the strategy of the BR player cannot influence whether they get the selection benefit or not, for this depends only on the opponent's strategy. So the only thing they can influence is whether there is successful coordination. In contrast IBA players imitate the action yielding the maximal payoffs in their neighborhood: often this is Hawk (see the evidence for Fact 7.2). It is as if they assume that their opponent will adjust to them, to facilitate this maximal payoff of selection and coordination benefits. Therefore, when a BR and an IBA player are matched, then it is often as if the two play a sequential Hawk-Dove game in which the IBA player moves first and the BR player moves second, which leads to successful coordination and to an additional selection benefit for the IBA player. In this, the adaptiveness of the BR strategy is both its strength and its weakness. It is more likely to receive the coordination benefit, especially if used against some predictable opponent (in the sense that he does not often change his action), but is therefore less likely to obtain the selection benefit if the opponent often plays Hawk. The strength of the IBA strategy is that it exploits any such adaptiveness, by his own lack of adaptiveness in often selecting Hawk regardless of the opponent. Note that if some IBA player receives the maximal payoffs (and thus plays Hawk), because he observes his own payoff, he will continue to play Hawk and receive maximal payoffs at least until he meets an IBA player or until he becomes a BR player. Also note that both BR players and IBA players are more likely to coordinate than BR players or IBA players among themselves, because of the predictability of the action choice by IBA players, and the adaptiveness by BR players.

We now consider some facts regarding the 'equilibrium level' of the share of BR players to which the dynamics converge. Fact 7.6 states that there is a unique level of the share of BR players in the population to which the dynamics converge, in the sense that the share of BR players approaches and remains close to that level. This level we call the equilibrium level. Facts 7.7 and 7.8 give some insight into this equilibrium level.

**Fact 7.6** *The dynamics converge to a unique equilibrium level of the share of BR players in the population if (i) there are initially at least 100 users ( $\frac{1}{9}$  of the population) of each strategy under the IBB rule; or (ii) if both strategies are available to the players and the long run rule is the BR rule.*

Let us consider the share of players that adopt the BR strategy in our model. In order to talk about the share of BR players, it is good to keep in mind how strategies are selected by the long run rule. The long run IBBR selects the strategy for a player which yields the highest average payoff in his neighborhood. Thus the popularity of a strategy is determined by the likelihood with which it gives the higher average neighborhood payoffs. So, roughly speaking what matters are the expected payoffs of a strategy and therefore (in the long run) the probabilities with which the selection benefit and the coordination benefit is acquired as well as the relative selection benefit.

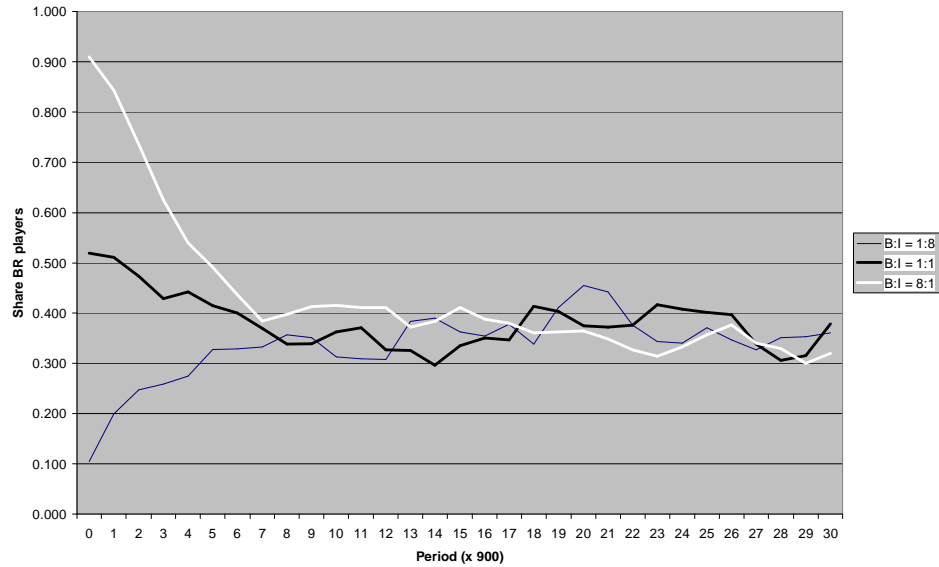


Figure 4: Time Series of three runs with different initial shares of BR players

In contrast the long run BRR selects for a player the strategy that prescribed the better action to him in his last interaction. The better action is equivalent with the action that yields the coordination benefit. To be precise: an IBA (BR) player switches to BR (IBA) if the action prescribed by the IBA (BR) strategy does not yield the coordination benefit, while the action that the BR (IBA) strategy would have prescribed would have given this benefit. So roughly speaking, what matters is the probability with which the coordination benefit is obtained in this model. Hence we would expect that the strategy which is, in absolute terms, better in acquiring the coordination benefit is also the more popular under BRR. Similarly, we expect that the strategy that is, in absolute terms, better in acquiring the selection benefit is more popular under IBRR than under BRR, if the relative selection benefit is at least one.

**Fact 7.7** *Under IBRR, the equilibrium level is weakly decreasing in the relative selection benefit. Under BRR the relative selection benefit does not influence the equilibrium level.*

**Fact 7.8** *The BR strategy is more popular than the IBA strategy under BRR. Under IBRR it depends on the relative selection benefit.*

**Evidence for Fact 7.6.** Table 1 shows that the standard deviation in the final share of BR players is quite low. Moreover, comparing Simulations HD15 to HD18 with Simulations HD6 and HD10 shows that the initial share of BR

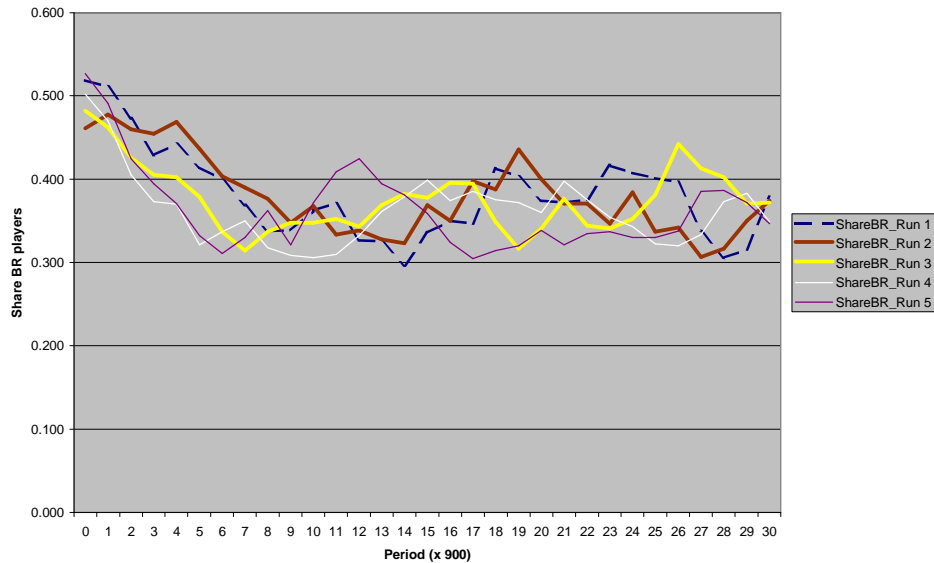


Figure 5: Time Series of 5 Runs of Simulation HD6.

players does not matter, at least if it is not lower than  $\frac{1}{9}$  or higher than  $\frac{8}{9}$ . Figure 4 shows some randomly chosen time series with different starting conditions and a relative selection benefit  $\frac{\alpha}{\epsilon} = 15$ . Observe that in the long run the share of BR players fluctuates around the equilibrium level (0.355) of BR players, regardless of the initial share. Note that the dynamics fluctuate around this equilibrium, instead of reaching and staying in it (see Figure 5). Simulations HD19 to HD24 show that this is also the case under BRR. Compare especially HD23 and HD24 with HD19.

Under IBBR, a strategy cannot be selected if it is not currently present in the population. Hence if the share of BR player becomes 1 or 0, then this share will remain the same for the rest of the simulation. Hence these are, strictly speaking, equilibria of the dynamics. Nonetheless, even for low initial shares (of about  $\frac{1}{9}$ ) and relatively high shares (of about  $\frac{8}{9}$ ) the simulations converge to the reported equilibrium levels. This, and the intuition that we develop below, suggests that the dynamics tend to go away from these 'extreme equilibria'. However, due to the randomness in the model, it is possible to end up in one of the extreme equilibria if the initial share is close enough to it. Therefore we need the restrictions in part (i) of Fact 7.6. Because under the BR rule a strategy can still be chosen if it is currently not present in the population, the dynamics can leave a state in which the share of BR players is 1 or 0. Hence the restrictions of (i) are not necessary under part (ii).

In the following paragraphs we will see why it is intuitive that such a unique equilibrium level exists. Consider first the model with IBBR. The existence of a unique equilibrium level to which the dynamics converge, indicates that for higher (lower) shares of BR players, the share of BR players is expected to decrease (increase). There are two reasons why this is the case.

First let us suppose that the (positive) probability with which an updating BR player (IBA player) switches to the IBA (BR) strategy does not depend on the share of BR players. It follows directly that the expected change in the share of BR players is negatively related to the share of BR players. The more BR players there are, the more likely it is that a BR player is allowed to update, and not an IBA player, and the more likely it is that a player switches from BR to IBA instead of the other way around. Hence the larger the share of BR players, the more likely it is to decrease.

Second we argue that the probability with which an updating player selects the IBA strategy is expected to increase in the share of BR players. Clearly this reinforces the first reason. A higher share of BR players implies that all players are more likely to be matched with a BR player. If IBA players are more likely to meet a BR player, they will receive both the selection and the coordination benefit with higher probability. For BR players a higher probability to meet a BR player implies a higher probability of getting the selection benefit, but a lower probability of getting the coordination benefit. As there is no reason to expect that the increase in the probability of receiving the selection benefit is larger for BR players than for IBA players, it is unlikely that the expected profits of BR players increase more than those of IBA players in response to an increased share of BR players.

Now consider the model with BRR. The first reason above clearly applies regardless of the long run rule, hence it applies here. Regarding the second reason, the strategy selection under BRR depends only on whether the coordination benefit is obtained, not on its size or the selection benefit. Since, as the share of BR players increases, BR players become less likely and IBA players more likely to obtain the coordination benefit, also the second reason applies here.

**Evidence for Fact 7.7.** See Simulations HD1 to HD24. Recall that Fact 7.5 stated that, around the equilibrium levels (in the long run), BR players are more likely to receive the coordination benefit, while IBA players are more likely to receive the selection benefit. Hence BR players benefit relatively much from the coordination benefit, and relatively little from the selection benefit. It is therefore logical that a decrease in the coordination benefit relative to the selection benefit (so an increase in the relative selection benefit) harms the average profits of BR players more than those of IBA players. Therefore it is intuitive that the probability with which the BR strategy in some neighborhood has the highest average payoffs also increases when the relative selection benefit decreases. Hence under IBBR the popularity of BR players is inversely related to the relative selection benefit. Under BRR payoffs as such do not matter. It matters only whether a best reply is played. Since the preference ordering of

players over the four possible outcomes does not depend on the relative selection benefit, a change in the relative selection benefit has no effect on the equilibrium level under the BR rule.

**Evidence for Fact 7.8.** Simulations HD19 to HD24 show this fact. Remember that under BRR it matters only which strategy or strategies prescribed the best reply in the last interaction. Hence the strategy which is best in obtaining the coordination benefit, becomes more popular. As we have seen before, see for instance Fact 7.5, the BR strategy is better at coordinating successfully. Hence the BR strategy is more popular under BRR.

In contrast, under IBBR payoffs do matter (see Simulations HD1 to HD18). Hence the actual payoffs play a role, which is a relative advantage to the IBA players provided that the relative selection benefit is not too small (look for example at Simulation HD12). Note that Simulation HD10 has a relative selection benefit equal to one, and that the equilibrium level (of the share of BR players) is lower there than in the simulations with BRR. This provides additional evidence for Fact 7.5.

To conclude we consider some facts regarding the level of clustering and the differences between local and global interaction. *Clustering* is defined as a distribution of strategies over the players such that randomly selected pairs of neighboring players are more likely to have the same strategies than randomly selected pairs of non-neighboring players. So if there is clustering, neighboring players tend to select the same strategies. If the opposite of clustering, *dispersion*, occurs, than neighboring players tend to select different strategies.

**Fact 7.9** *Under the IBBR rule there is clustering. Under the BR rule there is dispersion.*

**Fact 7.10** *With global interaction the equilibrium level (of the share of BR players) is (i) under the IBBR rule higher than with local interaction; (ii) under the BR rule lower than with local interaction.*

**Evidence for Fact 7.9.** The last column of Table 1 show whether there is clustering, that is, whether neighboring players tend to select the same strategies. From Simulations HD1 to HD12 it is apparent that under the IBBR there is clustering. This is not surprising since IBBR bases the selection of strategies on information from the neighborhood, and neighborhoods of adjacent (so neighboring) players overlap. So neighboring players base their strategy choice on correlated data sets. As a consequence there is clustering.

Simulations HD19 to HD22 show that under BRR the opposite of clustering occurs: dispersion. Neighboring players tend to select different strategies. We offer two reasons for this. First, the data sets on which neighboring players base their decision do not overlap, as they do with IBBR. Therefore one important reason for clustering is removed. Second, the predictability of IBA players makes it easy for matchings of BR and IBA players to coordinate, whereas BR-BR and IBA-IBA matchings are generally much less successful in coordination. Hence,

under BRR a player in the midst of BR players will more likely change from BR to IBA than vice versa. Among IBA players, an updating player will likely select the BR strategy. Dispersion is the result.

**Evidence for Fact 7.10.** This is apparent from the comparison of Simulations HD25, HD26 and HD27 with HD6, HD9 and HD21 respectively. To understand why this is the case, let us first rediscuss the evidence from Fact 7.6. There we found BR players are more likely to receive the selection benefit, but less likely to get the coordination benefit as a result of a larger probability of meeting BR players, while IBA gain both benefits with higher probability. Hence we expect that BR players benefit more from clustering under IBBR than IBA players. Moreover note that IBA-IBA matchings are less likely to result in coordination than BR-BR matchings. Thus also under BRR the IBA players suffer more from an increase in the clustering than the BR players.

Recognize that there is no clustering or dispersion possible with global interaction. Under IBBR this implies a 'decrease' in the level of clustering, and thus a relative benefit to the IBA players. Hence we see a lower equilibrium level with global interaction. Under BRR there is 'more clustering' with global than with local interaction. Hence there the equilibrium level is higher with global interaction.

## 4 Conclusions

In this paper we study a local interaction model where agents play a symmetric Hawk-Dove stage game against a randomly selected neighbor. We investigate by simulation which of two learning rules (BR and IBA) is most likely to be adopted by the population as their strategy.

In all of the simulations both strategies survived. Therefore it is not unlikely that agents use different 'learning rule strategies'. Because the dynamics in the action space never rest, the outcomes are determined by the long run result. We see that, at the equilibrium levels of BR players, the BR players are more likely to obtain the coordination benefit and that IBA players are more likely to get the selection benefit. As a consequence, the popularity of the BR strategy is decreasing in the relative selection benefit. In terms of actions, IBA players play Hawk often in local interaction and almost always under global interaction, because playing Hawk can result in the maximal payoff  $\alpha + \varepsilon$ . BR players tend to play Dove (the dove-like action) unless two BR players meet: then they tend to play Hawk (the hawkish action), because it pays off to choose a different action than their opponent.

When the long run rule is BRR instead of IBBR then the payoffs to strategies are irrelevant, except for the question which strategy prescribed a better reply. Hence the relative selection benefit is of no influence. The BR strategy is more popular under this rule than the IBA strategy, since at the IBBR equilibrium levels the BR strategy is better at coordination. Also, whereas under IBBR neighboring players tend to have the same strategy (clustering) they tend to

have different strategies (dispersion) under BRR. This is because the action choice of BR players fluctuates more than that of IBA players (who almost always play Hawk). This makes it easier for a BR player to coordinate with an IBA player than with another BR player. Thus players do better when their strategy is relatively unpopular in their neighborhood. Hence players should not learn as their neighbors do.

## 5 Appendix

The first column in Table 1 is an identification number for each simulation. The second column gives the relative selection benefit, while the third column provides, when needed, some additional information on the simulation. "BRR" is written there if the long run rule is BRR. It says "GI" if there is global interaction instead of local interaction, and it says "BR only" ("IBA only") if all players can only use the BR (IBA) strategy. Finally, "B:I =  $x:y$ " is written if initially players use the BR strategy with probability  $\frac{x}{x+y}$  instead of  $\frac{1}{2}$ .

The fourth column gives the average share of BR players in the population at the end of the runs, and the fifth column gives the standard deviation of this number. Columns 6 to 8 gives for all players, BR players and IBA players respectively, the frequency with which Dove was played on average in the last 9000 periods of each run. The last column provides some statistics regarding the tendency of neighboring players to have the same strategy. More specifically, it gives the average share of BR players among the neighbors of BR players. We say that clustering occurs if BR players have relatively many BR neighbors, thus if the share reported here is higher than the share of BR players in the total population.

In Table 1 we also present a few simulations with  $\varepsilon \geq \alpha > 0$ . For these parameters the game is not a HD game anymore. These few simulations indicate that the conclusions for the HD game also hold for stage games with this range of parameters.

## References

- [1] Anderlini, L. and A. Ianni (1996), "Path Dependence and Learning from Neighbors", *Games and Economic Behavior* 13, pp. 141-77.
- [2] Droste, E., C. Hommes and J. Tuinstra (2002), "Endogenous Fluctuations under Evolutionary Pressure in Cournot Competitions", *Games and Economic Behavior* 40(2), pp. 232-69.
- [3] Ellison, G. (1993), "Learning, Local Interaction, and Coordination", *Econometrica* 61(5), 1047-71.

- [4] Eshel, I., L. Samuelson and A. Shaked (1998), "Altruists, Egoists, and Hooligans in a Local Interaction Model", *American Economic Review* 88(1), pp. 157-79.
- [5] Fudenberg, D. and D.K. Levine (1998), *The Theory of Learning in Games*, MIT Press, Cambridge, Massachusetts.
- [6] Goyal, S. (1996), "Interaction Structure and Social Change", *Journal of Institutional and Theoretical Economics* 152, pp. 472-94.
- [7] Goyal, S. and M.C.W. Janssen (1997), "Non-Exclusive Conventions and Social Coordination", *Journal of Economic Theory* 77, pp. 34-57.
- [8] Janssen, M.C.W. (2000), "Imitation of Cooperation in Prisoner's Dilemma Games with Some Local Interaction", Tinbergen Institute Discussion Paper TI2000-019/1.
- [9] Josephson, J. (2001), "A Numerical Analysis of the Evolutionary Stability of Learning Rules", SSE/EFI Working Paper Series in Economics and Finance 474, Stockholm School of Economics.
- [10] Kandori, M., G. Mailath and R. Rob (1993), "Learning, Mutation and Long Run Equilibria in Games", *Econometrica* 61(5), pp. 29-56.
- [11] Kirchkamp, O. (1999), "Simultaneous Evolution of Learning Rules and Strategies", *Journal of Economic Behavior and Organization* 40, pp. 295-312.
- [12] Kirchkamp, O. (2000), "Spatial Evolution of Automata in the Prisoner's Dilemma", *Journal of Economic Behavior and Organization* 43, pp. 239-62.
- [13] Maynard Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.
- [14] Maynard Smith, J. and G.R. Price (1973), "The Logic of Animal Conflict", *Nature* 246, pp. 15-18.
- [15] Noailly, J., J.C.J.M. Van den Bergh and C.A. Withagen (2004), "Spatial Evolution of Social Norms in a Common-Pool Resource Game", mimeo.
- [16] Possajennikov, A. (2000), *Learning and Evolution in Games and Oligopoly Models*, Dissertation Series, Center for Economic Research, Tilburg University.
- [17] Samuelson, L. (1997), *Evolutionary Games and Equilibrium Stability*, MIT press, Cambridge, Massachusetts.
- [18] Tieman, A.F., H.E.D. Houba and G. van der Laan (2000), "On the Level of Cooperative Behavior in a Social Environment", *Journal of Economics* 71, pp. 1-30.

- [19] Vega-Redondo, F. (1996), *Evolution, Games, and Economic Behaviour*, Oxford University Press, New York.
- [20] Weibull, J.W. (1995), *Evolutionary Game Theory*, MIT Press, Cambridge, Massachusetts.
- [21] Young, H.P. (1998), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton University Press, Princeton, New Jersey.