Learning with fixed rules: The minority game

Willemien Kets*

March 7, 2011

Abstract

This paper gives a critical account of the literature on adaptive behavior in the minority game, a simple congestion game. The literature has proposed a model which differs markedly from many standard learning models in that players are endowed with a fixed subset of behavioral rules or response modes which map the observed history to actions. These rules need not have a behavioral interpretation or be derived from some form of optimizing behavior. Nonetheless, this model gives rise to behavior that is close to equilibrium behavior at the aggregate level. The individual-level behavior predicted by the model seems to capture some aspects of observed experimental behavior that are difficult to explain using standard models.

JEL classification: C73, C90.
Keywords: Learning, congestion games, experiments.

*Santa Fe Institute and Tilburg University. Address: 1399 Hyde Park Road, Santa Fe, NM 87501, USA. E-mail: willemien.kets@santafe.edu. Phone: +1-505-204-8012.

1 Introduction

Congestion games are ubiquitous in economics. In a congestion game, players use several facilities from a common pool (Rosenthal, 1973). The costs or benefits that a player derives from a facility depends on the number of users of that facility. A congestion game is therefore a natural game to model scarcity of common resources. Examples of such systems include vehicular traffic (Nagel et al., 1997), packet traffic in networks (Huberman and Lukose, 1997), and ecologies of foraging animals (DeAngelis and Gross, 1992). Similar coordination problems are encountered in market entry games (Selten and Güth, 1982).

The evolution of behavior in congestion games has been studied extensively. Many learning models give sharp predictions for these games, as congestion games form a subclasss of the well-studied class of potential games (Hofbauer and Hopkins, 2005; Hofbauer and Sandholm, 2002; Monderer and Shapley, 1996; Sandholm, 2001, 2010). There is, however, an experimental puzzle. At the aggregate level, players are remarkably successful at coordinating their actions, so that aggregate play is consistent with Nash equilibrium predictions (e.g Kahneman, 1988; Meyer et al., 1992; Selten et al., 2007). However, these regularities at the aggregate level conceal non-equilibrium behavior at the *individual* level (Meyer et al., 1992; Erev and Rapoport, 1998; Selten et al., 2007; Bottazzi and Devetag, 2004), something which is hard to explain using standard learning models.

This paper provides a survey of the literature on the minority game, a simple congestion game, which has proposed an alternative model. In a minority game, an odd number of players have to choose between two alternatives, and the ones that have chosen the least popular alternative (the "minority" action) receive the highest payoff. The learning model, inspired by the seminal paper of Arthur (1994) and largely developed within the physics literature, departs from the standard literature in assuming that players are endowed with a fixed subset of behavioral rules or *response modes* which map the observed history of play to an action for the next period. Importantly, these rules need not have a behavioral interpretation or be derived from some implicit optimizing procedure. In each period of play, a player chooses one of his response modes based on past performance; the response mode then selects an action for the next period as a function of the history of play. I relate this so-called fixed-rules learning model to the standard learning models in economics, and show how the assumption that players have fixed response modes can aid in explaining experimental findings.[1] In particular, not only does the model predict aggregate-level behavior that is close to equilibrium, it can also capture some aspects of observed individual-level behavior that are difficult to explain with other models.

After introducing the minority game and some related games in Section 2, I provide an overview of some experimental findings on learning in congestion games in Section 3. Section 4 introduces the learning model with fixed rules, and relates it to the wider literature on learning in games. Section 5 discusses the main predictions of the fixed-rules model, and relates them to the experimental evidence. Section 6 concludes by outlining possible directions for future research.

2 The minority game

The minority game is a game in which an odd number of players have to choose between two actions, e.g., going to a bar or stay home, buy or sell an asset, and so on. Players want to distinguish themselves from the crowd: their aim is to take a different action than the majority of players.

More precisely, the set of players is $N = \{1, \ldots, 2k+1\}$, where $k \in \mathbb{N}$. Each player $i \in N$ has two actions, labeled -1 and $+1$. Write $A_i = \{0, 1\}$ for the set of actions of player i. The utility of a player only depends on the number of players taking each action. That is, for each $b \in A_i$, there exists a function f_b mapping the number of players $m_b \in \{0, 1, \ldots, 2k+1\}$ taking action b to $\mathbb{R}$, so that the utility of a player i of action $a_i \in A_i$ given that the other players play according to $a_{-i} \in \times_{j \neq i} A_j$ is

$$u_i(a_i, a_{-i}) = f_{a_i}(|\{j \in N : a_j = a_i\}|).$$

As discussed by Tercieux and Voorneveld (2010), minority games are characterized by the assumption that a player always benefits from deviating from the majority to the minority: For all $m \in \{k+2, \ldots, 2k+1\}$,

$$f_a(m) < f_b(2k + 2 - m)$$

for $a, b \in \{-1, +1\}$ such that $a \neq b$. A common assumption is that the congestion effect is the same across alternatives:

$$f_{-1} = f_{+1}.$$

A commonly used form is $f_{-1}(m) = f_{+1}(m) = 1$ if $m \leq k$ and 0 otherwise (Challet and Zhang, 1997). In an alternative specification, players' payoffs decrease linearly in the number of players taking the same action. The exact form of the utility function has a limited impact on the predictions of the learning model (Li et al., 2000). The minority game is thus a congestion game (Rosenthal, 1973), with the payoffs to each action depending on the number of players taking that action. It is therefore a (finite exact) potential game (Monderer and Shapley, 1996).

Tercieux and Voorneveld (2010) and Kets and Voorneveld (2007) characterize the pure and mixed Nash equilibria of the game, respectively. The game has a large number of Nash equilibria (in fact a continuum). Most notably, in each pure Nash equilibrium, k players choose one alternative, and $k+1$ players the other. These equilibria are payoff-asymmetric, but not strict, as a deviation from a player from the majority action to the other alternative would shift the minority. There is a unique symmetric mixed Nash equilibrium in which all players randomize between both actions with probability 1/2. In addition, there are mixed equilibria in which some subset of the players randomize, while others choose a mixed strategy. In particular, there is a class of equilibria in which k players choose $a = -1$, k players choose $a = +1$, and the remaining player randomizes between both alternatives with any probability.

The minority game is closely related to two congestion games that have been studied experimentally, the market entry game and the route-choice game. In a market entry game, introduced by Selten and Güth (1982), players must decide whether to enter a market with a fixed capacity c or to stay out.[2] Players who enter the market receive a payoff that decreases in the number of entrants. The payoff to players who stay out of the market is typically taken to be constant. Like the minority game, the game generally has a large number of Nash equilibria, both in pure and in mixed strategies. An important difference between the market entry game and the minority game is that in the latter game, the payoffs of both actions depend on the number of players choosing that action, while in the market entry game, players can choose an action (staying out) whose payoffs are independent of the number of players choosing that action.

Route-choice games are closer to the minority game in that the payoffs of all actions are subject to congestion. In a route-choice game, players choose between two or more roads, where the payoffs of choosing a road decrease in the number of other players who choose that road (see Selten et al., 2007, and references therein). For the cost functions commonly studied in the literature, the equilibria of the game are such that players divide themselves over the roads so as to equalize travel times and thus payoffs. Consequently, the pure Nash equilibria of the game are payoff-symmetric and strict, in contrast with the minority game.

A priori it is not clear whether players can learn to play according to an equilibrium in these games, for a number of reasons. First, these games have a large number of equilibria, so that players face the difficulty of coordinating on the same equilibrium. As the equilibria in pure strategies cannot be Pareto-ranked or ordered in terms of risk-dominance, no particular pure-strategy Nash equilibrium can be singled out as being most salient (Schelling, 1960), so that without pre-play communication, players may not have enough information to implement a pure Nash equilibrium (cf. Menezes and Pitchford, 2006). Second, in congestion games, players may obtain asymmetric payoffs in equilibrium which may complicate attainment of equilibrium, as coordination cannot be achieved through tacit coordination based on historical precedent (cf. Meyer et al., 1992). Third, the pure Nash equilibria of the game require players to sort themselves across alternatives. This violates the common belief that in symmetric games, all rational players will evaluate the situation identically, and hence, make the same choices in similar situations (see Harsanyi and Selten, 1988, p. 73). The question whether players can learn to play according to an equilibrium in such congestion games, despite the apparent difficulties, has motivated a large number of experimental studies. The next section surveys some remarkable experimental results on these games.

3 Aggregate and individual behavior in congestion games

A robust experimental finding in market entry games is that subjects quickly learn to coordinate their play in the sense that aggregate play is close to equilibrium (Ochs, 1990, p. 169). For instance, Erev and Rapoport (1998) find that the number of entrants in a market entry game rapidly converges to the equilibrium value. Similarly, in their experiments on route-choice games, Selten et al. (2007) observe that the mean number of drivers on the different roads is very close to the equilibrium number. Similar experimental results have been reported for the minority game (Bottazzi and Devetag, 2007; Chmura and Pitz, 2006).

However, individual players generally do not play equilibrium strategies, even after a substantial learning period.[3] For instance, Erev and Rapoport (1998) observe large between- and within-subject variability which does not diminish with experience. In the context of route-choice games, Selten et al. (2007) reports that large fluctuations in individual behavior persist until the end of the experiment, as is the case in experiments on the minority game (Bottazzi and Devetag, 2007; Chmura and Pitz, 2006). In all cases, the hypothesis that fluctuations can be explained by a symmetric mixed Nash strategy equilibrium of the game can be rejected.

In many cases, it is not clear what type of behavioral rules subjects employ when they choose an action based on previous experience. For example, Selten et al. (2007) are unable to classify 42% of the subjects in terms of the behavioral rules they use in their route-choice experiments, while Zwick and Rapoport (2002) cannot classify the behavior of some 60% of their subjects in their experiments on the market entry game. This suggests that a significant fraction of subjects do not use rules that are readily explicable in terms of optimizing behavior or based on intuitive rules of thumb.

While the individual rules may not be easy to classify, the interplay between the different rules leads to behavior that is close to equilibrium. Some authors conjecture that subjects develop counteracting behavioral rules (Bottazzi and Devetag, 2007; Chmura and Pitz, 2006; Erev and Rapoport, 1998; Rapoport et al., 2000; Selten et al., 2007; Zwick and Rapoport, 2002). For example, in the experiments of Selten et al. (2007), subjects have to choose between two roads, with the payoffs of a given road decreasing in the number of players choosing that road. Selten et al. report that some subjects use a "direct" response mode, while other subjects use a "contrary" response mode. Subject who use the former response mode will switch roads if they experienced congestion in the last period, while subjects using the contrary response mode stick with their choice, as they expect other subjects to switch. More complicated rules are of course conceivable. At the aggregate level, play will be close to equilibrium as long as there is the right mix of behavioral rules, leading players to differentiate their choices. This does not mean that individual behavior conforms to a Nash equilibrium, however. In particular, there may be correlation in players' behavior across periods.

h_m	$s_{i,1}$	$s_{i,2}$	$s_{i,3}$	$s_{i,4}$
$(-1,-1,-1)$	$+1$	-1	-1	$+1$
$(-1,-1,+1)$	-1	-1	$+1$	-1
$(-1,+1,-1)$	$+1$	-1	-1	$+1$
$(-1,+1,+1)$	-1	$+1$	-1	$+1$
$(+1,-1,-1)$	$+1$	$+1$	$+1$	$+1$
$(+1,-1,+1)$	-1	-1	$+1$	$+1$
$(+1,+1,-1)$	-1	-1	-1	$+1$
$(+1,+1,+1)$	-1	$+1$	-1	$+1$

Table 1: An example of a subset of response modes for memory length $m = 3$ for some player i. For instance, if the history of outcomes is $(-1,-1,-1)$, then the behavioral rule $s_{i,1}$ prescribes action $a_i = +1$.

These experimental findings, and in particular the persistence of nonequilibrium choices at the individual level, cannot easily be explained by most learning models, which typically predict convergence to the pure Nash equilibria of such games or to Nash equilibria with at most one player who plays a mixed strategy.[4] The next section introduces the fixed-rule learning model, which departs from most other learning models by assuming that players have a fixed set of behavioral rules or response modes.

4 Learning with fixed rules

4.1 Model

At each period in time, an odd number of players play the minority game. After each round of play, they observe the aggregate play in that period. Each player is endowed with a fixed set of behavioral rules or response modes which they use to select an action for the next period. A response mode maps the recent history of play to an action for the next period. Which response mode a player uses depends on the performance of his response modes in past periods.

More formally, time is discrete and indexed by t. At each time t, the stage game is played. After each round of play t of the stage game, the players learn which of the two actions was chosen by the minority of players. That is, if $\ell \leq k$ players chose $a = +1$ in period t, with the other players choosing $a = -1$, then players receive the signal $M(t) = +1$. It is assumed that players have a limited memory: They only retain the sequence of the minority actions in the previous m rounds, where $m \in \mathbb{N}$. That is, each player observes the same history at time t, given by $h_m(t) := (M(\tau))_{\tau \in \{t-m, t-m+1, \ldots, t-1\}}$.

A behavioral rule or *response mode* s assigns to each history $h_m \in \{-1,+1\}^m$ an action

$a \in \{-1, +1\}$. That is, a response mode s prescribes which action $s(h_m(t)) \in \{-1, +1\}$ to take, for a given history of play $h_m(t)$ at time t. Note that response modes are stationary: $s(h_m(t))$ does not depend on t, other than through $h_m(t)$. Because there are 2^m possible signals h_m of length m, and two possible actions for each signal, there are 2^{2^m} response modes. These response modes need not have an intuitive interpretation; they are simply fixed rules that a player could use to select an action for the next period.

A key assumption in the model is that each player i is endowed with a random subset S_i of the set of all response modes, which is fixed across periods. All players are endowed with the same number of response modes; denote this number by n_S. For each player i, the response modes in S_i are drawn uniformly at random from the set of all response modes for histories of length m, independently across players. An example for $n_S = 4$ and $m = 3$ is given in Table 1.

After observing the current history, a player has to choose one of his response modes to determine his action for the next period. Which rule he chooses depends on the performance of the different response modes, as summarized by their virtual score. The virtual score of each response mode is updated after each time period, regardless of whether the response mode has been used in that periodor not. When a response mode would have correctly predicted the minority action, its virtual score is increased with the payoffs it would have earned, otherwise it is decreased with the same amount. More specifically, the *virtual score* that player i assigns to response mode $s_i \in S_i$ at time t is given by:

$$p_{i,t}(s_i) = p_{i,t-1}(s_i) - \left(\frac{s_i(h_m(t))}{2k+1}\right) \left[2(k - m_{s_i(h_m(t))}(t)) + 1\right]$$

where $m_a(t)$ is the number of players choosing action $a \in \{-1, +1\}$ in round t.

It is important to note that players do not take the effect of their action on the aggregate outcome into account. In determining the virtual score of a response mode, players only consider whether this response mode would have predicted the actual outcome correctly, not taking into account whether following this response mode would have affected the outcome. That is, suppose that at time t, player i chooses $a_i = -1$, and that the total number of players choosing this action is $m_{-1}(t) = k + 1$, i.e., -1 is the majority action. Then, $2(k - (k + 1)) + 1 = -1$ would be added to all response modes prescribing action $a_i = -1$, and $-(2(k-(k+1))+1) = +1$ would be added to all response modes prescribing $a_i = +1$. However, if player i would have chosen $a_i = +1$, the number of players choosing $a = +1$ would have been $k+1$, and $+1$ would have been the majority action.

The probability that a player chooses a response mode at a given time step is determined by its virtual score at that time, with the choice probabilities following the well-known logit choice rule. For $i \in N$, denote the response mode selected by player i at time t by $s_i(t)$. Then,

$$\forall s_i \in S_i : \quad \mathbb{P}(s_i(t) = s_i) = \frac{\exp\left[\beta\, p_{i,t}(s_i)\right]}{\sum_{s_j \in S_i} \exp\left[\beta\, p_{i,t}(s_j)\right]}.$$

The parameter β can be interpreted as the sensitivity of choice to marginal information. In the limiting case $\beta \to \infty$, players choose the response modes with the highest virtual score. Otherwise, every response mode is chosen with positive probability, with the probability increasing in the virtual score of the response mode.[5]

4.2 Discussion

In this section, I discuss two assumptions of the fixed-rules learning models that are nonstandard: the assumption that players are endowed with a random subset of response modes and the assumption that players update the virtual scores of response modes without taking into account the effect of that response mode on the game's outcome.

4.2.1 Fixed rules

The learning model discussed in the previous section departs from much of the literature on learning in assuming that each player is endowed with a fixed set of rules that determine his action as a function of the recent history of play, choosing the behavioral rule with the best past performance to select their action for the next period. By contrast, most learning models either assume that players have a single behavioral rule (such as "play a (myopic) best response to the recent history of play," as in fictitious play), or choose among *actions* based on past performance (as in reinforcement learning).

The assumption that players have a fixed set of different behavioral rules from which they choose is motivated by the work of Arthur (1994). In the game considered by Arthur, players need to decide whether to go to a bar or not. Going to the bar is only pleasant if it is not too crowded. Arthur suggests that players condition their decision to go on attendance levels in the previous weeks. For example, if the bar has been crowded for the last three weeks, they expect it to be crowded next week also. These mental models are mapped into actions: if a player expects the bar to be crowded, he will not go.

There is of course a large number of possible rules of this type, as illustrated by the long list of examples in Arthur (1994). An important question, therefore, is which response modes need to be included in the model. This is a nontrivial issue. As noted by Erev and Roth (1998, p. 873), while it is virtually impossible to include all possible behavioral rules, selecting specific rules bears the risk of "parameter fitting in a model with an enormous number of parameters."

The fixed-rules learning model therefore takes an agnostic approach, endowing players with a random subset of behavioral rules, so that no response mode is ruled out a priori. Players then choose the behavioral rule that performs best, given the behavioral rules of others. This allows behavioral rules, as well as actions, to coevolve. In a game such as the minority game, this may be a reasonable assumption. Whether a response mode is successful depends *only* on the response modes used by others. Conversely, *any* response mode, whether it has a sensible interpretation or not, will work if opponents use response modes that give rise to anti-correlated play. An agnostic approach as to which response modes players use may therefore be an elegant solution to the dilemma described by Erev and Rapoport, at least for this class of games.

This approach does raise a number of questions. First, one may ask why different players are endowed with a different set of response modes. A possible justification for such an assumption is that each player has different experiences prior to playing the game, and therefore has a different view as to which response modes are successful (cf. Aumann, 1997; Fudenberg and Levine, 1998). Second, and perhaps more importantly, one could ask why players only use a fixed number of response modes. It seems that individual players have an incentive to increase the number of response modes they use, to obtain an advantage over other players.[6] Even if it is reasonable to assume that players only use a fixed subset of simple "rules of thumb" (cf. Ellison and Fudenberg, 1993), rather than considering all 2^{2^m} response modes, it seems likely that over the longer run, players will adapt the set of rules they use. It is not clear what effect this would have on the evolution of play.

4.2.2 The law of simulated effect and simple decision rules

The fixed-rules learning model assumes that players also update the virtual score of strategies that are not used in that period. The assumption that players also consider the payoffs to strategies or response modes not played seems to be a reasonable one. Camerer and Ho (1999) argue on the basis of theoretical arguments as well as experimental evidence that players obey not only the "law of *actual* effect," but also the "law of *simulated* effect," meaning that in reinforcement, not only payoffs from strategies that are actually used count, but also foregone payoffs from strategies not played.[7]

Learning models that are based on the law of simulated effect have sometimes been criticized as being too demanding in terms of players' rationality. Not only do players need to to know the payoff structure of the game and the actions of other players, they also have to calculate their payoffs for the hypothetical case in which they chose a different action. Even if players only need to know the aggregate action, as in the minority game, this may require too much rationality on the part of the players. One way to reconcile players' bounded rationality with the law of simulated effect is to assume that players do not take the effect of their own action on the global outcome into account. In that way, players can account for foregone payoffs of actions or response modes not used, without having to do complicated calculations. That is, players use simple decision rules, as opposed to clever decision rules, in the terminology of Sandholm (1998). This is indeed the route taken in the fixed-rules learning model, but as we will see, this assumption is not innocuous in the minority game.

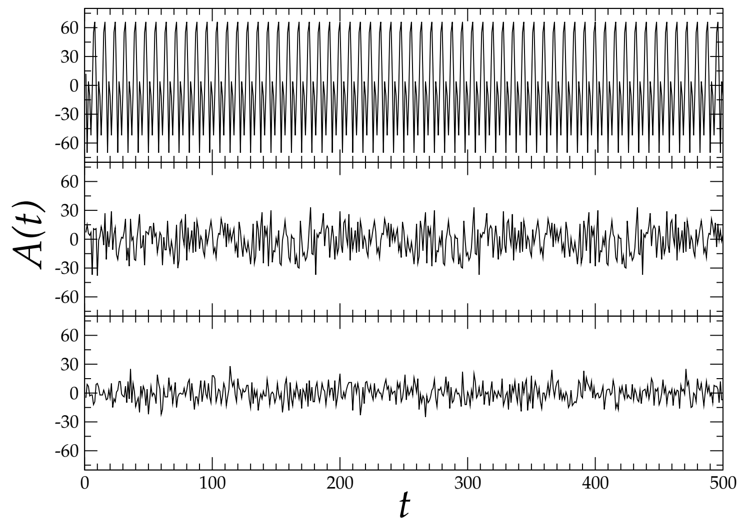

Figure 5.1: Time evolution of the aggregate action $A(t)$, where the number of players is $|N| = 301$ and each player has $n_S = 2$ response modes. Panels correspond to memory length $m = 2, 7, 15$ from top to bottom. Figure taken from Moro (2003).

5 Theoretical predictions and experimental evidence

This section provides an overview of the main results on the evolution of play in the minority game, assuming that players' adaptive behavior is described by the fixed-rules learning model. Some results in the literature are obtained analytically, others by simulations. In the simulations, a given number of agents is endowed with a random subset of response modes, and results are obtained by averaging over different assignments of response modes. In the first two sections, I characterize the behavior of the model in terms of social efficiency and informational efficiency, and show that the two are intimately related in the current learning model. In Section 5.3, I discuss how the predictions of the model can be understood in terms of the formation of groups that use counteracting response modes.

5.1 Aggregate play and volatility

An important question is whether the fixed-rules model can reproduce the experimental finding that aggregate play is close to equilibrium. In any equilibrium of the minority game, each alternative will be chosen (on average) by approximately half of the players–if not, a player in the majority could gain by switching to the minority action. Simulations of the fixed-rule learning model indeed show that the aggregate action $A(t) := \sum_{i \in N} a_i(t)$ as a function of time keeps fluctuating around 0 (corresponding to the state where each action is chosen by half of the players), as can be seen in Figure 5.1. As the game is symmetric, the time average of $A(t)$ will be 0 in the long run (e.g., Challet and Zhang, 1997).

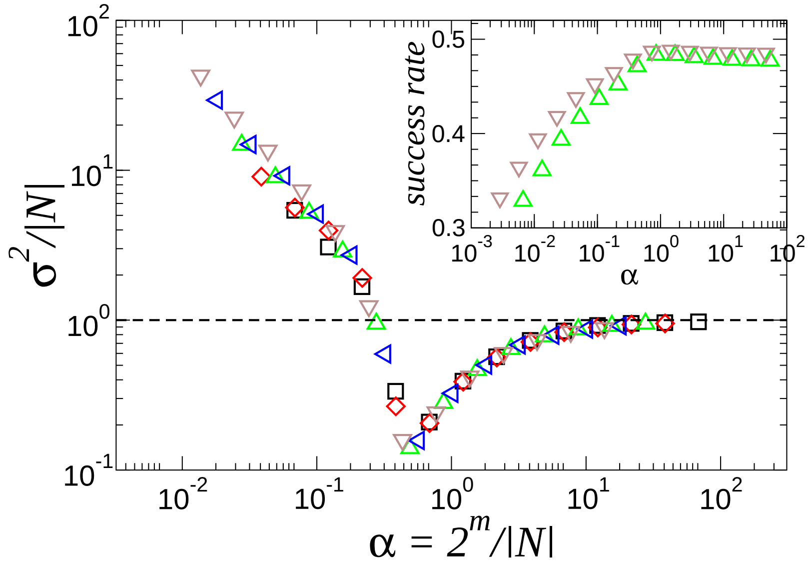

Figure 5.2: The time-averaged volatility, scaled by the number of players $|N|$ as a function of α for $n_S = 2$ response modes per player and different number of players $|N| := 101, 201, 301, 501, 701$ (□, ◊, △, ◁, ▽, respectively). Inset: players' average success rate as function of α (not discussed here; see Moro (2003) for a discussion). Figure taken from Moro (2003).

As can be seen in Figure 5.1, the variance or *volatility* of the aggregate action depends on players' memory length m. The volatility is an important variable in the present setting, as it is negatively related to players' success at coordinating their actions (as measured by total payoffs). Intuitively, large fluctuations imply that the size of the minority is only small, so that high volatility is associated with low total payoffs. It turns out that the volatility is in fact a nonmonotonic function of the memory length. Figure 5.2 shows the time-averaged volatility, denoted by σ^2, scaled by the number of players $|N| = 2k + 1$, as a function of $\alpha := 2^m/|N|$, for a fixed number of response modes $n_S = 2$. Section 5.3 discusses why it is natural to view the volatility as a function of α.

As can be seen from the figure, the behavior of the (time-averaged) volatility as a function of α is complex. In the limit of large α, the volatility converges to the volatility exhibited in the symmetric mixed Nash equilibrium. For smaller memory lengths (α small), total payoffs are much lower. In fact, the volatility is of order $(2k+1)^2$, so that the size of the group of players choosing the minority action is much smaller than k. However, for intermediate values of α, volatility is much lower than in the symmetric mixed Nash equilibrium. This suggests that players can somehow coordinate their play to reduce the volatility relative to the symmetric mixed Nash equilibrium. Coordination is not complete, however; while players come close to the level of volatility under a pure Nash equilibrium at the value of α at which volatility is at its minimum, they never reach it.

What prevents players from reaching the maximal level of coordination? It turns out that players continuously switch between the response modes they use, which makes that aggregate play does not settle down. What drives this behavior? Interestingly, this continuous switching is driven by the assumption that players update the virtual scores of all their response modes, whether they were used or not, not taking into account what the effect of their own action would have been on the collective outcome for response modes not used. While one might think that with a large number of players, the effect of one player's action will be limited, this is not the case for the minority game. Recall that response modes that are not used are rewarded if they prescribe the minority action for that period, even if playing that action would have tipped the minority to the other side. The response mode that was actually used of course does not have that advantage. This makes that players keep switching between response modes: over time, a response mode that is not played for some time will gather sufficiently many virtual points so as to be selected to be played, thus losing its advantage, until another response mode takes over again. Indeed, if players correct for this bias by allocating a small additional reward to the response mode that they currently use, players use the same response mode in every period (in the limit $\beta \to \infty$), and play converges to one of the pure Nash equilibria of the game, provided that the memory length m is sufficiently small (Marsili et al., 2000).

5.2 Information and efficiency

As discussed in the previous section, players seem to be able to coordinate reasonably well for some parameter value. How does this coordination come about? The only way players can coordinate is through the history of play. The question is thus whether the past history of play contains information that can help players coordinate their action. The information content of the history of play, or the degree of predictability can be measured by (Challet and Marsili, 1999):

$$H := \frac{1}{|\mathcal{H}_m|} \sum_{\nu \in \mathcal{H}_m} \langle A(t+1) \mid h_m(t) = \nu \rangle^2,$$

where $\langle A(t+1)|h_m(t) = \nu\rangle$ is the time average of the aggregate action conditional on a given history of play (for t large), and ν runs over the possible histories $\mathcal{H}_m$ of length m. Loosely speaking, H measures the information in the time series of $A(t)$. If $A(t+1)$ and $h_m(t)$ are independent, then $H = 0$. If $H > 0$, then the signal $A(t)$ contains information.

The way players choose their response modes in the fixed-rules learning model leads them to minimize the degree of predictability (Marsili et al., 2000). It is again the parameter α which determines how successful players are at doing that, as can be seen in Figure 5.3. Interestingly, the system changes from an informationally efficient phase with low aggregate payoffs ($H = 0$, σ^2 large) to an information-rich phase with high aggregate payoffs ($H > 0$, σ^2 small) at a critical value α_c of α, and α_c is exactly the value where the volatility is at its minimum (Figure 5.2).

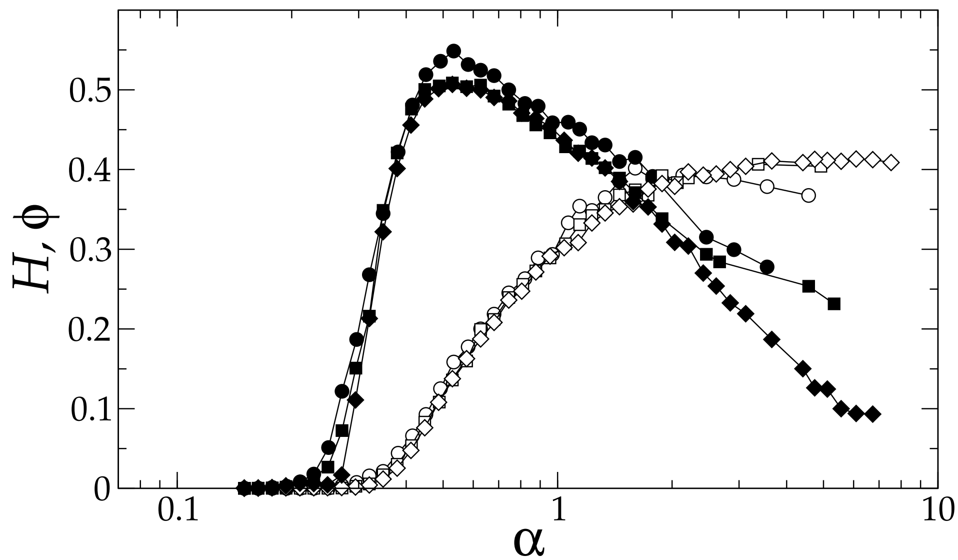

Figure 5.3: Information H (open symbols) and fraction of frozen players ϕ (full symbols, not discussed here; see Moro, 2003) as a function of $\alpha = 2^m/|N|$ for $n_S = 2$ response modes per player and memory length $m = 5, 6, 7$ ($\circ$, $\square$, and $\diamond$, respectively). Figure taken from Moro (2003).

This transition between the informationally efficient and the information rich phase, or equivalently between the phase with low aggregate payoffs and the phase with high aggregate payoffs, is central to the current learning model. At this transition, there is a qualitative change in collective behavior, while the principles behind the behavior of individual players remain unchanged. For all values of α, players try to outsmart each other, but for low values of α, they are on average less successful. The next section discusses the interpretation of α.

5.3 Response modes and their antagonists

The previous section shows that the qualitative behavior of the learning model is primarily a function of $\alpha = 2^m/|N|$, where m is players' memory length and $|N|$ is the number of players (for a fixed number of response modes n_S). Why does this parameter drive the behavior of the model?

Intuitively, the minority rule forces players to differentiate their actions. As there are 2^{2^m} response modes for $|N|$ players, one would expect that players succeed in differentiating if $|N|$ is much smaller than 2^{2^m}, and be unsuccessful when $|N|$ exceeds 2^{2^m}, which means that one would expect a qualitative change in behavior when $|N|$ is of the order 2^{2^m}, rather than of order 2^m, as observed.

The reason why it is the ratio between 2^m and $|N|$ that drives behavior rather than that between 2^{2^m} and $|N|$ is that the important variable is not how many response modes players can use, but how many *different* response modes they have. Say that two response modes are sufficiently anti-correlated if they prescribe a different action after at least half of the histories of play (of length m). Challet and Zhang (1998) have shown that for each response mode, the number of response modes that are sufficiently anti-correlated with it is proportional to 2^m. This means that the qualitative behavior of the model is driven by the ratio between 2^m and $|N|$.

This leads us to an intuitive interpretation of the model's results in terms of the interaction of groups using different response modes. Take two response modes s and $\bar{s}$ which describe a different action after every possible history of length m, i.e., s and $\bar{s}$ are anti-correlated. Suppose N_s players use the response mode s in a given time period, while $N_{\bar{s}}$ players use the anti-correlated response mode $\bar{s}$. If N_s is approximately equal to $N_{\bar{s}}$ for all anti-correlated pairs $(s, \bar{s})$ of response modes, then the actions of players effectively cancel out and the volatility will be small.

Whether the actions of players using a certain response mode s and its antagonist $\bar{s}$ can cancel out depends on the dimension of the space of response modes is fixed by the parameter m. When the number of players is large relative to the number of response modes that are available (α small), then players are forced to use response modes that are positively correlated, so that volatility will be large. For intermediate values of α, players use response modes that are either sufficiently anti-correlated, so that volatility is minimal. When m is very large relative to the number of players, the number of players using each response mode will only be small, so that players act more or less independently and the volatility will be close to that under the symmetric mixed Nash equilibrium (Moro, 2003). However, aggregate payoffs are still higher than under the benchmark of the symmetric mixed Nash equilibrium, as there always exist pairs of players that follow anti-correlated response modes (Challet and Zhang, 1998).

5.4 Comparison with experimental results

Can we relate this to the experimental results in Section 3? As noted above, the aggregate behavior is close to Nash play, in the sense that on average, the fraction of players choosing each alternative is close to 1/2. The idea that players achieve this by using counteracting response modes seems intuitive. Indeed, Selten et al. (2007) find that players sort in groups in the route-choice game, with one group of subjects switching roads if the road they chose in the previous period was congested ("direct response mode"), another group of subjects sticking to their choice of road if it was congested ("contrary response mode"), and subjects whose behavior is harder to classify. The number of players using the direct response mode was not close to the number of players using the contrary response mode, however, and indeed, large fluctuations in play around Nash play persisted until the end of the experiment. Also, Bottazzi and Devetag (2007) find that there is considerable heterogeneity in players' behavior in their experiments on the minority game. They show that it is not the heterogeneity per se which determines the players' success in coordinating, rather, it is the interaction between these different behavioral rules that make that players can successfully coordinate their play.

While the fixed-rules learning model thus demonstrates how aggregate Nash play and individual nonequilibrium behavior can be reconciled in a natural way, there are two important caveats. First, as noted earlier, there is a large fraction of subjects whose behavior cannot easily be classified in terms of response modes (Selten et al., 2007; Zwick and Rapoport, 2002). This could mean that subjects use response modes that may not have an intuitive interpretation (and are thus not recognized by the experimenters) but that nevertheless perform well, but it could also mean that their behavior is driven by an entirely different mechanism. A systematic study of the different response modes used by experimental subjects seems needed. Indeed, Zwick and Rapoport (2002) conclude that there is a need "to re-orient research on interactive decision making to individual differences, identify patterns of behavior shared by subsets of players ..., and then attempt to account for aggregate behavior in terms of the behavior of the clusters of players that form these aggregates."

Second, a key prediction of the fixed-rules model is that players continue to switch between response modes (see the discussion in Section 5.1). This has not been tested experimentally; the few experiments that have tried to classify subjects' response modes only look whether subjects use a *stationary* response mode. Testing whether and how subjects switch between response modes seems challenging, as changes in actions and in response modes cannot easily be separated.

6 Concluding remarks

While adaptive behavior in congestion games is well understood theoretically, a number of experimental findings seem hard to explain using standard learning models. This paper surveys an alternative learning model, whose distinguishing feature is that players are endowed with a random set of behavioral rules or response modes. It is shown how models of this type can aid in reconciling seemingly conflicting experimental results in congestion games. In particular, it offers an intuitive explanation of how aggregate play can be close to Nash equilibrium behavior in congestion games, while individual play does not conform to any equilibrium. In the model, players naturally sort into groups that use counteracting response modes, which means that aggregate play will be close to the equilibrium prediction, while individual behavior can be very far from equilibrium.

There are a number of important open questions. First, as argued in the previous section, more experimental tests are needed to see whether the model provides a good description of experimental play in congestion games. In particular, a better understanding of individual behavior and changes in behavior is needed. On the theoretical front, it seems natural to apply the model to other congestion games. An important question is to what extent the prediction that players keep switching response modes holds up in other games. Finally, as argued in Section 4.2, it seems natural to allow players to adapt the set of response modes they use over time, for example allowing them to experiment with different response modes at a low rate, to replace response modes that have not been performing well. How this would affect predictions is an open question.

Acknowledgements

I am indebted to Ginestra Bianconi, George Ehrhardt, Doyne Farmer, Matteo Marsili, Esteban Moro, Jan Potters, Dolf Talman, Mark Voorneveld, Eduardo Zambrano and two anonymous referees for helpful comments and suggestions. In addition, I would like to thank Esteban Moro for his kind permission for reproducing some of the figures from Moro (2003).

Notes

[1]This paper does not aim to provide a comprehensive survey of the extensive literature on the minority game, as it is impossible to justice to the extensive literature here. See Moro (2003); Challet et al. (2004) or Coolen (2005) for an introduction to the field. Papers in economics on the minority game include Bottazzi and Devetag (2007), Chmura and Pitz (2004), and Renault et al. (2005). Blonski (1999) and Kojima and Takahashi (2007) study learning in games very similar to the minority game.

[2]See Ochs (1999) for a survey of the experimental literature on the market entry game. Kahneman (1988) proposed a similar game.

[3]An exception is Duffy and Hopkins (2005) who find that subjects coordinate on one of the pure Nash equilibria of the market entry game after a large number of rounds when they are given feedback on others' choices.

[4]See Duffy and Hopkins (2005) and Kets and Voorneveld (2007) for a discussion of the predictions of different learning models for market entry games and minority games, respectively.

[5]The model is thus very similar to that analyzed by Cominetti et al. (2010). The main difference is that in that model, each player tracks the past performance of every strategy, which is mapped into a vector that specifies the probability with which the player chooses each strategy in the next period. Here, players choose response modes based on their past performance, and the response mode in turn determines the choice of action.

[6]Indeed, players in the minority game obtain higher payoffs when they use a larger set of response modes, ceteris paribus (Marsili et al., 2000).

[7]A similar idea underlies models of regret-based learning (Foster and Vohra, 1999; Hart and Mas-Colell, 2000; Young, 2004).

References

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review 84*, 406–411.

Aumann, R. (1997). Rationality and bounded rationality. In S. Hart and A. Mas-Colell (Eds.), *Cooperation: Game-Theoretic Approaches*, Volume 155 of *NATO ASI Serie F*. Berlin: Springer.

Blonski, M. (1999). Anonymous games with binary actions. *Games and Economic Behavior 28*, 171–180.

Bottazzi, G. and G. Devetag (2004). Coordination and self-organization in minority games: Experimental evidence. In M. Gallegati, A. Kirman, and M. Marsili (Eds.), *The Complex Dynamics of Economic Interaction: Essays in Economics and Econophysics*, Volume 521 of *Lecture Notes in Economics and Mathematical Systems*. Berlin: Springer-Verlag.

Bottazzi, G. and G. Devetag (2007). Competition and coordination in experimental minority games. *Journal of Evolutionary Economics 17*, 241 – 275.

Camerer, C. and T.-H. Ho (1999). Experience-weighted attraction learning in normal form games. *Econometrica 67*, 827–874.

Challet, D. and M. Marsili (1999). Phase transition and symmetry breaking in the minority game. *Physical Review E 60*, R6271–R6274.

Challet, D., M. Marsili, and Y.-C. Zhang (2004). *Minority Games: Interacting Agents in Financial Markets*. Oxford: Oxford University Press.

Challet, D. and Y. C. Zhang (1997). Emergence of cooperation and organization in an evolutionary game. *Physica A 246*, 407–418.

Challet, D. and Y. C. Zhang (1998). On the minority game: Analytical and numerical studies. *Physica A 256*, 514–532.

Chmura, T. and T. Pitz (2004). Minority game - Experiments and simulations of traffic scenarios. *Bonn Econ Discussion Papers 23/2004*.

Chmura, T. and T. Pitz (2006). Successful strategies in repeated minority games. *Physica A 363*, 477–480.

Cominetti, R., E. Melo, and S. Sorin (2010). A payoff-based learning procedure and its application to traffic games. *Games and Economic Behavior 70*, 71–83.

Coolen, A. (2005). *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents*. Oxford: Oxford University Press.

DeAngelis, D. and L. Gross (1992). *Individual-Based Models and Approaches in Ecology: Populations, Communities, and Ecosystems*. New York: Chapman and Hall.

Duffy, J. and E. Hopkins (2005). Learning, information, and sorting in market-entry games: Theory and evidence. *Games and Economic Behavior 51*, 31–62.

Ellison, G. and D. Fudenberg (1993). Rules of thumb for social learning. *Journal of Political Economy 101*, 612 – 643.

Erev, I. and A. Rapoport (1998). Coordination, "magic," and reinforcement learning in a market entry game. *Games and Economic Behavior 23*, 146–175.

Erev, I. and A. Roth (1998). Predicting how people play games: Reinforcement learning in experimental games with unique mixed strategy equilibria. *American Economic Review 88*, 848–881.

Foster, D. and R. Vohra (1999). Regret in the on-line decision problem. *Games and Economic Behavior 29*, 7–35.

Fudenberg, D. and D. Levine (1998). *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Harsanyi, J. C. and R. Selten (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Hart, S. and A. Mas-Colell (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica 68*, 1127–1150.

Hofbauer, J. and E. Hopkins (2005). Learning in perturbed asymmetric games. *Games and Economic Behavior 52*, 133–152.

Hofbauer, J. and W. H. Sandholm (2002). On the global convergence of stochastic fictitious play. *Econometrica 70*, 2265–2294.

Huberman, B. and R. Lukose (1997). Social dilemmas and internet congestion. *Science 277*, 535–537.

Kahneman, D. (1988). Experimental economics: A psychological perspective. In R. Tietz, W. Albers, and R. Selten (Eds.), *Bounded Rational Behavior in Experimental Games and Markets*, pp. 11–18. New York: Springer-Verlag.

Kets, W. and M. Voorneveld (2007). Congestion, equilibrium and learning: The minority game. *Discussion Paper, Tilburg University*.

Kojima, F. and S. Takahashi (2007). Anti-coordination games and dynamic stability. *International Game Theory Review 9*, 667–688.

Li, Y., A. VanDeemen, and R. Savit (2000). The minority game with variable payoffs. *Physica A 284*, 461 – 477.

Marsili, M., D. Challet, and R. Zecchina (2000). Exact solution of a modified El Farol's bar problem: Efficiency and the role of market impact. *Physica A 280*, 522–553.

Menezes, F. M. and R. Pitchford (2006). Binary games with many players. *Economic Theory 28*, 125–143.

Meyer, D. J., J. B. Van Huyck, R. C. Battalio, and T. R. Saving (1992). History's role in coordinating decentralized allocation decisions. *Journal of Political Economy 100*, 292–316.

Monderer, D. and L. Shapley (1996). Potential games. *Games and Economic Behavior 14*, 124–143.

Moro, E. (2003). The minority game: An introductory guide. In E. Korutcheva and R. Cuerno (Eds.), *Advances in Condensed Matter and Statistical Physics*. New York: Nova Science Publishers.

Nagel, K., S. Rasmussen, and C. Barrett (1997). Network traffic as a self-organized critical phenomenon. In F. Schweitzer (Ed.), *Self-organization of Complex Structures: From Individual to Collective Dynamics*, pp. 579. London: Gordon and Breach.

Ochs, J. (1990). The coordination problem in decentralized markets: An experiment. *Quarterly Journal of Economics 105*, 545 – 559.

Ochs, J. (1999). Coordination in market entry games. In D. Budescu, I. Erev, and R. Zwick (Eds.), *Games and Human Behavior: Essays in Honor of Amnon Rapoport*, pp. 143–172. Mahwah, NJ: Erlbaum.

Rapoport, A., D. A. Seale, and E. Winter (2000). An experimental study of coordination and learning in iterated two-market entry games. *Economic Theory 16*, 661 – 687.

Renault, J., S. Scarlatti, and M. Scarsini (2005). A folk theorem for minority games. *Games and Economic Behavior 53*, 208 – 230.

Rosenthal, R. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory 2*, 65–67.

Sandholm, W. H. (1998). Simple and clever decision rules for a model of evolution. *Economics Letters 61*, 165–170.

Sandholm, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory 97*, 81–108.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press.

Schelling, T. (1960). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Selten, R. and W. Güth (1982). Equilibrium point selection in a class of market entry games. In M. Diestler, E. Fürst, and G. Schwödiauer (Eds.), *Games, Economic Dynamics, and Time Series Analysis.* Wien-Würzburg: Physica-Verlag.

Selten, R., M. Schreckenberg, T. Pitz, T. Chmura, and S. Kube (2007). Commuters' route choice behavior. *Games and Economic Behavior 58*, 394 – 406.

Tercieux, O. and M. Voorneveld (2010). The cutting power of preparation. *Mathematical Methods of Operations Research 71*, 85–101.

Young, H. P. (2004). *Strategic Learning and its Limits.* Arne Ryde Memorial Lectures. Oxford, UK: Oxford University Press.

Zwick, R. and A. Rapoport (2002). Tacit coordination in a decentralized market entry game with fixed capacity. *Experimental Economics 5*, 253 – 272.