

Congestion, equilibrium and learning: The minority game*

Willemien Kets[†] Mark Voorneveld[‡]

July 3, 2009

Abstract

The minority game is a simple congestion game with two actions and an odd number of players. Players want to choose the action that is chosen by the minority of players. We characterize the set of equilibria of the game. As the set of equilibria is large, a natural question is which equilibrium will be selected by the players. Since the pure Nash equilibria of the game are not strict and lead to payoff asymmetry among players, intuitively one would predict that players play according to the unique symmetric mixed strategy equilibrium. However, we show that many standard learning processes do not predict convergence to that equilibrium, and that in fact these processes converge to distinct equilibria.

JEL classification: C72, D83

Keywords: Learning, congestion games, replicator dynamic, perturbed best response dynamics, quantal response equilibria, best-reply learning

*We thank Matteo Marsili and Wieland Müller for comments and suggestions. Financial support from the Netherlands Organization for Scientific Research (NWO) and the Wallander/Hedelius Foundation is gratefully acknowledged.

[†]Corresponding author. Santa Fe Institute and CentER, Tilburg University, The Netherlands. Address: Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM, 87501, USA. E-mail: willemien.kets@santafe.edu. Phone: +1 505 946 2782.

[‡]Dept. of Economics, Stockholm School of Economics, Sweden and Dept. of Econometrics and Operations Research, Tilburg University, The Netherlands.

1 Introduction

Congestion games are ubiquitous in economics. In a congestion game (Rosenthal, 1973), players use several facilities from a common pool. The costs or benefits that a player derives from a facility depend on the number of users of that facility. A congestion game is therefore a natural game to model scarcity of common resources. Applications include commuters choosing travel modes or routes (Rosenthal, 1973; Selten et al., 2007), firms deciding whether to enter a market (Selten and Güth, 1982), routing problems in computer networks (Nisan et al., 2007), and more mundane problems like crowding in local bars (Arthur, 1994).

Congestion games are also interesting from a theoretical point of view. In congestion games, players need to coordinate to differentiate. This seems to be more difficult than coordinating on the same action, since any commonality of expectations is broken up. For instance, when commuters have to choose between two roads, A and B , and all believe that the others will choose road A , nobody will choose that road to avoid traffic jams, invalidating beliefs. The sorting of players predicted in the pure-strategy Nash equilibria of such games 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). Moreover, in congestion games, players may obtain asymmetric payoffs in equilibrium. This may complicate attainment of equilibrium, as coordination cannot be achieved through tacit coordination based on historical precedent (cf. Meyer et al., 1992). Finally, congestion games often have many equilibria, so that players also face the difficulty of coordinating on the same equilibrium.

Therefore, it is of interest to know what type of behavior game theory predicts in such games. In this paper, we characterize the equilibria of the minority game, a simple congestion game based on the El Farol bar problem of Arthur (1994), and we study the limiting behavior of a number of well-known learning processes for this game. In the minority game, an odd¹ number of players choose between two *ceteris paribus* identical alternatives. Congestion is costly, so players prefer the alternative chosen by the smallest number of players.

Because the pure-strategy Nash equilibria of the game are not strict and lead to payoff

¹The minority game can also be analyzed for an even number of players (when the definitions are suitably modified, since the minority may not be well defined). However, for typical payoff specifications, the pure-strategy equilibria (in which the number of players choosing both actions is the same) are strict, and there is no payoff asymmetry. Hence, the coordination problem will not be as severe as in the current setting.

asymmetry among players, an intuitive prediction would be that players play according to the unique symmetric mixed strategy equilibrium (cf. Menezes and Pitchford, 2006). However, we show that many standard learning processes—such as the replicator dynamic, perturbed best response dynamics and best reply learning models with limited memory—do not converge to that equilibrium. In fact, the convergence results of these processes are decidedly different. The learning processes we study typically converge to a pure Nash equilibrium or to an equilibrium in which one player plays a completely mixed strategy. While pure equilibria and equilibria with one player using a mixed strategy are observationally similar to the symmetric mixed strategy profile when the number of players is large (as is typically the case in congestion games),² these equilibria may differ substantially in terms of expected payoffs. In the equilibria with at least one mixer, there is a positive probability that one alternative is severely overcrowded. When the payoffs of an action fall sharply in the number of players choosing that action, this means that the expected payoffs of the mixed equilibrium are considerably lower than the payoffs of a pure equilibrium, even if the number of players is large (cf. Menezes and Pitchford, 2006).

The current paper builds on the literature on learning in congestion games and more generally learning in potential games (e.g. Hofbauer and Hopkins, 2005; Hofbauer and Sandholm, 2002; Sandholm, 2001, 2007). Papers that study learning in games similar to the game considered here include Blonski (1999), Franke (2003), and Kojima and Takahashi (2004). Most of these papers focus on the predictions of a single learning model or class of models,³ while we compare predictions from different learning models. Moreover, while most results are obtained for games with either a small number or a continuum of players, we characterize the equilibria of the game and the limiting behavior of different learning processes for any (odd) number of players.

The minority game has been studied by a number of authors in economics. Renault et al. (2005) study repeated play in the game. Bottazzi and Devetag (2007) and Chmura and Pitz (2006) study the game experimentally. In addition, the game has been studied extensively in the physics literature; see Kets (2007) for a discussion on the relation between this literature and the literature in experimental economics and learning. The minority

²In the pure Nash equilibria of a minority game with $2k + 1$ players, exactly k of them choose one alternative, and $k + 1$ the other. In the equilibria with one mixer, there are k players who choose the first alternative and k players the second with probability 1, and the remaining player mixes between the two alternatives with arbitrary probability. As in the symmetric mixed strategy equilibrium, in which all players chooses each alternative with equal probability, players divide themselves more or less evenly over the two alternatives.

³Kojima and Takahashi (2004) is a notable exception.

game belongs to the class of anonymous games (Feddersen and Pesendorfer, 1997; Kalai, 2004; Mailath and Postlewaite, 1990; Schmeidler, 1973). In this class of games, payoffs only depend on the number or share of other players taking a certain action. Of particular interest is the work of Blonski (1999) and Menezes and Pitchford (2006) who study games closely related to ours. The focus of these papers is however very different from ours. Menezes and Pitchford (2006) study a binary-action congestion game where one of the actions is superior in terms of payoffs. They ask under what conditions an increase in the number of players can lead players to choose the inferior action with higher probability in the symmetric mixed equilibrium. Blonski (1999) characterizes the equilibria in a binary action game with a continuum of players and studies monotone learning dynamics in these games. By contrast, the focus of the current paper is on the predictions of different learning processes for an anonymous congestion game with a finite number of players.

The outline of this paper is as follows. In Section 2, we define the game and characterize its Nash equilibria. In Section 3, we characterize the set of stationary states and the set of asymptotically stable states under the replicator dynamic. In Section 4, we characterize the set of stationary states under the perturbed best-response dynamics. In Section 5, we characterize the limiting behavior in the minority game under best-reply learning processes with limited memory. Section 6 concludes. Proofs are contained in the appendices.

2 The minority game

2.1 Basic definitions

The minority game is a simple binary-action game where players want to choose the “minority” action. In line with the literature on this topic, we assume that the number of players is odd, so that the minority is well defined. That is, the set of players is $N = \{1, \dots, 2k + 1\}$, where $k \in \mathbb{N}$. Each player $i \in N$ has a set of pure strategies $A_i = \{-1, +1\}$: agents have to choose between two alternatives. The set of mixed strategies of player i is denoted by $\Delta(A_i)$. We denote a mixed strategy profile by $\alpha \in \times_{i \in N} \Delta(A_i)$, and we use the standard notation $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(A_j)$ to denote a strategy profile of players other than $i \in N$.

In the minority game, the payoffs to a player of choosing a certain action only depend on the number of players choosing the same action. The utility function of a player i taking action a_i can therefore be written as

$$u_i(a, a_{-i}) = f_{a_i}(|\{j \in N : a_j = a_i\}|), \quad (2.1)$$

where

$$f_{a_i} : \{1, \dots, 2k + 1\} \rightarrow \mathbb{R}$$

is a function that indicates for each $n \in \{1, \dots, 2k + 1\}$ the payoff $f_a(n)$ to a player choosing a_i when the total number of players choosing a_i equals n . Payoffs are extended to mixed strategies in the usual way.

The function f_a can have several forms. In the literature on the minority game (e.g. Challet et al., 2004), it is typically assumed that payoffs to an action are proportional to the “excess proportion” of players choosing that action. Specifically, it is assumed that

$$f_{-1}(m) = f_{+1}(m) = \frac{2(k - m) + 1}{2k + 1}$$

for all $m \in \{1, \dots, 2k + 1\}$. We make the more general assumptions that congestion is costly:

$$\text{[Mon]} \quad f_{-1} \text{ and } f_{+1} \text{ are strictly decreasing functions,}$$

and that the congestion effect is the same across alternatives:

$$\text{[Sym]} \quad f_{-1} = f_{+1}.$$

We refer to a player who uses a mixed strategy that puts positive probability on both pure strategies a *mixer*. A player that puts full probability mass on the alternative -1 is called a (-1) -*player*; similarly, a player that puts full probability mass on the alternative $+1$ is called a $(+1)$ -*player*.

2.2 Nash equilibria

Throughout this section, we fix $k \in \mathbb{N}$ and consider a minority game with $2k + 1$ players. We characterize the set of Nash equilibria of this game. The pure Nash equilibria are easy to characterize:

Proposition 2.1 [Tercieux and Voorneveld (2005)] *A pure strategy profile is a Nash equilibrium if and only if one of the alternatives -1 or $+1$ is chosen by exactly k of the $2k + 1$ players.*

It remains to characterize the game’s Nash equilibria with at least one mixer. A useful insight is that all mixers play the same strategy:

Lemma 2.2 *Let $\alpha \in \times_{i \in N} \Delta(A_i)$ be a Nash equilibrium. All mixers use the same strategy: for each $i, j \in N$, if $\alpha_i, \alpha_j \notin \{(1, 0), (0, 1)\}$, then $\alpha_i = \alpha_j$.*

Since all mixers use the same strategy and player labels are irrelevant by [Sym] (if α is a Nash equilibrium, so is every permutation of α), a Nash equilibrium with mixers can be summarized by its *type* (ℓ, r, λ) , where $\ell, r \in \{0, 1, \dots, 2k+1\}$ denote the number of players choosing pure strategy -1 or $+1$, respectively, and $\lambda \in (0, 1)$ is the probability with which the remaining $m(\ell, r, \lambda) := (2k+1) - (\ell + r) > 0$ mixers choose -1 . For convenience, we simply write $m := m(\ell, r, \lambda)$. It is important to note that the full distribution of player choices matters, not just the expected number of players choosing each option.

This notation allows us to write expected payoffs in a convenient way. Denote the expected payoff to a player choosing -1 by $v_{-1}(\ell, r, \lambda)$; $v_{+1}(\ell, r, \lambda)$ is defined similarly. In particular, if one of the mixers in (ℓ, r, λ) deviates to the pure strategy -1 , there are $m-1$ mixers left. To obtain expected payoffs, notice that the probability that $s \in \{0, \dots, m-1\}$ of these mixers choose -1 is $\binom{m-1}{s} \lambda^s (1-\lambda)^{m-1-s}$. Hence, the expected payoffs when one mixer deviates to -1 are given by:

$$v_{-1}(\ell+1, r, \lambda) = \sum_{s=0}^{m-1} \binom{m-1}{s} \lambda^s (1-\lambda)^{m-1-s} f_{-1}(\ell+1+s). \quad (2.2)$$

Similarly, when a mixer deviates to $+1$, we obtain

$$\begin{aligned} v_{+1}(\ell, r+1, \lambda) &= \sum_{s=0}^{m-1} \binom{m-1}{s} \lambda^s (1-\lambda)^{m-1-s} f_{+1}((r+1) + (m-1-s)) \\ &= \sum_{s=0}^{m-1} \binom{m-1}{s} \lambda^s (1-\lambda)^{m-1-s} f_{+1}(r+m-s). \end{aligned} \quad (2.3)$$

Using this notation, the Nash equilibria with at least one mixer can be characterized as follows.

Proposition 2.3

(a) (Characterization of equilibrium) *Let $\ell, r \in \{0, 1, \dots, 2k+1\}$ be such that $\ell+r < 2k+1$. Let $\lambda \in (0, 1)$. A strategy profile of type (ℓ, r, λ) is a Nash equilibrium if and only if*

$$v_{-1}(\ell+1, r, \lambda) = v_{+1}(\ell, r+1, \lambda). \quad (2.4)$$

- (b) **(Equilibria with one mixer)** *There exist equilibria with exactly one mixer. These equilibria are of type (k, k, λ) with arbitrary $\lambda \in (0, 1)$, i.e., the mixer uses an arbitrary mixed strategy, whereas the remaining $2k$ players are spread evenly over the two pure strategies.*
- (c) **(Equilibria with more than one mixer)** *Let $\ell, r \in \{0, 1, \dots, 2k + 1\}$ be such that $\ell + r \leq 2k - 1$. There is a Nash equilibrium of type (ℓ, r, λ) if and only if $\max\{\ell, r\} < k$. The corresponding probability $\lambda \in (0, 1)$ solving (2.4) is unique.*

This characterization of the game's Nash equilibria with mixers has several implications:

- (i) There are no Nash equilibria where the number of mixers is two, since in that case, $\max\{\ell, r\} \geq k$.
- (ii) Substitution in (2.4) gives that a strategy profile in which the number of (-1) -players is equal to the number of $(+1)$ -players and the remaining players mix with probability $1/2$, i.e., a profile of type $(t, t, 1/2)$ with $t \in \{0, \dots, k\}$, is a Nash equilibrium.

Having characterized the set of Nash equilibria, we now establish that there is in fact a continuum of equilibria. Specifically, by changing the probability with which a mixer chooses each actions one can move from any equilibrium with at most one mixer to any other such equilibrium:

Proposition 2.4 *The set of Nash equilibria with at most one mixer is connected.*

This result will be useful in deriving results for the replicator dynamic in the next section.

Given that there is a wide variety of Nash equilibria, a natural question is which equilibrium will be selected. In the following sections, we study the predictions of several standard learning processes.

3 The replicator dynamic

The replicator dynamic is a simple evolutionary process that assumes roughly speaking that more successful strategies will become more prevalent over time (e.g. Weibull, 1995). No assumptions about players' rationality or cognitive sophistication are made. It is used to model learning by unsophisticated, myopic players, who may not have any knowledge of the game. In this section, we investigate the stationary and stable states of this process.

To study the replicator dynamic for the minority game, we assume that there is a set $N = \{1, \dots, 2k + 1\}$ of populations, where each population is the unit interval $[0, 1]$. The

populations represent the $2k + 1$ player positions in the minority game. All agents in a population are initially programmed to some pure strategy. Hence, each population can be divided into two subpopulations (one of which may contain no agents), one for each of the pure strategies in the minority game. A *population state* is a vector $\alpha = (\alpha_1, \dots, \alpha_{2k+1})$ in the polyhedron of mixed-strategy profiles, where for each $i \in N$, α_i is a point in the simplex $\Delta(A_i)$, representing the distribution of agents in population i across the different pure strategies. The vector $\alpha_i \in \Delta(A_i)$ thus represents the state of population i , with $\alpha_i(a_i)$ denoting the proportion of agents programmed to play the pure strategy $a_i \in A_i$.

Time is continuous and indexed by t . Agents—one from each population—are continuously drawn uniformly at random from these populations to play the minority game. Suppose payoffs represent the effect of playing the game on an agent’s fitness, measured as the number of offspring per time unit, and that each offspring inherits its single parent’s strategy. Then, the growth rate $\dot{\alpha}_i(a_i)/\alpha_i(a_i)$ of a pure strategy a_i in population i is equal to the difference in payoffs of the pure strategy and the current average payoffs for the population:

$$\dot{\alpha}_i(a_i) = \alpha_i(a_i)(u_i(a_i, \alpha_{-i}) - u_i(\alpha_i, \alpha_{-i})). \quad (3.1)$$

Hence, the population shares of strategies that do better than average will grow, while the shares of the other strategies will decline. This system of differential equations (one for each pair of population and action) defines the (continuous time multipopulation) *replicator dynamic*. It is easy to see that the subpopulations associated with the pure best replies to the current population state have the highest growth rates.

The system of differential equations (3.1) defines a continuous *solution mapping* $\xi : \mathbb{R} \times (\times_{i \in N} \Delta(A_i)) \rightarrow \times_{i \in N} \Delta(A_i)$ which assigns to each time t and each initial state α^0 the population state $\xi(t, \alpha^0)$. The (solution) *trajectory* through a population state α^0 is the graph of the solution mapping $\xi(\cdot, \alpha^0)$.

A population state α is a *stationary state* of the replicator dynamics (3.1) if and only if for each population i every pure strategy a_i that is used by some agents in the population gives the same payoffs. In that case, $\dot{\alpha}_i(a_i) = 0$ for each population i and each action a_i . Let $S = \{\alpha \in \times_{j \in N} \Delta(A_j) \mid \forall i \in N, \forall a_i \in A_i : \dot{\alpha}_i(a_i) = 0\}$ be the set of stationary states. By definition, if $\alpha \in S$, then for each population i , agents belonging to a population i either use a pure strategy or—if $\alpha_i(a_i) \in (0, 1)$ for $a_i \in A_i$ —the payoffs to the two pure strategies are equal: $u_i(+1, \alpha_{-i}) = u_i(-1, \alpha_{-i})$. Using the proof of Lemma 2.2, we see that if i and j are two populations in which both pure strategies are played by a positive share of the population, the population shares for the two strategies are the same in i and j ; formally, if $\alpha_i(a), \alpha_j(a) \in (0, 1)$ for action a , then $\alpha_i(a) = \alpha_j(a) =: \alpha(a)$. In the latter case,

the proof of Proposition 2.3(c) establishes that this α solving (2.4) is uniquely determined by the number of populations in which all agents play the pure strategy -1 or the pure strategy $+1$. Hence, the set of stationary states can be partitioned into three subsets:

S_1 : the stationary states $\alpha \in \times_{i \in N} \Delta(A_i)$ that correspond to the connected set of Nash equilibria with at most one mixer;

and a finite collection of isolated stationary states, namely

S_2 : the stationary states $\alpha \in \times_{i \in N} \Delta(A_i)$ that correspond to Nash equilibria with more than one mixer;

S_3 : the stationary states $\alpha \in \times_{i \in N} \Delta(A_i)$ that correspond to nonequilibrium profiles of some type (ℓ, r, λ) , where

$$\begin{cases} \ell, r \in \{0, \dots, 2k + 1\}, \\ \ell + r \leq 2k + 1, \\ \text{if } \ell + r < 2k + 1, \text{ then } \lambda \in (0, 1) \text{ uniquely determined by (2.4)}. \end{cases}$$

It remains to study the stability properties of these stationary states. We consider Lyapunov stability and asymptotic stability. Roughly speaking, a population state is Lyapunov stable if no small change in the population shares can lead the replicator dynamics away from the population state, while a population state is asymptotically stable if it is Lyapunov stable and any sufficiently small change in the population shares results in a movement back to the original population state. Formally, a population state $\alpha \in \times_{i \in N} \Delta(A_i)$ is *Lyapunov stable* if every neighborhood B of α contains a neighborhood B^0 of α such that $\xi(t, \alpha^0) \in B$ for every $\alpha^0 \in B^0 \cap \times_{i \in N} \Delta(A_i)$ and $t \geq 0$. It is *asymptotically stable* if it is Lyapunov stable, and, in addition, there exists a neighborhood B^* of α such that

$$\lim_{t \rightarrow \infty} \xi(t, \alpha^0) = \alpha$$

for each initial state $\alpha^0 \in B^* \cap \times_{i \in N} \Delta(A_i)$.

The analysis relies on the existence of a Lyapunov function for the replicator dynamic in the minority game. Tercieux and Voorneveld (2005), using Thm. 3.1 in Monderer and Shapley (1996), show that a minority game is a potential game. Formally, there exists a real-valued function U on the pure strategy space such that for each $i \in N$, each $a_{-i} \in \times_{j \in N \setminus \{i\}} A_j$, and all $a_i, b_i \in A_i$:

$$u_i(a_i, a_{-i}) - u_i(b_i, a_{-i}) = U(a_i, a_{-i}) - U(b_i, a_{-i}). \quad (3.2)$$

The function U is called the *potential function*. Taking expectations, (3.2) can be extended to mixed strategies. We find that the payoff difference in (3.1) equals the corresponding change in the potential. Hence, the replicator dynamic can be rewritten as:

$$\dot{\alpha}_i(a_i) = \alpha_i(a_i)(U(a_i, \alpha_{-i}) - U(\alpha_i, \alpha_{-i})) \quad (3.3)$$

for each player i and each of his actions a_i . This makes the potential U a Lyapunov function of the replicator dynamic. More precisely:

Proposition 3.1 *The potential function U of the minority game is a strict Lyapunov function for the replicator dynamic: for each solution trajectory $(\alpha(t))_{t \in [0, \infty)}$, $dU(\alpha(t))/dt \geq 0$ with equality exactly in the stationary states.*

We are now ready to investigate the stability of the classes of stationary states we identified:

Proposition 3.2 *The stationary states corresponding to Nash equilibria with at most one mixer in S_1 are asymptotically stable under the replicator dynamic. Stationary states in S_2 and S_3 are not Lyapunov stable.*

Hence, if the replicator dynamic adequately represents players' learning process, play will converge to a pure Nash equilibrium or to one of infinitely many Nash equilibria with one mixer.

As stated above, the replicator dynamic is primarily used to study learning by unsophisticated players who may not have any knowledge of the game. To investigate whether these results still hold if we assume that players are somewhat sophisticated, we turn to learning processes in the next two sections that assume that players play a best response to beliefs about other players' play. In Section 4, we study a learning process where players choose a best response to their beliefs based on past play of their opponents, but sometimes make mistakes. Section 5 analyze two learning processes in which players always play a best response to their beliefs, but beliefs are only based on recent past play.

4 Stochastic fictitious play

4.1 Perturbed best-response dynamics

Under stochastic fictitious play (e.g. Hofbauer and Hopkins, 2005; Hofbauer and Sandholm, 2002; Hopkins, 2002), players repeatedly play a normal form game. They choose best replies to their beliefs about other players' actions based on the time average of past play,

with their payoff function being subject to random shocks. That is, players play a best response on the basis of a perturbed payoff function. This may represent random changes in preferences; alternatively, it could model players who make small mistakes when trying to play a best response. This process defines the perturbed best-response dynamic.

More specifically, the state variable at time $t = 1, 2, \dots$ is a vector $Z^t \in \times_{i \in N} \Delta(A_i)$, where the i th component Z_i^t denotes the time average of player i 's past play up to time t . Players' initial choices are arbitrary pure strategies; in later periods players best-respond to their beliefs Z^t , after their payoffs have been subjected to random shocks. That is, for each $i \in N$, let $(\varepsilon_i^a)_{a \in A_i}$ be a vector of payoff disturbances. The vector of payoff disturbances is independent and identically distributed across players and over time. Let $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(A_j)$ be a belief. The probability that player i chooses action a_i is equal to the probability that

$$u_i(a_i, \alpha_{-i}) + \varepsilon_i^{a_i} \geq u_i(b_i, \alpha_{-i}) + \varepsilon_i^{b_i}$$

for $b_i \neq a_i$. A reasonable assumption is that the perturbations are Gumbel-distributed. Gumbel-distributed payoff perturbations correspond to control costs of the relative entropy form and have been extensively studied, both theoretically and experimentally. The *perturbed best-response dynamic* associated with Gumbel-distributed perturbations with parameter $\beta > 0$ is then defined by:

$$\dot{\alpha}_i(a_i) = \frac{\exp \beta u_i(a_i, \alpha_{-i})}{\exp \beta u_i(+1, \alpha_{-i}) + \exp \beta u_i(-1, \alpha_{-i})} - \alpha_i(a_i) \quad (4.1)$$

By Proposition 4.1 of Hofbauer and Sandholm (2002), the process in (4.1) has a strict Lyapunov function that can be expressed in terms of the potential function and the control cost functions. For each $i \in N$, let α_i denote the probability with which player i chooses the action $a_i = -1$. Then, the Lyapunov function for the process in (4.1) is defined by:

$$\alpha \in \times_{i \in N} \Delta(A_i) : \quad V(\alpha) = U(\alpha) - \frac{1}{\beta} \sum_{i \in N} [\alpha_i \log(\alpha_i) + (1 - \alpha_i) \log(1 - \alpha_i)], \quad (4.2)$$

where U is the potential function. Since control cost functions of the relative entropy form satisfy the smoothness conditions of Proposition 4.2 of Hofbauer and Sandholm (2002), it follows that:

Proposition 4.1 *The collection of stationary states and recurrent points of the process in (4.1) coincide.*

Theorem 6.1(iii) of Hofbauer and Sandholm (2002) now implies that the perturbed best-response dynamic converges to these stationary states. The set of stationary states coincides with the set of logit quantal response equilibria of the minority game (McKelvey

and Palfrey, 1995). When the perturbation terms go to zero, we obtain Nash equilibria. The set of Nash equilibria that are the limit points of a sequence of logit quantal response equilibria is generally hard to characterize.⁴ In the next section, we characterize this set for the three-player minority game.

4.2 Stationary points for the three-player minority game

Consider the three-player minority game with $f_{-1} = f_{+1} = f$ strictly decreasing in the number of users. As it involves a simple rescaling of functions satisfying [Mon] and [Sym], we may without loss of generality set $f(2) = 0$ and $f(1) - f(3) = 1$. A potential function for the game is then given in Figure 4.1. Throughout this section, Nash equilibria are

$$\begin{array}{cc}
 & \begin{array}{cc} -1 & +1 \end{array} \\
 \begin{array}{c} -1 \\ +1 \end{array} & \begin{array}{|c|c|} \hline -1 & 0 \\ \hline 0 & 0 \\ \hline \end{array}
 \end{array}
 \quad
 \begin{array}{cc}
 & \begin{array}{cc} -1 & +1 \end{array} \\
 \begin{array}{c} -1 \\ +1 \end{array} & \begin{array}{|c|c|} \hline 0 & 0 \\ \hline 0 & -1 \\ \hline \end{array}
 \end{array}$$

Figure 4.1: A potential function of the 3-player minority game

denoted by $(p, q, s) \in [0, 1]^3$, where p, q, s are the probabilities with which player 1, 2, and 3, respectively, choose -1 . Then, the Nash equilibria of the game are $(1/2, 1/2, 1/2)$ and $(1, 0, \lambda)$ for some $\lambda \in [0, 1]$, and permutations of these (see Section 2.2).

Given parameter $\beta \geq 0$, the conditions for a logit quantal response equilibrium (QRE) become:

$$p = \frac{1}{1 + \exp -\beta(1 - q - s)}, \quad (4.3)$$

$$q = \frac{1}{1 + \exp -\beta(1 - p - s)}, \quad (4.4)$$

$$s = \frac{1}{1 + \exp -\beta(1 - p - q)}. \quad (4.5)$$

Given $\beta \geq 0$, we denote a logit QRE in which player 1, 2 and 3 play -1 with probability p, q, s by (p, q, s, β) . We now characterize the set of Nash equilibria that are the limit of a sequence of quantal response equilibria when $\beta \rightarrow \infty$.

Proposition 4.2 *Let $(p(\beta_n), q(\beta_n), s(\beta_n), \beta_n)_{n \in \mathbb{N}}$ be a sequence of logit quantal response equilibria: $\beta_n \rightarrow \infty$ and for each $n \in \mathbb{N}$, the quadruple $(p(\beta_n), q(\beta_n), s(\beta_n), \beta_n)$ solves*

⁴As the set of Nash equilibria is not finite, we cannot apply Corollary 6.6 of Benaim (1999) to characterize the subset of Nash equilibria to which the stochastic process (4.1) converges.

equations (4.3)-(4.5). A Nash equilibrium (p, q, s) is the limit of such a sequence if and only if one of the following conditions hold:

- (a) (p, q, s) is a pure Nash equilibrium,
- (b) (p, q, s) is a Nash equilibrium with exactly one mixer who mixes uniformly,
- (c) $(p, q, s) = (1/2, 1/2, 1/2)$.

Proposition 4.2 thus characterizes the set of stationary points of the perturbed best response dynamics (4.1) for the three-player minority game.

5 Best-reply learning with limited memory

In this section, we consider discrete time learning models in which players choose best replies to beliefs that are supported by observed play in the recent past. That is, players have limited memory, in the sense that they do not take into account the full history of play when forming their beliefs, as under stochastic fictitious play.

A prominent model in the class of best-reply models with limited memory is the learning model of Hurkens (1995). In this model, players choose any action that is a best reply to some belief over other players' actions that is consistent with their recent past play. The limiting behavior of this learning process is easy to characterize. Hurkens shows that the Markov processes defined by his learning process eventually settle down in so-called minimal curb sets (Basu and Weibull, 1991). Minimal curb sets are product sets of pure strategies containing all best responses against beliefs restricted to the recommendations to the remaining players. Unfortunately, this does not provide a sharp prediction in the minority game. As shown by Tercieux and Voorneveld (2005), the unique minimal curb set in the minority game consists of the entire strategy space. That is, over time, all players will keep on choosing both actions.

We therefore consider the alternative model of Kets and Voorneveld (2008). As in the model of Hurkens (1995), it is assumed that players play a best response to beliefs over others' play supported by recent past play. In addition, players display a so-called recency bias: when there are multiple best replies to a given belief, a player chooses the best reply that he played most recently. The behavioral economics literature provides several motivations for the common observation that agents appear somewhat unwilling to deviate from their recent choices, e.g., the formation of habits (cf. Young, 1998) or the use of rules of thumb (cf. Ellison and Fudenberg, 1993).

Kets and Voorneveld show that play converges to one of the minimal prep sets of the game under this learning process. Minimal prep sets (Voorneveld, 2004) are a set-valued solution concept for strategic games that combines a standard rationality condition, stating that the set of recommended strategies to each player must contain at least one best reply to whatever belief he may have that is consistent with the recommendations to the other players, with players' aim at simplicity, which encourages them to maintain a set of strategies that is as small as possible. Tercieux and Voorneveld (2005) show that the minimal prep sets of the minority game and the pure Nash equilibria of the game coincide. Hence, under the learning model of Kets and Voorneveld, play in the minority game converges to one of the pure Nash equilibria of the game.

In both the model of Hurkens (1995) and Kets and Voorneveld (2008), players need to recall a sufficiently long period of play in order for play to converge. We now turn to the question what this lower bound on players' memory is. More specifically, suppose players remember actions that were chosen during the past T periods. A memory length of $T = 1$ is clearly insufficient for a best-reply learning process with limited memory to converge. If players chose an action profile in the previous period that is not a pure Nash equilibrium, then some action, say -1 , was chosen by more than $k + 1$ players. Hence, all players chooses the unique best reply $+1$ in the current period, and consequently the unique best reply -1 to this in the next period, and so on, with action profiles forever cycling between these two extremes. However, we show that a memory length $T = 2$ is sufficient for the learning process of Kets and Voorneveld (2008) to convergence to a pure Nash equilibrium in the minority game.

When the memory length T is equal to 2, the process is a Markov chain with state space $H = \{(a^1, a^2) \mid a^1, a^2 \in A^{2k+1}\}$, where a *history* $h = (a^1, a^2) \in H$ indicates that the $2k + 1$ players remember that they chose action profile a^1 one period ago and a^2 two periods ago. A *transition probability function* is a function $P : H \times H \rightarrow [0, 1]$, where $P(h, h') \in [0, 1]$ is the probability of moving from state $h \in H$ to state $h' \in H$ in one period and $\sum_{h' \in H} P(h, h') = 1$ for all $h \in H$. We do not need to specify exact probabilities: for the convergence result, only sign restrictions are needed.

Moving from $h = (a^1, a^2)$ to $h' = (b^1, b^2)$ in one period means that h' is obtained from h after one more round of play, i.e., by appending a new profile of most recent actions. Formally:

[P1] $h' = (b^1, b^2)$ is a successor of $h = (a^1, a^2)$, i.e., $b^2 = a^1$.

Moreover, by moving from $h = (a^1, a^2)$ to $h' = (b^1, b^2)$, the processes in Kets and Voorn-

eveland (2008) require that each player $i \in N$ chooses a best reply to a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(\{a_j^1, a_j^2\})$ with support in the product set of actions chosen in the previous $T = 2$ periods, whenever possible sticking to the most recent best reply. In games with just two actions, the latter simply means that you continue playing as you did in the previous round, unless that action is no longer a best reply to your current belief. Formally:

[P2] For each $i \in N$, b_i^1 is a best reply to some belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(\{a_j^1, a_j^2\})$. Moreover, $b_i^1 = a_i^1$ if and only if a_i^1 is a best reply to α_{-i} .

Proposition 5.1 *Consider a Markov chain on H with transition probability function P , where, for all states $h, h' \in H$, it holds that $P(h, h') > 0$ if and only if [P1] and [P2] are true. This Markov process eventually settles down in a pure Nash equilibrium.*

Some remarks are in order. First, note that, due to the symmetry of the minority game, a minor revision of the proof indicates that convergence to pure Nash equilibria can also be established if the only thing players remember from the past two periods is what they chose themselves and *how many others* did so. This comes at the expense of a more complex notation and a larger deviation from that of Kets and Voorneveld (2008).

Second, the result that the lower bound on players' memory length is $T = 2$ indicates that the requirement on memory length in Kets and Voorneveld (2008) for general games can be decreased significantly in specific cases. Although the convergence result in Kets and Voorneveld (2008) for the entire class of finite strategic games also applies, we include an explicit proof for the case of the minority game here: the structure of a minority game allows us to give a considerably shorter proof of the convergence result for this specific game, and allows us to derive a much sharper bound on the memory length.

6 Concluding remarks

Though congestion games are apparently simple, game-theorists' understanding of play in such games is far from complete. In this paper, we study the minority game, a simple congestion game in which players have to choose between two actions, and want to select the action chosen by the minority. Despite the simple structure of the game, it is not directly obvious what kind of behavior game theory would predict. We show that the game has a large number of Nash equilibria, including a unique symmetric equilibrium (in mixed strategies). As the other equilibria lead to payoff asymmetry, one may expect this equilibrium to be focal. However, we show that many standard learning processes do not

converge to this equilibrium. In fact, the learning processes we study predict convergence to equilibria that differ considerably in terms of expected payoffs.

The characterization of equilibria and convergence results of learning processes in the current paper provide an interesting starting point for experimental work, by providing a benchmark against which subjects' behavior can be evaluated. Interesting questions—which have not been addressed in the experimental literature to date on the minority game (Bottazzi and Devetag, 2007; Chmura and Pitz, 2006)—include whether behavior of experimental subjects converges to one of the Nash equilibria of the game, and which learning process best describes subjects' behavior. Experiments on the closely related market entry game have sometimes shown puzzling results (Ochs, 1990), with aggregate play being close to the Nash equilibrium prediction, but with individual play not in line with any Nash equilibrium. It would be of interest to see how experimental play in the minority game and market entry games compare, as these games are very similar in terms of equilibria and in terms of the predictions of learning processes (Duffy and Hopkins, 2005), but may be very different from a behavioral point of view.

Appendix A Proofs

A.1 Proof of Lemma 2.2

By [Sym], the 2×2 subgame played by two mixers i (row player) and j (column player) given the strategy profile of the remaining players is of the form

$$\begin{array}{cc} & \begin{array}{cc} -1 & +1 \end{array} \\ \begin{array}{c} -1 \\ +1 \end{array} & \begin{array}{|cc|} \hline x, x & y, z \\ \hline z, y & w, w \\ \hline \end{array} \end{array}$$

where, for instance, y is the payoff to the player choosing -1 if the other player chooses $+1$ and the remaining players stick to the mixed strategy profile $(\alpha_k)_{k \in N \setminus \{i, j\}}$. By [Mon], a player is better off if the other chooses differently, i.e., $x < y$ and $z > w$. Let $p, q \in (0, 1)$ denote the equilibrium probability with which player i and j , respectively, choose -1 . In equilibrium, each player must be indifferent between playing $+1$ and playing -1 :

$$\begin{aligned} px + (1 - p)y &= pz + (1 - p)w, \\ qx + (1 - q)y &= qz + (1 - q)w. \end{aligned}$$

Subtracting the latter expression from the former yields

$$(p - q)(x - y) = (p - q)(z - w).$$

As $x < y$ and $z > w$, this can only hold if $p = q$. Since mixers i and j were chosen arbitrarily from the set of mixers, this implies that all mixers use the same strategy.

A.2 Proof of Proposition 2.3

(a): Condition (2.4) says that a mixer is indifferent between choosing -1 , thereby raising ℓ to $\ell + 1$ and obtaining payoff $v_{-1}(\ell + 1, r, \lambda)$, or choosing $+1$, thereby raising r to $r + 1$ and obtaining payoff $v_{+1}(\ell, r + 1, \lambda)$. Hence, (2.4) is a necessary condition for Nash equilibrium.

To establish sufficiency, it remains to show that also players using a pure strategy—if there are such players, i.e., if $\ell + r \geq 1$ —choose a best reply. Suppose $\ell \geq 1$. The payoff to a (-1) -player is $v_{-1}(\ell, r, \lambda)$, while a unilateral deviation to $+1$ yields $v_{+1}(\ell - 1, r + 1, \lambda)$. However:

$$v_{-1}(\ell, r, \lambda) \geq v_{-1}(\ell + 1, r, \lambda) \tag{A.1}$$

$$= v_{+1}(\ell, r + 1, \lambda) \tag{A.2}$$

$$\geq v_{+1}(\ell - 1, r + 1, \lambda). \tag{A.3}$$

Inequality (A.1) uses [Mon]: conditioning on the behavior of one of the $m := m(\ell, r, \lambda) > 0$ mixers, write

$$v_{-1}(\ell, r, \lambda) = \lambda v_{-1}(\ell + 1, r, \lambda) + (1 - \lambda)v_{-1}(\ell, r + 1, \lambda).$$

Then

$$\begin{aligned} v_{-1}(\ell, r, \lambda) - v_{-1}(\ell + 1, r, \lambda) &= (1 - \lambda) [v_{-1}(\ell, r + 1, \lambda) - v_{-1}(\ell + 1, r, \lambda)] \\ &= (1 - \lambda) \sum_{s=0}^{m-1} \binom{m-1}{s} \lambda^s (1 - \lambda)^{m-1-s} [f_{-1}(\ell + s) - f_{-1}(\ell + 1 + s)] \\ &\geq 0 \end{aligned}$$

by [Mon]. Inequality (A.3) follows similarly and (A.2) is simply condition (2.4). So if $\ell \geq 1$, (-1) -players choose a best reply. Similarly, if $r \geq 1$, $(+1)$ -players choose a best reply.

(b): Let $\lambda \in (0, 1)$. Substitution in (2.4) and [Sym] yield that strategy profiles of type (k, k, λ) are Nash equilibria:

$$v_{-1}(k + 1, k, \lambda) = f_{-1}(k + 1) = f_{+1}(k + 1) = v_{+1}(k, k + 1, \lambda).$$

Conversely, consider a Nash equilibrium of type (ℓ, r, λ) with exactly one mixer: $\ell + r = 2k$. We establish that $\ell = r$. Suppose not. Without loss of generality, $\ell > r$. Since $\ell + r = 2k$, this implies $\ell \geq k + 1$ and $r \leq k - 1$. The expected payoff to a (-1) -player is

$$\lambda f_{-1}(\ell + 1) + (1 - \lambda)f_{-1}(\ell),$$

while deviating to +1 would yield

$$\lambda f_{+1}(r+1) + (1-\lambda)f_{+1}(r+2).$$

Since $\ell+1 > r+1$, $\ell \geq r+2$, and $\lambda \in (0,1)$, it follows from [Sym] and [Mon] that a (-1)-player would benefit from unilateral deviation, contradicting the assumption that the profile of type (ℓ, r, λ) is a Nash equilibrium. Hence, $\ell = r$.

(c): Without loss of generality, $\ell \geq r$, so $\max\{\ell, r\} = \ell$. Let $m = (2k+1) - (\ell+r) \geq 2$ be the number of mixers. By substitution, $\ell < k$ if and only if $\ell+1 < r+m$. To prove (c), it therefore remains to establish three things.

First, if $\ell+1 < r+m$, there is a $\lambda \in (0,1)$ solving (2.4). To see this, use $\ell \geq r$ to find that $\ell+m > r+1$. By [Sym] and [Mon], it follows that

$$\begin{aligned} v_{-1}(\ell+1, r, 0) &= f_{-1}(\ell+1) > f_{+1}(r+m) = v_{+1}(\ell, r+1, 0), \\ v_{-1}(\ell+1, r, 1) &= f_{-1}(\ell+m) < f_{+1}(r+1) = v_{+1}(\ell, r+1, 1). \end{aligned}$$

By the Intermediate Value Theorem (applied to $v_{-1}(\ell+1, r, \cdot) - v_{+1}(\ell, r+1, \cdot)$), there exists $\lambda \in (0,1)$ solving (2.4): there is a Nash equilibrium of type (ℓ, r, λ) .

Second, this $\lambda \in (0,1)$ solving (2.4) is unique. By (2.2), $v_{-1}(\ell+1, r, \cdot)$ is the expectation of a strictly decreasing function of a binomial stochastic variable. By stochastic dominance,⁵ this makes $v_{-1}(\ell+1, r, \cdot)$, the left-hand side of (2.4), strictly decreasing in λ . Similarly, by (2.3), the right-hand side of (2.4) is strictly increasing in λ . Hence, the functions $v_{-1}(\ell+1, r, \cdot)$ and $v_{+1}(\ell, r+1, \cdot)$ intersect at most once. By the previous step, as long as $\ell+1 < r+m$, they intersect at least once, establishing uniqueness.

Third, if $\ell+1 \geq r+m$, there is no $\lambda \in (0,1)$ solving (2.4). To see this, notice that the inequality implies

$$\ell+m > \dots > \ell+2 > \ell+1 \geq r+m > r+m-1 > \dots > r+1,$$

so by [Sym] and [Mon]:

$$f_{-1}(\ell+m) < \dots < f_{-1}(\ell+2) < f_{-1}(\ell+1) \leq f_{+1}(r+m) < f_{+1}(r+m-1) < \dots < f_{+1}(r+1).$$

Substitution in (2.2) and (2.3) yields that

$$v_{+1}(\ell, r+1, \lambda) > v_{-1}(\ell+1, r, \lambda)$$

for all $\lambda \in (0,1)$: there is no solution to (2.4).

⁵The Bernoulli distribution is trivially monotonic in the success probability with respect to first-order dominance; since stochastic dominance is closed under convolution, the binomial distribution is also monotonic in the success probability with respect to first-order stochastic dominance.

A.3 Proof of Proposition 2.4

In a Nash equilibrium with exactly one mixer, the completely mixed strategy is arbitrary. Letting the probability go to zero or one, this line piece of Nash equilibria in the strategy space has a pure Nash equilibrium as its end point. Hence, to show connectedness, it suffices to show that for each pair of pure Nash equilibria, there is a chain of pure Nash equilibria differing in exactly one coordinate connecting them.

So let x and y be distinct pure Nash equilibria. By Proposition 2.1, the majority action, i.e., the action chosen by exactly $k + 1$ players in a given Nash equilibrium, is well-defined. We need to consider two cases. First, if this action is the same in x and y , without loss of generality -1 , then $x \neq y$ implies that the $(k + 1)$ -player majorities in x and y must be distinct. Let i be such a majority player, choosing -1 in x , but $+1$ in y . Second, if the majority action is different in x and y , without loss of generality -1 in x and $+1$ in y , then by definition of a majority, the $(k + 1)$ -player majorities in x and y have a nonempty intersection. Again, let i be a majority player choosing -1 in x , but $+1$ in y .

By construction, as i is a majority player, the path of Nash equilibria in which i increases the probability of playing the action $+1$ from 0 to 1 connects x to another pure Nash equilibrium x^* with $x_i \neq x_i^* = y_i$ and $x_j^* = y_j$ for all $j \neq i$, i.e., with a strictly smaller Hamming distance to y (recall that the Hamming distance between two finite-dimensional vectors is the number of coordinates in which they differ).

As the strategy vectors have only finitely many coordinates and we can reduce the Hamming distance between pure Nash equilibria by the procedure above, the result now follows by induction.

A.4 Proof of Proposition 3.1

Suppressing time indices for ease of notation, direct calculation gives

$$\begin{aligned}
\dot{U}(\alpha) &= \sum_{i \in N} \sum_{a_i \in A_i} U(a_i, \alpha_{-i}) \dot{\alpha}_i(a_i) \\
&= \sum_{i \in N} \sum_{a_i \in A_i} \alpha_i(a_i) (U(a_i, \alpha_{-i}) - U(\alpha_i, \alpha_{-i})) U(a_i, \alpha_{-i}) \\
&= \sum_{i \in N} \sum_{a_i \in A_i} (\alpha_i(a_i) U(a_i, \alpha_{-i})^2 - U(\alpha_i, \alpha_{-i})^2) \\
&= \sum_{i \in N} (\mathbb{E}_{\alpha_i} [U(a_i, \alpha_{-i})^2] - (\mathbb{E}_{\alpha_i} [U(a_i, \alpha_{-i})])^2) \\
&= \sum_{i \in N} \text{Var}_{\alpha_i} U(a_i, \alpha_{-i}) \\
&\geq 0,
\end{aligned}$$

with equality if and only if all variances are zero, i.e., if and only if α is a stationary point of the replicator dynamics.

A.5 Proof of Proposition 3.2

To see that the set S_1 of states associated with the collection of Nash equilibria is asymptotically stable, note that S_1 is the set of global maxima of U : the potential U in (3.2) was extended to mixed strategies by taking expectations, so U achieves a global maximum in a pure strategy profile which, again by (3.2), is a pure Nash equilibrium. By symmetry, all pure Nash equilibria are global maxima of U and so are equilibria with exactly one mixer. Other strategy profiles are not global maxima of U : they are not Nash equilibria or, if they are, they involve more than one mixer, in which case they put positive probability also on pure strategy profiles that are not Nash equilibria and consequently not global maxima of U . This connected set of global maxima of the Lyapunov function U is asymptotically stable (Weibull, 1995, Thm. 6.4).

We show that elements of S_2 are not Lyapunov stable; the case for points in S_3 is similar. Let $\alpha^* \in S_2$, i.e, α^* is a population state corresponding to a Nash equilibrium with more than one mixer. Suppose it is Lyapunov stable. Since it is an isolated point of the collection of stationary states, there is a neighborhood B of α^* whose closure contains only the stationary state α^* : $\text{cl}(B) \cap S_2 = \{\alpha^*\}$. By Lyapunov stability, as long as the initial state $\alpha(0)$ lies in a sufficiently small neighborhood B' of α^* , the entire solution trajectory $(\alpha(t))_{t \in [0, \infty)}$ remains in B .

Let $i \in N$ be one of the mixers in the Nash equilibrium α^* . Since i is indifferent between his two pure strategies and the potential U measures payoff differences, it follows that

$$U(\alpha^*) = U(-1, \alpha_{-i}^*) = U(+1, \alpha_{-i}^*).$$

Consequently, $U(\gamma_i, \alpha_{-i}^*) = U(\alpha^*)$ for all mixed strategies γ_i of player i . For $\gamma_i \neq \alpha_i^*$ sufficiently close to α_i^* , it follows that $(\gamma_i, \alpha_{-i}^*) \in B'$. Hence, the entire solution trajectory $(\gamma(t))_{t \in [0, \infty)}$ with $\gamma(0) := (\gamma_i, \alpha_{-i}^*)$ remains in B . Since its starting point is not stationary, Proposition 3.1 implies that the Lyapunov function U strictly increases along the trajectory, until it may reach a stationary state. Let $\gamma^* \in \times_{j \in N} \Delta(A_j)$ be a limit point of the trajectory $(\gamma(t))_{t \in [0, \infty)}$: there is a strictly increasing sequence of time points $t_m \rightarrow \infty$ such that $\lim_{m \rightarrow \infty} \gamma(t_m) \rightarrow \gamma^*$. Such a limit point exists and has to be a stationary point (Lemma A.1 of Sandholm, 2001, p. 104). Since $\text{cl}(B) \cap S_2 = \{\alpha^*\}$ and the trajectory lies in B , it follows that $\gamma^* = \alpha^*$. But then $\lim_{m \rightarrow \infty} U(\gamma(t_m)) = U(\alpha^*) = U(\gamma(0))$, contradicting that the Lyapunov function is increasing along the trajectory. Hence, α^* cannot be Lyapunov stable. For $\alpha^* \in S_3$, we can proceed similarly. As the state does not correspond to a Nash equilibrium, some i can profitably deviate slightly (to remain inside B'), so the remaining trajectory must increase the potential, but still have α^* as its limit point.

A.6 Proof of Proposition 4.2

The only Nash equilibria not covered by (a), (b), and (c) are those with one player, say player 1, choosing -1 , one player, say player 2 choosing $+1$, and player 3 mixing with probability $\lambda \in (0, 1) \setminus \{\frac{1}{2}\}$.

Suppose, to the contrary, that such an equilibrium is the limit of a sequence of logit QRE $(p(\beta_n), q(\beta_n), s(\beta_n), \beta_n)_{n \in \mathbb{N}}$ where $\beta_n \rightarrow \infty$ and $(p(\beta_n), q(\beta_n), s(\beta_n), \beta_n)$ solves equations (4.3) to (4.5) for a logit QRE. In the selected equilibrium, both the (-1) -player and the $(+1)$ -player choose their unique best response. By Lemma 3 in Turocy (2005, p. 251), $\beta_n(1 - p(\beta_n)) \rightarrow 0$ and $\beta_n q(\beta_n) \rightarrow 0$. Substituting this in the logit QRE condition (4.5) for the third player gives that

$$s(\beta_n) = \frac{1}{1 + \exp -\beta_n(1 - p(\beta_n) - q(\beta_n))} \rightarrow \frac{1}{2},$$

contradicting the assumption that $\lim_{n \rightarrow \infty} s(\beta_n) = \lambda \neq 1/2$.

It remains to show that the classes of equilibria in the proposition are indeed limits of a sequence of logit QREs.

(a): By symmetry, it suffices to show that the pure Nash equilibrium $(p, q, s) = (1, 1, 0)$ is the limit of a sequence of logit QREs.

Step 1: For each $\beta > 4$ there is a logit QRE (p, q, s, β) with $p = q \in (1/2, 1)$, and $s < 1/2$.

Proof of Step 1: Based on conditions (4.3) - (4.5) for a logit QRE and the substitution $p = q$, define for all $\beta > 0$ and $p \in [1/2, 1]$:

$$s(p, \beta) := \frac{1}{1 + \exp[-\beta(1 - 2p)]},$$

$$f(p, \beta) := \frac{1}{1 + \exp[-\beta(1 - p - s(p, \beta))]}.$$

Let $\beta > 4$. We show that there is a solution $p^* \in (1/2, 1]$ to the equation $p = f(p, \beta)$. Substitution in (4.3) - (4.5) yields that $(p, q, s, \beta) = (p^*, p^*, s(p^*, \beta), \beta)$ is a logit QRE with the desired properties. Notice that

$$\begin{aligned} \frac{\partial f(p, \beta)}{\partial p} &= \frac{-\beta \exp(-\beta(1 - p - s(p, \beta)))}{(1 + \exp(-\beta(1 - p - s(p, \beta))))^2} \left(1 + \frac{\partial s(p, \beta)}{\partial p}\right), \\ &= \frac{-\beta \exp(-\beta(1 - p - s(p, \beta)))}{(1 + \exp(-\beta(1 - p - s(p, \beta))))^2} \left(1 + \frac{-2\beta \exp(-\beta(1 - 2p))}{(1 + \exp(-\beta(1 - 2p)))^2}\right). \end{aligned}$$

Since $f(1/2, \beta) = 1/2$ and

$$\frac{\partial f(1/2, \beta)}{\partial p} = -\frac{\beta}{4} \left(\frac{2 - \beta}{2}\right) > 1$$

for $\beta > 4$, it follows that $f(p, \beta) > p$ for p slightly larger than $1/2$. Moreover, $f(1, \beta) < 1$. Hence, by the Intermediate Value Theorem applied to $f(\cdot, \beta)$, we find that $f(p^*, \beta) = p^*$ for some $p^* \in (1/2, 1)$.

Step 2: Let $\beta_0 > 4$ and let $p_0 \in (1/2, 1)$ solve $f(p_0, \beta_0) = p_0$. This is possible by Step 1. The function $f(p_0, \cdot)$ is strictly increasing on $[\beta_0, \infty)$.

Proof of Step 2: By definition of f , it suffices to show that the derivative of

$$\beta \mapsto \beta(1 - p_0 - s(p_0, \beta)), \quad \beta \in [\beta_0, \infty)$$

is positive. This derivative equals

$$-\beta \frac{\partial s(p_0, \beta)}{\partial \beta} + 1 - p_0 - s(p_0, \beta). \tag{A.4}$$

Using $p_0 > 1/2$ and the definition of s , it follows that $\partial s(p_0, \beta)/\partial \beta < 0$, i.e., the function $s(p_0, \cdot)$ is strictly decreasing on $[\beta_0, \infty)$. Moreover, as $p_0 = f(p_0, \beta_0) > 1/2$, it follows from

the definition of f that $1 - p_0 - s(p_0, \beta_0) > 0$. As $s(p_0, \cdot)$ is decreasing, this implies that $1 - p_0 - s(p_0, \beta) > 0$ for each $\beta \in [\beta_0, \infty)$. Therefore, the expression in (A.4) is positive.

Step 3: The pure Nash equilibrium $(p, q, s) = (1, 1, 0)$ is the limit of a sequence of QREs.

Proof of Step 3: Let $\beta_0 > 4$ and consider a QRE (p_0, q_0, s_0, β_0) as in Step 1. Set $\beta_1 = \beta_0 + 1$. By Step 2, $p_0 = f(p_0, \beta_0) < f(p_0, \beta_1)$. Moreover, $f(1, \beta_1) < 1$. By the Intermediate Value Theorem applied to the function $f(\cdot, \beta_1)$, there is a $p_1 \in (p_0, 1)$ with $p_1 = f(p_1, \beta_1)$. Hence, there is a QRE (p_1, q_1, s_1, β_1) with

$$\begin{aligned} p_1 &= q_1 &= f(p_1, \beta_1) &> p_0, \\ s_1 &= s(p_1, \beta_1) \\ \beta_1 &= \beta_0 + 1 \end{aligned}$$

Repeating this construction allows us to define a sequence $(p_n, q_n, s_n, \beta_n)_{n \in \mathbb{N}}$ of solutions to (4.3) - (4.5) satisfying the conditions of Step 1 and with $\beta_n \rightarrow \infty$ and $(p_n)_{n \in \mathbb{N}}$ strictly increasing.

As $(p_n, q_n, s_n)_{n \in \mathbb{N}}$ is a sequence in the compact strategy space, we may assume without loss of generality that the sequence converges. Its limit (p, q, s) must be a Nash equilibrium (McKelvey and Palfrey, 1995). As $(p_n)_{n \in \mathbb{N}}$ is a strictly increasing sequence in $(1/2, 1)$ and $(s_n)_{n \in \mathbb{N}}$ is a sequence in $(0, 1/2)$, it must be $p = q > 1/2$ and $s \leq 1/2$. The only Nash equilibrium of the game with these properties is $(p, q, s) = (1, 1, 0)$.

(b): By symmetry, it suffices to show that the Nash equilibrium $(p, q, s) = (1, 0, 1/2)$ is the limit of a sequence of logit QREs. The steps are similar to those in (a). Therefore, the proof is kept short.

Step 1: For each $\beta > 4$ there is a logit QRE (p, q, s, β) with $p \in (1/2, 1)$, $q = 1 - p$, $s = 1/2$.

Proof of Step 1: Let $\beta > 4$. Based on the substitution $q = 1 - p$ and $s = 1/2$ in condition (4.3) for a logit QRE, define

$$g(p, \beta) := \frac{1}{1 + \exp[\beta(1/2 - p)]}.$$

We show that there is a solution $p^* \in (1/2, 1)$ to the equation $p = g(p, \beta)$. Substitution in (4.3) - (4.5) yields that $(p, q, s, \beta) = (p^*, 1 - p^*, 1/2, \beta)$ is a logit QRE with the desired properties. Notice that

$$\frac{\partial g(p, \beta)}{\partial p} = \frac{\beta \exp[\beta(1/2 - p)]}{(1 + \exp \beta(1/2 - p))^2}.$$

Since $g(1/2, \beta) = 1/2$ and $\partial g(1/2, \beta)/\partial p = \beta/4 > 1$, it follows that $g(p, \beta) > p$ for p slightly larger than $1/2$. Moreover, $g(1, \beta) < 1$, so the Intermediate Value Theorem implies that $g(p^*, \beta) = p^*$ for some $p^* \in (1/2, 1)$.

Step 2: For each $p_0 \in (1/2, 1)$, the function $g(p_0, \cdot)$ is strictly increasing on $(0, \infty)$.

Proof of Step 2: Immediate from the definition of g .

Step 3: The Nash equilibrium $(p, q, s) = (1, 0, 1/2)$ is the limit of a sequence of logit QREs.

Proof of Step 3: Reasoning as in the proof of step 3 in part (a) allows us to construct a sequence $(p_n, q_n, s_n, \beta_n)_{n \in \mathbb{N}}$ of solutions to (4.3) - (4.5) satisfying the conditions of Step 1 and with $\beta_n \rightarrow \infty$ and $(p_n)_{n \in \mathbb{N}}$ strictly increasing. As $(p_n, q_n, s_n)_{n \in \mathbb{N}}$ is a sequence in the compact strategy space, we may assume without loss of generality that the sequence converges. Its limit (p, q, s) must be a Nash equilibrium (McKelvey and Palfrey, 1995). As $(p_n)_{n \in \mathbb{N}}$ is a strictly increasing sequence in $(1/2, 1)$, $q_n = 1 - p_n$ and $s_n = 1/2$ for all $n \in \mathbb{N}$, it must be $p > 1/2, q = 1 - p, s = 1/2$. The only Nash equilibrium of the game with these properties is $(p, q, s) = (1, 0, 1/2)$.

(c): It follows by substitution that $(p, q, s, \beta) = (1/2, 1/2, 1/2, \beta)$ is a logit QRE for all $\beta \geq 0$. Consequently, the Nash equilibrium $(p, q, s) = (1/2, 1/2, 1/2)$ is the limit of a sequence of logit QREs with $\beta \rightarrow \infty$.

Appendix B Proof of Proposition 5.1

Let $h_0 = (a^1, a^2) \in H$ and distinguish two cases:

Case 1: a^1 is a pure Nash equilibrium. By [P2], the players will react with positive probability to the belief that everybody plays as in a^1 . Each player's most recent best reply is to continue playing as in a^1 , so the process moves with positive probability to the history $h_1 = (a^1, a^1)$. From here on, the only feasible belief based on the past two periods is that the players play a^1 and the most recent best reply implies that they will continue to play a^1 : the process stays in state h_1 and play has converged to a pure Nash equilibrium.

Case 2: a^1 is not a pure Nash equilibrium. By Proposition 2.1, some alternative, say -1 , was chosen by a set $S \subseteq N$ of players with $|S| > k + 1$. Each player's unique best response to a^1 is therefore to choose $+1$. By [P2], the process moves with positive probability to state $h_1 = ((+1, \dots, +1), a^1)$. Let $a^* \in A^{2k+1}$ be a pure Nash equilibrium where $k + 1$ members of S choose $+1$ and the others choose -1 . Again using [P2], the process moves with positive probability from h_1 to $h_2 = (a^*, (+1, \dots, +1))$:

- For each of the selected $k + 1$ members of S , $+1$ is the unique best reply to the belief drawn from the past two periods that at least $k + 1$ other players from S will choose -1 .

- For each of the remaining k players, -1 is the unique best response to the belief that all other players will continue to play last period's profile $(+1, \dots, +1)$.

Notice that history h_2 belongs to case 1.

Hence, regardless of the initial state h_0 , the Markov process moves with positive probability to an absorbing state where the players continue to play one of the game's pure Nash equilibria. As the Markov process is finite and the initial state was chosen arbitrarily, this will eventually happen with probability one (Kemeny and Snell, 1976): play eventually settles down in a pure Nash equilibrium.

References

- Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review* 84, 406–411.
- Basu, K. and J. Weibull (1991). Strategy subsets closed under rational behavior. *Economics Letters* 36, 141–146.
- Benaïm, M. (1999). Dynamics of stochastic algorithms. In J. Azéma, M. Emery, M. Ledoux, and M. Yor (Eds.), *Séminaire de Probabilités XXXIII*, Volume 1709 of *Lecture Notes in Mathematics*, pp. 1 – 68. Berlin: Springer-Verlag.
- Blonski, M. (1999). Anonymous games with binary actions. *Games and Economic Behavior* 28, 171–180.
- Bottazzi, G. and G. Devetag (2007). Competition and coordination in experimental minority games. *Journal of Evolutionary Economics* 17, 241 – 275.
- Challet, D., M. Marsili, and Y.-C. Zhang (2004). *Minority Games: Interacting Agents in Financial Markets*. Oxford: Oxford University Press.
- Chmura, T. and T. Pitz (2006). Successful strategies in repeated minority games. *Physica A* 363, 477 – 480.
- 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.

- Feddersen, T. and W. Pesendorfer (1997). Voting behavior and information aggregation in elections with private information. *Econometrica* 65, 1029–1058.
- Franke, R. (2003). Reinforcement learning in the El Farol model. *Journal of Economic Behavior and Organization* 51, 367–388.
- Harsanyi, J. C. and R. Selten (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.
- 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.
- Hopkins, E. (2002). Two competing models of how people learn in games. *Econometrica* 70, 2141–2166.
- Hurkens, S. (1995). Learning by forgetful players. *Games and Economic Behavior* 11, 304–329.
- Kalai, E. (2004). Large robust games. *Econometrica* 72, 1631–1666.
- Kemeny, J. G. and J. L. Snell (1976). *Finite Markov Chains*, Volume 40 of *Undergraduate Texts in Mathematics*. Berlin: Springer-Verlag.
- Kets, W. (2007). The minority game: An economics perspective. Working Paper 2007-46, CentER, Tilburg University.
- Kets, W. and M. Voorneveld (2008). Learning to be prepared. *International Journal of Game Theory* 37, 333–352.
- Kojima, F. and S. Takahashi (2004). Anti-coordination games and dynamic stability. Working Paper, Harvard University. To appear in *International Game Theory Review*.
- Mailath, G. J. and A. Postlewaite (1990). Asymmetric information bargaining problems with many agents. *Review of Economic Studies* 57, 351–367.
- McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior* 10, 6–38.

- 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.
- Nisan, N., T. Roughgarden, E. Tardos, and V. V. Vazirani (Eds.) (2007). *Algorithmic Game Theory*. Cambridge: Cambridge University Press.
- Ochs, J. (1990). The coordination problem in decentralized markets: An experiment. *Quarterly Journal of Economics* 105, 545 – 559.
- 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. (2001). Potential games with continuous player sets. *Journal of Economic Theory* 97, 81–108.
- Sandholm, W. H. (2007). *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press. Forthcoming.
- Schmeidler, D. (1973). Equilibrium points of nonatomic games. *Journal of Statistical Physics* 7, 295–300.
- 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 (2005). The cutting power of preparation. *SSE/EFI Working Paper* 583.

- Turocy, T. (2005). A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior* 51, 243 – 263.
- Voorneveld, M. (2004). Preparation. *Games and Economic Behavior* 48, 403 – 414.
- Weibull, J. W. (1995). *Evolutionary Game Theory*. Cambridge, MA: MIT Press.
- Young, H. P. (1998). *Individual Strategy and Social Structure*. Princeton, NJ: Princeton University Press.