Learning in Perturbed Asymmetric Games

Josef Hofbauer*
Department of Mathematics
University College London
London WC1E 6BT
jhofb@math.ucl.ac.uk
http://mailbox.univie.ac.at/Josef.Hofbauer/

Ed Hopkins
Edinburgh School of Economics
University of Edinburgh
Edinburgh EH8 9JY, UK
E.Hopkins@ed.ac.uk
http://homepages.ed.ac.uk/ehk/

May, 2004

Abstract

We investigate the stability of mixed strategy equilibria in 2 person (bimatrix) games under perturbed best response dynamics. A mixed equilibrium is asymptotically stable under all such dynamics if and only if the game is linearly equivalent to a zero sum game. In this case, the mixed equilibrium is also globally asymptotically stable. Global convergence to the set of perturbed equilibria is shown also for (rescaled) partnership games, also known as potential games. Lastly, mixed equilibria of partnership games are shown to be always unstable under all dynamics of this class.

Journal of Economic Literature classification numbers: C72, D83.

Keywords: Games, Learning, Best Response Dynamics, Stochastic Fictitious Play, Mixed Strategy Equilibria, Zero Sum Games, Potential Games.

*We would like to thank David Leslie, Bill Sandholm, two referees and an associate editor for helpful comments. Errors remain ours.

1 Introduction

There has been a long history of interest in the question as to whether it is possible to learn to play Nash equilibria. The first and most enduring learning process, fictitious play, provided initially optimistic results for zero sum and 2×2 games (see Fudenberg and Levine (1998) for a recent survey). However, it has also been known since Shapley's (1964) famous example that there are games with a unique mixed strategy equilibrium which is unstable under fictitious play and other learning processes. The question as to how typical is instability remains open, as there are few general results. This shortage of clear theoretical predictions has become increasingly relevant as researchers have begun to try to fit a stochastic version of fictitious play and reinforcement learning to data from experiments, see for example, Cheung and Friedman (1997), Erev and Roth (1998) and Camerer and Ho (1999).

This paper analyses the properties of the perturbed best response dynamics that underlie models of stochastic or smooth fictitious play. We here produce definitive results on when it is possible to learn mixed strategy equilibria in two person games under such a learning process. Our results are partly positive, in that we provide results on global convergence for two classes of games introduced by Hofbauer and Sigmund (1998), rescaled zero sum games, and rescaled partnership games. Together they include almost all 2×2 games and many of the games which have been subject to theoretical or experimental investigation in the recent literature on learning. However, our results are also negative in that we show that the set of games with mixed equilibria stable for all perturbed best response dynamics is exactly the set of rescaled zero sum games. That is, mixed strategy equilibria in all other games are unstable for at least one form of the dynamic. We go on to discuss what these results imply for stochastic fictitious play and reinforcement learning.

Smooth or stochastic fictitious play is a learning process, first examined by Fudenberg and Kreps (1993), where players' payoffs are perturbed in the spirit of Harsanyi's (1973a) purification argument. Using techniques from stochastic approximation theory, Benaïm and Hirsch (1999) showed that the behaviour of stochastic fictitious play could be predicted by analysis of the associated continuous time perturbed best response dynamics. In a recent paper Ellison and Fudenberg (2000) examine these perturbed best response dynamics in 3×3 bimatrix games. They find that there are many games for which mixed equilibria are locally stable for some member of this class of dynamics, despite the many previous negative results on the stability of mixed equilibria under learning and evolution, particularly in asymmetric games. Indeed, they carry out a Monte Carlo experiment which suggests that this set of games forms a significant proportion of games that possess mixed equilibria.

The current paper differs from that of Ellison and Fudenberg (2000) on a number of levels. First, we extend the analysis of local stability to two-player bimatrix games of arbitrary size. Second, this paper also considers the global analysis of stability.

Thus, for example, we are able to show by the Lyapunov function method that stochastic fictitious play must converge in partnership games, local analysis reveals that mixed equilibria of partnership games are unstable, indicating convergence to a pure equilibrium. Third, whereas Ellison and Fudenberg derive the conditions for an equilibrium to be stable for some form of perturbed best response dynamic, we derive a criterion for stability under all perturbed best response dynamics. Indeed, as we will see, the set of dynamics we consider is also somewhat larger. Thus, our criterion for stability is in two ways more demanding than that of Ellison and Fudenberg.

The global approach is taken further in Hofbauer and Sandholm (2002). Lyapunov functions, similar to those used here, are used to show global convergence of stochastic fictitious play in a number of classes of symmetric evolutionary games including zero sum games, potential games and games with an interior evolutionarily stable strategy. Convergence is also proved for supermodular games. Thus, Hofbauer and Sandholm (2002) extends the global approach to the analysis of stochastic fictitious play, while this paper contains a blend of local and global analysis of deterministic dynamics. It is only this combination of approaches that enables the "if and only if" result that is found here.

This current paper analyses general two person games in strategic form. In 2×2 games, as Ellison and Fudenberg (2000) point out, the behaviour of the perturbed best response dynamics is clear and intuitive. Where the game is purely competitive and there exists only an equilibrium in mixed strategies, this learning process converges to the equilibrium. In other games, where there exist pure strategy equilibria, a mixed strategy equilibrium (if one exists) is unstable, and (generically) play converges to a pure equilibrium. For games with more than two strategies, a much wider range of behaviour is possible. However, rescaled zero sum and rescaled partnership games have sufficient structure that the simple behaviour found in 2×2 games is reproduced in games of arbitrary size. In particular, in rescaled zero sum games which are purely competitive, there is global convergence to a unique (often mixed) equilibrium. And in rescaled partnership games, games involving a strong element of coordination and/or common interest, generically there is convergence to a pure equilibrium.

2 Best Response Learning Dynamics

We consider learning in the context of two-player normal-form games. The games are asymmetric (in the evolutionary sense). That is, the players labelled 1 are drawn from a different "population" from the players labelled 2. For example, in the "Battle of the Sexes" game, players are matched so that a female always plays against a male. The first population choose from n strategies, the second population has m strategies available. Payoffs are determined by two matrices, A, which is $n \times m$, for the first population, and B, which is $m \times n$, for the second population. Let a_{ij} denote a typical element of A, and b_{ji} an element of B.

There are two principal reasons why one might be interested in (perturbed) best response dynamics. The first is that they describe the learning dynamics within a large population of agents who are randomly matched with opponents in continuous time. A model of this type is set out in Ellison and Fudenberg (2000). Second, and perhaps more importantly, we can apply best response dynamics to cases where each "population" consists of only one player. An example of this are the experiments reported in Erev and Roth (1998) where pairs of experimental subjects play the same game repeatedly over a large number of periods. Of course, what Erev and Roth argue is that the observed behaviour is best described by a stochastic learning model. However, results from the theory of stochastic approximation show that the asymptotic behaviour of such stochastic processes is linked to the behaviour of associated continuous time deterministic dynamics. In Section 6 of this paper, we discuss the implications of our results on the perturbed best response dynamics for stochastic fictitious play and reinforcement learning.

There are also two principal ways of constructing the perturbed best response dynamics. The first, due originally to Fudenberg and Kreps (1993), is that the decision maker payoffs' are subject to random shocks. This idea is familiar from the theory of random utility. The second, is that the agent faces a deterministic control cost of implementing a mixed strategy. For example, placing a probability of $1/n$ on each possible strategy may be "easier" than using a mixed strategy that implies that one or more strategies are played with a very low probability. See Mattson and Weibull (2002) for a formal treatment of this idea.

Under fictitious play, agents' beliefs about their opponents' actions are based on the past play of their opponents. Let $x \in S_n$, where S_n is the simplex $\{x = (x_1, ..., x_n) \in I\!R^n : \sum x_i = 1, x_i \geq 0, \text{ for } i = 1, ..., n\}$, be the vector of historic frequencies of the actions of the first player. That is, if up to some time t player 1 chooses the first of two strategies 30% of the time, then at time t, $x = (0.3, 0.7)$. As discussed above, it is also possible to think of this as the average past play of a whole population of player 1's. Let $y \in S_m$ be the vector of the historic frequencies of the choices of the second player. We write the interior of the simplex, that is where all strategies have positive representation, as $\operatorname{int} S_n$ and its complement, the boundary of the simplex as ∂S_n. We also make use of the tangent space of S_n, which we denote $I\!R^n_0 = \{\xi \in I\!R^n : \sum \xi_i = 0\}$

The best response dynamics in the asymmetric case are simply specified as

$$\dot{x} \in BR(y) - x,\ \dot{y} \in BR(x) - y \tag{1}$$

where $BR(y)$ is the set of all best responses of player 1 to y. Of course, $BR(y)$ is therefore typically not a function but a correspondence and so (1) does not represent a standard dynamical system, but a differential inclusion or a set–valued semi–dynamical system on $S_n \times S_m$. It is still possible to characterise its behaviour in several circumstances as Gilboa and Matsui (1991), Matsui (1992) and Hofbauer (1995, 2000) show. However, with small changes to our specification, we can obtain the smooth perturbed best response dynamics. This approach was pioneered by Fudenberg and Kreps (1993), Kaniovski and Young (1995), and Benaïm and Hirsch (1999).

Given fictitious play beliefs, if the first player were to adopt a strategy $p \in S_n$, and the second $q \in S_m$, they would expect payoffs of $p \cdot Ay$ and $q \cdot Bx$ respectively. Following Fudenberg and Levine (1999, p. 118 ff), we suppose payoffs are perturbed such that payoffs are in fact given by

$$\pi_1(p, y) = p \cdot Ay + \varepsilon v_1(p), \quad \pi_2(x, q) = q \cdot Bx + \varepsilon v_2(q), \tag{2}$$

where $\varepsilon > 0$. Here the function $v_1 : \operatorname{int} S_n \to I\!R$ is defined at least for completely mixed strategies $p \in \operatorname{int} S_n$ and has the following properties:

1. v_1 is strictly concave, more precisely its second derivative v_1'' is negative definite, i.e., $\xi \cdot v_1''(p)\xi < 0$ for all $p \in \operatorname{int} S_n$ and all nonzero vectors $\xi \in I\!R_0^n$.
2. The gradient of v_1 becomes arbitrarily large near the boundary of the simplex, i.e., $\lim_{p \to \partial S_n} |v_1'(p)| = \infty$.

One possible interpretation of the above conditions is that the player has a control cost to implementing a mixed strategy with the cost becoming larger nearer the boundary.[1] For example, in the game of matching pennies, the strategy $(\frac{1}{2}, \frac{1}{2})$ is probably easier to adopt than the strategy $(\frac{1}{999}, \frac{998}{999})$. In any case, these conditions imply that for each fixed $y \in S_m$ there is a unique $p = p(y) \in \operatorname{int} S_n$ which maximizes the perturbed payoff $\pi_1(p, y)$ for player 1. We assume that $v_2 : \operatorname{int} S_m \to I\!R$ possesses the same properties so that for each $x \in S_n$ there is a unique $q = q(x) \in \operatorname{int} S_m$ which maximizes $\pi_2(x, q)$. We can therefore replace the best reply correspondence $BR(x, y)$ with a 'perturbed best reply function' $(p(y), q(x))$. Typical examples of perturbation functions that satisfy these conditions are $v_1(p) = \sum_i \log p_i$ (as first used in the logarithmic games of Harsanyi (1973b)) and $v_1(p) = -\sum_i p_i \log p_i$.

Differentiating the perturbed payoff functions (2) with respect to first and second argument (which we will denote by ∂_1 and ∂_2) respectively, the first order conditions for a maximum will be $\partial_1 \pi_1(p(y), y) = \partial_2 \pi_2(x, q(x)) = 0$ or

$$\xi \cdot Ay + \varepsilon v_1'(p(y))\xi = 0 \quad \forall \xi \in I\!R_0^n \quad \text{and} \quad \eta \cdot Bx + \varepsilon v_2'(q(x))\eta = 0 \quad \forall \eta \in I\!R_0^m. \tag{3}$$

This could be written formally as

$$p(y) = (v_1')^{-1}(-\frac{Ay}{\varepsilon}), \quad q(x) = (v_2')^{-1}(-\frac{Bx}{\varepsilon}). \tag{4}$$

[1] See Mattson and Weibull (2002) for an alternative axiomatic approach to the control cost problem.

This shows that the perturbed best reply functions p and q are smooth. However, an explicit evaluation of p seems to be possible only in special cases, see (7) below.

The original formulation of stochastic fictitious play due to Fudenberg and Kreps (1993), see also Fudenberg and Levine (1998, p. 105 ff), involved a truly stochastic perturbation of payoffs. For example, one can replace (2) with

$$\pi_1(p, y) = p \cdot Ay + \varepsilon p \cdot \rho_1, \quad \pi_2(x, q) = q \cdot Bx + \varepsilon q \cdot \rho_2, \tag{5}$$

where each ρ_i is a vector of i.i.d. random variables with a fixed distribution function and a strictly positive and bounded density. Assume each player sees the realisation of her own perturbation, then chooses an action to maximise the perturbed payoff. Then, the probability that she will choose action i will be

$$p_i(y) = \Pr[\arg\max_j (Ay)_j + \varepsilon\rho_{1j} = i]. \tag{6}$$

This defines a smooth function $p(y)$ which approximates the best reply correspondence.

The relation between these two ways to perturb payoffs and smoothen best replies was clarified by Hofbauer and Sandholm (2002). They show that *any* perturbed best response function derived from a stochastic model (5) can also be derived from a deterministic optimisation problem of the form (2). For a well–known example, consider the exponential or logit choice rule, given in current notation with $\beta = \varepsilon^{-1}$ by

$$p_i(y) = \frac{\exp \beta (Ay)_i}{\sum_{j=1}^{n} \exp \beta (Ay)_j}. \tag{7}$$

This rule can be derived from the deterministic optimisation procedure (2) by setting $v_1(p) = -\sum p_i \log p_i$ and from the stochastic model (5) if the disturbance is given by the double exponential extreme value distribution. However, what Hofbauer and Sandholm (2002) also establish is that the converse does not hold. That is, there are perturbation functions (e.g. $v_1(p) = \sum_i \log p_i$ for $n \geq 4$) that satisfy the two conditions above but for which there is no corresponding stochastic perturbation even if the i.i.d. assumption is weakened. This implies that the best response functions derived from the deterministic procedure form a wider class than those arriving from the stochastic optimisation problem.

If we assume that within two large populations, there is a smooth adjustment toward the (perturbed) best response, we can write down the two population dynamics as

$$\dot{x} = p(y) - x, \quad \dot{y} = q(x) - y. \tag{8}$$

Equally, as discussed above, this system of differential equations can be used to predict the stochastic learning of individual agents and the long run behaviour of evolutionary models for games with randomly perturbed payoffs, see Hofbauer and Sandholm (2002).

We emphasize that we will be able to analyze the dynamics (8) despite the fact that the perturbed best reply functions $p(y), q(x)$ are not explicitly known — except in the prime example (7) — and study the stability properties of its equilibria for which there is no explicit formula. The equilibria of (8) occur at each intersection of the perturbed best response functions. In that sense, they are Nash equilibria of a game with the perturbed payoffs (2). They are also "quantal response equilibria" in the terminology of McKelvey and Palfrey (1995). In analogy to Harsanyi's (1973a) purification theorem, for each regular Nash equilibrium (x^*, y^*) and for each small $\varepsilon > 0$, there will be an associated perturbed equilibrium $(\hat{x}_\varepsilon, \hat{y}_\varepsilon)$ and $\lim_{\varepsilon \to 0}(\hat{x}_\varepsilon, \hat{y}_\varepsilon) = (x^*, y^*)$.[2]

In the case of zero sum games, the perturbed equilibrium will be shown to be unique (Theorem 3.2 below). In a two player zero sum game, either a player has a unique equilibrium strategy or she is indifferent between different strategies each of which guarantee her a payoff which is equal to the value of the game. The addition of noise will break this indifference. For example, if all strategies had the same value, then the unique equilibrium point is simply the point that maximises the perturbation function v. In the case of partnership games, multiple (perturbed) equilibria are possible. The stable equilibria are the maxima of a suitably perturbed potential function, see Theorem 3.3 below.

3 Results on Global Convergence

Note that the exact form of the functions $p(y), q(x)$ and hence the dynamic depends on the nature of the perturbation functions v_1, v_2. The fundamental question is therefore what results can be obtained which are independent of the exact form of v. It is known (see Benaïm and Hirsch (1999)) that for generic 2×2 games the global qualitative behavior of (8) does not depend on the perturbation function v and is the same as that of the best response dynamics (1). Using ideas of Hofbauer (2000) in a symmetric setting, we extend this result to higher dimensions for two important classes of games, games of conflict and games of coordination.[3]

Hofbauer and Sigmund (1998, p127-8) consider the following equivalence relation: the bimatrix game (A', B') is linearly equivalent to (or a rescaling of) the bimatrix game (A, B) if there exist constants c_j, d_i and $\alpha > 0, \beta > 0$ such that

$$a'_{ij} = \alpha a_{ij} + c_j, b'_{ji} = \beta b_{ji} + d_i. \tag{9}$$

[2]Both McKelvey and Palfrey (1995) and Binmore and Samuelson (1999) provide a more detailed analysis of the relationship between perturbed equilibria and Nash equilibria of the original unperturbed game.

[3]More recently, convergence results for stochastic fictitious play were obtained also for a third class of game, supermodular games, by Hofbauer and Sandholm (2002).

Then (A, B) is a rescaled zero sum game if there exists a rescaling such that $B' = -(A')^T$ and a rescaled partnership game if $B' = (A')^T$. Equilibrium points of games are unchanged under rescaling. That is, if (x^*, y^*) is a Nash equilibrium of the game (A, B), it is also of (A', B'). The corresponding perturbed equilibrium may change, however. To see this note that the first order condition (3), for example for the first population, becomes $\alpha\xi \cdot Ay + \varepsilon v_1'(x)\xi = 0$. In this model of perturbed payoffs, multiplying the payoff matrix by a positive factor is equivalent to reducing the noise parameter ε by an equivalent amount.

We start with the following simple characterization.

Lemma 3.1 *(A, B) is a rescaled partnership or zero sum game if and only if*

$$c\xi \cdot A\eta = \eta \cdot B\xi \quad \textit{for all } \xi \in I\!R_0^n, \eta \in I\!R_0^m \tag{10}$$

for some $c \neq 0$, where $c > 0$ for a rescaled partnership game and $c < 0$ for a rescaled zero sum game.

Proof: Hofbauer and Sigmund (1998, p128-9). QED

Note that in 2×2 games the condition (10) reduces to $ca_0 = b_0$ where $a_0 = a_{11} + a_{22} - a_{12} - a_{21}$ and similarly $b_0 = b_{11} + b_{22} - b_{12} - b_{21}$. This will be satisfied for some nonzero c provided neither a_0 or b_0 are zero. Or, in other words, almost all 2×2 games are either rescaled zero sum games or rescaled partnership games.[4] Thus the results we establish here include most existing results on 2×2 games as special cases.

For example, the class of rescaled zero sum games includes all 2×2 games with unique mixed strategy equilibria, which have been of particular interest both to theorists, e.g. Fudenberg and Kreps (1993), and to experimentalists, e.g. Erev and Roth (1998). Benaïm and Hirsch (1999) show that the equilibria of such games are globally asymptotically stable under the perturbed best response dynamics. This is now generalised to all rescaled zero sum games.

Theorem 3.2 *For any two person rescaled zero-sum game, the perturbed best response dynamic (8) has a unique rest point which is globally asymptotically stable.*

Proof: Consider the functions

$$V_1(x, y) = \pi_1(p(y), y) - \pi_1(x, y), \quad V_2(x, y) = h[\pi_2(x, q(x)) - \pi_2(x, y)]$$

[4]Ellison and Fudenberg (2000) make a similar observation, but using the terms "games of conflict" and "games of coordination".

where $h > 0$ is defined as $h = -1/c$ where c is the constant from (9) implied by the fact that (A, B) is a rescaled zero sum game. These functions are nonnegative and vanish together precisely at perturbed equilibria. Then define,

$$\begin{aligned} V(x,y) &:= V_1(x,y) + V_2(x,y) \\ &= \pi_1(p(y),y) - \varepsilon v_1(x) + h\pi_2(x,q(x)) - h\varepsilon v_2(y) + \varepsilon(-x \cdot Ay + hy \cdot Bx). \end{aligned}$$

V is a strictly convex function of (x, y) since $\pi_1(p(y), y) = \max_z \pi_1(z, y)$ is convex, being the maximum of linear functions in y, $-v_1$ is strictly convex in x and $-v_2$ is strictly convex in y, and the final term reduces to a linear function in x and y as the game is rescaled zero-sum. Hence the function V will attain its minimum value 0 at a *unique* point, which is the unique perturbed equilibrium of the rescaled zero sum game. The definition of $p(y)$ and $q(x)$ implies

$$\partial_1\pi_1(p(y),y) = 0, \quad \partial_2\pi_2(x,q(x)) = 0, \tag{11}$$

where ∂_1 and ∂_2 denote again the partial derivatives with respect to the first and second variable (within $\mathbb{R}^n_0$ and $\mathbb{R}^m_0$). Hence

$$\dot{V}_1 = \partial_1\pi_1(p,y)\dot{p} - \partial_1\pi_1(x,y)\dot{x} + \partial_2\pi_1(p,y)\dot{y} - \partial_2\pi_1(x,y)\dot{y}$$

Because of (11) we can rewrite this as

$$\begin{aligned} \dot{V}_1 &= (\partial_1\pi_1(p,y) - \partial_1\pi_1(x,y))\dot{x} + (p-x) \cdot A\dot{y} \\ &= \varepsilon(v_1'(p) - v_1'(x))(p-x) + (p-x) \cdot A(q-y). \end{aligned} \tag{12}$$

In a similar way, together with the application of Lemma 3.1, one obtains

$$\begin{aligned} \dot{V}_2 &= h\varepsilon(v_2'(q) - v_2'(y))(q-y) + h(q-y) \cdot B(p-x) \\ &= h\varepsilon(v_2'(q) - v_2'(y))(q-y) - (p-x) \cdot A(q-y) \end{aligned} \tag{13}$$

By the strict concavity of v_1 and v_2, $\dot{V} = \dot{V}_1 + \dot{V}_2 \leq 0$ follows, with equality only if $p = x$ and $q = y$. Thus V is a strict Lyapunov function, and Lyapunov stability and global attractivity of the unique perturbed equilibrium follows. QED

Partnership games are games of coordination and common interest. One obvious group of games within this class are the games sometimes called games of pure coordination which have positive entries on the diagonal and zero elsewhere. Another prominent example is the so-called stag hunt game, with two different pure equilibria, one pareto dominant, the other risk dominant. Partnership games have the property that the payoff for player 1 is equal to the payoff of player 2, or $x \cdot Ay = y \cdot Bx$.

Every local maximum (x^*, y^*) of the players' payoffs $x \cdot Ay$ is a Nash equilibrium, but not conversely, and the strict Nash equilibria correspond to the strict local maxima. Thus if, under some learning rule, payoffs are always rising in the unperturbed game, there will be convergence to the set of Nash equilibria. Analogously, we show in the next result that under the perturbed best response dynamics a suitable perturbation of the players' payoff is always rising out of equilibrium and hence orbits move toward the set of perturbed equilibria.

Theorem 3.3 *For any rescaled partnership game, each orbit of the perturbed best response dynamics (8) converges to the set of perturbed equilibria.*

Proof: Consider the function

$$U(x, y) = x \cdot A'y + \varepsilon v_1(x) + \beta\varepsilon v_2(y), \tag{14}$$

where $(A', (A')^T)$ is a rescaling of (A, B), and where, without loss of generality, the scaling factor α in (9) is set to one. $\beta > 0$ is the scaling factor for the second population. The first order conditions for the critical points of U in $S_n \times S_m$ will be $\partial_1 U(x, y) = \partial_2 U(x, y) = 0$ or

$$\xi \cdot A'y + \varepsilon v_1'(x)\xi = 0 \quad \forall \xi \in I\!R_0^n \quad \text{and} \quad \eta \cdot B'x + \beta\varepsilon v_2'(y)\eta = 0 \quad \forall \eta \in I\!R_0^m. \tag{15}$$

Comparison with the first order conditions (3) reveals that perturbed equilibria, $x = p(y)$, $y = q(x)$, form the critical points of the function U.

We have

$$\dot{U} = \dot{x} \cdot A'y + x \cdot A'\dot{y} + \varepsilon v_1'(x)\dot{x} + \beta\varepsilon v_2'(y)\dot{y}. \tag{16}$$

Note that from (3), $\xi \cdot A'y = \xi \cdot Ay = -\varepsilon v_1'(p)\xi$ and $x \cdot A'\eta = \beta\eta \cdot Bx = -\beta\varepsilon v_2'(q)\eta$ holds for all $\xi \in I\!R_0^n$ and $\eta \in I\!R_0^m$, so that for $\xi = \dot{x}$ and $\eta = \dot{y}$

$$\dot{U} = \varepsilon\left(v_1'(x) - v_1'(p)\right)(p - x) + \beta\varepsilon\left(v_2'(y) - v_2'(q)\right)(q - y). \tag{17}$$

Again by the strict concavity of v_1 and v_2, $\dot{U} \geq 0$ with equality only at $x = p(y)$ and $y = q(x)$. Hence every ω–limit set consists of perturbed equilibria. QED

The result is that all orbits of the dynamical system converge to the set of perturbed equilibria for any value of ε. Additionally, for small enough $\varepsilon > 0$, convergence will be to a perturbed strict equilibrium and never a perturbed mixed equilibrium, provided we restrict our attention to generic games and generic initial conditions. This is because mixed equilibria of rescaled partnership games will be saddle points under the perturbed best response dynamics (see Theorem 4.5 below). Hence, while they attract some part of the state space (their stable manifold), for generic initial conditions the dynamics will diverge from any mixed equilibrium.

The restriction to generic games is to exclude games with continua of Nash equilibria, which are indeed non-generic in the strategic form. Of course, extensive form games often give rise to strategic forms of the type we exclude. An example of this is the "Ultimatum Minigame" analysed by Gale, Binmore and Samuelson (1995). This is a rescaled partnership game with a continuum of Nash equilibria, which in the unperturbed game, i.e. with $\varepsilon = 0$, are local maxima for the potential function U. In games such as these a noisy learning model may have an attracting perturbed equilibrium which is not close to a pure equilibrium.

It is a reasonable question to ask whether these results on global convergence of the perturbed best response dynamics can be extended to games with more than two players. It has long been recognised that zero sum games with three or more players have a completely different competitive structure than those with two players, for example, the minimax theorem does not hold. A similar distinction seems to hold for dynamics and Benaïm and Hirsch (1999) have already shown that the perturbed best response dynamics cycle rather than converge in a 3-person matching pennies game. In contrast, partnership games naturally generalise to n players. A convergence result for multiplayer partnership/potential games is given in Hofbauer and Sandholm (2002).

4 A Converse Result

We now ask the question whether there are mixed equilibria of any other games (besides rescaled zero sum games) which are stable under all perturbed best response dynamics that can be constructed using the deterministic perturbation method. We find that there are not. The first step is to construct the linearisation of the perturbed best response dynamics.

Lemma 4.1 *We can write $dp(y)/dy$ as $\frac{1}{\varepsilon}Q_1A$ where Q_1 is a symmetric matrix positive definite with respect to $\mathbb{R}^n_0$. That is, $\xi \cdot Q_1\xi > 0$ for any nonzero $\xi \in \mathbb{R}^n_0$. Furthermore, $Q_1\mathbf{1} = 0$ where $\mathbf{1} = (1, 1, ...1)$.*

Proof: This result is essentially obtained from differentiating (3), with

$$Q_1 = -(v_1''(p(y)))^{-1}, \tag{18}$$

which is positive definite since $v_1''(p)$ is negative definite by assumption. See Hopkins (1999) for more details.[5] QED

Now, given Lemma 4.1 and the dynamics (8), the Jacobian taken at a perturbed equilibrium $(\hat{x}, \hat{y})$ will be

$$J = \begin{pmatrix} 0 & \frac{dp(y)}{dy} \\ \frac{dq(x)}{dx} & 0 \end{pmatrix} - I = \frac{1}{\varepsilon}\begin{pmatrix} Q_1(\hat{x}) & 0 \\ 0 & Q_2(\hat{y}) \end{pmatrix}\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} - I. \tag{19}$$

where Q_1 and Q_2, from Lemma 4.1, are both symmetric and positive definite with respect to $\mathbb{R}^n_0$ and $\mathbb{R}^m_0$ respectively. Note that since the vector field $(p(y)-x, q(x)-y)$ is in $\mathbb{R}^n_0 \times \mathbb{R}^m_0$ for $x \in S_n$ and $y \in S_m$, it is the properties of the linearisation on

[5] The analysis in Hofbauer and Sandholm (2002) shows that this result also follows from taking the Legendre transform of the convex function $-v_1$.

this subspace that matter. $I\!R^n$ can be decomposed into two orthogonal subspaces $I\!R_0^n$ and $I\!R_1^n = \{x \in I\!R^n : x = c\mathbf{1}, c \in I\!R\}$. When we look at a linearisation around an equilibrium of any dynamics on $S_n \times S_m$, the stability of that equilibrium will be determined by $n-1$ eigenvalues which refer to $I\!R_0^n$ and the $m-1$ eigenvalues which refer to $I\!R_0^m$, and not the two remaining eigenvalues which refer to $I\!R_1^n$ and to $I\!R_1^m$.

We can write the linearisation (19) somewhat more compactly as $J = \frac{1}{\varepsilon}Q(\hat{x}, \hat{y})H - I$, where Q is the matrix with Q_1 and Q_2 forming blocks on the diagonal and H is the matrix with A and B on the off-diagonal. It implies that if μ is an eigenvalue for QH, then (19) has an eigenvalue $\mu/\varepsilon - 1$. The condition for stability as ε becomes small is therefore that QH must not have an eigenvalue with positive real part. However, as the matrix QH has a zero trace, the only way to satisfy the condition for stability is to have all eigenvalues with real part zero. If this is the case, the eigenvalues of the Jacobian $QH/\varepsilon - I$ will all have real part negative. It is this stable case that must be identified.

We do this by determining a condition for instability and show that only rescaled zero sum games fail to satisfy it. In essence, because we consider the whole class of perturbed dynamics, we have a free choice of the matrices Q_1, Q_2, the only restrictions being that they are positive definite and symmetric. Therefore, for almost any mixed equilibrium, we can choose Q_1, Q_2 such that it will be unstable.

Lemma 4.2 *If there exists $\xi \in I\!R_0^n$ and $\eta \in I\!R_0^m$ such that*

$$\xi \cdot A\eta > 0 \text{ and } \eta \cdot B\xi > 0, \tag{20}$$

then there are symmetric matrices Q_1, Q_2, positive definite on R_0^n and R_0^m respectively, and $Q_i\mathbf{1} = 0$ such that the product

$$QH = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix} \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$$

has at least one positive eigenvalue.

Proof: Let (ξ, η) be such that (20) holds. Write $H(\xi, \eta) = (A\eta, B\xi) = (x, y)$. Then, $\xi \cdot x > 0$ and $\eta \cdot y > 0$. Because $\xi \cdot x > 0$, we can find an orthonormal basis of vectors $\{z_0, ..., z_{n-1}\}$ with $z_0 = \mathbf{1} = (1, 1, ..., 1)$ such that $\xi = \sum \alpha_i z_i$, $x = \sum \beta_i z_i$ with $\alpha_i \beta_i > 0$ for $i = 1, ..., n-1$ (that is, as $\xi \cdot x > 0$, the angle between x and ξ is less than $90°$, and so it is possible to find a basis such that the two vectors are in the same orthant). Then there exists a unique symmetric matrix Q_1, positive definite with respect to $I\!R_0^n$ with $\{z_0, ..., z_{n-1}\}$ as its eigenvectors such that $Q_1 z_i = \alpha_i/\beta_i z_i$ for $i = 1, ..., n-1$ and $Q_1 z_0 = 0$. Consequently, $Q_1 x = \xi$. Similarly, let $Q_2 y = \eta$. Then $QH(\xi, \eta) = \lambda(\xi, \eta)$, where $\lambda = 1 > 0$. QED

Clearly there exists a counterpart to Lemma 4.2, that if $\eta \cdot B\xi < 0$ and $\xi \cdot A\eta < 0$, for some ξ, η, then we can find a Q such that QH has an eigenvalue with real part negative. Therefore, for QH to have all eigenvalues with real part zero, it must be true that

$$(\xi \cdot A\eta)(\eta \cdot B\xi) \leq 0, \text{ for all } \xi \in I\!R_0^n,\ \eta \in I\!R_0^m. \tag{21}$$

Of course, this condition is satisfied if either A or B are zero matrices. But if we assume that the bimatrix game (A, B) has an isolated mixed equilibrium, then this trivial case is excluded, and it follows that the two bilinear forms are proportional, and the game is rescaled zero sum. This is the essence of the following.

Theorem 4.3 *Let $(x^*, y^*) \in \operatorname{int}(S_n \times S_m)$ be an isolated interior equilibrium of the bimatrix game (A, B). If for all strictly concave disturbance functions v_1, v_2 satisfying the assumptions in Section 2 the perturbed equilibrium $(\hat{x}_\varepsilon, \hat{y}_\varepsilon)$ is locally stable for the perturbed best response dynamics (8) for arbitrarily small $\varepsilon > 0$ then the game (A, B) is a rescaled zero sum game.*

Proof: If $(x^*, y^*) \in \operatorname{int}(S_n \times S_m)$ be an isolated interior equilibrium of the bimatrix game (A, B), then it is a regular equilibrium and $n = m$. It was shown in (Hofbauer and Sigmund, 1998, Theorem 11.4.2, p135-6) that the condition (21) holds if and only if the equilibrium $(\hat{x}, \hat{y})$ is a Nash–Pareto pair if and only if (A, B) is a rescaled zero sum game. If it does not hold, by Lemma 4.2 there exist symmetric positive definite matrices Q_1, Q_2 such that the matrix QH has a positive eigenvalue. Next, find strictly concave functions v_1, v_2 such that $Q_1 = -(v_1''(x^*))^{-1}$ and $Q_2 = -(v_2''(y^*))^{-1}$ at the equilibrium (x^*, y^*). Then, for small $\varepsilon > 0$, the linearisation of the perturbed best response dynamics, equal to $Q(\hat{x}_\varepsilon, \hat{y}_\varepsilon)H/\varepsilon - I$ by (19), will have a positive eigenvalue too. QED

Now, the above theorem says that only equilibria of rescaled zero sum games can be stable for all dynamics of the form (8). However, this leaves open the possibility that for other games an isolated fully mixed equilibrium may be stable for some specifications of the noise function v and not for others. While this is not possible for 2×2 games, this happens indeed for an open set of $n \times n$ games if $n \geq 3$, see Ellison and Fudenberg (2000) for some explicit examples with $n = 3$.

We proceed with our examination of the linearisation of the perturbed best response dynamics, with a result on the calculation of the eigenvalues of a matrix of the form (19), see e.g. Hofbauer and Sigmund (1998, p.118).

Lemma 4.4 *The eigenvalues of any matrix of the form*

$$H = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix},$$

with blocks of zeros on the diagonal and A is $n \times m$ and B is $m \times n$, are the square roots of the eigenvalues of AB.

Proof: The eigenvalue equation is $H(x,y) = \lambda(x,y)$. This equation can be decomposed into $Ay = \lambda x$ and $Bx = \lambda y$. Premultiplying the second equation by A, the result is $ABx = \lambda Ay$ or, substituting from the first equation, $ABx = \lambda^2 x$. That is, if λ is an eigenvalue of H then λ^2 is an eigenvalue of AB. If $m \neq n$, then the rank of H is at most $2\min(n,m)$, and it is the nonzero eigenvalues of H which are the square roots of the nonzero eigenvalues of AB. QED

We have already established the global stability of the unique perturbed equilibrium in rescaled zero sum games. However, in some applications, it is not sufficient that an equilibrium is stable, it is also required to be hyperbolic.[6] Hence we note the following, which extends a generic local result of Ellison and Fudenberg (2000, Proposition 8) for 3×3 zero sum games.

Theorem 4.5 *Any Jacobian matrix of form (19) evaluated at $(\hat{x},\hat{y})$, the unique perturbed equilibrium of a rescaled zero sum game under the dynamics (8), will have all eigenvalues with real part negative.*

Proof: By Lemma 4.4, the nonzero eigenvalues of $Q(\hat{x},\hat{y})H$ will be square roots of the eigenvalues of Q_1AQ_2B, assuming without loss of generality that $n \leq m$. If the game (A,B) is a rescaled zero sum game, then Q_1AQ_2B will have the same eigenvalues with respect to eigenvectors in $I\!R^n_0$ as $cQ_1AQ_2A^T$ with $c < 0$. The matrix $P = AQ_2A^T$ is symmetric and positive semidefinite with respect to $I\!R^n_0$. The product of a symmetric positive definite matrix with a symmetric positive semidefinite matrix has all eigenvalues real and nonnegative (see for example, Hines, 1980). Hence, the eigenvalues of cQ_1P are less than or equal to zero. The nonzero eigenvalues of the matrix QH are therefore the purely imaginary square roots of the negative eigenvalues of Q_1AQ_2B. The eigenvalues of the linearisation (19) as a whole, will be equal to the eigenvalues of $Q(\hat{x}_\varepsilon,\hat{y}_\varepsilon)H/\varepsilon$ minus one and hence all have real part negative. QED

The next result complements Theorem 3.2 and shows that the mixed equilibria of rescaled partnership games are always saddle points under the perturbed best response dynamics.

Theorem 4.6 *Let E be a isolated interior equilibrium of a rescaled partnership game. Then the corresponding perturbed equilibrium is unstable for (8) for all admissible disturbance functions v_1, v_2 and all small $\varepsilon > 0$.*

Proof: As we are considering a rescaled partnership game, the symmetric matrix $P = AQ_2A^T$ used in the proof to the previous theorem is now positive definite with

[6] An example where this condition is required is Theorem 2 in Benveniste et al. (1990, pp107-8) which deals with convergence of stochastic processes with a constant step size. This could be used to analyse stochastic fictitious play when agents place greater weight on more recent experience.

respect to $\mathbb{R}^n_0$. Hence, cQ_1P has $n-1$ positive eigenvalues, as now $c > 0$. The matrix QH has therefore $n-1$ positive and $n-1$ negative eigenvalues, these being the square roots of the positive eigenvalues of Q_1AQ_2B. The eigenvalues of the linearisation (19) as a whole, will be equal to the eigenvalues of $Q(\hat{x}_\varepsilon, \hat{y}_\varepsilon)H/\varepsilon$ minus one. Now, as the mixed equilibrium (x^*, y^*) is in the interior of $S_n \times S_n$, by Lemma 4.1 and property 1 of the perturbation functions given in Section 2, $\lim_{\varepsilon\to 0} Q_1(\hat{x}_\varepsilon) = Q_1(x^*)$ will be positive definite, and similarly for $\lim_{\varepsilon\to 0} Q_2(\hat{y}_\varepsilon)$.[7] Hence, for a small $\varepsilon > 0$, the absolute value of the eigenvalues of QH/ε will be sufficiently large such that $QH/\varepsilon - I$ will have at least one positive eigenvalue. QED

There is another criterion for stability established by Ellison and Fudenberg (2000) for the 3×3 case. A game is symmetric if, in current notation, $A = B$. In generic symmetric games, where both players have the same perturbation function, any mixed equilibrium is unstable. The genericity assumption rules out one notable exception, zero sum games, which may be symmetric, but as we have seen, do not have unstable mixed equilibria.

Theorem 4.7 *Let E be a isolated interior equilibrium of a generic symmetric game with $A = B$. Then if the two disturbance functions v_1, v_2 are the same the corresponding perturbed equilibrium is unstable for all perturbed dynamics (8) for sufficiently small $\varepsilon > 0$.*

Proof: Since $A = B$ and the two perturbation functions are the same, then $\hat{x} = \hat{y}$ and consequently, when evaluated at $(\hat{x}, \hat{y})$, $Q_1 = Q_2$. So, $Q_1AQ_2B = (Q_1A)^2$. This implies that the eigenvalues of the matrix Q_1AQ_2B will be the square of the eigenvalues of Q_1A. Given that in turn the eigenvalues of QH are the square roots of the eigenvalues of Q_1AQ_2B, then if λ is an eigenvalue of Q_1A, then both λ and $-\lambda$ are eigenvalues of QH. Hence, QH generically will have at least one eigenvalue with positive real part.[8] Hence, for a small $\varepsilon > 0$, the absolute value of the eigenvalues of QH/ε will be sufficiently large such that $QH/\varepsilon - I$ will have at least one positive eigenvalue. QED

What can we say about mixed Nash equilibria with less than full support, that is, where players only place a positive probability on some of their strategies? The analysis of this case turns out to be surprisingly difficult. Suppose that there is a symmetric 3×3 game with a mixed equilibrium (x^*, y^*) that only involves 2 strategies. In the neighbourhood of the Nash equilibrium, as ε becomes small, we would expect $p(y)$ to approach x^* which places positive weight on only two strategies. So, the

[7] Note that in contrast each $\lim_{\varepsilon\to 0} Q_i(\hat{x}_\varepsilon)$ may be a zero matrix if $(\hat{x}_\varepsilon, \hat{y}_\varepsilon)$ corresponds to a pure Nash equilibrium. This means that the $-I$ term in J would dominate and a perturbed strict equilibrium, for example, would be stable.

[8] The principal exception is when $A = B = -A^T$, that is the game is zero sum. In that case, the eigenvalues of Q_1A will be entirely imaginary. This is not generic, as for some small perturbation of A, Q_1A will have some eigenvalue with non-zero real part.

perturbed best response dynamics for the 3×3 game seems to become "close" to the dynamics for the 2×2 game for which (x^*, y^*) is a fully mixed equilibrium and we would expect the equilibrium to have the same stability properties under both. The problem is we have to consider the behaviour of $Q_1(\hat{x}_\varepsilon)/\varepsilon$ as ε approaches zero. This may explode rather than converge.

Thus, in order to be able to analyse the local stability of such equilibria we would need the following properties of the behaviour of the matrix functions Q_1, Q_2 on the boundary of the simplex to hold. Let $(A, B)|_{K_1 \times K_2}$ be the smaller game where only strategies from the sets K_1 and K_2 are available to player 1 and 2 respectively. A Nash equilibrium $E = (x^*, y^*)$ is *regular* if it is isolated and $x^* \cdot Ay^* > (Ay^*)_i$ for $i \notin K_1$ and $y^* \cdot Bx^* > (Bx^*)_j$ for $j \notin K_2$.

Assumption A Let $E = (x^*, y^*)$ be a regular Nash equilibrium of (A, B) with support (K_1, K_2) and $(\hat{x}_\varepsilon, \hat{y}_\varepsilon)$ be corresponding perturbed equilibria for small $\varepsilon > 0$. Then, first, we assume that the limits $\lim_{\varepsilon \to 0} Q_1(\hat{x}_\varepsilon)$ and $\lim_{\varepsilon \to 0} Q_2(\hat{y}_\varepsilon)$ exist and are finite and second, if $i \notin K_1$, that is $\lim_{\varepsilon \to 0} \hat{x}_{\varepsilon,i} = 0$, then

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} Q_{1,ij}(\hat{x}_\varepsilon) = 0 \text{ for all } j. \tag{22}$$

and a similar property holds for Q_2.

In effect what is required is that as ε approaches zero, the linearisation of the perturbed best response dynamics approach that we would obtain from examining the same dynamics on the smaller game $(A, B)|_{K_1 \times K_2}$. This seems reasonable, but (22) will only hold if each $Q_{1,ij}$ approaches zero at a sufficiently fast rate, a technical property which is very difficult to establish for the general case. Thus, while we imagine that this assumption is met by most perturbation functions satisfying the assumptions of Section 2, we only demonstrate that this is the case for the exponential choice rule (7).

Lemma 4.8 *If $v_1(p) = -\sum p_i \log p_i$ and $v_2(q) = -\sum q_i \log q_i$ then Assumption A is satisfied.*

Proof: First, if $v_1(p) = -\sum p_i \log p_i$ then

$$Q_{1,ii} = p_i(1 - p_i), \quad Q_{1,ij} = -p_i p_j \tag{23}$$

so that clearly the first part of Assumption A is satisfied. Second, let $E = (x^*, y^*)$ be a regular equilibrium. Let $P(p) = p_i \prod_{j=1}^n p_j^{-x_j^*} \geq p_i$ for some $i \notin K_1$. Then, given the particular functional form (7), we have $p_i/p_j = \exp((Ay)_i - (Ay)_j)/\varepsilon)$, for all j. And we can write

$$P(p) = \prod_{j=1}^n \left(\frac{p_i}{p_j} \right)^{x_j^*} = \exp\left(-\frac{1}{\varepsilon}(x^* \cdot Ay - (Ay)_i) \right).$$

Note that from the regularity of the equilibrium E, there is some neighbourhood Y of y^*, where $x^* \cdot Ay - (Ay)_i > c > 0$ for all $y \in Y$. So, $p_i(\hat{y}_\varepsilon)/\varepsilon \leq P(p(\hat{y}_\varepsilon))/\varepsilon \leq (1/\varepsilon)\exp(-c/\varepsilon)$, which approaches zero as ε goes to zero. Finally, together with (23) this shows that the second part of Assumption A holds. QED

Of course, if (A, B) is a rescaled partnership game, then $(A, B)|_{K_1 \times K_2}$ is also. It is relatively easy to establish that partially mixed equilibria of symmetric games and rescaled partnership games are also saddlepoints and hence unstable under the exponential version of the perturbed best response dynamics.

Theorem 4.9 *Let $(x^*, y^*) \notin \mathrm{int}(S_n \times S_m)$ be a regular mixed equilibrium with support (K_1, K_2) of either a rescaled partnership game or a generic symmetric game ($A = B$). If the perturbation functions are $v_1(p) = -\sum p_i \log p_i$ and $v_2(q) = -\sum q_i \log q_i$, the perturbed equilibrium $(\hat{x}_\varepsilon, \hat{y}_\varepsilon)$ is unstable for the perturbed best response dynamics (8) for sufficiently small $\varepsilon > 0$.*

Proof: Since (x^*, y^*) is a regular mixed equilibrium, K_1 and K_2 have the same number of elements, and this number must be at least two. Consider the perturbed best response dynamics for the game $(A, B)|_{K_1 \times K_2}$ on $S_{K_1} \times S_{K_2}$. The linearisation, which we write $Q_K H_K/\varepsilon - I$, at a fully mixed equilibrium will have at least one positive eigenvalue if the game is a rescaled partnership game (Theorem 4.6) or if it is symmetric (Theorem 4.7). Now, returning to the dynamics on $S_n \times S_m$, under Assumption A, $\lim_{\varepsilon \to 0} Q(\hat{x}_\varepsilon, \hat{y}_\varepsilon)H$ will have a mixture of zero eigenvalues and eigenvalues of $Q_K H_K$ and therefore at least one positive eigenvalue. Thus, for ε sufficiently small, $Q(\hat{x}_\varepsilon, \hat{y}_\varepsilon)H/\varepsilon - I$ will have one too. QED

5 Deterministic versus Stochastic Perturbations

The purpose of this section is to clarify our results. We were able to show in the previous section that, for any mixed equilibrium of a game that is not rescaled zero sum, there is at least one form of the perturbed best response dynamics under which it is unstable. However, as we saw in Section 2, there are two ways of constructing perturbed best response functions. The proof of our result relies upon the wider set of perturbed best response functions that arise from the deterministic method. How different would our results be if we could only employ functions constructed from stochastic perturbations?

Hofbauer and Sandholm (2002, Proposition 2.2) point out perturbed best response functions constructed using the deterministic method form a larger class. This is because there are functions in that class that fail to satisfy a condition which does hold for all functions constructed by the stochastic method. Note that both types of perturbed best reply functions (4), (6) take the form $p(y) = C(Ay)$ where the function C maps expected payoffs to choice probabilities. Then, for any best response function constructed by the stochastic method, it must be that

$$(-1)^k \frac{\partial^k C_{i_0}(u)}{\partial u_{i_1} .... \partial u_{i_k}} > 0 \tag{24}$$

for each $k = 1, ..., n-1$ and each set of $k+1$ distinct indices $\{i_0, i_1, ... i_k\}$. The linearisation (19) only involves first order partial derivatives and so the only implication of the condition (24) for the linearisation is that $Q_{1,ij} < 0$ must hold for $i \neq j$, and similarly for Q_2. That is, the off-diagonal elements of the matrices Q_1, Q_2 in the stochastic model must be negative. In the deterministic model the requirement is only that Q_1, Q_2 are positive definite with respect to $I\!R^n_0$ and symmetric, with no restriction on the sign of the off-diagonal elements.

What was essential in establishing our result and in particular Lemma 4.2 was the free choice of Q_1 and Q_2 from the entire set of positive definite symmetric matrices. If we considered only the perturbed best response dynamics constructed from stochastic perturbations, the set of possible positive definite matrices available to play with is strictly smaller. It is therefore a logical possibility that, for some game (A, B), Lemma 4.2 fails to hold if we restrict our attention to matrices with negative off-diagonals. On the other hand, the set of available matrices is still very large. It is our conjecture that Lemma 4.2 still holds, though we have not been able to establish a proof. At least, we have been unable to identify a game, which is not rescaled zero sum, that has a mixed equilibrium which is stable for all stochastic perturbations.

6 Relation to Other Results

Perturbed best response dynamics are not the only model of learning in games. One reasonable question therefore is whether the results we have produced here apply to other forms of learning. If they do not then we would be much less confident about whether our current results can predict actual behaviour in experimental games. Luckily, as we will see, our current results are broadly in line with those obtained for other learning models.

The model most closely linked to the perturbed best response dynamics is stochastic fictitious play. Imagine two agents, who play the same game repeatedly at discrete time intervals. Classical fictitious play is based on the principle that each individual plays a best response to the past play of her opponent. Stochastic fictitious play modifies the original model slightly by assuming a random choice of action in each period, with the probabilities given by the perturbed best response functions $p(y)$ and $q(x)$. In particular, one can think of players' payoffs being subject to stochastic form of perturbation, as outlined in Section 2. Further details of this model can be found in Fudenberg and Levine (1998, Chapter 4), Benaïm and Hirsch (1999), and Hofbauer and Sandholm (2002). However, it should be apparent that the expected motion of the above stochastic process is a discrete time form of the perturbed best response dynamics (8). Or, in the terminology of stochastic approximation, the perturbed best response dynamics represent the mean value ODE associated with stochastic fictitious play. This means that stochastic fictitious play and the perturbed best response dynamics will have (essentially) the same asymptotic behaviour. This link is particularly strong when the differential equation is shown to be globally convergent. Thus, the global stability results in Section 3 can also be applied to study the long run behavior of evolutionary models for games with randomly perturbed payoffs, as demonstrated in Hofbauer and Sandholm (2002). Moreover, the results, Theorem 4.6 and Theorem 4.9, that mixed strategy equilibria of rescaled partnership games are unstable can be used in the analysis of stochastic fictitious play. See Duffy and Hopkins (2001) for an example of this approach.

The main rival to stochastic fictitious play as a model of human learning has been reinforcement or stimulus response learning, see Erev and Roth (1998). There is still much debate over which of the two best describe human learning behaviour. It has been known for some time that the expected motion of the Erev-Roth model of reinforcement learning is related to the evolutionary replicator dynamics.[9] These dynamics in the asymmetric case can be written

$$\dot{x} = R(x)Ay,\ \dot{y} = R(y)Bx \tag{25}$$

where $R(x)$ is the matrix function with $x_i(1 - x_i)$ on the diagonal and $-x_i x_j$ on the off-diagonal. It is therefore also a symmetric matrix positive definite with respect to $I\!R^n_0$.

Hofbauer and Sigmund (1998, Ch 11) show that in rescaled partnership games the replicator dynamics converge. This can be used to show that reinforcement learning converges also. A result of this type can be found in Duffy and Hopkins (2001) combined with an experimental test of its predictions. The linearisation of (25) at a mixed equilibrium (x^*, y^*) can be calculated as

$$\begin{pmatrix} R(x^*) & 0 \\ 0 & R(y^*) \end{pmatrix} \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}. \tag{26}$$

There is an obvious similarity with the linearisation of the perturbed best response dynamics (19). The similarity is even greater in the case of exponential perturbed best response function (7), as then Q is actually identical to R. This, together with Theorem 4.6, implies that the mixed strategy equilibrium of any rescaled partnership game will be unstable under the replicator dynamics. This result can be used to show that reinforcement learning will not converge to mixed strategy equilibria in rescaled partnership games (again see Duffy and Hopkins, 2001).

[9]The actual details are quite complicated and are analysed in Hopkins (2002), which also has several results on the relationship between reinforcement learning and stochastic fictitious play.

The case of rescaled zero sum games is more complicated. The absence from (26) of the additional $-I$ term present in (19) implies that, for example, at an interior equilibrium of a rescaled zero sum game, the linearisation will have all eigenvalues with zero real part. Indeed, Hofbauer and Sigmund (1998, Theorem 11.2.5) show that such equilibria are only Lyapunov (neutrally) stable under the replicator dynamics. Hofbauer (1996) shows that while there is an open set of 3×3 games for which mixed equilibria have all eigenvalues purely imaginary, at least two algebraic relations between the payoffs are needed for Lyapunov stability. On the other hand, the family of rescaled zero sum games has codimension 3 in the set of all 3×3 games (that is, in effect, they are non-generic). Hofbauer (1996) conjectures that an isolated interior equilibrium can be Lyapunov stable for the replicator dynamics only for rescaled zero sum games. Recently, Beggs (2001) has shown that despite the neutral stability of mixed equilibria under the replicator dynamics, reinforcement learning does converge in zero sum games.

Fictitious play is obviously linked to the best response dynamics. See Hofbauer (1995, 2000) for some general results on this connection. It has been known for some time that fictitious play converges in zero sum games. More recently, Monderer and Shapley (1996) proved that fictitious play converges also in partnership games which they call games of identical interest. Krishna and Sjöström (1998) show, for games with certain cyclic structure, that mixed equilibria are generically unstable under (continuous time) fictitious play. More recently, Echenique and Edlin (2002) prove a very general result on the instability of mixed strategy equilibria under fictitious play-like learning processes in games with strategic complements. In contrast, Sandholm (2003) shows that if one allows a sufficiently broad set of permissable perturbations, any mixed strategy equilibrium can be stable under the resulting perturbed dynamic. Hofbauer (1995) conjectures that if a mixed equilibrium is asymptotically stable for fictitious play or the best response dynamics (1), then the game is a rescaled zero sum game. Our Theorem 4.3 is obviously related to these two conjectures.

7 Conclusions

We have examined the stability of mixed strategy equilibria in general two player games under the perturbed best response dynamics linked to smooth or stochastic fictitious play. We found that such mixed strategy equilibria are stable for all dynamics of this class if and only if the game is a rescaled zero sum game. In that case, the equilibrium is globally stable. We also show that in rescaled partnership games, mixed strategy equilibria are unstable for the whole class of perturbed dynamics.

The result that mixed strategy equilibria are stable only in games that are linearly equivalent to zero sum games seems strong. However, the claim is that it is only in these rescaled zero sum games that there can be mixed equilibria that are stable under all perturbed best response dynamics. There will be other games that possess mixed equilibria that are stable for some dynamics in this class. Indeed, as Ellison and Fudenberg (2000) find, these may constitute a significant proportion of the total. However, for each of these we can find at least one perturbed best response dynamic for which it would be unstable.

It is a somewhat subjective question as to what is the more important criterion, "stable for some" or "stable for all". It is complicated further by the fact that "all", that is the set of admissible perturbed dynamics, has two possible definitions. As recently discovered by Hofbauer and Sandholm (2002), the set of perturbed best response dynamics constructed by the deterministic control cost method is strictly larger than that arising from the stochastic perturbation of payoffs. Thus, our demonstration that all mixed equilibria, excluding those of rescaled zero sum games, are unstable under at least one form of these dynamics, is a claim that they fail a very strict test. The argument in favour of the approach we take here is that it allows relatively sharp results. Furthermore, given there is some uncertainty about how people learn, there must be even greater uncertainty about what form of payoff perturbation is appropriate.[10] Consequently, results that do not depend on the exact form of the perturbation function are particularly valuable.

Furthermore, the differences between the stochastic and the deterministic methods are difficult to pin down. First, the two models coincide in a wide number of cases. The considerable success of the quantal response equilibrium literature, starting with McKelvey and Palfrey (1995), in explaining significant deviations of observed play from Nash equilibrium has largely used the logit choice rule, which is consistent with both the deterministic and stochastic models. Second, the set of cases for which the two models differ is very difficult to characterize. For example, it is a logical possibility that there is a class of games, that are not rescaled zero sum, which have mixed strategy equilibria that are stable under all perturbed best response dynamics consistent with stochastic perturbations. Given our present results, these mixed equilibria could not be stable for all dynamics consistent with deterministic perturbations and therefore for these games the two models give a substantially different prediction. However, not only do we not have a characterization of this set, we do not have a single example of such a game,[11] or any known case where the two models offer different predictions.

In conclusion, rescaled zero sum games and rescaled partnership games are games with a very strong structure and hence their stability properties under learning are largely independent of the learning process used. What is perhaps more unexpected is that in games without the strong structure of pure competition present in rescaled zero sum games, it is always possible to find a reasonable learning process that will

[10]Empirical studies that have tried to fit learning models to experimental data have found that individual behaviour is highly heterogeneous, for example, Cheung and Friedman (1997).

[11]Indeed, our conjecture is that the set is in fact empty.

diverge from any mixed strategy equilibrium.

References

Beggs, A., 2001. On the convergence of reinforcement learning. Working paper, University of Oxford.

Benaïm, M., 1999. Dynamics of stochastic algorithms, in *Séminaire de Probabilités XXXIII*, J. Azéma et al. Eds, Berlin: Springer-Verlag.

Benaïm, M., Hirsch, M.W., 1999. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games Econom. Behav.* **29**, 36-72.

Benveniste, A., M. Métivier, and P. Priouret. 1990. *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag, Berlin

Binmore, K., Samuelson, L., 1999. Evolutionary drift and equilibrium selection. *Rev. Econom. Studies* **66**, 363-393.

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

Cheung, Y-W., Friedman, D., 1997. Individual learning in normal form games: some laboratory results. *Games Econom. Behav.* **19**, 46-76.

Duffy J., Hopkins, E., 2001. Learning, information and sorting in market entry games: theory and evidence, forthcoming *Games Econom. Behav.*

Echenique, F., Edlin, A., 2002. Mixed equilibria in games of strategic complements are unstable. Forthcoming *J. Econom. Th.*

Ellison, G., Fudenberg, D., 2000. Learning purified mixed equilibria. *J. Econom. Th.* **90**, 84-115.

Erev, I., Roth, A.E., 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Amer. Econ. Rev.* **88**, 848-881.

Fudenberg, D., Kreps, D., 1993. Learning mixed equilibria. Games Econom. Behav. **5**, 320-367.

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

Fudenberg, D. Levine, D., 1999. Conditional universal consistency, *Games Econom. Behav.*, **29**, 104-130.

Gale, J., Binmore, K., Samuelson, L., 1995. Learning to be imperfect: the ultimatum game. *Games Econom. Behav.* **8**, 56-90.

Gilboa, I., Matsui, A., 1991. Social stability and equilibrium. *Econometrica* **59**, 859-867.

Harsanyi, J., 1973a. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *Int. J. Game Theory* **2**, 1-23.

Harsanyi, J., 1973b. Oddness of the number of equilibrium points: a new proof. *Int. J. Game Theory* **2**, 235-250.

Hines, W.G.S., 1980. Three characterizations of population strategy stability. *J. Applied Prob.*, **17**, 333-340.

Hofbauer, J., 1995. Stability for the best response dynamics. Working paper, University of Vienna.

Hofbauer, J., 1996. Evolutionary dynamics for bimatrix games: a Hamiltonian system? *J. Math. Biology* **34**, 675-688.

Hofbauer, J., 2000. From Nash and Brown to Maynard Smith: equilibria, dynamics and ESS. *Selection* **1**, 81-88.

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

Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK.

Hopkins, E., (1999). A note on best response dynamics. *Games Econom. Behav.* **29**, 138-150.

Hopkins, E., 2002. Two competing models of how people learn in games. *Econometrica* **70**, 2141-2166.

Kaniovski, Y., Young, P., 1995. Learning dynamics in games with stochastic perturbations. *Games Econom. Behav.* **11**, 330–363.

Krishna, V., Sjöström, T., 1998. On the convergence of fictitious play. *Mathematics of Operations Research* **23**, 479-511.

Matsui, A., 1992. Best response dynamics and socially stable strategies. *J. Econom. Theory* **57**, 438-362.

Mattson, L-G., Weibull, J.W., 2002. Probabilistic choice and procedurally bounded rationality. *Games Econom. Behav.* **41**, 61-78.

McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. *Games Econom. Behav.* **10**, 6-38.

Monderer, D., Shapley, L., 1996. Fictitious play property for games with identical interests. *J. Econom. Theory* **68**, 258-65.

Sandholm, W.H., 2003. Evolution in Bayesian Games II: Applications. Working paper, University of Wisconsin.

Shapley, L., 1964. Some topics in two person games. In Dresher, M. et al. (Eds.), *Advances in Game Theory.* Princeton University Press, Princeton.