Edinburgh School of Economics

Discussion Paper Series

Number 53

Learning in Perturbed Asymmetric Games

Josef Hofbauer (Univerity of Vienna)
Ed Hopkins (University of Edinburgh)

Date
September 2000

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://www.ed.ac.uk/schools-departments/economics

THE UNIVERSITY *of* EDINBURGH

Learning in Perturbed Asymmetric Games

Josef Hofbauer
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Vienna, Austria
Josef.Hofbauer@univie.ac.at

Ed Hopkins*
Department of Economics
University of Edinburgh
Edinburgh EH8 9JY, UK
E.Hopkins@ed.ac.uk

1 September, 2000

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 games of identical interest). Some applications of these results to stochastic learning models are given.

Journal of Economic Literature classification numbers: C72, D83.

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

---

*This paper was completed while Hopkins was visiting the Department of Economics, University of Pittsburgh, which he thanks for its hospitality.

1 Introduction

This paper analyses the properties of the perturbed best response dynamics that underlie models of stochastic or smooth fictitious play. Recently researchers have begun to try to fit stochastic learning models to data from experiments, see for example, Cheung and Friedman (1997), Erev and Roth (1998) and Camerer and Ho (1999). Though there has been some success in explaining the data, these attempts are running ahead of theory in that the properties of these models are not well understood. We here produce definitive results on when it is possible to learn mixed strategy equilibria under stochastic fictitious play. In particular, we provide results on global convergence for two classes of games, rescaled zero sum games, and rescaled partnership games. Together they include many of the games which have been subject to theoretical or experimental investigation in the recent literature on 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 purification argument. Using techniques from stochastic approximation theory, this has been studied by Benaïm and Hirsch (1999) in terms of the associated continuous time perturbed best response dynamics. In a recent paper Ellison and Fudenberg (2000) examine the perturbed best response dynamics in 3×3 bimatrix games. They find that there are many games for which mixed equilibria are 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.

Here, we undertake an analysis of the general case. We show that the set of games with mixed equilibria stable for all perturbed best response dynamics is exactly the set of games called rescaled zero sum games by Hofbauer and Sigmund (1998). We show global convergence for these games, and instability of mixed equilibria in all other games. In rescaled partnership games, also known as games of identical interest, it is possible to obtain a stronger instability result: mixed equilibria are unstable for all dynamics of this class.

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.

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. Results from the theory of stochastic approximation show that the asymptotic behaviour of a perturbed form of the best response dynamics and stochastic fictitious play, that is, fictitious play where payoffs are randomly perturbed, are closely linked. The exact connections are set out in Fudenberg and Levine (1998, Chapters 2 and 4). Thus, although we largely confine our attention to deterministic continuous time dynamics, as shown in Section 5 below, our results may be used for predicting the behaviour of stochastic fictitious play, a stochastic process unfolding in discrete time.

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 $\{x = (x_1, ..., x_n) \in \mathbb{R}^n : \sum x_i = 1, x_i \geq 0, \text{ for } i = 1, ..., n\}$, 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.

If $\dot{x} = f(x)$ is a differential equation on the $n-1$ dimensional simplex S_n such that the solutions $x(t)$ remain in S_n, the vector field $f(x)$ must be in the tangent space $\mathbb{R}^n_0 = \{\xi \in \mathbb{R}^n : \sum \xi_i = 0\}$ of S_n. Note that $\mathbb{R}^n$ can be decomposed into two orthogonal subspaces $\mathbb{R}^n_0$ and $\mathbb{R}^n_1 = \{x \in \mathbb{R}^n : x_i = x_j \text{ for all } i, j\}$. When we look at a linearisation around an equilibrium of such a dynamical system, the stability of that equilibrium will be determined by $n-1$ eigenvalues which refer to $\mathbb{R}^n_0$, which has dimension $n-1$, and not the nth eigenvalue which refers to $\mathbb{R}^n_1$.

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

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

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 subject it to detailed analysis as Hofbauer (1995) shows. 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 (1998, Chapter 4), 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., $\lim_{p \to \partial S_n} |v_1'(p)| = \infty$.

Typical examples are $v_1(p) = \sum_i \log p_i$ and $v_1(p) = -\sum_i p_i \log p_i$. We assume that v_2 possesses similar properties. 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, and similarly 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' $(p(y), q(x))$.[1] (Brief reflection reveals that any other specification of the noise function v would not guarantee that (2) actually had a unique maximum in $\operatorname{int} S_n$ and thus for every y, there would be a unique best reply.) Though clearly in the spirit of Harsanyi's (1973) purification argument, it does employ a somewhat narrower specification in that the perturbations to player 1's payoffs depend only on her strategy and not on that of player 2. That is, v_1 is a function of p but not of y. One argument in favour of the current specification is that the shocks to payoffs are private information and therefore independent of the play of the other player.

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

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

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 (6) below.

[1] The original formulation of stochastic fictitious play due to Fudenberg and Kreps (1993) involved a truly stochastic perturbation of payoffs. Hofbauer and Sandholm (2000) show that any perturbed best response function derived from such a stochastic problem can also be derived from a deterministic optimisation problem of the form considered here.

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

$$\dot{x} = p(y) - x, \quad \dot{y} = q(x) - y. \tag{5}$$

Equally, as discussed above, this system of differential equations can be used to predict the stochastic learning of individual agents. See Section 5.

Let (x^*, y^*) be a regular Nash equilibrium. We note that from Harsanyi's (1973) purification theorem, when ε, the level of perturbation, is small, there is an associated equilibrium of the perturbed best response dynamics (5), which we label $(\hat{x}, \hat{y})$ and that $\lim_{\varepsilon \to 0}(\hat{x}, \hat{y}) = (x^*, y^*)$. Such a perturbed equilibrium is also a "quantal response equilibrium" in the terminology of McKelvey and Palfrey (1995). When a Nash equilibrium of the original game forms part of a continuum of equilibria, the relation between Nash and perturbed equilibria is no longer one-to-one. Under perturbation, a continuum of equilibria may be replaced by a single (or possibly several) perturbed equilibrium or may disappear entirely. See Binmore and Samuelson (1999) for a more complete discussion of this issue.

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.

We note in passing two particular functional forms that have been used in the recent literature. First, an exponential form

$$p_i(y) = \frac{\exp \beta (Ay)_i}{\sum_{j=1}^n \exp \beta (Ay)_j}, \tag{6}$$

which can be obtained by setting $v_1(p) = -\sum p_i \log p_i$. Second,

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

In both cases, $\beta = \varepsilon^{-1}$. The latter form has the advantage that for $\beta = 1$, this rule approximates a reinforcement learning rule (see Camerer and Ho, 1999; Erev and Roth, 1998), whereas for stochastic fictitious play, one imagines ε to be small and hence β large. Therefore empirical estimates of β would seem to give an indication of whether the behaviour of experimental subjects is better described by reinforcement learning or stochastic fictitious play. However, the second form, despite its apparent similarity to the exponential form, has the disadvantage that it is not in fact consistent with the model of maximisation of perturbed payoffs described here. There is no disturbance function v which would generate such a rule (see Section 4 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 (5) does not depend on the perturbation function v and is the same as that of the best response dynamics (1). We extend this result to higher dimensions for two important classes of games, games of conflict and games of coordination.

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

$$a'_{ij} = \alpha a_{ij} + c_j, b'_{ji} = \beta b_{ji} + d_i. \tag{8}$$

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. However, the relationship between equilibria of the original game and its rescaling is still clearly one-to-one.

We start with the following simple characterization.

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

$$c\xi \cdot A\eta = \eta \cdot B\xi \quad \text{for all } \xi \in I\!R^n_0, \eta \in I\!R^m_0 \tag{9}$$

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

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

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). The following global stability result extends a generic local result of Ellison and Fudenberg (2000, Proposition 8) for 3×3 zero sum games.

Theorem 3.2 *For any two person rescaled zero-sum game, the perturbed best response dynamic (5) has a unique rest point which is globally asymptotically stable.*

Proof: Consider the functions

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

where $h > 0$ is defined as $h = -1/c$ where c is the constant from (8) 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,

$$V(x,y) := V_1(x,y) + V_2(x,y). \tag{10}$$

V_1 (and similarly V_2) is a convex function of (x,y) since $\pi_1(x,y)$ is concave and $\pi_1(p(y),y) = \max_z \pi_1(z,y)$ is convex, being the maximum of linear functions in y. Moreover, V_1 is strictly convex in x and V_2 is strictly convex in y. This shows that V attains its minimum value 0 in a *unique* point, which is the unique perturbed equilibrium of the rescaled zero sum game. The definition of $p(y), q(x)$ implies

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

where ∂_1 and ∂_2 denote again the partial derivatives with respect to the first and second variable (within $I\!R_0^n$ and $I\!R_0^m$). 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 9, 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 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 $q = y$. ∎

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 the players' perturbed payoff is always rising out of equilibrium and hence moving toward the set of perturbed equilibria.

Theorem 3.3 *For any rescaled partnership game, each orbit of the perturbed best response dynamics (5) 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 (8) 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, $x = p$, $y = q$, 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), $\xi \cdot A'y = \xi \cdot Ay = -\varepsilon v_1'(p)$ and $x \cdot A'\eta = \beta\eta \cdot Bx = -\beta\varepsilon v_2'(q)$ 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 $x = p$ and $y = q$. Hence every ω–limit set consists of perturbed equilibria. ∎

As a consequence, we can expect, that in generic games, for small enough $\varepsilon > 0$, the ω–limit set of a generic initial condition will be a perturbed strict equilibrium.

The restriction to generic games would be to exclude games with connected components 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 could converge to a perturbed equilibrium which is not close to to a pure equilibrium.

The restriction to generic initial conditions would be necessary since mixed equilibria of rescaled partnership games will be saddlepoints under the perturbed best response dynamics (see Theorem 4.5 below) and therefore attract some part of the state space (their stable manifold).

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. 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 $\frac{1}{\varepsilon}Q_1 A$ where Q_1 is a symmetric matrix positive definite with respect to $I\!R_0^n$. That is, $\xi \cdot Q_1 \xi > 0$ for any nonzero $\xi \in I\!R_0^n$.*

Proof: This result is obtained simply 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 (1999a) for more details. ∎

We remark in passing that this Lemma shows why the perturbed best response function (7) cannot arise from a maximisation problem. Clearly, if $Q_1 = -(v_1'')^{-1}$, Q_1 must be symmetric. However, the equivalent matrix arising from differentiating (7) is not. In any case, given Lemma 4.1 and the dynamics (5), 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 $I\!R_0^n$ and $I\!R_0^m$ respectively. We can write this 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.

Note that the evolutionary replicator dynamics for the first population in the asymmetric case can be written $\dot{x} = R(x)Ay$, where R is also a symmetric matrix positive definite with respect to $I\!R_0^n$. The linearisation at a mixed equilibrium can therefore be written $R(x^*, y^*)H$. There is an obvious similarity with the linearisation of the perturbed best response dynamics.[2] However, the additional $-I$ term in the case of (19) is crucially important. It is responsible for the asymptotic stability under the perturbed best response dynamics as opposed to the neutral stability under the replicator dynamics of a mixed equilibrium of a rescaled zero sum game.

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 should have no

[2] The similarity is even greater in the case of exponential perturbed best response function (6), as then Q is actually identical to R. See Hopkins (1999a, b) for links between the two dynamics.

eigenvalues with positive real part. However, looking at the matrix QH, as it has a zero trace, there are two possible structures for its eigenvalues. First, they are a mixture of positive and negative summing to zero. This case is clearly unstable. Second, all eigenvalues have real part zero. In this 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 first determine a condition for instability.

Lemma 4.2 *There exist positive definite symmetric Q_1, Q_2 such that the following 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 if there exists $\xi \in I\!R^n_0$ and $\eta \in I\!R^m_0$ such that

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

Proof: Let (ξ, η) be such that (20) holds. Write $H(\xi, \eta) = (x, y)$. Then, $\xi \cdot x > 0$ and $\eta \cdot y > 0$. Because $\xi \cdot x > 0$, there is a positive definite symmetric Q_1, such that $Q_1 x = \xi$. Similarly, let $Q_2 y = \eta$. Then $QH(\xi, \eta) = \lambda(\xi, \eta)$, where $\lambda = 1 > 0$. ∎

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

$$(\xi \cdot A\eta)(\eta \cdot B\xi) \leq 0, \text{ for all } \xi \in I\!R^n_0,\ \eta \in I\!R^m_0. \tag{21}$$

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 $(x^*, y^*) \in \mathrm{int}(S_n \times S_m)$ 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 (5) for arbitrarily small $\varepsilon > 0$ then the game (A, B) is a rescaled zero sum game.*

Proof: 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 (18) holds at $(\hat{x}, \hat{y})$. 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. ∎

Now, the above theorem says that only equilibria of rescaled zero sum games can be stable for all dynamics of the form (5). 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 now ask which sort of mixed equilibria are definitely unstable for all such dynamics. We start with a preliminary result on the calculation of the eigenvalues of a linearisation 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 (for $m \neq n$, 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). ∎

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. Hence we note the following.

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 (5), 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 $\mathbb{R}^n_0$ as $cQ_1AQ_2A^T$ with $c < 0$. The matrix $P = AQ_2A^T$ is symmetric and positive definite with respect to $\mathbb{R}^n_0$. The product of two symmetric positive definite matrices has all eigenvalues real and positive (see for example, Hines, 1980). Hence, cQ_1P has $n - 1$ negative eigenvalues, the other zero eigenvalue corresponding to $\mathbb{R}^n_1$. The nonzero eigenvalues of the matrix QH are therefore the $2(n-1)$ purely imaginary square roots of the negative eigenvalues of $Q_1 A Q_2 B$. 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. ∎

The next result complements Theorem 3.2 and shows that the mixed equilibria of rescaled partnership games are always saddlepoints 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 (5) for all admissible disturbance functions v_1, v_2 and all small $\varepsilon > 0$.*

Proof: We define $P = AQ_2A^T$, as in the proof to the previous theorem, but now cQ_1P has $n-1$ positive eigenvalues, as we are considering a rescaled partnership game and $c > 0$. 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, $\lim_{\varepsilon\to 0} Q_1(\hat{x}_\varepsilon)$ will be positive definite, and similarly for $\lim_{\varepsilon\to 0} Q_2(\hat{y}_\varepsilon)$.[3] 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. ∎

Of course, for $n \geq 3$ there are a large number of games with mixed strategy equilibria which are neither rescaled zero sum or partnership games. Indeed this point highlights one difference in our approach from that of Ellison and Fudenberg (2000). They are principally concerned with the conditions under which a 3×3 game possesses a mixed equilibria which is stable under some perturbed best response dynamic. Furthermore, they carry out a Monte Carlo experiment which suggests that this set of games forms a significant proportion of games that possess mixed equilibria.[4] It is a somewhat subjective question as to what is the more important criterion, "stable for some" or "stable for all", but the argument in favour of the approach we take here is, first, that it allows relatively sharp results. Second, given there is some uncertainty about how people learn, there must be even greater uncertainty about what form of payoff perturbation is appropriate. Consequently, results that do not depend on the exact form of the perturbation function are particularly valuable.

[3]Note that in contrast each $Q_i^*(\hat{x}_\varepsilon)$ 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.

[4]Indeed, Binmore and Samuelson (2000) and Ely and Sandholm (2000) both show, albeit in quite different contexts, that, if the repertoire of possible perturbations is sufficiently rich, that mixed strategy equilibria even of partnership games can stabilised.

5 Stochastic Fictitious Play

The advantage of global stability results is that they can be usefully applied via the theory of stochastic approximation to stochastic learning models. In this section, we carry out such a transfer, giving an example of how the results we have obtained can be used to analyse the discrete time process known as stochastic fictitious play. Imagine two agents, who play the same game repeatedly at discrete time intervals, indexed by t. The probabilities with which the first player chooses between his actions in period t are given by $p(y_t)$, where p is the perturbed best reply function as defined before and y_t is the vector of empirical frequencies of her opponent's past actions. Similarly, player 2 chooses according to probabilities $q(x_t)$.

Then, one can calculate,

$$x_t - x_{t-1} = \frac{p(y_{t-1}) - x_{t-1} + \eta_t^1}{t},\ y_t - y_{t-1} = \frac{q(x_{t-1}) - y_{t-1} + \eta_t^2}{t}, \tag{22}$$

where η_t^1, η_t^2 are random variables with expectation conditional on x_{t-1} and y_{t-1} equal to zero. Further details of this model can be found Fudenberg and Levine (1998, Chapter 4) and Benaïm and Hirsch (1999), but it should be apparent that the expected motion of the above stochastic process is a discrete time form of the perturbed best response dynamics, with a step size of $1/t$. We give the following two theorems, which extend the existing results of Fudenberg and Kreps (1993) and Benaïm and Hirsch (1999) on 2×2 games.

Theorem 5.1 *For all stochastic fictitious play processes of form (22), in any rescaled zero sum game,* $\lim_{n\to\infty}(x_n, y_n) = (\hat{x}, \hat{y})$ *with probability one. That is, the learning process converges to the unique perturbed equilibrium.*

Proof: This follows from our Theorem 3.2 and Theorem 3.3 of Benaïm and Hirsch (1999). ∎

Similarly, it is possible to transfer our deterministic results on rescaled partnership games.

Theorem 5.2 *All stochastic fictitious play processes of form (22) in generic rescaled partnership games converge to the set of perturbed equilibria.*

Proof: This follows from our Theorem 3.3 and Corollary 6.6 of Benaïm (1999), which applies if the number of equilibrium points is countable. This will be the case in generic games. ∎

6 Comparison with Other Results

Rescaled zero sum games and rescaled partnership games were analysed with respect to the evolutionary replicator dynamics by Hofbauer and Sigmund (1998, Ch11). They prove a convergence result for partnership games similar to the one found here. However, the mixed strategy equilibria of rescaled zero sum games are only 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 further conditions on the higher order terms 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. Hofbauer (1996) conjectures that an isolated interior equilibrium can be Lyapunov stable for the replicator dynamics only for rescaled zero sum games.

Fictitious play is obviously linked to 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. 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.

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). Hopkins (1999b) shows that reinforcement learning and the perturbed best response dynamics analysed here can both be modelled as forms of noisy replicator dynamic. In particular, there exists a form of reinforcement learning which has the same expected motion as the exponential form of stochastic fictitious play and hence has identical local stability properties.

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 without the strong structure of pure competition present in zero sum games, convergence to mixed strategy equilibrium is extremely unlikely.

References

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 Econ. Behav.*, **29**, 36-72.

Binmore, K., Samuelson, L. (1999) "Evolutionary drift and equilibrium selection," *Rev. Econ. Studies*, **66**, 363-393.

Binmore, K., Samuelson, L. (2000) "Evolution and mixed strategies," forthcoming *Games Econ. Behav.*

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 Econ. Behav.*, **19**, 46-76.

Ellison, G., Fudenberg, D. (2000) "Learning purified mixed equilibria," *J. Econ. Th.*, **90**, 84-115.

Ely, J.C., Sandholm, W.H. (2000) "Evolution with diverse preferences," working paper, University of Wisconsin.

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 Econ. Behav.*, **5**, 320-367.

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

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

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

Hines, W.G.S. (1980). "Three characterizations of population strategy stability," *J. Appl. 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," forthcoming *Selection.*

Hofbauer, J., Sandholm, W.H. (2000). "Disturbance-based models of evolution and learning in games," working paper.

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

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

Hopkins, E. (1999b). "Two competing models of how people learn in games," working paper.

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

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

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

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