ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

LEARNING IN BAYESIAN GAMES WITH BINARY ACTIONS

A.W. Beggs

Number 232

April 2005

Manor Road Building, Oxford OX1 3UQ

Learning in Bayesian Games with Binary Actions

A. W. Beggs
Wadham College
Oxford
OX1 3PN
UK

alan.beggs@economics.ox.ac.uk

Current Version: April 2005

Abstract

This paper consider a simple adaptive learning rule in Bayesian games where players employ threshold strategies. Global convergence results are given for supermodular games and potential games.

Keywords: bayesian games, learning, binary actions, passive stochastic approximation

JEL Nos: C72, D83

I am grateful to seminar participants at Oxford for helpful comments.

1. Introduction

Bayesian games are widely used in Economics, for example in the study of auctions and public goods provision. Rather little is known about whether players are likely to learn to play the equilibria involved in such games. This paper studies a simple adaptive learning model in the context of binary action Bayesian games and gives conditions for its convergence to equilibrium.

The class of Bayesian games with binary actions is restrictive but is widely used in practice. It includes for example, but is not restricted to, many of the applications of the theory of global games stimulated by the work of Carlsson and van Damme (1993) — see for example the survey by Morris and Shin (2003). It has the advantage that the form of the equilibria is fairly simple: in the examples studied here players play one action below a certain threshold, the other above it. If players cannot learn to play equilibrium in such games, it is unlikely they can do so in more complex ones.

Learning in a Bayesian game is not simple. A player must learn to play a strategy which specifies an action for each possible signal he may receive. Moreover in attempting to learn about other players, he is likely to be able only to observe his own past signals and their actions, not their strategies. One might imagine that the player would attempt to infer his opponents' strategies from past data but, especially if the space of signals is continuous, this is a non-trivial inference problem. Furthermore one might not even wish to assume that the player knows the distribution from which he and his opponents draw their signals.

This paper examines a simpler adaptive rule. In the current environment optimal strategies are specified by a threshold. It is assumed that each player employs such a threshold and adjusts in the direction suggested by past play. Even such a rule is not entirely straightforward as he must estimate the direction in which to move. The paper shows that this can be achieved by a simple scheme.

Two features of the learning rule merit highlighting. Firstly, it is perhaps in the spirit of 'directional learning', as suggested by Selten and Buchta (1998). Players do not pick best responses at each instant but rather move in a direction that will improve their payoffs. Given the difficulties of even computing the other player's strategy, let alone a best response, this does not seem unreasonable.

Secondly, the ideas in the paper may also be linked to those 'similarity', as studied for example by Gilboa and Schmeidler (2003). Players only observe one signal at a time. In order to estimate payoffs at signals they do not observe, they must use observations of 'similar' signals, weighting them by their proximity to the signal at which payoffs are to be estimated.

The paper gives global convergence results for games with binary actions and strategic complementarities. It also demonstrates convergence in potential games with monotone equilibria. It also gives estimates for the rate of convergence. The proofs use ideas from the theory of passive stochastic approximation introduced by Härdle and Nixdorf (1987).

One could of course take the view that players are fully rational and so can easily compute the equilibrium. This is the viewpoint taken in the literature on global games. Experimental work by Cabrales et al. (2002) and Heinemann et al. (2004) throws doubt on this. In the experiments reported players do not leap immediately to equilibrium.

Learning in games has of course been widely studied. A good survey can be found in Fudenberg and Levine (1998). Most of this work has studied learning in environments without private information. Much of it has also restricted attention to normal-form games, where strategies are observable. Neither feature is appropriate here.

An exception is the work on perturbed fictitious play, for example of Fudenberg and Kreps (1993) or Benaïm and Hirsch (1999). This work studies convergence to equilibria of normal-form games, using Harsanyi's interpretation of mixed equilibria as Bayesian equilibria. In these games, however, perturbations to payoffs are independent and the signal one player receives tells him nothing about the likely action of another player. By contrast the latter inference is a key feature of the equilibria studied here. With independent perturbations the only feature a player needs to estimate is the probability with which his opponents choose their actions, which he can do by looking at their frequency in past play. Here however the player needs to know how his signal alters the probabilities of different actions being played.[1]

Some of the work on convergence to Rational Expectations equilibrium (see for example Bray (1982)) can be interpreted as studying convergence to a Bayesian equilibrium. The environment there, however, is non-strategic. Dekel et al. (2004) discuss what learning might imply about beliefs of players in Bayesian games but do not discuss any specific learning rules or convergence.

The paper proceeds as follows. Section 2 gives two examples and Section 3 outlines the general environment studied. Section 4 explains the learning rule. Section 5 gives the main convergence results. It also gives results on the rate of convergence. Section 6 discusses extensions. Section 7 concludes.

2. Two Examples

This section gives two simple examples of the class of games considered.

Example 1 *(A Public Goods Game)*

Consider the game in Figure 1. It can be thought of as a simple public goods game with voluntary contributions of the kind studied by Palfrey and Rosenthal (1994). There are two players and each player can either contribute (action 0) and not (action 1). If both players contribute then the project goes ahead yielding a gross payoff of 1 to each player, otherwise it does not and they each receive zero. The cost of contribution to player i is t_i, so net payoffs are as shown in Figure 1. Each player observes only his own cost level.

[1] Benaïm and Hirsch (1999) do allow for correlation in perturbations but they assume that players ignore it when choosing actions, so in this case the system does not converge to a true Bayesian equilibrium.

Suppose that t_1 and t_2 are drawn from a distribution with joint distribution function $F(t_1, t_2)$ and density function f. Let the support of F be $[\underline{t}_1, \overline{t}_1] \times [\underline{t}_2, \overline{t}_2]$. $\infty \geq \overline{t}_i > 1$ and $-\infty \leq \underline{t}_i < 0$ for $i = 1, 2$. Negative values of costs are allowed and may be taken to reflect 'altruism'.

Let $F_j(t_j|t_i)$ be the cumulative distribution function of t_j given t_i. It is straightforward to check that if $F_j(t_j|t_i)$ is decreasing in t_i, that is the higher a player's own cost the higher the other player's cost is likely to be, then there is an interior equilibrium of the form:

$$\begin{cases} \text{Player } i \text{ contributes} & \text{if } t_i < k_i \\ \text{Player } i \text{ does not contribute} & \text{if } t_i > k_i \end{cases} \tag{1}$$

k_i is determined the pair of equations

$$k_i = F_j(k_j|k_i) \qquad j \neq i, \quad i = 1, 2 \tag{2}$$

It natural to ask whether players can learn to play such an equilibrium. Two features of this model are important in the future analysis: firstly that $F_j(t_j|t_i)$ is decreasing in t_i, secondly that the relative payoff to not contributing (action 1) is higher if the other player does not contribute and increasing in own cost. The second is a supermodularity assumption. The first assumption is clearly true if costs are independently distributed. This not necessary, however, and the current model allows for one player's costs to be informative about the other's.

Example 2 *(Global Games)*

Another example of the class of games considered is the global games considered by Carlsson and van Damme (1993) and Morris and Shin (2003). Consider the following simple example drawn from Morris and Shin (2003) shown in Figure 2.

When payoffs are commonly observed, the unique equilibrium is for both players to play action 0 if $\theta < 0$ and action 1 $\theta > 1$. For $0 < \theta < 1$, there are multiple equilibria: both playing 0 is an equilibrium, so is both playing 1 and there is a mixed equilibrium.

The literature on global games considers the case when each player only observes a noisy signal of his payoffs. Morris and Shin (2003) consider the case when player i observes $t_i = \theta + u_i$ when u_i is Normal with mean zero and variance σ^2 with u_1 and u_2 independent. In the case when players have an improper prior for θ (the uniform density on the real line), they show that there is a unique equilibrium where player i plays

$$\begin{cases} \text{Action } 0 & \text{if } t_i < \frac{1}{2} \\ \text{Action } 1 & \text{if } t_i > \frac{1}{2} \end{cases} \tag{3}$$

Various authors have extended these results (see Morris and Shin (2003) for a survey) and others have criticised their robustness (see for example Hellwig (2003)). This is not the focus here. For the purposes of this paper the interesting feature of the model is that it is a Bayesian game where players pursue a simple cutoff strategy: play action 1 if their signal exceeds some value k_i.

The form of the rule is simple yet one might ask how easy it is for players to learn to use such a rule. The arguments given by Carlsson and van Damme (1993) and Morris and Shin (2003) are largely deductive in nature. Experiments in Cabrales et al. (2002) and Heinemann et al. (2004) give ambiguous support to the model. Players do not seem to converge immediately to the solution implied by deductive reasoning. They do, however, seem to employ threshold strategies.

One might wonder whether naïve players can learn to play such an equilibrium. Morris and Shin (2003) note that if the game in Figure 2 is considered as a normal-form game it is is supermodular in the sense of Milgrom and Roberts (1990). It follows from the results in the latter paper that if each player observes each other's choice of cutoff at each stage, many simple learning rules will converge to equilibrium. The assumption that cutoffs are observable seems doubtful. The current paper asks whether players can learn to play the equilibrium when they simply observe their own signals and payoffs and the other player's actions.

Two features in this example will play an important role in the later analysis. In the first place, the gain to playing action 1 rather than action 0 is increasing in θ for a fixed action of the other player and for fixed θ is higher if the other player takes action 1 rather than action 0. In other words, payoffs are supermodular. In the second place, the higher the signal one player observes the more likely it is that the other player has observed a high signal. Together these features imply that each player will use a cutoff strategy. Supermodularity will also be used to prove convergence of the learning rule studied.

3. The Model

This section describes the general class of games to be studied.

Assumption 1 *There are two players* $i = 1, 2$.

This assumption will be relaxed in Section 6.

Assumption 2 *Each player has a binary action set* $A_i = \{0, 1\}$, $i = 1, 2$.

Attention is limited to binary action games. The implications of relaxing this assumption are discussed in Section 6.

As is standard in Bayesian games it is assumed that each player has certain possible types, belonging to T_i for player i. In the examples of Section 2, the type of a player is the signal he receives. Payoffs may also depend on the state of nature, which lies in T_0. In the second example of Section 2 the state of nature was given by θ, which had an improper prior distribution. In this section it is assumed that the distribution of types is described by a probability measure and so an improper prior is not allowed. The qualitative features of the example are, however, preserved with a proper but highly spread out prior. The second example of Section 2 fits into the current framework with this qualification.

The following assumptions are largely technical.

Assumption 3 *The set of types T_i equals the real line, $i = 1, 2$.*

In Example 1 of Section 2 it may be slightly more natural to assume:

Assumption 3′ *The set of types T_i is a finite interval $[\underline{t}_i, \bar{t}_i]$, $i = 1, 2$.*

The learning rule studied needs to be slightly modified in this case but, as discussed in Section 5, the convergence results still hold.

Assumption 4 *The marginal distribution of types on $T_1 \times T_2$ has a density, f, with respect to Lebesgue measure which is continuous and strictly positive.*

The assumption that f is strictly positive guarantees that observations are always possible at every point.

The realised payoff to player i is $y_i(a, t_0, t)$, where a are the chosen actions, t_0 is the state of nature and t is the vector of players' types. $\Delta y_i(a_j, t_0, t)$ denotes the difference in payoffs to player i from taking action 1 rather than action 0. For most of the paper T_0 is not of direct interest. Assumptions will therefore be made directly on y_i and on the the expected payoff to actions conditional on the players' signals, denoted by $\pi_i(a_1, a_2, t_1, t_2)$. It is assumed that the latter is well-defined:

Assumption 5 *For each i and all a_1, a_2, $\pi_i(a_1, a_2, t_1, t_2)$ is well-defined, finite and is integrable with respect to f.*

Let

$$\Delta\pi_i(a_j, t_i, t_j) = \pi_i(1, a_j, t_i, t_j) - \pi_i(0, a_j, t_i, t_j) \tag{4}$$

denote the incremental payoff to taking action 1 rather than zero.

Assumption 6 *For each i, there exist finite $\overline{\tau}_i$ and $\underline{\tau}_i$ together with $l_i > 0$ such that $E\left(\Delta\pi_i(a_j, t_i, t_j)|t_i\right) < -l_i$ for all $t_i < \underline{\tau}_i$ and $E\left(\Delta\pi_i(a_j, t_i, t_j)|t_i\right) > l_i$ for all $t_i > \overline{\tau}_i$, for all a_j.*

Assumption 6 amounts to the assumption that for low enough signals action 0 is the best action to play for a player regardless of the other player's actions and for high enough signals action 1 is the best action to play. It guarantees that only cutoffs within a bounded range need be considered and so the learning model remains bounded. In the case of Assumption 3′ boundedness is automatic. In this case Assumption 6 implies that equilibrium cutoffs are strictly interior to the interval of types.

Assumption 7 *$\Delta\pi_i(a_j, t_i, t_j)$ is strictly increasing in a_j and increasing in t_i and t_j for $i = 1, 2$, $j \neq i$*

Assumption 7 expresses the assumption that payoffs are supermodular in actions and types (given that the action space is binary). The assumption that $\Delta\pi_i$ is strictly increasing in a_i is used in the proof of convergence of the algorithm.

Assumption 8 *Let $F_i(t_j|t_i)$ be the conditional distribution function of t_j given t_i. F_i is decreasing in t_i for $i = 1, 2$.*

Assumption 8 implies that a higher signal for player i makes it more likely that player j has received a higher signal, in the sense of first-order stochastic dominance. Assumptions 7 and 8 imply (see for example Topkis (1998) Lemma 3.9.1(b) and Corollary 3.9.1(a)) that

Lemma 1 *$E\left[\Delta\pi_i(a_j, t_i, t_j)|t_i\right]$ is increasing in t_i for each i, where E denotes expectation.*

A *strategy* for player i is a map $\sigma_i: T_i \longrightarrow A_i$. It follows from Lemma 1 that each player always has an increasing best response to an increasing strategy and so, under Assumptions 1–8, there is an equilibrium in monotonic strategies (see for example Athey (2001) Lemma 3). There may be other, non-monotonic, equilibria.

An increasing strategy can be characterised by a *cutoff* k: play action 1 if $t > k$, action 0 otherwise. A monotonic equilibrium can therefore be characterised by a pair of cutoffs, k_1 and k_2 for each player.

In the remainder of the Section some more technical assumptions are gathered for convenience. The reader who is not interested in these details may skip directly to Section 4.

In Section 4, the marginal payoff to player i if he receives signal k_i will be of interest. This is given for player 1 by

$$H_1(k_1, k_2) = \int_{-\infty}^{k_2} f(k_1, y)\Delta\pi_1(k_1, y, 0)\, dy + \int_{k_2}^{\infty} f(k_1, y)\Delta\pi_1(k_1, y, 1)\, dy \qquad (5)$$

and similarly for $i = 2$. Note that this is not the same as payoff to player i conditional on receiving i. The term marginal is used in analogy to the term marginal distribution.

Assumption 9 *For each i, H_i is bounded and is continuously differentiable in k_1 and k_2. H_i has a bounded first derivative with respect to k_i.*

The boundedness of the first derivative implies that H_i satisfies a Lipschitz condition, so that observations at nearby points are a good guide to behaviour at k_i. The smoothness of H_i is required in the analysis of the differential equation.

Assumption 9 can derived from more primitive conditions on $\Delta\pi_i$ and f. For example boundedness follows if $f\Delta\pi_i$ is bounded for each a_j by an integrable function of t_2 (note that $\Delta\pi_i$ need not be bounded). Smoothness can be deduced if $f\Delta\pi$ and its derivatives are dominated by a suitable integrable function (see for example Billingsley (1995) Theorem 16.8). Continuous differentiability in the other player's cutoff, k_j, merely requires continuity of $f\Delta\pi_i$ in t_j. On the other hand, these primitive conditions are superfluous if Assumption 9 holds, so it seems preferable to state it directly.

In the case of Example 1

$$H_i(k_1, k_2) = f_i(k_i)k_i - F_i(k_j|k_i)f_i(k_i) \qquad (6)$$

f_i denotes the marginal density of t_i. Assumption 9 will therefore hold if f and its first derivatives are bounded and go fast enough to zero as k_1 and k_2 become large. It is easy to check that Assumption 9 holds in the second example of Section 2 if it is assumed that θ has a proper prior following a Normal distribution.

For the rate of convergence proofs a slightly stronger condition is required:

Assumption 9′ *For each i, H_i has the properties in Assumption 9 and in addition has a second derivative with respect to k_i which is bounded and continuous in k_1 and k_2.*

Again this is satisfied in Example 2 if θ has a Normal prior and under mild conditions in Example 1.

It also assumed that second moments exist:

Assumption 10 *For each i, Δy_i has a finite second moment for all a_j, $j \neq i$.*

Note that this assumption is on the unconditional distribution of payoffs. It, and boundedness of H_i in Assumption 9, are required for technical reasons in the convergence proofs.

In the rate of convergence proof a stronger assumption will be required. Let M_i be the marginal expected value of the squared difference in payoffs to actions to player i when he receives signal k_i and his opponent uses cutoff k_j:

$$M_1(k_1,k_2) = \int_{-\infty}^{k_2} f(k_1,y)E\left(\Delta y_1^2(t,0)|k_1,y\right)\,dy + \int_{k_2}^{\infty} f(k_1,y)E\left(\Delta y_1^2(t,1)|k_1,y\right)\,dy \tag{7}$$

M_i is the analogous function to H_i for the squared differences in payoffs.

Assumption 10′ *(a) For each i, Δy_i has a finite fourth moment for all a_j, $j \neq i$.*

(b) For each i, M_i is a Lipschitz continuous function of k_i

All the Assumptions of this Section are satisfied in Example 1 of Section 2 if appropriate moments exist and f is smooth enough and tends to zero fast enough as types tend to infinity. Similarly they are satisfied in Example 2 if it is modified so that instead of having an improper prior, θ has a Normal prior distribution.

4. The Learning Rule

The idea pursued is that players pursue a simple adaptive rule and adjust their cutoffs.

The game is played repeatedly at stages $n = 1, 2, \ldots$. At each stage each player observes their own type at that stage, $t_i(n)$, but not the type of the other player. Types are drawn at each stage from the distribution specified in Section 2 and draws are independent between stages. Each player chooses an action according to their current cutoff, $k_i(n)$. After each stage each player observes the payoffs that each action would have achieved if he had played it. They do not, however, observe the payoff they would have received if they had had another type.

In the context of the Example 2 of Section 2, the assumptions on what is observed amount to the assumption that after each stage each player observes the aggregate state, θ, and the action of the other player. The player can therefore calculate the payoff he would have achieved if he had taken a different action. Although he can observe θ he cannot calculate the payoff he would have received if he had had a different type as the other player might then also have had a different type — types may be correlated — and so taken a different action. The type of the other player is assumed not to be observed, though in some games it may be deduced from payoffs.

Formally, recall that the payoff of action a to player i depends on the state of nature, types and the action taken by each player: $y_i(a, a_j, t_0, t_1, t_2)$. It is assumed that at stage n player i observes his own type t_i and the value of y_i for different values of a. Denote the latter by $\tilde{y}_i(a, n)$ and the realised difference in payoffs between the two actions 0 and 1 by $\Delta\tilde{y}_i(n) = \tilde{y}_i(1, n) - \tilde{y}_i(0, n)$. Player i does not observe, or at least does not make direct use of, the values of t_0, t_j or a_j $(j \neq i)$ at stage n.

As noted in the Introduction, one might attempt to estimate how the expected payoff to each action depends on current type, that is the map $t_i \longmapsto E(\Delta\Pi_i|t_i)$. This is potentially complex and also unnecessary. As noted in Section 2, the experiments of Cabrales et al. (2002) and Heinemann et al. (2004) suggest that although players do not leap to equilibrium they do employ threshold strategies. Given this, all that needs to be estimated for each player is the optimal cutoff, that is the point where $E(\Delta\Pi_i|t_i) = 0$

Intuitively this is straightforward to do. One simply looks at the difference in expected payoffs at the current cutoff $k_i(n)$ — see Figure 3. If these are positive, action 1 yields a higher payoff than action 0 when the signal is $k_i(n)$, so the cutoff should be lowered, and similarly if they are negative. Of course one can only observe realised not expected payoffs but this should not matter in the long run.

More serious is the fact that the player does not control his type. He may wish to estimate payoffs at $k_i(n)$ but his realised type may be completely different. The obvious solution is to use information data on payoffs at the realised type as an estimate of those at $k_i(n)$ but since payoffs at types far from $k_i(n)$ may be far from those at $k_i(n)$ these should be given less weight. If payoffs are smooth enough then nearby points give a reasonable guide to payoffs at $k_i(n)$.

The following adjustment rule is therefore considered. At each stage player i observes the realised difference in payoffs between actions $\Delta\tilde{y}_i$. He adjusts k_i according to the rule

$$k_i(n+1) = k_i(n) - \frac{\alpha_n}{h_n} K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \tag{8}$$

Recall that $\Delta\tilde{y}_i$ denotes the realised difference in payoffs between the two actions at stage n. $K(\ )$, the kernel function, weights observations according to their distance from $k_i(n)$ in line with the intuition above.

This can also be interpreted in terms of 'similarity'. $K(\ )$ can be thought of as measuring the degree of similarity between the current observation and $k_i(n)$. Observations are weighted according to their similarity. Further discussion of the relationship between kernel estimation and similarity can be found in Gilboa and Schmeidler (2003).

α_n governs the extent of adjustment of cutoffs at each stage. If the system is to converge then this needs to become small, so that random shocks in payoffs do not always perturb it away from equilibrium.

h_n governs the extent to which more weight is put on nearby actions. Using observations on payoffs at points other than $k_i(n)$ introduces a distortion into the estimates and this needs to be eliminated if the system is to converge to the true equilibrium. Making h_n small achieves this. α_n, h_n and $K(\;)$ are assumed the same for both players but this is not important.

To understand the model, suppose that the player were only to revise his cutoff when he receives a signal equal to his current cutoff. (8) would then become

$$k_i(n+1) = k_i(n) - \mathbf{1}\left(t_i(n) = k_i(n)\right)\alpha_n \Delta\pi_i\left(t_i(n)\right) \tag{9}$$

where $\mathbf{1}$ takes the value 1 if the equality holds, zero otherwise.

If h_n is small, then almost no weight is put on observations that are not near to $k_i(n)$, so the system is close to this.

If α_n is small, the system changes slowly so it is plausible its behaviour can be well described by

$$\dot{k}_i = -f_i(k_i)\Delta\tilde{\pi}_i(k_1, k_2) \tag{10}$$

where f_i is the marginal density of t_i. $\Delta\tilde{\pi}_i(k_1, k_2)$ denotes the expected difference in payoffs between actions 0 and 1 to player i conditional on his having type k_i and the other player adopting cutoff k_j ($j \neq i$). f_i gives the probability density of receiving a signal equal to k_i. This intuition will be formalised in the next section.

The assumption that the density of f is strictly positive implies that the zeroes of this system occur precisely at the points where the conditional payoffs to each action are equal, that is at equilibrium cutoffs.

The scheme in (8) is much simpler than trying to estimate the entire dependence of payoffs on types. It requires, though, a modicum of sophistication and so players cannot be completely naïve if they are to converge to the true equilibrium. A scheme of this form is often referred to as 'passive stochastic approximation' and was first studied by Härdle and Nixdorf (1987). The adjective 'passive' is employed as, in contrast to ordinary stochastic approximation, agents cannot control the signals they observe.

Note that minus the right-hand side of (10) can also be expressed in the case $i = 1$ as

$$H_i(k_1, k_2) = \int_{-\infty}^{k_2} f(k_1, y)\Delta\pi_1(k_1, y, 0)\,dy + \int_{k_2}^{\infty} f(k_1, y)\Delta\pi_1(k_1, y, 1)\,dy \tag{11}$$

and similarly for $i = 2$.

5. Convergence Results

Section 5.1 gives conditions for convergence to equilibrium. Section 5.2 discusses the rate of convergence.

5.1 Convergence Results

In order to guarantee convergence to equilibrium some regularity conditions are required. The Kernel function K is assumed to satisfy the following standard properties:

Assumption 11 *(a) $\int K(u)\,du = 1$, (b) $\int uK(u)\,du = 0$, (c) $\int u^2K(u)\,du < \infty$, (d) K is bounded.*

Assumption 11(a) simply states that the total weight available adds up to one. (b) states that the mean weight given to observations is zero and (c) that its variance is finite. These are fairly standard assumptions in non-parametric estimation. Other variations are possible. See for example Härdle (1990) for discussion.

Using observations other than at the current cut-off is essential if it is to be adjusted (the probability of an observation exactly at the cut-off is zero). On the other hand, using observations far away from the cut-off may distort the estimates. This suggests letting h_n become smaller as n rises but not too fast, otherwise too few observations will be used for a reliable estimate. In addition, as time passes information has accumulated and so one might expect later observations to be weighted less.

The scheme considered is therefore

Assumption 12 *(a) $\sum_n \alpha_n = \infty$, (b) $\sum_n \alpha_n h_n < \infty$, (c) $\sum_n \frac{\alpha_n^2}{h_n^2} < \infty$.*

Assumption 12 is a generalisation of those found in the stochastic approximation literature. (a) guarantees that the process cannot become stuck away from the optimal cut-off because of initial bad luck. (b) and (c) express the assumption that h_n goes to zero but not too fast in comparison with α_n, which also tends to zero. They guarantee that random effects wash out in the long run. Assumption 12 is satisfied by for example by $\alpha_n = 1/n$ and $h_n = 1/n^\gamma$ with $0 < \gamma < 1/2$.

The following convergence result holds:

Theorem 1 *Under Assumptions 1–12*

(a) If the game has a unique equilibrium, play converges to it almost surely.

(b) If there are multiple equilibria, then the set of limit points of (8) is contained in the set of equilibria almost surely. If equilibria are isolated then play converges to one of them almost surely.

The proof is in the Appendix. An argument similar to that in Benaïm and Hirsch (1999) shows that the limiting behaviour of the system can be analysed using (10). The convergence results follow from the fact that the assumptions of supermodularity imply that

$$\frac{\partial H_i}{\partial k_j} < 0 \qquad j \neq i \tag{12}$$

In other words, if j raises his cutoff this lowers the relative payoff to playing action 1 rather than 0 at k_i (and so i should raise his cutoff). It follows that (10) is a cooperative differential equation (see for example Smith (1995)). The results follow from this and an argument of Benaïm (2000) (see Appendix).

Assumption 6 is used to show that the k_i remain bounded (see the proof of Theorem 1 in the Appendix). One could instead make Assumption $3'$ and assume directly that each type space is compact. In this case, though, one would need to modify the learning rule slightly to ensure that it cannot be driven outside the type space by a large random error and then become stuck there (no signals are received outside the type space). If it is assumed that players know the lowest and highest signals they can receive, then a natural modification would be to let

$$\tilde{k}_i(n+1) = k_i(n) - \frac{\alpha_n}{h_n} K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \tag{13}$$

and set

$$k_i(n+1) = \begin{cases} \tilde{k}_i(n+1) & \text{if } \tilde{k}_i(n+1) \in [\underline{t}_i, \overline{t}_i] \\ \text{nearest endpoint of } [\underline{t}_i, \overline{t}_i] & \text{otherwise} \end{cases} \tag{14}$$

This modification ensures that the process never leaves $T_1 \times T_2$. Under Assumption 6, the vector field given by (10) points inwards on the boundary on $T_1 \times T_2$. A standard argument, see for example Kushner and Yin (1997) Chapter 5 Theorem 2.1, then shows that the long-run behaviour of the process is still given by that of (10). The rest of the argument is as in Theorem 1. One therefore has:

Corollary *Theorem 1 holds if Assumption 3 is replaced by Assumption $3'$ and the learning rule is modified to (14).*

In standard stochastic approximation, one can show under some assumptions , see for example Pemantle (1990) or Brandiere and Duflo (1996), that the system cannot converge to linearly unstable equilibria. These results are not, however, directly applicable here. The extra error terms induced by using observations at nearby points means that the conditions on the noise they assume need not hold.[2] It is though plausible that the result holds.

5.2 Rate of Convergence

Even if a learning model converges to equilibrium, the relevance of equilibrium play may be limited if it does so very slowly. The rate of convergence may be measured in different ways. In this sub-section the approach, familiar in econometrics, is taken that that this is measured by the factor necessary to re-scale the deviations of observations from the equilibrium values so that they converge to a standard normal distribution.

In 2×2 games Kaniovski and Young (1995) show that stochastic fictitious play converges to a perturbed mixed equilibrium at rate $n^{1/2}$. In the current context, the

[2] Specifically, Pemantle (1990) assumes martingale difference innovations. Brandiere and Duflo (1996) allow for an additional error term which is square summable. These conditions need not hold here.

fact that players must estimate payoffs from points near their current cutoff means that a slower rate of convergence is obtained. The result below is a straightforward generalization of Theorem 2 of Härdle and Nixdorf (1987), who consider the one-dimensional case. The proof may found in the Appendix.

Denote the current vector of cutoffs by $k(n)$. $\Rightarrow$ represents convergence in distribution.

Theorem 2 *Let Assumptions 1–8, 9′, 10′ and 11 hold and let k^* be an equilibrium. Assume that $a_n = 1/n$ and $h_n = 1/n^\gamma$. If the real parts of the eigenvalues of the Jacobian matrix $J = -\left(\frac{\partial H_i}{\partial x_j}\right)$ are strictly less than $-(1-\gamma)/2$ at k^*, then conditional on the event that $k(n)$ converges to k^**

(a) if $1/2 > \gamma > 1/5$,

$$n^{(1-\gamma)/2}\left(k(n) - k^*\right) \Rightarrow N(0, \Sigma) \tag{15}$$

where Σ is the unique solution to the equation

$$\left(\frac{(1-\gamma)}{2}I + J\right)\Sigma + \Sigma\left(\frac{1-\gamma}{2}I + J\right)' = -S \tag{16}$$

with S a diagonal matrix with entries $\left(\int K^2(u)\,du\right) f_i(k_i)E\left(\Delta\tilde{y}_i^2|k_i\right)$ on the diagonal.

(b) If $\gamma = 1/5$, then the same result holds, except that the limiting multivariate distribution on the right-hand side of (15) has non-zero mean $(-J - \frac{(1-\gamma)}{2})^{-1}b$, where b has elements $-\left(\int u^2 K(u)\,du\right)\frac{1}{2}\frac{\partial^2 H_i}{\partial k_i^2}$.

The statement that the result is conditional on the event that $k(n)$ converges to k^* allows for the possibility of multiple equilibria to which the system may converge. The result states that if the system converges to k^*, it does so at rate $n^{(1-\gamma)/2}$.

The best rate of convergence is $n^{2/5}$, obtained when $\gamma = 1/5$. In this case a bias term appears in the limiting distribution. This feature is standard in non-parametric estimation (see for example Härdle (1990)). With ordinary stochastic approximation, see for example Kushner and Yin (1997), convergence would be at rate $n^{1/2}$ if the eigenvalues of the Jacobian had real part less than $-1/2$. The result of Kaniovski and Young (1995) follows from this.

The assumptions of Theorem 2 are not particularly realistic. It is not claimed that players are likely to choose the optimal sequence h_n. It merely shows that even if they do, convergence is likely to be slower than one might expect in stochastic fictitious play.

To interpret the conditions, note that if the eigenvalues of J have real parts less than zero, k^* is locally stable under the dynamic (10). The condition in the Theorem is therefore stronger than local stability. To interpret the condition further note that the right-hand side of (10) is

$$-f_i(k_i^*)\Delta\tilde{\pi}_i(k_1^*, k_2^*) \tag{17}$$

At equilibrium $\Delta\tilde{\pi}_i(k_1^*, k_2^*) = 0$, so J at k^* has elements

$$J = (J_{ij}) = \left(-f_i(k_i^*)\frac{\partial\Delta\tilde{\pi}_i}{\partial k_j}\right) \tag{18}$$

Two factors which are perhaps unexpected influence the size of the eigenvalues of J. Firstly, note that the smaller f_i, other things equal, the smaller they will be. This reflects the nature of passive stochastic approximation. If signals at or near the equilibrium point are rarely observed then cutoffs will only be adjusted rarely when the system is in the neighbourhood of equilibrium. Convergence is therefore likely to be slow. Secondly, the absolute size of payoffs matters. Considered as a normal form game in k_1 and k_2, the slope of reaction functions only depends on the ratio of the values of $\frac{\partial \Delta \tilde{\pi}_i}{\partial k_j}$ and $\frac{\partial \Delta \tilde{\pi}_i}{\partial k_i}$. If players used some version of the best response dynamic, therefore, absolute sizes of payoffs would be unimportant. In the current algorithm, however, players adjust their cutoff in response to differences in payoffs. If these are small, convergence is likely to be slow.

Other measures of the rate of convergence exist. Nazin et al. (1992) for example consider the mean squared error and show how suitable choice of Kernel can improve the rate of convergence if the underlying function is smooth enough. Such sophisticated schemes do not appear suitable for models of boundedly rational learning.

6. Extensions

This section considers extensions of the above results (a) to the case of more than two players, (b) to the case of more than two actions, (c) to other payoff functions.

6.1 More than two players

Theorem 1 extends straightforwardly to the case of more than two players, so long as there is a unique equilibrium. The Assumptions of Section 2 generalise in an obvious way to the m player case.[3]

Theorem 3 *Suppose there are $m \geq 2$ players and the corresponding versions of Assumptions 2–12 hold. If the game has a unique equilibrium, play converges to it almost surely.*

If there are multiple equilibria, it is no longer clear that play must converge to one of them. The special feature of the two player case is that the state space is two dimensional. In this case play must converge to the equilibrium set in a cooperative system. Benaïm (2000) conjectures in the context of ordinary stochastic approximation that this also holds in higher dimensions for systems satisfying (12). If this is correct, then Theorem 1 would generalise fully.

Theorem 2 generalises immediately to the current setting, as it is conditional on convergence to equilibrium.

6.2 More than two actions

The results of the paper can in principle be extended to the case when there are more than two possible actions available.

[3] In Assumptions 5–8, interpret a_j and t_j and so on as the vectors of the other players' actions with the obvious vector order.

Suppose for example that there are three actions. Under analogous assumptions to those in Section 3, there would be a range of realisations of types under which player 1 would play action 1, some action 2, others action 3. The natural way to extend the learning algorithm is to suppose that player 1 revises his cutoff between actions 1 and 2, k_1^{12}, and between actions 2 and 3, k_1^{23} in the manner of (8).

There are two difficulties. In the first place, comparisons between adjacent actions may not be enough. For example, for some strategies of the other player, it may be undesirable to ever play action 2. In this case one would switch directly from action 1 to 3 at some cutoff. To deal with this one would also need to introduce the cutoff between actions 1 and 3, k_1^{13}, as well. In principle, this is fine but if for some cutoffs of the other player it is optimal to play strategies 1,2 and 3 and others only 1 and 3, at the points where action 2 drops out there may be a lack of smoothness in payoff functions (more precisely the analogues of the H_i).

Secondly, even if the first difficulty does not arise, if (8) is applied independently to k_1^{12} and k_1^{23}, there is no guarantee that payoff realisations will ensure that $k_1^{12}(n)$ is always less than $k_1^{23}(n)$. One would need to constrain the algorithm in some way to ensure this is true. This complication perhaps makes the learning less attractive in this case.

Subject to these provisos the results could be extended. If there are more than two players, then the state space will have more than two dimensions so as in Theorem 3 convergence to equilibrium could only be guaranteed if equilibrium is unique.

6.3 Other Payoff Functions

Supermodularity is used in two ways in the analysis in previous sections: (a) to guarantee the existence of monotone equilibria, (b) to ensure convergence of the learning algorithm. There are a number of other assumptions which would deliver (a), as discussed extensively by Athey (2001)). For example, it could be assumed that payoff functions are log-supermodular in types and actions and that types are affiliated (see for example Athey (2001). The analysis of (b) does, however, rest rather strongly on supermodularity, as it guarantees that (10) is a cooperative system.

It is possible, nevertheless, to extend the result somewhat. To see this note that (10) is essentially a gradient algorithm. If one defines $\tilde{\pi}_i(k_1, k_2)$ to be the (unconditional) expected payoff to player i if he employs cutoff k_i and the other player employs cutoff k_j, then it is straightforward to check (cf. the expression for H_i in (11)), that (10) is equivalent to

$$\dot{k}_i = -\frac{\partial \tilde{\pi}_i}{\partial k_i} \tag{19}$$

The same is true if there are more than two players.

Suppose now that, for each realization of types, the players play a potential game in the sense of Monderer and Shapley (1996). That is there is a function U and functions ϕ_i such that $y_i(t, a) = U(t, a) + \phi_i(t, a_{-i})$ for all i, where a_{-i} denotes the actions of players other than i. In other words, players receive identical payoffs up to an additive factor which does not affect the relative payoffs of their actions. (19) then becomes

$$\dot{k}_i = -\frac{\partial V}{\partial k_i} \tag{20}$$

where V is the expected value of U as a function of the cutoffs. In this case, V acts as a Lyapounov function for (10) and, under some additional technical assumptions, this implies convergence of the learning algorithm to equilibrium.

In order to ensure that a potential game has an equilibrium in monotone strategies one of course needs to impose assumptions on U. As noted above, it would for example be sufficient to assume that U is log-supermodular in types and actions and that types are affiliated. More generally, it is enough to assume that the following single crossing-property, which is a weakening of Lemma 1, is satisfied. Let $E\Delta\tilde{U}(t_i)$ denote the expected incremental payoff to choosing action 1 rather than 0 when player i receives signal t_i given the strategies of the other players (suppressed for ease of notation). Recall that an increasing strategy is characterised by a cutoff.

Assumption 13 *For each i, if other players use increasing strategies, then $E\Delta\tilde{U}(t_i) \geq 0$ implies $E\Delta\tilde{U}(s_i) \geq 0$ for $s_i > t_i$ and $E\Delta\tilde{U}(t_i) > 0$ implies $E\Delta\tilde{U}(s_i) > 0$ for $s_i > t_i$.*

This Assumption, together with the others made, guarantees that an equilibrium in monotone strategies exists (see Athey (2001) Lemma 3). It also implies that any rest point of (20) is an equilibrium.

To ensure convergence V must be smooth enough. Now in the two-player case, for example, V can be written as[4]

$$\begin{aligned}&\int_{-\infty}^{k_1}\int_{-\infty}^{k_2} U(t,(0,0))f(t)\,dt_1dt_2 + \int_{k_1}^{\infty}\int_{-\infty}^{k_2} U(t,(1,0))f(t)\,dt_1dt_2\\ &+\int_{-\infty}^{k_1}\int_{k_2}^{\infty} U(t,(0,1))f(t)\,dt_1dt_2 + \int_{k_1}^{\infty}\int_{k_2}^{\infty} U(t,(1,1))f(t)\,dt_1dt_2\end{aligned} \tag{21}$$

Smoothness follows from standard results on integrals. V will for example be C^2 if $U(t,a)f(t)$ is continuous in t for each a and $\frac{\partial(U(t,a)f(t))}{\partial t_i}$ is dominated by an integrable function of t for each i and a (see for example Billingsley (1995) equation 17.5 and Theorem 16.8).

Theorem 4 *Suppose there are $m \geq 2$ players and the corresponding versions of Assumptions 2–6 and 9–13 hold and that V is C^m.*

(a) If the game has a unique equilibrium, play converges to it almost surely.

(b) If there are multiple equilibria, then the set of limit points of (8) is contained in the set of equilibria almost surely. If equilibria are isolated then play converges to one of them almost surely.

The proof can be found in the Appendix.[5] The existence of a smooth Lyapounov function means that the restriction on the dimension state space required when payoffs are supermodular for (b) to hold is not needed.

[4]With a slight abuse of notation as t_0 is implicitly integrated out.

[5]If V is C^2, H_i is C^1 but the rest of Assumption 9 is not redundant.

7. Conclusion

This paper has considered a simple set of learning rules in Bayesian games with binary actions and has shown convergence under some conditions. This provides some support for the view that equilibrium will be played in such games. On the other hand even in this simple setting players require a modicum of sophistication to learn how to play. Whether they are likely to be able to do so in more complex environments is a question for future research.

Appendix

Proof of Theorem 1

The proof of Theorem 1 proceeds in three stages. First some preliminary properties of the disturbances are established. Secondly, it is shown that the system almost surely remains bounded. Finally, it is shown that given this it must converge to an equilibrium.

Step 1: Preliminary Properties

In what follows i denotes a particular player and j the other player. (8) can be re-written as

$$k_i(n+1) = k_i(n) - \alpha_n H_i + \alpha_n v_i(n) + \alpha_n u_i(n) \tag{22}$$

where

$$v_i(n) = H_i - \int \frac{1}{h_n} K\left(\frac{t_i - k_i(n)}{h_n}\right) H_i(t_i, k_j(n))\, dt_i \tag{23}$$

and

$$u_i(n) = \int \frac{1}{h_n} K\left(\frac{t_i - k_i(n)}{h_n}\right) H_i(t_i, k_j(n))\, dt_i - \frac{1}{h_n} K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \tag{24}$$

$\Delta\tilde{y}_i(n)$ is dominated by the maximum of $\Delta y_i(0, t_0, t_1, t_2)$ and $\Delta y_i(1, t_0, t_1, t_2)$ at time n. The latter are identically and independently distributed (between periods) and have finite first and second moments by Assumption 10. Furthermore by Assumption 11, K is bounded. It follows that $u_i(n)$ has finite expectation for each n and i. Let $\mathcal{F}_{n-1}$ denote the sigma-field generated by events up to and including time $n-1$. It follows that

$$E\left(u_i(n) | \mathcal{F}_{n-1}\right) = 0 \tag{25}$$

and

$$E\left(u_i(n)^2 | \mathcal{F}_{n-1}\right) \leq \frac{C}{h_n^2} \tag{26}$$

It follows, using Assumption 12 on h_n and α_n, that $\sum_n \alpha_n u_i(n)$ is an L^2-bounded martingale and so almost surely

$$\sum_n \alpha_n u_i(n) \qquad \text{converges} \tag{27}$$

Also it follows directly from (26) that

$$\sum_n \alpha_n^2 E\left(u_i(n)^2 | \mathcal{F}_{n-1}\right) < \infty \tag{28}$$

Next Assumption 9 implies that H_i satisfies a Lipschitz condition. A simple change of variables in (23) using this and the properties of K in Assumption 11 shows that for some constant L_i,

$$|v_i(n)| \leq h_n L_i \tag{29}$$

and so

$$\sum_n \alpha_n v_i(n) < \infty \qquad and \qquad \sum_n \alpha_n^2 E\left(v_i(n)^2|\mathcal{F}_{n-1}\right) < \infty \tag{30}$$

Step 2: Boundedness

Let $w_i(n) = u_i(n) + v_i(n)$. By (25), (26) and (29), $E\left(w_i(n)|\mathcal{F}_{n-1}\right) = O(h_n)$ and $E\left(\alpha_n^2 w_i(n)^2|\mathcal{F}_{n-1}\right) = O(\frac{\alpha_n^2}{h_n^2} + \alpha_n h_n)$ (using $\alpha_n^2 h_n^2 \leq \alpha_n h_n$ for large enough n from Assumption 12). $O(g)$ is used to mean that a random variable is of order at most g almost surely. Using (29), the boundedness of H_i from Assumption 9 and the inequality $ab \leq (a^2+b^2)/2$

$$E\left(k_i(n+1)^2|\mathcal{F}_n\right) \leq (1+\beta_n)k_i(n)^2 - 2\alpha_n k_i(n)H_i + \gamma_n \tag{31}$$

where $\beta_n = O(\alpha_n^2 + \alpha_n h_n)$ and $\gamma_n = O(\alpha_n^2/h_n^2 + \alpha_n^2 + \alpha_n h_n)$. Hence $\sum_n \beta_n < \infty$ and $\sum_n \gamma_n < \infty$ almost surely. Now from Assumptions 2 and 6, $k_i(n)H_i > 0$ if $|k_i| \geq \tau_i$, where $\tau_i = \max\{|\underline{\tau}_i|, |\overline{\tau}_i|\}$. It follows from the Robbins-Siegmund Lemma on non-negative supermartingales (see for example Duflo (1997) Theorem 1.3.12), that if $|k_i(n)| \geq \tau_i$ for all n, then $k_i(n)$ and $\sum_n k_i H_i$ converge almost surely. If $\lim_n |k_i(n)| \geq \tau_i$, then this contradicts convergence of $\sum_n k_i H_i$. For recall that $H_i = f_i(k_i)\Delta\tilde{\pi}_i(k_1,k_2)$. It follows from Assumptions 4 and 6 that if $\lim_n |k_i(n)| \geq \tau_i$, $k_i H_i$ remains bounded away from zero, a contradiction. It follows that k_i must enter the region $[-\tau_i, \tau_i]$. Repeating the argument on exit from this region shows it must enter it infinitely often.

A standard argument now shows that for some R_i, $z_i(n)$ lies in $[-R_i, R_i]$ for all but finitely n almost surely. Pick $R_i > \tau_i$. It was shown in the previous paragraph that the region $[-\tau_i, \tau_i]$ is entered infinitely often. On the other hand as shown above, $\sum_n \alpha_n u_i(n)$ and $\sum_n \alpha_i v_i(n)$ converge. Also from Assumption 9, $k_i H_i$ is bounded on $|k_i| \leq \tau_i$. On $\tau_i \leq |k_i| \leq R_i$, $k_i H_i > 0$. It follows from (22) that $k_i(n)^2$ cannot be driven outside the region $k_i^2 \leq R_i^2$ infinitely often. For details see Duflo (1996) Theorem 3.II.2 or Delyon (1996) Theorem 1.

Applying this argument for each i shows that almost surely the system lies in the compact set $[-R_1, R_1] \times [-R_2, R_2]$ for all but finitely many stages.

Step 3: Convergence

Step 2 shows that almost surely $k(n) = (k_1(n), k_2(n))$ is bounded. Also from (27) and (30), $\sum_n \alpha_n u_i(n)$ and $\sum_n \alpha_n v_i(n)$ converge almost surely. Theorems 4.1, 5.3 and and Corollary 6.11 of Benaïm (1999) show that almost surely the set of limit points of $k(n)$ is (a) a non-empty, compact, connected set invariant under the flow of (10), (b) is contained in the chain recurrent set of (10). For the definition of chain recurrence see Benaïm (1999).

The remarks after the statement of Theorem 1 show that (10) is cooperative and irreducible. The argument of Theorem 4.3 of Benaïm (2000) shows that as the state space is two-dimensional, the chain recurrent set of (10) consists solely of equilibria (see also Hirsch (1999). In Benaïm (2000) it is assumed that the vector field is C^2 but it only need be C^1 for this part of his results. The C^2 assumption is used to show non-convergence to unstable equilibria.) This proves the result.

Proof of Theorem 2

The argument follows the proof of Theorem 2 of Härdle and Nixdorf (1987), which treats the one-dimensional case, closely. That convergence can be considered conditional on the event that $k(n)$ converges to k^* follows from a standard argument (see for example Theorem 4.III.5 of Duflo (1996) and its discussion). In vector notation one can write (22) as

$$k(n+1)-k^* = (1 - B(n)/n)\,(k(n) - k^*) + n^{-(1+(1-\gamma))/2}\upsilon(n) + n^{-1-(1-\gamma)/2}\sigma(n) \quad (32)$$

where

$$\upsilon_i(n) = h_n^{-1/2} E\left\{ K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \middle| \mathcal{F}_{n-1} \right\} - h_n^{-1/2} K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \quad (33)$$

and

$$\sigma_i(n) = n^{(1-\gamma)/2} \left\{ H_i - \frac{1}{h_n} E\left[K\left(\frac{t_i(n) - k_i(n)}{h_n}\right) \Delta\tilde{y}_i(n) \middle| \mathcal{F}_{n-1} \right] \right\} \quad (34)$$

and $\{B(n)\}$ is a sequence of matrices converging to $\left(\frac{\partial H_i}{\partial k_j}\right)$ such that $B(n)\,(k(n) - k^*) = H(k(n))$. Such a sequence exists if $k(n)$ converges to k^* since H is differentiable at k^*. Note that there is a typo in Härdle and Nixdorf (1987) and the power of n multiplying υ_n is as here not in the corresponding term there.

Using Assumptions $9'$ and 11 it is straightforward to show that $\sigma_i(n)$ converges to zero if $\gamma > 1/5$ and that if $\gamma = 1/5$ it converges to a vector with elements $-\frac{1}{2}\frac{\partial^2 H_i}{\partial k_i^2}(\int u^2 K(u)\,du)$. The sign is the opposite one to that given in Härdle and Nixdorf (1987) but it is easily checked that this is correct. On the other hand, $E(\upsilon(n)|\mathcal{F}_{n-1}) = 0$ and it follows from Assumptions $10'$ and 11 that $E(\upsilon_1(n)\upsilon_2(n)|\mathcal{F}_{n-1})$ converges to zero and $E(\upsilon_i^2(n)|\mathcal{F}_{n-1})$ converges to $f_i(k_i)E(\Delta\tilde{y}_i^2|k_i)(\int K^2(u)\,du)$. As in Härdle and Nixdorf (1987) one can verify the other moment conditions necessary to appeal to Remark 2.8 of Berger (1986), which yields the result.

Proof of Theorem 4

The argument is exactly as in Theorem 1 except that in the last paragraph of Step 3 one appeals to the argument of Corollary 6.7 of Benaïm (1999) to show that under the smoothness assumption made on V, the chain recurrent consists solely of equilibria (note that m is the dimension of the state space).

References

Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69:861–889.

Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. In *Seminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 1–68. Springer Verlag, Berlin.

Benaïm, M. (2000). Convergence with probability one of stochastic approximation algorithms whose average is cooperative. *Nonlinearity*, 13:601–616.

Benaïm, M. and Hirsch, M. (1999). Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, 29:36–72.

Berger, E. (1986). Asymptotic behaviour of a class of stochastic approximation procedures. *Probability Theory and Related Fields*, 71:517–22.

Billingsley, P. (1995). *Probability and Measure*. John Wiley and Sons, New York, third edition.

Brandiere, O. and Duflo, M. (1996). Les algorithmes stochastiques contournent-ils les pieges. *Ann. Inst. Henri Poincaré*, 32:395–427.

Bray, M. (1982). Learning, estimation and the stability of rational expectations. *Journal of Economic Theory*, 26:318–339.

Cabrales, A., Nagel, R., and Armenter, R. (2002). Equilibrium selection through incomplete information in coordination games: An experimental study. mimeo, Universitat Pompeu Fabra.

Carlsson, H. and van Damme, E. (1993). Global games and equilibrium selection. *Econometrica*, 61:989–1018.

Dekel, M., Fudenberg, D., and Levine, D. (2004). Learning to play bayesian games. *Games and Economic Behavior*, 46:282–303.

Delyon, B. (1996). General results on the convergence of stochastic algorithms. *IEEE Transactions on Automatic Control*, 41:1245–55.

Duflo, M. (1996). *Algorithmes Stochastiques*. Springer Verlag, Berlin.

Duflo, M. (1997). *Random Iterative Models*. Springer Verlag, Berlin.

Fudenberg, D. and Kreps, D. (1993). Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367.

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

Gilboa, I. and Schmeidler, D. (2003). Inductive inference: An axiomatic approach. *Econometrica*, 71:1–26.

Härdle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, Cambridge.

Härdle, W. and Nixdorf, R. (1987). Nonparametric sequential estimation of zeroes and extrema of regression functions. *IEEE Transactions in Information Theory*, 33:367–372.

Heinemann, F., Nagel, R., and Ockenfels, P. (2004). The theory of global games on test: Experimental analysis of coordination games with public and private information. *Econometrica*, 72:1583—99.

Hellwig, C. (2003). Public information, private information and the multiplicity of equilibria in coordination games. *Journal of Economic Theory*, 107:191–222.

Hirsch, M. (1999). Chain transitive sets for smooth strongly monotone dynamical systems. *Dynamics of Continuous, Discrete and Impulsive Systems*, 5:529–543.

Kaniovski, Y. and Young, P. (1995). Learning dynamics with stochastic perturbations. *Games and Economic Behavior*, 11:330–363.

Kushner, H. and Yin, G. (1997). *Stochastic Approximation Algorithms and Applications.* Springer Verlag, New York.

Milgrom, P. and Roberts, J. (1990). Rationalizabilty, learning and equilibrium in games with strategic complementarities. *Econometrica*, 58:1255–1277.

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

Morris, S. and Shin, H. (2003). Global games: Theory and applications. In Dewatripont, M., Hansen, L., and Turnovsky, S., editors, *Advances in Economics and Econometrics*, Cambridge. Cambridge University Press.

Nazin, A., Polyak, B., and Tsybakov, A. (1992). Optimal and robust algorithms for passive stochastic approximation. *IEEE Transactions on Information Theory*, 38:1577–83.

Palfrey, T. and Rosenthal, H. (1994). Repeated play, cooperation and coordination: An experimental study. *Review of Economic Studies*, 61:545–65.

Pemantle, R. (1990). Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability*, 18:698–712.

Selten, R. and Buchta, J. (1998). Experimental sealed bid first price auction with directly observed bid functions. In Budescu, I., Erev, I., and Zwick, R., editors, *Games and Human Behavior, Essays in Honor of Amnon Rapoport*, pages 79–102. Lawrence Erlbaum Associates, Mahwah, NJ.

Smith, H. (1995). *Monotone Dynamical Systems.* American Mathematical Society, Providence, RI.

Topkis, D. (1998). *Supermodularity and Complementarity.* Princeton University Press, Princeton, NJ.

Figures

	0	1
0	$1-t_1, 1-t_2$	$-t_1, 0$
1	$0, -t_2$	$0, 0$

Figure 1

	0	1
0	$0, 0$	$0, \theta-1$
1	$\theta-1, 0$	θ, θ

Figure 2

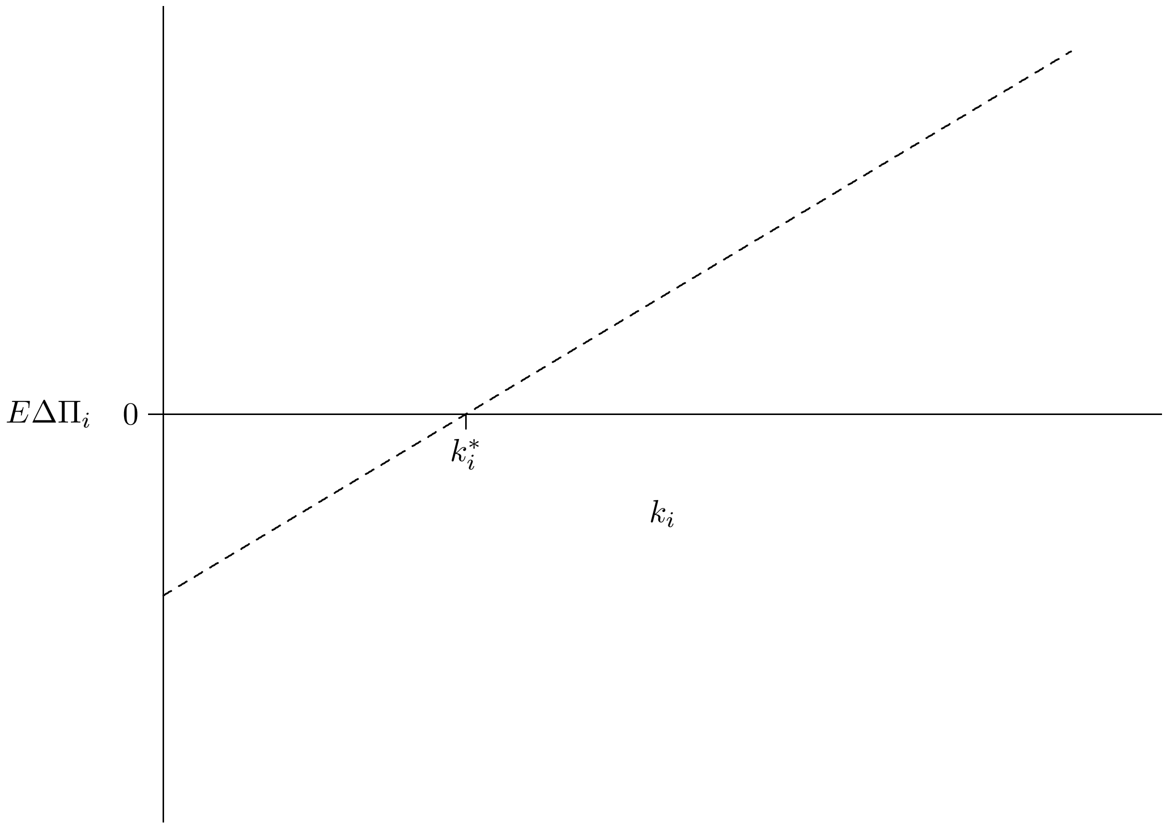

Figure 3