

*ISSN 1471-0498*



**DEPARTMENT OF ECONOMICS  
DISCUSSION PAPER SERIES**

**LEARNING IN MONOTONE BAYESIAN GAMES**

**A.W. Beggs**

**Number 737  
January 2015**

**Manor Road Building, Manor Road, Oxford OX1 3UQ**

# Learning in Monotone Bayesian Games

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

`alan.beggs@economics.ox.ac.uk`

Current Version: December 2014

## Abstract

This paper studies learning in monotone Bayesian games with one-dimensional types and finitely many actions. Players switch between actions at a set of thresholds. A learning algorithm under which players adjust their strategies in the direction of better ones using payoffs received at similar signals to their current thresholds is examined. Convergence to equilibrium is shown in the case of supermodular games and potential games.

*Keywords:* bayesian games, monotone strategies, learning, stochastic approximation, supermodular games

*JEL Nos:* C72, D83

---

The model of normal-form learning in Section 5 originally appeared in a 2011 working paper ‘Regularity and Stability in Monotone Bayesian Games’, which was principally concerned with other issues, but has been separated out from it to form part of this paper focussed on learning.

# 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. Beggs (2009) proposed a simple model of learning but limited attention to games with binary actions. This paper extends the analysis to games with many actions. It uses recent results of Benaïm and Faure (2012) to show that in supermodular games play converges to equilibrium under weak assumptions. Convergence is also shown in potential games.

In a Bayesian game each player receives a private signal before choosing between actions. This signal may convey information about his and other players' payoffs and therefore indirectly about their likely choice of action. Learning in such a game is in principle very complex. This paper shows that convergence to equilibrium can be achieved by a simple adaptive rule.

The paper studies learning where the actions players take are ordered and are increasing functions of their signal, which is assumed one-dimensional. It follows that a player's strategy can be represented by a collection of thresholds at which he switches between actions. It is assumed that each player adjusts their thresholds in the direction suggested by past play. It is shown that this can be achieved by a simple scheme.

The scheme proposed is a generalization to many actions of that suggested in the case of two actions by Beggs (2009). As noted in that paper the rule can be seen as encapsulating two ideas. Firstly, it can be seen as example 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. Secondly, it incorporates ideas of '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 closeness to the signal at which payoffs are to be estimated.

The main technical tool of the paper is stochastic approximation. The learning rule studied is an example of passive stochastic approximation, introduced

by Härdle and Nixdorf (1987). Recent results of Benaïm and Faure (2012) on non-convergence to unstable sets in perturbed cooperative systems, adapted to apply to the current environment, play a key role in the analysis.

Convergence to equilibrium is shown in supermodular games satisfying a concavity property. Convergence is also shown in potential games.

The literature on learning in Bayesian games is fairly small. The literature on learning in stochastic fictitious play considers learning in normal form games where players's payoffs are subject to random perturbations in the spirit of Harsanyi — see for example Fudenberg and Kreps (1993) or Benaïm and Hirsch (1999). These perturbed games can be regarded as Bayesian games but it is assumed that the perturbations are independent, which makes learning much more straightforward. Benaïm and Hirsch (1999) allow for correlation in perturbations but assume that players ignore it when choosing actions so the system does not converge to a true Bayesian equilibrium.

Steiner and Stewart (2007) consider a model similar in spirit to that of Beggs (2009). They also employ a rule under which players used payoffs at nearby signals are used to estimate the signals at the current one but do not employ a threshold rule. They restrict attention to global games. Convergence is however to a perturbed version of equilibrium rather than a true Bayesian equilibrium as in the current paper. The published version Steiner and Stewart (2008) differs slightly in that the game is drawn randomly each period but players both observe perfectly the game.

Section 2 lays out the general framework. Section 3 states preliminary results characterizing equilibrium conditions in the static game. In particular it shows that the equilibria of the game can be represented as the solution to a variational inequality. Section 4 gives examples of the environment studied. Section 5 presents a model of normal-form learning to motivate the principal model of extensive form learning outlined in Section 6. Section 7 contains the main results on convergence in supermodular and potential games. Section 8 concludes.

## 2. The Static Game

This section describes the underlying static game. Attention is restricted to pure strategy Bayesian equilibria.

There are  $p$  players and each observes a signal  $t_i$ ,  $i = 1, \dots, p$ , drawn from a set  $T_i$ . Payoffs may also depend on the state of nature which belongs to the set  $T_0$ . Let  $T = T_0 \times T_1 \times \dots \times T_p$ . After observing his signal each player then takes an action  $a_i$  drawn from a set  $A_i$ . Let  $A = A_1 \times \dots \times A_p$ . Nature takes no action.

**Assumption 1**  $T_i$  is the interval  $[0, 1]$  for all  $i$ .

**Assumption 2** The signals have a joint density  $f$  on  $T$  with respect to Lebesgue measure which is continuous in  $t$  and strictly positive.

The assumption that all players share the same type space is without loss of generality and simplifies the notation.  $F$  denotes the cumulative distribution function of  $f$ .

**Assumption 3**  $A_i$  is a strictly ordered finite set with  $m(i) + 1$  elements,  $0, \dots, m(i)$ , for  $i = 1, \dots, p$ .

Each player  $i$  has a bounded measurable payoff function  $U_i(a, t)$  where  $a \in A$  and  $t \in T$ .

**Assumption 4**  $U_i(a, t)$  is jointly continuous in  $t$  for each  $a \in A$ ,  $i = 1, \dots, p$ .

For the convergence analysis stronger assumptions will be needed

**Assumption 5**  $U_i$  is continuously differentiable in  $t$  for each  $i$  and  $a$ .  $f$  is continuously differentiable.

**Assumption 6**  $U_i$  is twice continuously differentiable in  $t$  for each  $i$  and  $a$  and  $f$  is twice continuously differentiable.

A strategy for player  $i$  is a measurable function  $\sigma_i : T_i \rightarrow A_i$ . It is said to be increasing if  $\sigma_i(t'_i) \geq \sigma_i(t_i)$  for all  $t_i$  and  $t'_i$  with  $t'_i \geq t_i$ .

Let  $F(t_{-i} | t_i)$  denote be the conditional distribution of signals in  $T_{-i} = \prod_{j \neq i} T_j$  given  $t_i$ ,  $f(t_{-i} | t_i)$  be their density and  $f_i(t_i)$  be the marginal density of  $t_i$ . If opponents use a measurable strategy  $\sigma_j : T_j \rightarrow A_j$ ,  $j \neq i$  then  $i$ 's interim

expected payoff conditional on receiving signal  $i$  and taking action  $a_i$  is, denoting by  $\sigma_{-i}$  the vector of opponents strategies,

$$V_i(a_i, t_i; \sigma_{-i}) = \int_{t_{-i}} U_i(a_i, \sigma_{-i}(t_{-i}), t) dF(t_{-i} | t_i)$$

$\sigma_{-i}$  will often be suppressed from the notation.

Player  $i$ 's ex ante expected payoff from using strategy  $i$  is

$$W_i(\sigma_i, \sigma_{-i}) = \int_T U_i(\sigma_i(t_i), \sigma_{-i}(t_{-i}), t) dF = \int_{t_i} V_i(\sigma(t_i), t_i; \sigma_{-i}) f_i(t_i) dt_i$$

A set of strategies  $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*)$  is an equilibrium if for each  $i$ ,  $W_i(\sigma_i^*, \sigma_{-i}^*) \geq W_i(\sigma_i, \sigma_{-i}^*)$  for all strategies  $\sigma_i$  of player  $i$ .

Recall that a real-valued function  $g(x, t)$  of two variables on two ordered sets satisfies *single crossing* in  $(x, t)$  if

$$g(x', t) \geq (>) g(x, t) \Rightarrow g(x', t') \geq (>) g(x, t')$$

for all  $x' > x$  and  $t' > t$

**Assumption 7** *If all players apart from  $i$ ,  $i = 1, \dots, n$ , use increasing strategies  $\sigma_j: T_j \rightarrow A_j$  then  $V_i$  satisfies single crossing in  $(a_i, t_i)$ .*

If Assumption 7 holds then if all players apart from  $i$  use an increasing strategy,  $i$  has a best reply which is increasing (see Athey (2001) and Milgrom and Shannon (1994)). Existence of equilibrium under the assumptions above follows from the general results of Athey (2001).

Recall that a function  $\phi(x, y)$  is (strictly) supermodular in variables  $x$  and  $y$  if it has (strictly) increasing differences, that is

$$\phi(x', y') - \phi(x, y') (>) \geq \phi(x', y) - \phi(x, y)$$

if  $x' \geq x$  and  $y' \geq y$ . In other words, the gain to increasing  $x$  is increasing in  $y$ . A function is log-supermodular if its log is supermodular. A set of random variables will be said to be *affiliated* if they have a joint density which is log-supermodular.

As noted in Athey (2001) two important sufficient conditions for Assumption 7 are

**Definition 1** A game is (strictly) supermodular with increasing beliefs if for all  $i$ , (i)  $U_i(a, t)$  is (strictly) supermodular in  $a$  and supermodular in  $(a_i, t_j)$ ,  $j = 1, \dots, n$  and (ii)  $\int_S dF(t_{-i}|t_i) dt_i$  is increasing for all sets whose indicator functions are increasing.

Affiliation is sufficient for (ii), see Athey (1996) and Milgrom and Weber (1982).

**Definition 2** A game is log-supermodular with affiliated beliefs if (i) for all  $i$ ,  $U_i(a, t)$  is non-negative log-supermodular in  $(a, t)$  and (ii) types are affiliated.

If payoffs are supermodular then the gain to increasing action for player  $i$  increases as the actions of the other players increase and signals increase.

If all other players use increasing strategies under Definition 1  $V_i(a_i, t_i)$  is supermodular in  $(a_i, t_i)$  and under Definition 2  $V_i(a_i, t_i)$  is log-supermodular, so Assumption 7 is satisfied.

### 3. Equilibrium Conditions

This section shows that the conditions for equilibrium of the static game can be written as the solution to a variational inequality describing the first-order conditions for each agent. This representation is all that is required for what follows but some interpretation is given for completeness. The latter is not essential for the following sections and may be skipped. These characterizations first appeared in Beggs (2011), where a more detailed discussion can be found.

As described by Athey (2001) a monotone strategy be specified by a vector of cutoffs. For player  $i$ ,  $k_j^i$  denotes the signal at which he first plays action  $j$ , that is for  $1 \leq j \leq m(i)$ ,  $k_j^i = \sup\{t_i \in [0, 1] \mid \sigma_i(t_i) < j\}$ . The dummy cutoffs  $k_0^i \equiv 0$  and  $k_{m(i)+1}^i \equiv 1$  are sometimes convenient for notational purposes. The strategy set for player  $i$  is the set

$$\Sigma_i \equiv \{k^i \mid 0 \leq k_1^i \leq k_2^i \leq \dots \leq k_{m(i)}^i \leq 1\}$$

Let  $\Sigma = \Sigma_1 \times \dots \times \Sigma_p$ .

Note that player 1's payoff conditional on receiving signal  $t_1$  and taking action  $a_1$ ,  $V_1$  as defined in Section 2, can be written as:

$$V_1(a_1, t_1) = \sum_{l_2=0}^{m(2)} \cdots \sum_{l_p=0}^{m(p)} \int_{t_2=k_{l_2}^2}^{k_{l_2+1}^2} \cdots \int_{t_p=k_{l_p}^p}^{k_{l_p+1}^p} U_i(a_1, a_{l_2}, \dots, a_{l_p}, t) dF(t_{-i} | t_i) \quad (1)$$

and similarly for other actions and players.

For any pair of actions  $j$  and  $l$  let  $\Delta U_i(j, l, a_{-i}, t) = U_i(l, a_{-i}, t) - U_i(j, a_{-i}, t)$  be the payoff in switching from  $j$  to  $l$  and let  $\Delta V_i(j, l, t_i)$  denote the expected payoff to switching from action  $j$  to  $l$  conditional upon receiving signal  $t_i$ :

$$\Delta V_i(j, l, t_i) = \int_{t_{-i}} \Delta U_i(j, l, \sigma_{-i}(t_{-i}), t) dF(t_{-i} | t_i) \quad (2)$$

A player's ex ante payoff function as a function of his cutoffs can be written, given the strategies of the other players  $k^{-i}$ , as

$$W_i(k^i; k^{-i}) = \sum_{j=0}^{m(i)} \int_{k_j^i}^{k_{j+1}^i} V_i(a_j, t_i) f_i(t_i) dt_i \quad (3)$$

Under Assumptions 1–4  $W_i$  is continuously differentiable in  $k^i$  with

$$\frac{\partial W_i}{\partial k_j^i} = -f_i(k_j^i) \Delta V_i(j-1, j, k_j^i) \quad (4)$$

Denote the gradient vector of  $W_i$  with respect to  $k^i$  by  $\nabla W_i$ .

Recall that for a closed convex set  $X$ , the **Normal cone** at  $x \in X$  is the set  $N_X(x) = \{y : \langle y, z - x \rangle \leq 0 \quad \forall z \in X\}$ . If  $\psi$  is a function,  $x \in X$  is said to satisfy the **Variational inequality**  $VI(\psi, X)$  if  $-\psi(x) \in N_X(x)$ .<sup>1</sup>

The first-order necessary conditions for  $k^i$  to maximize  $W_i$  can be written as

$$\nabla W_i(k^i) \in N_{\Sigma_i}(k^i) \quad (5)$$

That is any direction of increase of  $W_i$  must point outside  $\Sigma_i$  (see Borwein and Lewis (2006) Proposition 2.1). Equivalently  $k^i$  solves the variational inequality  $VI(-\nabla W_i, \Sigma_i)$  given  $k^{-i}$ .

Since the normal cone to a product set is the product of the normal cones, the first-order necessary conditions for equilibrium are, letting  $\nabla W = (\nabla W^1, \dots, \nabla W^p)$ ,

$$\nabla W(k^*) \in N_{\Sigma}(k^*) \quad (6)$$

---

<sup>1</sup>See Facchinei and Pang (2003) for an extensive treatment of variational inequalities.

In other words  $k^*$  solves the variational inequality  $VI(-\nabla W, \Sigma)$ . Note that  $\nabla W$  is simply the product of the vectors  $\nabla W_i$  and not a true derivative vector.

**Lemma 1** *Under Assumptions 1–4 and 7,  $k^i$  maximizes  $W_i$  given  $k^{-i}$  if and only if (5) holds.  $k^* = (k^{*1}, \dots, k^{*P})$  is an equilibrium vector of cutoffs if and only if (6) holds.*

In other words, the first-order necessary conditions for optimality are also sufficient under Assumption 7. This will follow from the equivalent characterization given below. In the case when payoffs are supermodular, sufficiency is clear if the marginal density of signals is uniform, since then (4)  $W_i$  is concave in  $k^i$  as  $\Delta V_i(j, j-1, t_i)$  is increasing in  $t_i$ . By a suitable transformation it can always be assumed that the marginal distribution of signals is uniform (see for example Nelsen (2006)). In general, Assumption 7 implies that the derivative of  $W_i$  with respect to each cutoff changes sign only once and this turns out to be enough for sufficiency.

An equivalent but more intuitive characterization can be obtained by noting that in order to maximise  $W_i$ , player  $i$ 's action must maximise his interim payoff  $V_i(a, t)$  for almost every signal  $t_i$ .

If an action is played over an interval of positive measure then it will be said to be **active**, otherwise it is inactive. That is, action  $j$  is active if and only if  $k_{j+1}^i - k_j^i > 0$ . Call two active actions **adjacent** if there is no active action lying between them.

Under Assumptions 1–4,  $\Delta V_i$  is continuous in  $k^i$ , and Assumption 7 yields the following simple characterization of best replies:

**Lemma 2** *If other players use increasing strategies then under Assumptions 1–4 and 7 a set of cutoffs for  $k^i$  for player  $i$  correspond to a strategy which is an optimal response if and only if:*

(a) *If  $j$  and  $l$  are adjacent active actions then*

(i)

$$\Delta V_i(j, l, k_l^i) = 0$$

(ii) *For all  $j'$ , with  $j \leq j' < l$ ,*

$$\Delta V_i(j, j', k_l^i) \leq 0$$

(b) If  $j$  is the least active action then for all  $l \leq j$

$$\Delta V_i(l, j, k_0^i) \geq 0 \quad (7)$$

(c) If  $j$  is the greatest active action then for all  $l \geq j$

$$\Delta V_i(j, l, k_{m(i)+1}^i) \leq 0 \quad (8)$$

These conditions are equivalent to (5).

Condition (a)(i) simply expresses the fact that the player must be indifferent at the switch point between two active actions. Condition (a)(ii), (b) and (c) rule out payoffs being raised by using a currently inactive action. Note that if an action is not used then an active action must come into use at exactly the same cutoff.

The sufficiency of these conditions follows directly from single crossing (Assumption 7). If  $j < l$  are indifferent at  $k_j^i$  then by single crossing  $j$  must be (weakly) better for all lower signals and  $l$  weakly better for all higher signals. Iterating this argument yields that  $j$  is (weakly) better than all active actions in the range in which it is active. A similar argument implies that  $l$  is better than any inactive action in this range as well. The equivalence to (5) is shown in the Appendix.

## 4. Examples

This section presents three examples of the environment studied.

### Example 1

	0	1
0	0, 0	$0, -2 + 3t_2$
1	$-2 + 3t_1, 0$	$-1 + 3t_1, -1 + 3t_2$

**Figure 1**

Consider the coordination game in Figure 1.  $t_1$  and  $t_2$  will be referred to as the types or signals of players 1 and 2 respectively. Players privately observe

their own type before choosing actions. Note that action 0 is dominant for low values of  $t_i$  and action 1 for high values.

$t_1$  and  $t_2$  take values between 0 and 1 and are drawn from a distribution with joint distribution function  $F(t_1, t_2)$  and density function  $f$ .

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  $t_i$  for every  $t_j$ , that is the higher a player's own type the higher the other player's type is likely to be, then there is at least one interior equilibrium of the form:

$$\begin{cases} \text{Player } i \text{ plays 0} & \text{if } t_i < k^i \\ \text{Player } i \text{ plays 1} & \text{if } t_i > k^i \end{cases} \quad (9)$$

$k^i$  is determined by the pair of equations

$$3k^i = 1 + F_j(k^j|k^i) \quad j \neq i, \quad i = 1, 2 \quad (10)$$

It is natural to ask whether players can learn to play such an equilibrium. After the game has been played once, if each player can observe the action of the other he can compute the ex post difference in payoffs to actions 0 and 1. In the case of player  $i$ , this will be

$$\begin{cases} 3t_i - 2 & \text{if player } j \text{ played 0} \\ 3t_i - 1 & \text{if not} \end{cases} \quad (11)$$

At the cutoff,  $k^i$ , the difference in realized payoffs should have expectation zero ex ante. This suggests raising the cutoff if action 0 has higher realized payoff than action 1. Since the observed signal,  $t_i$ , may not equal the equilibrium cutoff, this suggests revising the cutoff less if  $t_i$  is far from  $k^i$  as behavior at  $t_i$  may be a poor guide to behavior at  $k^i$ . Beggs (2009) analyzed a formal model on these lines in the case of binary actions. A general model is presented in Section 7 of this paper.

Two features of this example are worthy of note: firstly that  $F_j(t_j|t_i)$  is decreasing in  $t_i$ , secondly that the gain to playing action 1 rather than action 0 is increasing in own type and the action played by the other player, that is payoffs are supermodular. These assumptions guarantee that there is an equilibrium in which player's actions are increasing in their type and so their strategies can

be characterized by a cutoff. They are will also be of use in the convergence analysis.

**Example 2** ( *A Coordination Game* )

One can also consider games with more than two actions:

	0	1	2
0	6, 6	5, $5 + \theta_2$	3, $3 + 2\theta_2$
1	$5 + \theta_1$ , 5	$7 + \theta_1$ , $7 + \theta_2$	$6 + \theta_1$ , $6 + 2\theta_2$
2	$3 + 2\theta_1$ , 3	$6 + 2\theta_1$ , $6 + \theta_2$	$8 + 2\theta_1$ , $8 + 2\theta_2$

**Figure 2**

Each player receives a signal  $t_i$ , which has a positive continuous density on  $[0, 1]$  and is independent of the other player's signal.  $\theta_i = \lambda \left( t_i - \frac{1}{2} \right)$ , where  $\lambda$  is large positive parameter. Note that as in the previous example the gain from switching from a lower action to a higher action is increasing in the other player's action and the player's own signal. Signals are independent, so beliefs about the other player's type are unaffected by the signal. Again there will be an equilibrium in which actions are increasing functions of the players' signals.

It is straightforward to see that, for large enough  $\lambda$ , each player will choose action 0 if their signal is low, action 1 for medium values and action 2 if it is high. As in the previous example, equilibrium can be specified by a thresholds for each player: at  $k_1^i$  player  $i$  switches from action 0 to action 1 and at  $k_2^i$  player  $i$  switches from action 1 to 2.

In a general game, not all actions need be played. For example a player might play action 0 for low signals and then switch directly to action 2 rather than playing action 1 for intermediate values. The structure of the game here rules this out. Note that action 0 is dominant for low signals and action 2 for high signals. In addition the gain to raising action is decreasing in action: the gain to switching from 0 to 1 is less than the gain from switching to 1 to 2 whatever the signal received or the action taken by the other player. These assumptions will also play a role in some of the convergence analysis.

**Example 3** (*Global Games*)

Consider the following modification of Example 1:

	0	1
0	0, 0	0, $-2 + \theta$
1	$-2 + \theta, 0$	$-1 + \theta, -1 + \theta$

**Figure 1**

$\theta$  is a parameter drawn from a continuous distribution. It corresponds to  $t_0$  in Section 2 if it is appropriately transformed to take values in  $[0, 1]$ . Suppose each player does not observe  $\theta$  before choosing actions but receives a signal  $t_i = \theta + \epsilon_i$ , where  $\epsilon_i$  are noise terms independent of each other and  $\theta$ . Under mild assumptions on the distributions this falls into the class of games studied in this paper.<sup>2</sup> The gain to raising action is increasing in the other player's action and in  $\theta$ , so payoffs are supermodular.

As in Example 1, there will exist an equilibrium in cutoff strategies. If  $\theta$  and the action of the other player are observed after the game is played, cutoffs can be revised as outlined in Example 1. This is an example of the kind of game studied in the literature on global games (see for example Morris and Shin (2003)).

## 5. Normal Form Learning

The rest of the paper considers learning. It is assumed that the game outlined in Section 2 is played repeatedly over periods  $n = 1, 2, 3, \dots$  with signals drawn independently across periods.

The equilibrium conditions in the previous section suggest a simple model of normal-form learning.<sup>3</sup> Consider the following algorithm with  $k(n)$  the value of  $k$  in period  $n$ :

$$k(n+1) = \Pi_{\Sigma} \circ \tilde{h}(k(n)) \tag{12}$$

where

---

<sup>2</sup>If  $\theta$ ,  $t_1$  and  $t_2$  all have log-concave densities they they affiliated, which is sufficient (see Section 2).

<sup>3</sup>This model originally appeared in Beggs (2011).

$$\tilde{h}_j^i(k^1, \dots, k^p) = k_j^i - \alpha_n f_i(k_j^i) \Delta V_i(j-1, j, k_j^i) \quad (13)$$

and  $\Pi_\Sigma$  denotes orthogonal projection onto  $\Sigma$ .

In this algorithm each player adjusts the cutoff between neighboring actions up and down according as the payoff difference between them is negative or positive (a higher cutoff meaning play the lower action until a higher signal is reached).  $\alpha_n > 0$  measures the degree of adjustment. The term  $f_i(k_j^i)$  implies that players adjust payoffs less at signals which occur less often, perhaps because this is less urgent. This raw procedure may lead to non-monotone cutoffs so it is assumed that players instead pick the nearest point which is monotone, applying the projection  $\Pi_\Sigma$ , which is the product of the projections on individual  $\Sigma_i$ .

In this form it is clear that each player is essentially using a projected gradient algorithm since:

$$\tilde{h}_j^i = k_j^i - \alpha_n \frac{\partial W_i}{\partial k_j^i}$$

In the case where the players are playing a potential game this algorithm can be shown to converge to equilibrium. The game is a potential game if there is a function  $\Upsilon$  and functions  $\phi_i$ ,  $i = 1, \dots, q$ , such that for each  $i$

$$U_i(a, t) = \Upsilon(a, t) + \phi_i(a_{-i}, t)$$

In other words all players are maximising a common objective function plus a function independent of player  $i$ 's choices. Note that  $\phi_i$  does not affect  $\partial W_i / \partial k_j^i$ . It is straightforward to check that Example 1 is a potential game.

If one lets  $\bar{\Upsilon}(k)$  be the expected value of  $\Upsilon$  as a function of cutoffs (cf. the definition of  $W_i$ ) then the algorithm can be written as

$$k(n+1) = \Pi_\Sigma(k(n) - \alpha_n \nabla \bar{\Upsilon}(k(n)))$$

In other words, the overall learning process can be regarded as a projected gradient algorithm. It is well known that such processes converge to stationary points of  $\bar{\Upsilon}$ , under mild assumptions, and under single-crossing these are equilibria.

**Theorem 1** *Under Assumptions 1–5 and 7 if the game is a potential game and  $\alpha_n = M$  for all  $n$ , then if  $M$  is sufficiently small, any limit point of (12) is an equilibrium. In particular, if equilibrium is unique then (12) converges to it.*

This is a standard result for projected gradient algorithms. The proof can be found in the Appendix. One can also allow the step-size ( $\alpha_n$ ) to be variable as in the next section.

This is of course a model of normal-form learning. In particular it assumes that the cutoffs, or equivalently the strategies, of other players are observed, so that expected payoffs can be calculated. The next section considers an extension of this model to the case where cutoffs are unobserved.

## 6. Unobserved Cutoffs

If cutoffs are unobserved the rule of the previous is not feasible. The adjustment rule

$$k_j^i(n+1) = k_j^i(n) - \alpha_n f_i(k_j^i(n)) \Delta V_i(j-1, j, k_j^i(n)) \quad (14)$$

requires the cutoffs and so the strategies of the other agents to be known. A natural idea is to estimate expected payoffs by the realized payoffs in any period. The actual signal received may, however, be far from the cutoff whose payoffs it is desired to adjust.

Beggs (2009) suggested, in the case of binary actions, circumventing these difficulties by using the rule

$$k_j^i(n+1) = k_j^i(n) - \alpha_n K \left( \frac{k_j^i(n) - t_i(n)}{h_n} \right) \Delta U_j^i(n) \quad (15)$$

where  $\Delta U_j^i(n)$  is the realized difference in payoffs between actions  $j-1$  and  $j$  at time  $n$ . Denote this rule in vector form by  $k^i(n+1) = h(k^i(n))$ .

$t_i(n)$  is the signal received by player  $i$  in period  $n$ .  $K$  is a weighting or Kernel function, which determines how much weight is given to the current observation. If the signal received is far from the cutoff being adjusted, then little weight will be attached to it. As noted this rule can be interpreted as reflecting an idea of ‘similarity’: cutoffs are adjusted more if the signal received is close or ‘similar’ to them. Such a system is an example of passive stochastic approximation, suggested first by Härdle and Nixdorf (1987).

In terms of the examples in Section 4, the proposed learning rules requires the player to observe the realized payoffs for all his actions at stage  $n$ . That

is the row player, say, observes the payoffs in the column corresponding to the action chosen by the other player. In Examples 1 and 2, he can calculate these if he knows the form of the payoff matrix and observes the action played by the other player, since he knows his own type each period. In Example 3 this will be the case if he in addition observes the state of nature,  $\theta$ .

$K$  will be assumed to satisfy the standard assumptions for a kernel:

**Assumption 8** (a)  $\int K(u) du = 1$ , (b)  $\int uK(u) du = 0$ , (c)  $K$  has compact support, (d)  $K$  is bounded.

$K$  is assumed the same for all players to simplify the notation but this is not essential.

$\alpha_n$  and  $h_n$  are adjustment parameters.  $\alpha_n$  represents the extent to which the current cutoff is adjusted and plays the usual role in stochastic approximation: it must tend to zero to eliminate the influence of noise but not too quickly that the system can become stuck far from the optimum.  $h_n$  controls the extent to which weight is put on observations far from the cutoff. As  $n$  becomes large,  $h_n$  becomes small, putting more and more weight near to the cutoff.

They will be assumed to satisfy

**Assumption 9** (a)  $\sum_n \alpha_n = \infty$ , (b)  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ , (c)  $\sum_n \frac{\alpha_n^2}{h_n^2} < \infty$ .

If cutoffs are adjusted independently then random shocks may mean that they need not be monotonic. To circumvent this it is assumed the agent chooses the nearest monotonic cutoffs:

$$k^i(n+1) = \Pi_{\Sigma_i} h(k^i(n)) \quad (16)$$

The projection may be thought a little unnatural and assumptions will be made later which imply that it is not used asymptotically.

(15) can be rewritten as a modified version of (14)

$$k_j^i(n+1) = k_j^i(n) - \alpha_n f_i(k_j^i) \Delta V_i(j-1, j, k_j^i(n)) + \alpha_n (u_j^i(n) + b_j^i(n)) \quad (17)$$

The term  $f_i(k_j^i)$  reflects the fact that cutoffs are adjusted only when the signals are close to  $k_j^i$ .  $u_j^i(n)$  is a noise term given by

$$u_j^i(n) = E \left( \frac{1}{h_n} K \left( \frac{k_j^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) \middle| \mathcal{F}_{n-1} \right) - \frac{1}{h_n} K \left( \frac{k_j^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) \quad (18)$$

where  $\mathcal{F}_{n-1}$  denotes the history of the process up to time  $n$ .  $b_j^i(n)$  is a bias term given by

$$b_j^i(n) = f_i(k_j^i) \Delta V_i(j-1, j, k_j^i(n)) - E \left( \frac{1}{h_n} K \left( \frac{k_j^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) \middle| \mathcal{F}_{n-1} \right) \quad (19)$$

reflecting the fact that observations at signals other than the current cutoff are used for updating.

In vector form this can be written as

$$k(n+1) = k(n) + \alpha_n G(k(n)) + \alpha_n (u(n) + b(n)) \quad (20)$$

and (16) as

$$k(n+1) = \Pi_\Sigma \left( k(n) + \alpha_n G(k(n)) + \alpha_n (u(n) + b(n)) \right) \quad (21)$$

where

$$G_{ij}(k) = -f(k_j^i) \Delta V(j-1, j, k_j^i) \quad (22)$$

Standard results, see Kushner and Yin (1997), show the asymptotic properties of (20) are closely related to those of the differential equation

$$\dot{k} = G(k) + L_\Sigma \quad (23)$$

where  $L_\Sigma$  is an impulsive term which is non-zero only when the process is on the boundary of  $\Sigma$  and serves to keep the process within  $\Sigma$ .

Equivalently they are related to those of the projected differential equation

$$\dot{k} = \pi(k, G(k)) \quad (24)$$

where

$$\pi(k, v) = \lim_{\delta \rightarrow 0} \frac{\Pi_\Sigma(k + \delta v) - k}{\delta} \quad (25)$$

$\pi$  projects  $v$  onto the tangent cone to  $\Sigma$  at  $k$ . The latter is the closure of the set of feasible directions at  $k$ :  $T_\Sigma(k) = \text{cl}\{\lambda(y - k), y \in \Sigma, \lambda \geq 0\}$ . Equivalently  $T_\Sigma(k)$  is the polar of the Normal cone to  $\Sigma$  at  $k$ :  $T_\Sigma(k) = \{y : y'x \leq 0, x \in N_\Sigma(k)\}$  (see Rockafellar and Wets (1998) Chapter 6).

A solution to (24) is an absolutely continuous function such that (24) holds for almost all  $t$ . Dupuis and Nagurney (1993) show the equivalence of the formulations (20) and (24) and prove the uniqueness of solutions (existence follows from earlier results, for example Henry (1973)).

**Lemma 3** *Under Assumptions 1-5 and 7-9 the set of limit points of  $k(n)$  under (20) is almost surely a compact, connected set invariant under the flow of (24), or equivalently of (23). Moreover the limit set is contained in the internally chain recurrent set of (24).*

The proof is in the Appendix and is immediate from standard results. A set is internally chain recurrent (or transitive) if it contains no proper attractor (see Benaïm (1999)).

The rest points of (24) are precisely those of the underlying game:

**Lemma 4**  *$k^*$  is a stationary point of (24), or equivalently (23), if and only if it is an equilibrium of the game.*

This follows from Lemma 1 of Dupuis and Nagurney (1993), for example, who note that the zeros of (24) correspond to solutions of the variational inequality  $VI(\Sigma, -G)$ , that is by Lemma 1 of this paper are equilibria of the game.

The next section provides conditions under which the limit set of the learning model is contained in the set of equilibria of the game.

## 7. Convergence Results

This section provides the principal convergence results. In Section 7.1 convergence in supermodular games is studied. In section 7.2 potential games are examined.

### 7.1 Supermodular Games

As noted in the previous section the convergence of the stochastic algorithm is closely related to the asymptotic behavior of the deterministic system

$$\dot{k} = G(k) + L_{\Sigma} \quad (26)$$

where  $L_{\Sigma}$  is a reflection term. In the absence of the latter term, this system defines a cooperative dynamical system if the game is supermodular, that is

$$\frac{\partial G_{\alpha}}{\partial k_{\beta}} \geq 0 \quad \alpha \neq \beta \quad (27)$$

using single indices for simplicity. Equivalently the Jacobian matrix  $DG$  has non-negative entries.

This follows since

$$G_{ij} = -f(k_j^i)\Delta V_i(j-1, j, k_j^i, k^{-i}) \quad (28)$$

which is increasing in others' cutoffs (and independent of a player's own other cutoffs). If other players raise their cutoffs the benefit to raising own cutoffs is increased if payoffs are supermodular. If payoffs are strictly supermodular then the Jacobian matrix  $DG$  is in addition irreducible since then the benefit to increasing cutoffs is strictly increasing in others' cutoffs.

Projection may not preserve the cooperative nature of the system. If however the vector field  $G$  points inward on the boundary then the projection term can be dropped so the limiting differential equation is cooperative. In addition the projection is almost surely not used for large  $n$ , removing a less attractive feature of the algorithm.

It will be assumed that each player's lowest action (0) is dominant if his signal is low enough and his highest action ( $m(i)$  for player  $i$ ) is dominant if his signal is high enough:

**Assumption 10** *There exist  $\underline{t} > 0$  and  $\bar{t} < 1$  such that for all  $i$  (a) if  $t_i < \underline{t}$ ,  $U_i(0, a_{-i}, t_i, t_i) > U_i(a_i, a_{-i}, t_i, t_{-i})$  for all  $a_i \neq 0$ ,  $a_{-i}$  and  $t_{-i}$  and (b) if  $t_i > \bar{t}$ ,  $U_i(m(i), a_{-i}, t_i, t_{-i}) > U_i(a_i, a_{-i}, t_i, t_{-i})$  for  $a_i \neq m(i)$ ,  $a_{-i}$  and  $t_{-i}$ .*

It will also be assumed that actions display a diminishing return (or concavity) property<sup>4</sup>, that is the gain to raising actions is decreasing in the level of the action.

**Assumption 11** *For all  $i$ ,  $\Delta U_i(j-1, j, a_{-i}, t)$  is strictly decreasing in  $j$  for all  $a_{-i}$  and  $t$ .*

These two assumptions guarantee that the vector field defined by  $G$  points inward on the boundary and in particular equilibria cannot occur there. Assumption 10 guarantees that action 0 must be played if signals are low enough and the highest action must be played if they are high enough. Cutoffs are therefore bounded away from 0 and 1. Assumption 11 guarantees that it can never be desirable to leap from (say) action 1 to 3 and omit action 2. If switching from 1

---

<sup>4</sup>Krishna (1992) and Berger (2008) make a similar assumption in the context of fictitious play.

to 2 would lower payoffs then Assumption 11 implies that switching from 2 to 3 certainly would. It follows that the monotonicity constraints are never binding. These assumptions are satisfied in Example 2.

**Lemma 5** *Under Assumptions 10–11  $L_\Sigma$  is identically zero in (26) and the limit set of any trajectory of (24) lies strictly in the interior of  $\Sigma$ . In particular, any equilibrium lies strictly in the interior of  $\Sigma$ .*

The other ingredient in the convergence result is the recent paper of Benaïm and Faure (2012). They consider, inter alia, the Robbins-Munro algorithm

$$x(n+1) = x(n) + \frac{1}{n} (H(x(n)) + v(n)) \quad (29)$$

where  $H$  is globally Lipschitz and

$$E(v(n)|\mathcal{F}_{n-1}) = 0, \quad \sup_n E(\|v(n)\|^2) < \infty \quad (30)$$

They show that  $x(n)$  cannot converge to a linearly unstable set (more precisely a normally hyperbolic repulsive set)  $\Gamma$  of the differential equation

$$\dot{x} = H(x) \quad (31)$$

provided that the noise is non-degenerate in a neighborhood  $U$  of  $\Gamma$ . More precisely they assume that there is a continuous mapping  $Q(x)$  from  $U$  to the set of positive-definite matrices such that for  $x(n-1) \in U$ ,

$$E(v(n)v(n)'|\mathcal{F}_{n-1}) = Q(x(n)) \quad (32)$$

They also assume that

$$E(\|v(n)\|^{2p}|\mathcal{F}_{n-1}) \quad (33)$$

is almost surely bounded for some  $p > 1$ .

They prove their result by using a diffusion approximation. They apply the result to the convergence of stochastic fictitious play in supermodular games. This result is significantly more general than earlier results, for example of Pemantle (1990) who considered only isolated equilibria. It is also easier to adapt to the current model.

In the current context

$$k(n+1) = k(n) + \alpha_n G(k(n)) + \alpha_n (u(n) + b(n)) \quad (34)$$

where  $b(n)$  has non-zero mean. Moreover  $u(n)$  has unbounded variance (see (18)). The results of Benaïm and Faure (2012) cannot therefore be directly applied. It is, however, straightforward to modify them to apply to the current case by using a diffusion approximation for passive stochastic approximation (cf. Härdle and Nixdorf (1987)). In order to do so, one further assumption will be made:

**Assumption 12**  $\alpha_n = 1/n$  and  $h_n = 1/n^\eta$ , where  $1/5 < \eta < 1/2$ .

This assumption guarantees that one can make a diffusion approximation using a suitable scaling and the bias term can be neglected. The details are somewhat technical and are relegated to the Appendix.

**Theorem 2** *If the game is strictly supermodular with increasing beliefs, then under Assumptions 1–6, 8 and 10–12 the limit set of  $\{k(n)\}$  is almost surely an ordered arc of equilibria that is not linearly unstable. In particular if equilibrium is unique, play converges to it almost surely.*

The assumption of strict supermodularity is used to guarantee that the vector field defined by  $G$  is irreducible as well as co-operative. In addition it guarantees that the noise in the process is non-degenerate in the neighborhood of any limit set of the process using variants on standard results for the variance matrix of kernel regression (cf. Schuster (1972)).

**Corollary 1** *Under the assumptions of Theorem 2, if in addition  $f$  and  $U_i$  are real analytic for all  $i$ , then play converges almost surely to a Nash equilibrium.*

This follows from a result of Jiang (1991) which states that a real-analytic cooperative vector field cannot have an ordered arc of equilibria.

If there is a unique equilibrium, then Benaïm and Faure (2012)'s results are not needed as it follows from Theorem 3.3 of Benaïm (2000) that the internally chain recurrent set consists solely of this equilibrium. One therefore does not need the precise step size assumptions in Assumption 12.

**Theorem 3** *If the game is strictly supermodular with increasing beliefs and has a unique equilibrium, then under Assumptions 1-5 and Assumptions 8-11, the limit set of  $\{k(n)\}$  is almost surely the equilibrium set of cutoffs.*

## 7.2 Potential Games

In the case of potential games one obtains an immediate generalization of Theorem 1. As in the previous section let  $\bar{\Upsilon}(k)$  be the expected value of the potential. Let  $\mathcal{E}$  denote the equilibrium set.

**Theorem 4** *Under Assumptions 1-5 and 7-9 if the game is a potential game with potential function  $\Upsilon$  then the set of limit points of (12) is contained in the set of equilibria if  $\bar{\Upsilon}(\mathcal{E})$  has empty interior. In particular, if equilibrium is unique then play converges to it.*

The condition on  $\bar{\Upsilon}(\mathcal{E})$  holds in particular if the set of equilibria is finite. It also holds if  $\bar{\Upsilon}$  is smooth enough by Sard's theorem.

## 8. Conclusion

This paper has studied a simple model of learning in Bayesian games with monotone equilibria. Convergence has been shown in the case of supermodular and potential games. Learning in other Bayesian games is a promising topic for future exploration.

## Appendix

### *Proof of Lemma 1*

$\Sigma_i = \{k : 0 \leq k_1 \leq \dots \leq k_{m(i)} \leq 1\}$  is the monotone non-negative cone bounded above by 1 (see for example Boyd and Vandenberghe (2004) exercise 2.33), so the result can be deduced from standard ones, but a proof is sketched for completeness.  $y \in N_{\Sigma_i}(k)$  if and only if  $y'(x - k) \leq 0$  for all  $x \in \Sigma_i$ .  $y'(x - k) = \sum_{l=1}^{m(i)} y_l w_l$ , where  $w_l = x_l - k_l$ . Summation by parts implies that

$$\sum_{l=1}^{m(i)} w_l y_l = \sum_{l=1}^{m(i)} (w_l - w_{l+1}) \left( \sum_{l'=1}^l y_{l'} \right) \quad (35)$$

with  $w_{m(i)+1} = 0$ . Using this implies that  $y \in N_{\Sigma}(k)$  if and only if

(a) If  $j$  and  $j'$  are adjacent active actions then

(i)  $\sum_{l=j}^{j'} y_l = 0$ ,

(ii) for any  $j''$  with  $j < j'' < j'$ ,  $\sum_{l=j}^{j''} y_l \geq 0$ ,

(b) If  $j$  is the least active action, then  $\sum_{l=1}^j y_l \leq 0$ ,

Similarly

(c) If  $j$  is the greatest active action, then  $\sum_{l=1}^j y_j \geq 0$ .

This implies that, the vector  $\nabla W_i = (-f(k_j^i) \Delta V_i(j-1, j, k_j^i))$  lies in  $N_{\Sigma_i}(k^i)$  if and only if the conditions in the Lemma hold since  $\Delta V_i(j, j', k_j^i) = \sum_{l=j+1}^{j'} \Delta V_i(l-1, l, k_l^i)$  and  $f(k_l^i) = f(k_{l+1}^i)$  if  $k_l^i = k_{l+1}^i$ .

### *Proof of Theorem 1*

According to Rusczyński (2006) Proposition 6.1 if  $X$  is a compact convex set and  $h$  a continuously differentiable function whose gradient  $\nabla h$  satisfies a Lipschitz condition with constant  $L$  on  $K$  ( $\|\nabla h(x) - \nabla h(y)\| \leq L\|x - y\| \forall x, y \in X$ ) then any limit point of the algorithm

$$x_{n+1} = \Pi_X(x_n - \alpha_n \nabla h(x_n))$$

with  $\alpha_n = s$  for all  $k$  where  $0 < s < 1/L$  is a stationary point. Under Assumptions Assumptions 1–6  $\bar{\Upsilon}$  is twice continuously differentiable in  $k$  (using the fundamental theorem of calculus and dominated convergence theorem) and since  $\Sigma$  is compact satisfies a Lipschitz condition, so the result is immediate.

### Proof of Lemma 3

Under Assumption 1–5,  $G$  is Lipschitz (see (1)). Moreover from (18),  $E(u_j^i | \mathcal{F}_{n-1}) = 0$  and  $E((u_j^i)^2 | \mathcal{F}_{n-1}) \leq \frac{M}{h_n^2}$  for some constant  $M$ . It follows from Assumption 9 that  $\sum_n \alpha_n u(n)$  is an  $L^2$ -bounded martingale and therefore convergent. Moreover from (19)  $b_j^i(n) \leq M'h_n$  for some constant  $M'$ , using the change of variables  $s = (k_j^i - t_i(n))/h_n$  and the properties of  $K$  in Assumption 8. Moreover  $k(n)$  is bounded, so in particular its second moment is uniformly bounded.  $\sum_n \alpha_n^2 < \infty$  by Assumption 9.

It follows from Theorems 5.2.3 and 5.2.5 of Kushner and Yin (1997) (see also Theorems 4.1, 5.3 and 6.11 of Benaim (1999)) that the set of limit points of  $k(n)$  is almost surely a compact, connected set invariant under the flow of (23) and is internally chain connected.

### Proof of Lemma 5

If  $X = \{x : a'_j x \leq b_j \ \forall j\}$  is a polyhedral set, then by Theorem 6.46 of Rockafellar and Wets (1998), the tangent cone at  $x \in X$  is  $T_X(x) = \{y : a'_j y \leq 0, j \in J\}$  where  $J$  is the set of indices corresponding to the set of constraints which bind, that is for which  $a'_j x_j = b_j$ .

Applying this to  $\Sigma_i = \{k : -k_1 \leq 0, k_1 - k_2 \leq 0, \dots, k_{m(i)} - k_{m(i)-1} \leq 0, k_{m(i)} \leq 1\}$ , shows that the tangent cone of  $\Sigma_i$  is the set of vectors  $y$  such that

$$k_0^i = 0 \Rightarrow y_0 \geq 0 \quad (36)$$

$$k_j^i = k_{j+1}^i \Rightarrow y_j \leq y_{j+1} \quad j = 1, \dots, m(i) - 1 \quad (37)$$

$$k_{m(i)} = 1 \Rightarrow y_{m(i)} \leq 0 \quad (38)$$

Under Assumptions 10 and 11 the vector  $y = (-\Delta V_i(j-1, j, k_j^i))$  satisfies all the constraints above with strict inequality:

$$k_0^i = 0 \Rightarrow \Delta V_i(0, 1, k_0^i) < 0 \quad (39)$$

$$k_j^i = k_{j+1}^i \Rightarrow \Delta V_i(j-1, j, k_j^i) > \Delta V_i(j, j+1, k_j^i) \quad j = 1, \dots, m(i) - 1 \quad (40)$$

$$k_{m(i)} = 1 \Rightarrow \Delta V_i(m(i)-1, m(i), k_{m(i)}^i) > 0 \quad (41)$$

and hence belongs to the interior of the tangent cone to  $\Sigma_i$ . Assumption 10 implies (39) and (41) and Assumption 11 implies (40).

Since  $f(k_j^i) = f(k_{j+1}^i)$  if  $k_j^i = k_{j+1}^i$ , the vector  $y' = (-f(k_j^i)\Delta V_i(j-1, j, k_j^i))$  also belongs to the interior of the tangent cone to  $\Sigma_i$ . Since  $T_\Sigma(k) = T_{\Sigma_1}(k^1) \times \dots \times T_{\Sigma_p}(k^p)$ ,  $G$  belongs to the interior of the tangent cone of  $\Sigma$  for all  $k$ , so is inward-pointing on the boundary.

It follows that the projection  $\pi(k, G)$  always equals  $G$ , so  $L_\Sigma$  vanishes. As  $G$  points inwards on the boundary any trajectory starting on the boundary of  $\Sigma$  immediately enters the interior, so any equilibria and limit sets lie in the interior.

### *Proof of Theorem 2*

From Lemma 3, the limit set of  $k(n)$  is almost surely an invariant set of (23). From Lemma 5, the vector field  $G$  points inward on the boundary so any internally chain recurrent set of (23) lies strictly within  $\Sigma$ . It follows that the reflection term can be neglected asymptotically and that any limit set of  $k(n)$  is an internally chain-recurrent set of the flow of  $G$ . The recursion studied is from (20)

$$k(n+1) = k(n) + \alpha_n G(k(n)) + \alpha_n (u(n) + b(n)) \quad (42)$$

$G$  is cooperative and irreducible and  $C^2$ . Theorem 4.2 of Benaïm and Faure (2012) implies that any internally chain-recurrent set of  $G$  is either an ordered arc of equilibria or linearly unstable (more precisely a normally-repulsive hyperbolic set — see definition 3.1 of Benaïm and Faure (2012)). The result will follow provided the result of Benaïm and Faure (2012) on nonconvergence to linearly unstable sets can be extended to this context. The rest of the proof checks this.

It is straightforward to check using (18) and (19) that under Assumptions 1–6

$$E(u(n)|\mathcal{F}_{n-1}) = 0, \quad E(\|u(n)\|^2|\mathcal{F}_{n-1}) = O\left(\frac{1}{h_n}\right) \quad (43)$$

and

$$E(\|b(n)\||\mathcal{F}_{n-1}) = O(h_n^2) \quad (44)$$

Benaïm and Faure (2012) consider a  $X(t) \in \mathbb{R}^d$  continuous-time process adapted to some filtration  $\mathcal{F}_{t \geq 0}$ . Let  $\Phi_t$  be the flow corresponding to the differential equation  $\dot{x} = H(x)$ . A continuous function  $\chi(x)$  is said to **asymptotic pseudo-trajectory** for  $\Phi$  if

$$\lim_{t \rightarrow \infty} d_\chi(t, T) = 0 \quad (45)$$

where

$$d_\chi(t, T) = \sup_{h \in [0, T]} \|\chi(t+h) - \Phi_h(\chi(t))\| = 0$$

They show that if  $X(t)$  satisfies the following two hypotheses (Hypotheses 2.2 and 3.6 of Benaïm and Faure (2012) respectively)

**Hypothesis 1** *There is a map*

(i) *For any  $\delta > 0$ ,  $T > 0$*

$$P\left(\sup_{s \geq t} d_X(s, T) \geq \delta | \mathcal{F}_t\right) \leq \omega(t, \delta, T) \quad (46)$$

(ii)  $\lim_{t \rightarrow \infty} \omega(t, \delta, T) = 0$

**Hypothesis 2** *There is a function  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that*

(i) *For any open ball  $O \in \mathbb{R}^d$ , there exists  $c > 0$  such that*

$$\liminf_{t \rightarrow \infty} P\left(\frac{X(t+1) - \Phi_1(X(t))}{\gamma(t)} \in O\right) > c$$

*almost surely*

(ii) *There exists  $a > 0$  such that*

$$\limsup_{t \rightarrow \infty} \omega\left(t, a\sqrt{\gamma(t)}, T\right) < 1$$

then (Theorem 3.9 of Benaïm and Faure (2012)) if  $\Gamma$  is a linearly unstable set  $P(X(t) \rightarrow \Gamma) = 0$ . As noted in remark 3.7 of Benaïm and Faure (2012) it suffices that (ii) of Hypothesis 2 holds in a neighborhood  $U$  of  $\Gamma$ .

A sufficient condition Hypothesis 1 is that there exist a function  $r : \mathbb{R}^3 \rightarrow \mathbb{R}_+$  such that

$$P\left(\sup_{s \geq t} d_X(s, T) \geq \delta | \mathcal{F}_t\right) \leq \int_t^{t+T} r(s, \delta, T) ds \quad (47)$$

where

$$\int_0^\infty r(s, \delta, T) ds < \infty \quad (48)$$

In this case  $\omega(t, \delta, T) = \int_t^\infty r(s, \delta, T) ds$ .

Consider the discrete-time process

$$x(n+1) = x(n) + \frac{1}{n} (H(x(n)) + v(n)) \quad (49)$$

(42) falls into this framework with  $x(n) = k(n)$ ,  $H = G$  and  $v(n) = u(n) + b(n)$  under Assumption 12.  $x_n$  and  $v_n$  will be used interchangeably with  $x(n)$  and  $v(n)$  for notational convenience.

Let  $\alpha_n = \frac{1}{n}$  and set  $\tau_n = \sum_{l=1}^n \alpha_l$ . Consider the interpolated process

$$X(\tau_l + r) = x_l + r \frac{x_{l+1} - x_l}{\alpha_{l+1}} \quad l = 0, 1, 2, \dots \quad r \in [0, \alpha_{l+1}[ \quad (50)$$

It will be shown that  $X(t)$  satisfies Hypothesis 1 and Hypothesis 2 when  $X(t)$  is derived from (42) and the limit set of  $X(t)$  lies in a neighborhood  $U$  of a linearly unstable set  $\Gamma$ .

#### *Verification of Hypothesis 1*

Benaïm and Faure (2012) note that it follows from Benaïm (1999) (Proposition 4.4) and Benaïm (2000) (the claim in Lemma 3.5) that Hypothesis 1 holds for the interpolated process when  $v(n)$  satisfies (30) with  $r(s, \delta, T) = B \frac{\hat{\gamma}}{\delta^2}$ , for large enough  $s$  and a constant  $B$ , where  $\hat{\gamma}(\tau_l + s) = \alpha_{l+1} \quad s \in [0, \alpha_{l+1})$ .

It is straightforward to check that the arguments in Benaïm (1999) and Benaïm (2000) can be modified in the current to context to show that Hypothesis 1 holds with

$$r(s, \delta, T) = B \frac{\bar{\gamma}}{\delta^2} \quad (51)$$

with

$$\bar{\gamma}(\tau_l + s) = \frac{\alpha_{l+1}}{h_{l+1}} \quad s \in [0, \alpha_{l+1}) \quad (52)$$

This follows since  $\sum_n \alpha_n u(n)$  is a martingale with conditional variance of order  $\sum_n \frac{\alpha_n^2}{h_n}$  and  $b(n)$  converges to zero.

#### *Verification of Hypothesis 2*

In the case of the Robbins-Munro process, Benaïm and Faure (2012), prove Hypothesis 2 by using a diffusion approximation. Let  $m_n = \sup\{s \in \mathbb{N} : \tau_s \leq n\}$ ,  $s_n = m_{n+1} - m_n$ ,  $t_j^n = \tau_{m_n+j} - \tau_{m_n}$  and  $t_n = t_{s_n}^n$ . Let  $\{\mathcal{G}_n\}_n$  be the  $\sigma$ -algebra  $\{\mathcal{F}_{m_n}\}_n$ .

Benaïm and Faure (2012) note that to prove Hypothesis 2 in the case of the discrete-time interpolated process  $x_n$  it is enough to find a positive vanishing sequence  $\{\gamma(n)\}$  and a  $\mathcal{G}_n$ -adapted process  $Y_n$  such that

(i) For any  $\beta > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \left\| \frac{x_{m_{n+1}} - \Phi_{t_n}(x_{m_n})}{\gamma(n)} - Y_{n+1} \right\| > \beta \middle| \mathcal{G}_n \right) = 0$$

(ii) For any open ball  $O \in \mathbb{R}^d$ , there exists  $c > 0$  such that

$$\liminf_{n \rightarrow \infty} P(Y_{n+1} \in O | \mathcal{G}_n) > c$$

almost surely

(iii) There exists  $a > 0$  such that

$$\limsup_{n \rightarrow \infty} \omega \left( n, a\sqrt{\gamma(n)}, T \right) < 1$$

They consider the sequence

$$Y_{n+1} = \sum_{i=1}^{s_n} y_i^n \quad (53)$$

where

$$y_j^n = \frac{\alpha_{m_n+j}}{\sqrt{\gamma(n)}} \left( \prod_{s=j+1}^{s_n} \left( I + DH \left( \Phi_{t_{s-1}^n} \right) (x_{m_n}) \right) \right) v_{m_n+j} \quad (54)$$

with  $\gamma(n) = \sum_{s=1}^{s_n} \alpha_{m_n+s}^2$

Here the same sequence is considered (with  $H$  corresponding to  $G$  in the current paper) but with

$$\gamma(n) = \sum_{s=1}^{s_n} \alpha_{m_n+s}^2 / h_{m_n+s} \quad (55)$$

Equivalently

$$y_j^n = \frac{\tilde{\alpha}_{m_n+j}}{\sqrt{\gamma(n)}} \left( \prod_{s=j+1}^{s_n} \left( I + DH \left( \Phi_{t_{s-1}^n} \right) (x_{m_n}) \right) \right) \tilde{v}_{m_n+j} \quad (56)$$

where  $\tilde{\alpha}_n = \alpha_n / \sqrt{h_n}$  and  $\tilde{v}_n = \sqrt{h_n} v_n$ .

It is straightforward to check that Benaïm and Faure (2012)'s arguments for (i) can be easily modified for the current context. (iii) follows immediately from the definition of  $\gamma(n)$  in (55) together with (51) and (52).

To prove point (ii) Benaïm and Faure (2012) appeal to a Central Limit Theorem for triangular arrays due to Hall and Heyde (1980), in a form given by Chow and Teicher (1998) Theorem 2 in Section 9.5:

**Theorem A1** (Hall-Heyde) For any  $n \geq 1$  let  $s_n$  be a positive integer and  $(\Omega_n, \mathcal{F}^n, P_n)$  a probability space. Consider an increasing sequence of  $\sigma$ -algebras  $\mathcal{F}_1^n \subset \dots \subset \mathcal{F}_{s_n}^n \subset \mathcal{F}^n$  and a family  $\{y_j^n\}_{j=1, \dots, s_n}$  of  $(\mathcal{F}_j^n)$ -adapted random variables such that

- (a)  $\sum_{j=1}^{s_n} E(y_j^n | \mathcal{F}_{j-1}^n) \xrightarrow{p} 0$ ,  $\sum_{j=1}^{s_n} \|E(y_j^n | \mathcal{F}_{j-1}^n)\|^2 \xrightarrow{p} 0$ ,
- (b)  $\sum_{j=1}^{s_n} E_n \left( \|y_j^n\|^2 1_{\|y_j^n\| > \epsilon} | \mathcal{F}_{j-1}^n \right) \xrightarrow{dist} 0 \quad \forall \epsilon > 0$
- (c) there exists a positive  $\mathcal{F}_1^n$  adapted sequence  $w_n$  such that

$$\sum_{j=1}^{s_n} E \left( y_j^n (y_j^n)' | \mathcal{F}_{j-1}^n \right) - w_n \xrightarrow{p} 0$$

- (d) There exists a positive random matrix  $w$  defined on some probability space  $(\Omega, \mathcal{F}, P)$  such that

$$\sum_{j=1}^{s_n} E \left( y_j^n (y_j^n)' | \mathcal{F}_{j-1}^n \right) - w \xrightarrow{p} 0$$

then the sequence  $(y_n)_n$ , with  $y_n = \sum_{j=1}^{s_n} y_j^n$ , converges in distribution to a random variable  $y$  defined on  $(\Omega, \mathcal{F}, P)$  and whose characteristic function is given by  $E \left( e^{-(1/2)t'wt} \right)$ .

These hypotheses are verified below:

*Verification of (a)*

In the case of Benaïm and Faure (2012), the  $y_j^n$  are a martingale, so (a) is automatic. Here Assumption 12 guarantees that the bias term can be asymptotically neglected so (a) holds.

In more detail, under Assumption 6,  $f$  and  $U$  are  $C^2$ , so using a second-order Taylor expansion yields that  $E(\|b(n)\|) \leq Lh_n^2$  for some constant  $L$  using Assumption 8. Hence

$$\left\| \sum_{j=1}^{s_n} E(y_j^n | \mathcal{F}_{j-1}^n) \right\| \leq \frac{L}{\sqrt{\gamma(n)}} \sum_{j=1}^{s_n} \tilde{\alpha}_{m_n+j} \sqrt{h_{m_n+j}} h_{m_n+j}^2 \quad (57)$$

By Assumption 12  $\sum_{n=1}^{\infty} h_n^5$  is convergent, so an application of the Cauchy-Schwartz inequality yields that the right hand-side tends to zero since  $\gamma(n) = \sum_{j=1}^{s_n} \tilde{\alpha}_{m_n+j}^2$ . The remainder of (a) is verified similarly.

Verification of (b)

(b) is verified in the same way as Benaïm and Faure (2012). Note that

$$E(v_j^i(n))^4 = \frac{1}{h_n^4} EK^4 \left( \frac{k_j^i(n) - t_i(n)}{h_n} \right) (\Delta U_j^i(n))^4 \quad (58)$$

Using the change of variables  $s_i = (k_j^i(n) - t_i(n))/h_n$  and the boundedness of  $K$  and  $\Delta_j^i U(n)$  shows that

$$E(v_j^i(n))^4 \leq \frac{C}{h_n^3} \quad (59)$$

Applying Hölder's inequality as in Benaïm and Faure (2012) yields that

$$\begin{aligned} E\left(\|y_j^n\|^2 1_{\|y_j^n\| > \epsilon} | \mathcal{F}_{j-1}^n\right) &\leq E\left(\|y_j^n\|^4 | \mathcal{F}_{j-1}^n\right)^{1/2} P\left(\|y_j^n\|^4 > \epsilon^4 | \mathcal{F}_{j-1}^n\right)^{1/2} \\ &\leq \frac{1}{\epsilon^2} E\left(\|y_j^n\|^4 | \mathcal{F}_{j-1}^n\right) \\ &\leq \frac{e^{4\|H\|}}{\gamma(n)^2 \epsilon^2} \alpha_{m_n+j}^4 E\|v(m_n+j)\|^4 \\ &\leq \frac{D}{\gamma(n)^2} \frac{\alpha_{m_n+j}^4}{h_{m_n+j}^3} \end{aligned} \quad (60)$$

for some constant  $D$  depending on  $\epsilon$ . Hence

$$\sum_{j=1}^{s_n} E\left(\|y_j^n\|^2 1_{\|y_j^n\| > \epsilon} | \mathcal{F}_{j-1}^n\right) \leq \sup D \frac{\alpha_{m_n+j}^2}{\gamma(n) h_{m_n+j}^2}$$

which tends to zero as  $n \rightarrow \infty$  as  $\gamma(n) \rightarrow \infty$  and  $\alpha_n^2/h_n^2 \rightarrow 0$  by Assumption 12.

The remaining details are as in Benaïm and Faure (2012).

Verification of (c) and (d)

(c) and (d) follow from (32) in Benaïm and Faure (2012).

The same argument can be applied here provided it is shown that there is a continuous mapping  $Q(x)$  from  $U$ , where  $U$  is neighborhood of the global attractor of  $X$ , to the set of positive-definite matrices such that for  $k_{n-1} \in U$ ,

$$Q_n(k(n)) := E(\tilde{v}(n)\tilde{v}(n)' | \mathcal{F}_{n-1}) \rightarrow Q(k(n)) \quad (61)$$

uniformly in  $n$ . All except the positive-definiteness of  $Q$  is clear.

From (a) the bias terms can be neglected and will be omitted in the expressions below.

$$\begin{aligned}
& E \left( \tilde{v}_j^i(n) (\tilde{v}_{j'}^i(n))' | \mathcal{F}_{n-1} \right) = \\
& \frac{1}{h_n} E \left( K \left( \frac{k_j^i - t_i(n)}{h_n} \right) K \left( \frac{k_{j'}^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) \Delta U_{j'}^i(n) | \mathcal{F}_{n-1} \right) - \\
& \frac{1}{h_n} E \left( K \left( \frac{k_j^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) | \mathcal{F}_{n-1} \right) E \left( K \left( \frac{k_{j'}^i - t_i(n)}{h_n} \right) \Delta U_{j'}^i(n) | \mathcal{F}_{n-1} \right)
\end{aligned} \tag{62}$$

Using the change of variables  $s_i = (k_j^i(n) - t_i(n))/h_n$ , yields that if  $j = j'$  the second line is equal to

$$\left( \int K^2(s) ds \right) f(k_j^i) E \left( (\Delta U^i(j-1, j, t))^2 | k_j^i \right) \tag{63}$$

plus terms of order  $h_n$  and higher. The third line is of order  $h_n$  (using the same change of variables in both expectations). It follows that the diagonal elements convergence uniformly to those of the matrix  $Q$  whose diagonal entries are given by (63). The diagonal elements of  $Q$  are therefore strictly positive and bounded away from zero in the interior of  $\Sigma$  as the assumption of strict supermodularity implies that incremental payoffs are non-constant (and in particular not zero everywhere). The diagonal elements are also bounded above by continuity.

It is straightforward if tedious to check (cf. Schuster (1972)) that for large enough  $n$  the off-diagonal elements of  $Q_n$  converge to zero. If  $j \neq j'$  the compact support assumption in Assumption 8 and the fact that any internally-chain recurrent set is strictly in the interior of  $\Sigma$  implies that (62) tends to zero uniformly.

For covariances involving different players note that

$$\begin{aligned}
& E \left( \tilde{v}_j^i(n) (\tilde{v}_{j'}^{i'}(n))' | \mathcal{F}_{n-1} \right) = \\
& \frac{1}{h_n} E \left( K \left( \frac{k_j^i - t_i(n)}{h_n} \right) K \left( \frac{k_{j'}^{i'} - t_{i'}(n)}{h_n} \right) \Delta U_j^i(n) \Delta U_{j'}^{i'}(n) | \mathcal{F}_{n-1} \right) - \\
& \frac{1}{h_n} E \left( K \left( \frac{k_j^i - t_i(n)}{h_n} \right) \Delta U_j^i(n) | \mathcal{F}_{n-1} \right) E \left( K \left( \frac{k_{j'}^{i'} - t_{i'}(n)}{h_n} \right) \Delta U_{j'}^{i'}(n) | \mathcal{F}_{n-1} \right)
\end{aligned} \tag{64}$$

Using the change of variables  $s_i = (k_j^i(n) - t_i(n))/h_n$  and  $s_{i'} = (k_{j'}^{i'}(n) - t_{i'}(n))/h_n$  yields that both terms are of order  $h_n$ , so converge to zero uniformly.

Now

$$w_n = \sum_{j=1}^{s_n} E \left( y_j^n (y_j^n)' | \mathcal{F}_{j-1}^n \right) = \sum_{j=1}^{s_n} \frac{\tilde{\gamma}_{m_n+j}^2}{\gamma(n)} \Pi_{n,j} E(\tilde{v}_n \tilde{v}_n' | \mathcal{F}_{n-1}) \Pi_{n,j}' \quad (65)$$

where  $\Pi_{n,j} = \Pi_{l=j+1}^{s_n} (I + \gamma_{m_n+l} DH(\Phi_l^n(x_{m_n})))$  and if one defines  $w_n^*$  by

$$w_n^* = \sum_{j=1}^{s_n} \frac{\tilde{\gamma}_{m_n+j}^2}{\gamma(n)} \Pi_{n,j} Q(x_{m_n+j}) \Pi_{n,j}' \quad (66)$$

then  $w_n - w_n^*$  converges uniformly to zero and one can apply the argument of Benaïm and Faure (2012) applied to  $w_n^*$  to verify (d).

### *Proof of Theorem 3*

The proof is exactly the same as that of Theorem 2 except that as Theorem 3.3 of Benaïm (2000) implies that the only internally chain-recurrent set consists solely of the equilibrium cutoffs, one does not need to appeal to Benaïm and Faure (2012) and these steps can be omitted.

### *Proof of Theorem 4*

From Lemma 3 the set of limit points of (20) is a compact invariant subset of the flow of (24). Proposition 6.4 of Benaïm (1999) states that if  $\Lambda$  is an invariant set for a semi-flow  $\Phi$ ,  $V$  is a strict Lyapounov function for  $\Phi$  and  $V(\Lambda)$  has empty interior then the set of internally chain-recurrent points is contained in  $\Lambda$ .

The result here follows provided it is shown that  $\bar{\Upsilon}$  is a Lyapounov function for the flow of (24) and the equilibrium set  $\Lambda = \{k : \pi(k, G) = 0\} = \{k : \pi(k, \nabla \bar{\Upsilon}) = 0\} = \mathcal{E}$ , where  $\nabla$  is the gradient operator. The argument follows that in Bianchi and Jakubowicz (2013). Let  $\langle \cdot, \cdot \rangle$  denote inner product.

$$\frac{d}{dt} \bar{\Upsilon} = \left\langle \nabla \bar{\Upsilon}, \frac{dk}{dt} \right\rangle \quad (67)$$

From (24),

$$\frac{d}{dt} \bar{\Upsilon} = \left\langle \nabla \bar{\Upsilon}, \pi(k, \nabla \bar{\Upsilon}) \right\rangle \quad (68)$$

$\pi$  projects onto the tangent cone to  $\Sigma$  at  $k$  (see (25)). Any vector  $v$  can be decomposed into as  $v = \pi(k, v) + n$  where  $n$  is orthogonal to  $\pi(k, v)$  (see Rockafellar and Wets (1998)). Applying this result to (68),

$$\frac{d}{dt} \bar{\Upsilon} = \langle \pi(k, \nabla \bar{\Upsilon}), \pi(k, \nabla \bar{\Upsilon}) \rangle \quad (69)$$

so

$$\frac{d}{dt} \bar{\Upsilon} \leq 0 \quad (70)$$

and

$$\frac{d}{dt} \bar{\Upsilon} < 0 \quad k \notin \Lambda \quad (71)$$

which is the desired result.

## References

- Athey, S. (1996). Characterizing properties of stochastic objective functions. mimeo, MIT.
- Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69:861–889.
- Beggs, A. (2009). Learning in bayesian games with binary actions. *The B.E. Journal of Theoretical Economics: Advances in Theoretical Economics*, 9(1):Article 33.
- Beggs, A. (2011). Regularity and stability in monotone bayesian games. Technical report, University of Oxford.
- 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 Faure, M. (2012). Stochastic approximation, cooperative dynamics and supermodular games. *Annals of Applied Probability*, 22:2133–2164.
- 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, U. (2008). Learning in game with strategic complementarities revisited. *Journal of Economic Theory*, 143:292–301.
- Bianchi, P. and Jakubowicz, J. (2013). Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization. *IEEE Transactions on Automatic Control*, 58:391–405.
- Borwein, J. and Lewis, A. (2006). *Convex Analysis and Nonlinear Optimization*. Springer Verlag, New York, second edition.

- Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge.
- Chow, Y. and Teicher, H. (1998). *Probability Theory: Independence, Exchangeability and Martingales*. Springer Verlag, New York, third edition.
- Dupuis, P. and Nagurney, A. (1993). Dynamical systems and variational inequalities. *Annals of Operations Research*, 44:9–42.
- Facchinei, F. and Pang, J.-S. (2003). *Finite Dimensional Variational Inequalities and Complementarity Problems*. Two Volumes. Springer Verlag, New York.
- Fudenberg, D. and Kreps, D. (1993). Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367.
- Gilboa, I. and Schmeidler, D. (2003). Inductive inference: An axiomatic approach. *Econometrica*, 71:1–26.
- Hall, P. and Heyde, C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, New York.
- Härdle, W. and Nixdorf, R. (1987). Nonparametric sequential estimation of zeroes and extrema of regression functions. *IEEE Transactions in Information Theory*, 33:367–372.
- Henry, C. (1973). An existence theorem for a class of differential equations with multivalued right-hand sides. *Journal of Mathematical Analysis and Applications*, 41:179–186.
- Jiang, J. (1991). Attractors in strongly monotone flows. *Journal of Mathematical Analysis and Applications*, 162:210–222.
- Krishna, V. (1992). Learning in games with strategic complementarities. Technical report, Harvard University.
- Kushner, H. and Yin, G. (1997). *Stochastic Approximation Algorithms and Applications*. Springer Verlag, New York.
- Milgrom, P. and Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62:157–180.

- Milgrom, P. and Weber, R. (1982). A theory of auctions and competitive bidding. *Econometrica*, 50:1089–1122.
- 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: Theory and Applications, Proceedings of the Eighth World Congress of the Econometric Society*, pages 56–114, Cambridge. Cambridge University Press.
- Nelsen, R. (2006). *An Introduction to Copulas*. Springer, New York, 2nd edition.
- Pemantle, R. (1990). Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability*, 18:698–712.
- Rockafellar, R. T. and Wets, R. (1998). *Variational Analysis*. Springer Verlag, Berlin.
- Ruszczyński, A. (2006). *Nonlinear Optimization*. Princeton University Press, Princeton, NJ.
- Schuster, E. (1972). Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *Annals of Mathematical Statistics*, 43:84–8.
- 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.
- Steiner, J. and Stewart, C. (2007). Learning by similarity in coordination games. Technical report, CERGE.
- Steiner, J. and Stewart, C. (2008). Contagion through learning. *Theoretical Economics*, pages 431–458.