Tilburg University

Naive Bayesian Learning in 2 x 2 Matrix Games

Eichberger, J.; Haller, H.H.; Milne, F.

Publication date:
1990

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Eichberger, J., Haller, H. H., & Milne, F. (1990). *Naive Bayesian Learning in 2 x 2 Matrix Games*. (CentER Discussion Paper; Vol. 1990-71). CentER.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

CRM
CBM
R
for
8414
1990
mic Research

71

Discussion paper

* C I N O O 9 4 5 *

No. 9071

**NAIVE BAYESIAN LEARNING IN
2 X 2 MATRIX GAMES**

by J. Eichberger, H. Haller
and F. Milne

R20

518.55

December 1990

ISSN 0924-7815

NAIVE BAYESIAN LEARNING IN 2 X 2 MATRIX GAMES

J. Eichberger * , H. Haller ** , and F. Milne ***

October 1990

* Department of Economics, Faculty of Economics and Commerce,
University of Melbourne, Parkville, Victoria 3052, Australia

** Department of Economics, Virginia Polytechnical Institute,
Blacksburg, Virginia, U.S.A.

*** Department of Economics, Faculty of Economics and Commerce,
Australian National University, GPO Box 4, Canberra, ACT 2601,
Australia

ABSTRACT

This paper looks at sequences of two-player games where each player 'naively' assumes that his opponent follows a certain mixed strategy which can be learned by observing his behaviour. In each stage of the game players up-date their beliefs and choose an optimal response to the expected behaviour of the opponent. Provided that this learning process converges, the equilibrium beliefs form a Nash equilibrium. Hence, it is possible to derive a probability distribution on the set of Nash equilibria conditional on the players' prior beliefs. For prior distributions of the beta-distribution type, it can be shown that the learning process converges. Some examples illustrate how this fact can be used in games with multiple equilibria to obtain a probability distribution on the equilibrium set.

NAIVE BAYESIAN LEARNING IN 2 X 2 MATRIX GAMES

J. Eichberger, H. Haller, and F. Milne¹1. Introduction

The concept of a Nash equilibrium is the most common equilibrium concept used in social sciences, in particular in economics. In addition, in recent years, it has been applied successfully in biology to explain certain stable patterns of animal and plant populations (compare e.g. Maynard Smith (1982) and for a survey of the more recent developments in this field Hines (1987)). In these applications a Nash equilibrium is considered as the stable outcome of an evolutionary process. One feed-back of these developments to game theory as applied in social sciences has been a revival of interpretations of Nash equilibrium as the result of a learning process (compare e.g. Crawford (1988) and Crawford and Haller (1987)).

There are indeed two competing interpretations of Nash equilibrium: one takes a Nash equilibrium as the stable outcome of a rather unspecified learning process, while the other treats it as the only self-enforcing outcome of an unspecified negotiation process. The first of these interpretations is best exemplified by the description of equilibrium for a duopoly as set out already in Cournot (1838). According to this view duopolists choose the profit-maximising level of output taking the choice of the competitor as independent of their own action. An iteration of this process leads (if it converges) to an equilibrium where each agent has the profit-maximising choice of his output given the optimal choice of the other player, i.e. a Nash equilibrium results. Indeed as long as there is a unique Nash equilibrium of the game, this is the only natural stable situation no matter what the particular form of the adjustment process in beliefs is

¹We would like to thank Vince Crawford for helpful comments in the early stages of this paper. Hans Haller acknowledges the hospitality and financial support by CentER, KUB, and NWO.

(compare e.g. Moulin (1986) for a detailed analysis of Cournot-type adjustment processes). If there are multiple Nash equilibria, however, then the precise modeling of adjustment matters.

Since multiplicity of equilibria is a pervasive property of games and it is impossible to select among these potential outcomes of a game, another interpretation has been favoured in recent years. Presuming that agents can communicate but not commit themselves before the actual game starts, Nash equilibria are the only outcomes of such a pre-game negotiation process since they are self-policing in the sense that no agent would have an incentive to break away from such an agreement as long as every other player follows it. Note however that one would have to model the pre-game negotiations to predict which of the potential equilibria will be played in the end (compare R. Aumann (1987) as an example for this view).

It is characteristic for both of these interpretations that they point to the need of specifying some unmodeled part of the game to come to a precise prediction of the outcome of a game. Obviously, this need does not arise in games with a unique Nash equilibrium since in these cases any specification of adjustment or learning and any pre-game negotiation process will lead to the same final outcome. Put differently, a Nash equilibrium is a necessary condition for equilibrium behaviour of rational agents in any adjustment process or in any pre-game negotiation without commitment, but not a sufficient one.

Not surprisingly therefore, more than ten years of discussion have focused on the problem of "refinement" of Nash equilibrium (compare E. Van Damme (1987)) for an excellent survey of this literature). The main objective of this approach has been an attempt to impose "reasonable" additional constraints on the equilibria of general games to reduce the number of Nash equilibria preferably to one. In some special cases this attempt has proved to be extremely successful as in bargaining theory where, for the case of complete information, a subgame perfectness requirement reduces the number of Nash equilibria from a continuum to one (compare e.g. A. Rubinstein (1982)). Note however that this required explicit specification of all details of the bargaining process and that the resulting subgame perfect equilibrium depends crucially on these details (e.g. discount rates,

sequence of moves etc.). For games in general however, this approach was not capable of reducing the number of equilibria substantially, not even in bargaining games with complete information (see Fernandez and Glazer (1989), Haller (1986), and Haller and Holden (1990)).

The problem of multiplicity of equilibria is not only of theoretical interest. In nearly all applications of game theory where there is a sequence of games this problem has been formidable. If in a later stage of such a game multiplicity of equilibria prevails there is in general no way to find an optimal strategy in an earlier stage because this choice will be conditional on which equilibrium will obtain later. Examples of this type can be found particularly in Industrial Organisation applications (compare e.g. Tirole (1988, e.g. p. 345) or Diamond and Dybvig (1983) and Eichberger and Milne (1990)).

It is before this background that one has to consider the recent interest in the treatment of Nash equilibrium in biology. An evolutionary stable equilibrium has been interpreted in this context as the equilibrium of an evolutionary adjustment process. In contrast to game theory in general, there has been more emphasis on research in different evolutionary adjustment processes which might or might not converge to Nash equilibria (compare e.g. Hines (1987)). Therefore, one way to approach the problem of how to predict which Nash equilibrium will arise, instead of trying to reduce the number of equilibria, is to look for rather general learning processes which may lead to Nash equilibrium outcomes. This approach has been chosen by Harsanyi and Selten (1988) and the process which they consider is called the "tracing procedure". Though not very intuitive as a learning mechanism, it determines an adjustment process leading from some prior beliefs to a specific equilibrium. Hence, it is possible to associate different Nash equilibria with different prior beliefs.

A similar approach will be pursued in this paper for a simple but often used class of games, namely 2x2 matrix games. In contrast to Harsanyi and Selten (1988), however, this paper will focus on the rather intuitive notion of Bayesian learning. The basic idea focuses on the assumption that players who are uncertain about what strategy another player may choose will use experience from playing this (or a similar) game to predict the behaviour of

their opponent. Starting with any prior beliefs about the other player's behaviour, a player is assumed to use the information about the strategy choice of his opponent to update his beliefs about the opponent's future play.

The idea about learning from past strategy choices which is advanced in this paper implies of course a sequence of plays. In addition, it requires players who are uncertain about their opponents' behaviour. The uncertainty can be caused by the inability or unwillingness to agree on the selection of one of the several equilibria or by incomplete information concerning players' objectives. A further requirement of the learning concept forwarded here is that, despite their uncertainty, players assume that their opponents repeat certain (mixed) strategies which can be inferred from observation of their behaviour. Finally, players are required to be 'naive' in ignoring the possibility that their opponents might also learn from and respond to past strategy choices. Such naive learning seems justifiable where a player enters a social situation assuming that a certain equilibrium unknown to him has been established as a standard of behaviour.

One of the surprising results of this paper is the fact that such beliefs, if held by all players and therefore necessarily incorrect, will lead to Nash equilibrium behaviour which justifies these beliefs retrospectively. On the other hand, as outlined in the paper, the prior beliefs which agents hold influence greatly the probability of reaching one or the other of several equilibria of an underlying game with multiple Nash equilibria. It seems that such a scenario of players being uncertain about their opponents' behaviour is rather common in social science applications (compare in this context in particular Schelling (1980), chapter 4).

Recently, the idea of Bayesian learning has been used in different game theoretic contexts by Fudenberg and Kreps (1988), Canning (1989a, 1989b), Jordan (1990a, 1990b), Kalai and Lehrer (1990), and Milgrom and Roberts (1990). It is impossible to discuss these different approaches in detail here (compare Eichberger (1990) for a detailed comparison of the Canning (1989a) and Fudenberg and Kreps (1988) papers and the approach of this paper). However, a common theme of all these papers is the consideration of infinite repetitions of normal form games where agents learn in a similarly

naïve adaptive way as they do in this paper. Encouragingly, a number of fairly general results on the convergence of Bayesian learning to Nash equilibrium can be proved in a variety of different models. This indicates that the convergence result in this paper is not an isolated example, but rather represents a general property of such learning processes. On the other hand, the idea to use such a learning approach to obtain a probability distribution on the set of Nash equilibria provides a promising approach for applied work plagued by the multiplicity of Nash equilibria.

The following section will provide a characterisation of the Nash equilibrium set of 2×2 matrix games for arbitrary payoff parameters. Section 3 describes the Bayesian learning process considered and discusses the equilibrium concept. In addition, it provides results which allow to associate with any prior belief system a probability of reaching any of the possible Nash equilibria. In section 4, it is shown how these results can be used to obtain a probability of reaching each of several equilibria. The final section provides examples showing the dependence of these results on the type of prior distribution. The paper concludes with some remarks on open questions.

2. Classification of 2×2 games

Consider the following general form of a 2×2 matrix game:

		Player 2	
		L	R
Player 1	T	a_1, a_2	c_1, c_2
	B	b_1, b_2	d_1, d_2

Obviously, $S_1 = (T, B)$ and $S_2 = (L, R)$ are the sets of pure strategies for player 1 and 2 respectively. The set of mixed strategies for player 1 is given as $M_1 = \{(m_T, m_B) \in \mathbb{R}_+^2 \mid m_T + m_B = 1\}$. Similarly, one has M_2 as the set of mixed strategies of player 2. Since there are just two pure strategies for each player, one can identify each player's set of mixed strategies with the unit interval, i.e. $M_i = [0, 1]$ for $i = 1, 2$. Given this convention, $q_1 \in$

$[0,1]$ determines the mixed strategy $m_T = q_1$ and $m_B = 1 - q_1$ for player 1. Similarly, each $q_2 \in [0,1]$ defines a mixed strategy for player 2.

Note that throughout this paper pure strategies are identified with mixed strategies giving probability one to a pure strategy.

Proposition 2.1: Let $\Omega_i = a_i + d_i - b_i - c_i$ for $i=1,2$, $\theta_1 = (d_1 - c_1)$ and $\theta_2 = (d_2 - b_2)$ hold.

- (i) If $\theta_i \cdot (\Omega_i - \theta_i) > 0$ for $i=1,2$ and $\theta_1 \cdot \theta_2 < 0$ holds, then there exists no pure strategy equilibrium and one mixed strategy equilibrium, namely

$$(m_T^*, m_L^*) = (\theta_2 / \Omega_2, \theta_1 / \Omega_1) .$$

- (ii) If $\theta_i \cdot (\Omega_i - \theta_i) > 0$ for $i=1,2$ and $\theta_1 \cdot \theta_2 > 0$ holds, then there exist two pure strategy equilibria and one mixed strategy equilibrium, namely for $\theta_i > 0$, $i=1,2$,

$$(m_T^*, m_L^*) = (0,0) , (m_T^*, m_L^*) = (1,1) ,$$

$$(m_T^*, m_L^*) = (\theta_2 / \Omega_2, \theta_1 / \Omega_1) ,$$

and for $\theta_i < 0$, $i=1,2$,

$$(m_T^*, m_L^*) = (1,0) , (m_T^*, m_L^*) = (0,1) ,$$

$$(m_T^*, m_L^*) = (\theta_2 / \Omega_2, \theta_1 / \Omega_1) .$$

- (iii) If $\theta_i \cdot (\Omega_i - \theta_i) < 0$ for some i , $i=1,2$, and not $(\theta_i > 0, \theta_j = 0)$ or $(\theta_i < 0, \theta_j = \Omega_j)$ for $j \neq i$ holds, then there exists exactly one pure strategy Nash equilibrium and no mixed strategy equilibrium.

- (iv) In all other cases there exist either two pure strategy equilibria, one of which contains dominated strategies, or a continuum of mixed strategy equilibria containing at least one pure strategy equilibrium.

A proof for this proposition can be found in the appendix. Note that case (iv) is non-generic in the space of payoff parameters. Hence this proposition illustrates the well-known result that, in general, there are an odd number of locally unique Nash equilibria (compare Van Damme (1987), Theorem 2.6.2, p. 48).

3. Bayesian learning and Nash equilibrium

This section analyses the behaviour of a player who does not know what strategy his opponent will choose either because of incomplete information in regard to this player's payoff or in regard to what equilibrium this player will play (provided there are more than one).

If player i believes that his opponent pursues a certain strategy and if the game is played repeatedly, player i can use the observed behaviour of his opponent to learn about this player's behaviour even if he is ignorant as to his payoff. Note in particular that the sequence of games played is not considered as one supgame.

3.1 The learning process

A natural way to think about learning is a Bayesian learning process. According to this view, player i starts with a prior belief, a prior distribution, on the set of mixed strategies of player j , M_j , $j \neq i$. Player i chooses his strategy conditional on this prior belief and uses his observation of the opponent's actual strategy choice to update his beliefs.

Denote by $\Phi(M_i)$ the set of all probability distributions on the set M_i of player i 's mixed strategies. For $(m_T, m_B) = (q_1, 1-q_1)$ for player 1 and $(m_L, m_R) = (q_2, 1-q_2)$ for player 2, one can identify M_i with the unit interval for both players. Hence, $\Phi(M_i)$ can be represented by $\Phi([0,1])$, the set of probability distributions on $[0,1]$.

A belief of player i on player j 's strategy choice is a probability distribution $f \in \Phi([0,1])$. Suppose f_j^0 is the prior distribution of player i , reflecting his beliefs as to the mixed strategy used by his opponent. The expected mixed strategy of player j in the opinion of player i is given as

$$\mu_j^0 = \int_{[0,1]} q \, d f_j^0,$$

where μ_j^0 denotes the expected probability of player j using his first strategy given player i 's prior distribution f_j^0 . Note that $\mu_j^0 \in M_j$ holds. Hence, one can view the product measure of f_j^0 and μ_j^0 as a joint probability

distribution on $M_j \times S_j$. Information about a choice $s_j \in S_j$ can therefore be used to update the beliefs of player i in regard to player j 's strategy choice.

If player i learns that player j has played the pure strategy s_j , he can use this information to update his beliefs as follows:

for any measurable set A in $[0,1]$,

$$f_j^1(A|s_j) = \left(\int_{[0,1]} m(s_j) df_j^0 \right)^{-1} \cdot \left(\int_A m(s_j) df_j^0 \right)$$

where $m(s_j) = q_j$ if s_j was player j 's first strategy and $m(s_j) = 1 - q_j$ if s_j was player j 's second strategy. Provided player i 's prior beliefs f_j^0 are not concentrated on one of the end points of the unit interval, i.e. he does not believe with certainty that player j plays one of his pure strategies, this updating procedure is well-defined.

Using f_j^1 as new beliefs player i can determine his expectations about the mixed strategy played by player j , μ_j^1 , and choose his strategy accordingly. Every round of the game provides new information through observation of the opponent's behaviour. Therefore, the question arises whether this learning process converges and what equilibrium strategy choices will obtain in the limit.

Clearly, the choice of a prior distribution f_j^0 is crucial for this induced learning process. In this paper, it will be assumed that both players have a prior distribution of the beta distribution family. Such a distribution is determined by two positive parameters, α and β , as follows:

$$F(x|\alpha, \beta) = G(\alpha, \beta)^{-1} \cdot \int_0^x y^{\alpha-1} \cdot (1-y)^{\beta-1} dy \quad \text{for } x \in [0,1],$$

where $G(\alpha, \beta) = \int_0^1 y^{\alpha-1} \cdot (1-y)^{\beta-1} dy$ holds. The expected value of F is simply $\mu(\alpha, \beta) = \alpha/(\alpha+\beta)$. A property of this distribution which is useful and easy to check is the fact that $\mu(\alpha, \beta) = G(\alpha+1, \beta)/G(\alpha, \beta)$ holds.

Let (α_1, β_1) be the parameters characterising the prior belief of player 2 about player 1's choice of his first strategy, and similarly (α_2, β_2) the

parameters characterising player 1's prior beliefs. Hence player 1 expects player 2 to choose his first strategy L with probability $\alpha_2/(\alpha_2+\beta_2)$ and his strategy R with probability $\beta_2/(\alpha_2+\beta_2) = 1 - \alpha_2/(\alpha_2+\beta_2)$. Similarly player 2 expects player 1 to choose T with probability $\alpha_1/(\alpha_1+\beta_1)$ and B with probability $\beta_1/(\alpha_1+\beta_1)$.

Suppose that player 1 observes that player 2 chooses L. Then one gets the following updated distribution

$$F(x|L; \alpha_2, \beta_2) \\ = [G(\alpha_2, \beta_2)^{-1} \cdot \int_0^1 y^{\alpha_2} \cdot (1-y)^{\beta_2-1} dy]^{-1} \cdot [G(\alpha_2, \beta_2)^{-1} \cdot \int_0^x y^{\alpha_2} \cdot (1-y)^{\beta_2-1} dy]$$

for $x \in [0,1]$, with expected value $\mu(L; \alpha_2, \beta_2) = (\alpha_2+1)/(\alpha_2+\beta_2+1)$. Similarly, if R is observed $\mu(R; \alpha_2, \beta_2) = \alpha_2/(\alpha_2+\beta_2+1)$ must hold. The same argument leads to $\mu(T; \alpha_1, \beta_1) = (\alpha_1+1)/(\alpha_1+\beta_1+1)$ and $\mu(B; \alpha_1, \beta_1) = \alpha_1/(\alpha_1+\beta_1+1)$ as the updated expectation of player 2 in regard to player 1's choice of his strategy T.

As shown above the optimal mixed strategy response of any player 1 depends upon the mixed strategy choice of player j, $i \neq j$, i.e.

$$(3.1) \quad r_i(q_j) \in \begin{cases} (1) & \text{for } [\Omega_i q_j - \theta_i] > 0 \\ [0,1] & \text{for } [\Omega_i q_j - \theta_i] = 0 \\ (0) & \text{for } [\Omega_i q_j - \theta_i] < 0 \end{cases}$$

where Ω_i and θ_i are defined as in proposition 2.1. From the linearity of the payoff function it follows that beliefs about the opponent j's strategy choice affect player i's choice through their expected value only. Denote by $q_j(s_j^{t-1})$ the expected mixed strategy choice of player j after observing strategy s_j^{t-1} . Substituting these expectations in 3.1 for q_j yields the best reply of player i in period t after observing s_j^{t-1} .

Since for $[\Omega_i q_j - \theta_i] = 0$ the optimal choice of player i is indetermined, it will be assumed in this case that player i plays his first pure strategy with an arbitrary but constant probability $\sigma_i \in [0,1]$. Note that this arbitrariness does not matter for the behaviour of the learning process

(since it is a zero probability event), but has some implications for the definition and interpretation of equilibrium.

Denote by

$$k_1(s_1) = \begin{cases} 1 & \text{if } s_1 = T \\ 0 & \text{if } s_1 = B \end{cases} \quad k_2(s_2) = \begin{cases} 1 & \text{if } s_2 = L \\ 0 & \text{if } s_2 = R \end{cases}$$

the indicator functions for the first strategy of each player and by $n_i(t) = \sum_{r=1}^t k_i(s_1^r)$ for $i = 1, 2$. The expected mixed strategy of player i according to player j 's beliefs in period t after observation of $(s_1^r)_{r=1}^{t-1}$ can now be written in the following convenient form:

$$(3.2) \quad q_i(n_i(t-1); \alpha_i, \beta_i) = (\alpha_i + n_i(t-1)) / (\alpha_i + \beta_i + t-1).$$

Note that for $t = 1$ the expected mixed strategy is based on α_i and β_i , i.e. the prior beliefs, only, while for any $t > 1$ it depends on the whole history of observed strategy choices $(s_1^r)_{r=1}^{t-1}$.

Lemma 3.1 collects some useful properties of this learning process.

Lemma 3.1: For any $\alpha_i, \beta_i > 0$, let $q_i(t) = q_i(n_i(t-1); \alpha_i, \beta_i)$.

(i) If there exists an integer T such that $k_i(s_1^r) = 1$ ($k_i(s_1^r) = 0$) holds for all $r \geq T$, then $\lim_{t \rightarrow \infty} q_i(t) = 1$ ($= 0$).

(ii) $q_i(t) = q_i(t-1) + (\alpha_i + \beta_i + t-1)^{-1} \cdot [k_i(s_1^{t-1}) - q_i(t-1)]$.

Proof: Obvious. ■

Before proceeding to the next section a remark on the learning process is in order. The learning process induced by the beta distribution resembles the so-called "fictitious play" as introduced in the literature by Robinson (1951) and analysed by Brown (1951), Shapley (1964), and Rosenmüller (1971). Though the process was mainly used as an iterative method to derive the Nash equilibrium of a zero-sum game, some of the results from this literature carry over to our more general and differently motivated approach.

3.2. Nash equilibrium and learning

Given his prior beliefs α_i and β_i and after observing $(s_i^r)_{r=1}^{t-1}$, the best reply function of player j (a selection from the best reply correspondence 3.1) can be written as

$$(3.3) \quad r_j'(q_i(n_i(t-1); \alpha_i, \beta_i)) = \begin{cases} 1 & \text{for } [\Omega_j(\alpha_i + n_i(t-1)) / (\alpha_i + \beta_i + t-1) - \theta_j] > 0 \\ \sigma_j & \text{for } [\Omega_j(\alpha_i + n_i(t-1)) / (\alpha_i + \beta_i + t-1) - \theta_j] = 0 \\ 0 & \text{for } [\Omega_j(\alpha_i + n_i(t-1)) / (\alpha_i + \beta_i + t-1) - \theta_j] < 0. \end{cases}$$

Note that the best reply function has been written as the optimal probability of playing a player's first strategy.

In each round of the game, each player has beliefs based on past observations, chooses a mixed strategy which maximises his expected payoff, and observes the pure strategy actually played by his opponent. A sequence of games is therefore determined by the sequence

$$(q_1(n_1(t-1); \alpha_1, \beta_1), s_1^t, q_2(n_2(t-1); \alpha_2, \beta_2), s_2^t)_{t=1}^{\infty}.$$

Definition: For a given two-player 2x2 matrix game and prior beliefs $(\alpha_i, \beta_i)_{i=1,2}$, an *equilibrium with learning* is a point expectation for each player $(q_1^{\infty}, q_2^{\infty})$, with $q_i^{\infty} = \lim_{t \rightarrow \infty} q_i(n_i(t-1); \alpha_i, \beta_i)$ for $i=1,2$, such that

$$\begin{aligned} q_1^{\infty} &\in r_1(q_2^{\infty}), & \text{and} \\ q_2^{\infty} &\in r_2(q_1^{\infty}) & \text{hold.} \end{aligned}$$

According to this definition, an equilibrium with learning is characterised by limiting beliefs about the opponents' behaviour which coincide with actual behaviour. Hence, rational expectations about the opponent's strategy choice obtain in an equilibrium with learning.

Note however that the definition of an equilibrium with learning is in terms of the best response correspondence 3.1 and not the selection given in 3.3. This selection is discontinuous at the mixed strategy equilibrium strategy. Moreover, a player is hardly ever indifferent about his strategy choice throughout the game. Hence, the probability σ_1 has no role to play in the determination of limiting beliefs q_1^∞ . If q_1^∞ corresponds to a pure strategy equilibrium, then this fact will be learned by player j eventually. On the other hand, if q_1^∞ corresponds to a mixed strategy equilibrium, then it need not coincide with a best response σ_1 which has been selected from the correspondence at that point. This follows since a player may switch from one pure strategy to the other (thus, approximating the mixed strategy in relative frequency), but will never find it optimal to stick to his mixed strategy. As will be shown below, the learning process will eventually converge to the mixed strategy equilibrium, even if σ_1 is not a mixed equilibrium strategy.

The following proposition shows that it is sufficient for existence of an equilibrium with learning to show that learning actually converges.

Proposition 3.2: If (q_1^∞, q_2^∞) exists, then it is an equilibrium with learning.

Proof: The result follows immediately from the continuity of the reaction functions for all q_1^∞ satisfying $[\Omega_j \cdot q_1^\infty - \theta_j] \neq 0$. Assume therefore that $[\Omega_j q_1^\infty - \theta_j] = 0$ holds. From (3.1), player j is indifferent about his strategy. Hence, q_j^∞ is an optimal choice for player j . ■

The main question therefore concerns the existence of the limits (q_1^∞, q_2^∞)

This problem will be considered for the three generic cases of proposition 2.1 ((i), (ii), and (iii)).

Proposition 3.3: If there is a unique pure strategy Nash equilibrium, then a unique equilibrium with learning exists which coincides with the pure strategy Nash equilibrium.

Proof: W.l.o.g. assume $\theta_1(\Omega_1 - \theta_1) < 0$, $\theta_1 > 0$, and $\Omega_2 > \theta_2 \neq 0$. Then $(q_1, q_2) = (1, 1)$ is the unique Nash equilibrium. From (3.1) it is clear that the best

response of player 1 is constant for all periods t and equal to 1, i.e.

$r_1(n_2(t-1); \alpha_2, \beta_2) = 1$ for all t and all (α_2, β_2) . Observing that player 1 plays T in all periods, i.e. $n_1(t) = t$, player 2 expects $q_1(n_1(t-1); \alpha_1, \beta_1) = (\alpha_1 + t - 1) / (\alpha_1 + \beta_1 + t - 1)$ as the mixed strategy choice of player 1. Obviously,

$q_1^\infty = \lim_{t \rightarrow \infty} q_1(n_1(t-1); \alpha_1, \beta_1) = \lim_{t \rightarrow \infty} (\alpha_1 + t - 1) / (\alpha_1 + \beta_1 + t - 1) = 1$. Hence, $q_1^\infty = 1$

holds. On the other hand, since $\theta_2 / \Omega_2 < 1$ holds and $q_1(n_1(t-1); \alpha_1, \beta_1)$ is an increasing sequence which converges to 1, there exists $T \geq 0$ such that

$q_1(n_1(t-1); \alpha_1, \beta_1) > \theta_2 / \Omega_2$ for all $t \geq T$ holds. Therefore, for $t \geq T$, the best response of player 2 is to play 1, i.e. $r_2(n_1(t-1); \alpha_1, \beta_1) = 1$ holds.

This in turn implies $q_2(n_2(t-1); \alpha_2, \beta_2) = (\alpha_2 + n_2(T) + t - T - 1) / (\alpha_2 + \beta_2 + t - 1)$ for all $t \geq 0$. Thus, $q_2^\infty = \lim_{t \rightarrow \infty} q_2(n_2(t-1); \alpha_2, \beta_2) = 1$ follows. Consequently, $q_2^\infty = 1$

is true. Similar reasoning obtains in the other cases of games with a unique pure strategy equilibrium. ■

Proposition 3.4: If there is a unique mixed strategy Nash equilibrium, then a unique equilibrium with learning exists which coincides with the mixed strategy Nash equilibrium.

Proof: Rosenmüller (1971), Satz 2.2 and Satz 3.10. ■

Remark: Rosenmüller's Satz 3.10 is actually more general than proposition 3.4 above. In particular, it covers the result of proposition 3.5 too. The proof given below is however quite different and provides some insight in the dynamics of the adjustment.

Proposition 3.5: If there are two pure strategy Nash equilibria and one mixed strategy Nash equilibrium, then for each set of prior beliefs (α_i, β_i) , $i=1,2$, an equilibrium with learning exists.

Proof: Without loss of generality, assume that $(\Omega_1 - \theta_1) \cdot \theta_1 > 0$ and $\theta_1 > 0$ for $i = 1, 2$ hold, i.e.

$$d_1 - c_1 > 0, a_1 - b_1 > 0, d_2 - b_2 > 0, \text{ and } a_2 - c_2 > 0.$$

In this case, $(m_1^*, m_2^*) = (0, 0)$, $(m_1^*, m_2^*) = (1, 1)$, and $(m_1^*, m_2^*) = (\bar{m}_T, \bar{m}_L)$,

with $(\bar{m}_T, \bar{m}_L) = ((d_2 - b_2) / (a_2 + d_2 - b_2 - c_2), (d_1 - c_1) / (a_1 + d_1 - b_1 - c_1))$, are the three Nash equilibria.

Consider the following three cases:

- (i) For $\alpha_1/(\alpha_1+\beta_1) > (d_2-b_2)/(a_2+d_2-b_2-c_2)$
and $\alpha_2/(\alpha_2+\beta_2) > (d_1-c_1)/(a_1+d_1-b_1-c_1)$, both players will choose their first pure strategy with probability one from the beginning. Since, by lemma 3.1, $q_1(t) > q_1(t-1)$ follows for all t if the agent j observes agent i choosing his first strategy, both agents will choose their first strategy throughout the game. Hence, learning as well as the relative frequency of strategy choices converges to one, i.e. $q_1^\infty = 1$ holds for both players.

- (ii) An essentially identical argument shows that for

$$\alpha_1/(\alpha_1+\beta_1) < (d_2-b_2)/(a_2+d_2-b_2-c_2) \quad \text{and} \\ \alpha_2/(\alpha_2+\beta_2) < (d_1-c_1)/(a_1+d_1-b_1-c_1)$$

learning and relative frequency of strategy choices converge to zero, i.e. $q_1^\infty = 0$ must hold for both players.

- (iii) Suppose $\alpha_1/(\alpha_1+\beta_1) > (d_2-b_2)/(a_2+d_2-b_2-c_2)$
and $\alpha_2/(\alpha_2+\beta_2) < (d_1-c_1)/(a_1+d_1-b_1-c_1)$ holds in the beginning.
In this case,

$q_1(t) < q_1(t-1)$ as long as $q_1(t-1) > (d_2-b_2)/(a_2+d_2-b_2-c_2)$ and
 $q_2(t) > q_2(t-1)$ as long as $q_2(t-1) < (d_1-c_1)/(a_1+d_1-b_1-c_1)$ holds.

Since, by lemma 3.1, monotonic growth will lead to

$q_1(r_1) < (d_2-b_2)/(a_2+d_2-b_2-c_2)$ and $q_2(r_2) > (d_1-c_1)/(a_1+d_1-b_1-c_1)$ in finite time, three possibilities arise:

- for $r_1 > r_2$,
 $q_1(r_2) < (d_2-b_2)/(a_2+d_2-b_2-c_2)$ and $q_2(r_2) < (d_1-c_1)/(a_1+d_1-b_1-c_1)$
holds and convergence to $q_1^\infty = 0$ for both players follows by the same argument as under (ii);
- for $r_1 < r_2$,
 $q_1(r_1) > (d_2-b_2)/(a_2+d_2-b_2-c_2)$ and $q_2(r_1) > (d_1-c_1)/(a_1+d_1-b_1-c_1)$
holds and convergence to $q_1^\infty = 1$ for both players follows by the same argument as under (i);
- for $r_1 = r_2$,
 $q_1(r_1) < (d_2-b_2)/(a_2+d_2-b_2-c_2)$ and $q_2(r_1) > (d_1-c_1)/(a_1+d_1-b_1-c_1)$
holds. Hence, $q_1(t) > q_1(t-1)$ and $q_2(t) < q_2(t-1)$ follow for $t \geq r_1$. By the same argument as before this adjustment will lead to a reversal of the learning direction and, depending on who

changes first, to $q_1^\infty = 1$ for both players or to $q_1^\infty = 0$ for both players or to simultaneous reversal. Suppose that there is a subsequence of periods where both players change their behaviour simultaneously. Since, by lemma 3.1,

$$|q_1(t) - q_1(t-1)| < [\alpha_1 + \beta_1 + t - 1]^{-1}$$

holds, the alternating sequence must converge to

$$q_1^\infty = (d_2 - b_2) / (a_2 + d_2 - b_2 - c_2) \quad \text{and}$$

$$q_1^\infty = (d_1 - c_1) / (a_1 + d_1 - b_1 - c_1)$$

which corresponds to the mixed strategy equilibrium. ■

The preceding propositions show that in all generic cases an equilibrium with learning exists for prior distributions from the beta family. Though we strongly suspect that, for 2x2 matrix games, this property carries over to any type of prior distribution having $[0,1]^2$ as support, this needs further research. If the prior distributions are concentrated on some subset of $[0,1]^2$ this proposition is no longer true as the examples in section 5 demonstrate.

4. Selection of equilibria in terms of prior beliefs

The analysis of the previous section points to the possibility to select among different Nash equilibria in regard to the prior beliefs of the players. Put differently, will a particular set of prior beliefs give rise to a particular Nash equilibrium?

In this section, the case of games with more than one Nash equilibrium will be considered only. Without loss of generality, assume that $(\Omega_1 - \theta_1) \cdot \theta_1 > 0$ and $\theta_1 > 0$ for $i = 1, 2$ hold, i.e.

$$d_1 - c_1 > 0, \quad a_1 - b_1 > 0, \quad d_2 - b_2 > 0, \quad \text{and} \quad a_2 - c_2 > 0.$$

In this case, $(m_T^*, m_L^*) = (0, 0)$, $(m_T^*, m_L^*) = (1, 1)$, and $(m_T^*, m_L^*) = (\bar{m}_T, \bar{m}_L)$,

with $(\bar{m}_T, \bar{m}_L) = ((d_2 - b_2) / (a_2 + d_2 - b_2 - c_2), (d_1 - c_1) / (a_1 + d_1 - b_1 - c_1))$, are the three Nash equilibria of the game.

The existence results in section 3 are not sufficient to answer this question, since it is not obvious that learning leads to a unique equilibrium for arbitrary prior beliefs. Nevertheless, for some prior beliefs learning leads to a unique equilibrium. Note that the space of all prior beliefs for the class of beta distribution functions equals $\mathbb{R}_+^4$ and consider the following subsets:

$$A = \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in \mathbb{R}_+^4 \mid \alpha_1/(\alpha_1 + \beta_1) > \bar{m}_T, \alpha_2/(\alpha_2 + \beta_2) > \bar{m}_L\},$$

$$B = \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in \mathbb{R}_+^4 \mid \alpha_1/(\alpha_1 + \beta_1) < \bar{m}_T, \alpha_2/(\alpha_2 + \beta_2) < \bar{m}_L\}.$$

Proposition 4.1:

- (i) For $(\alpha_1, \beta_1, \alpha_2, \beta_2) \in A$, $q_i^\infty = 1$, $i = 1, 2$, follows.
- (ii) For $(\alpha_1, \beta_1, \alpha_2, \beta_2) \in B$, $q_i^\infty = 0$, $i = 1, 2$, follows.

Proof: This follows by the same argument as in (i) and (ii) of the proof of proposition 3.5. ■

The following examples will indicate that for other initial parameters a unique equilibrium with learning is no longer guaranteed.

Example 4.1 :

Consider the following game:

		Player 2	
		L	R
Player 1	T	1 , 1	0 , 0
	B	0 , 0	1 , 1

Obviously, (T,L) and (B,R) are the pure strategy equilibria and $(q_1, q_2) = (1/2, 1/2)$ is the mixed strategy equilibrium.

Suppose that $(\alpha_1, \beta_1) = (1, 4)$ and $(\alpha_2, \beta_2) = (4, 1)$ are the prior beliefs of the players leading to point expectations as to the other player's mixed strategy $(q_1(1), q_2(1)) = (1/5, 4/5)$. Given these beliefs, the game has player 1 choose T with probability one and player 2 choose R with

probability one. This leads to updated expectations of $(q_1(2), q_2(2)) = (2/6, 4/6)$ and a repeated choice of (T,R) etc. The consistent observation of player 1 choosing T makes player 2 revise his expectations of player 1's choice upwards. Similarly, player 1 revises his beliefs about player 2's strategy downwards. In round 4, one has $(q_1(4), q_2(4)) = (4/8, 4/8)$ and each player assumes the other to play the mixed strategy $(1/2, 1/2)$ which makes him indifferent between his choices. Note that learning has converged to the mixed strategy equilibrium in four steps. Nevertheless, this does not represent a stable equilibrium situation, since the choices of the agents will lead to renewed revision of beliefs.

In stage four each player picks his strategy according to (σ_1, σ_2) as indicated above. Hence, (T,L) is chosen with probability $\sigma_1 \cdot \sigma_2$ and (B,R) with probability $(1-\sigma_1) \cdot (1-\sigma_2)$. In addition, (T,R) and (B,L) are observed with probability $\sigma_1 \cdot (1-\sigma_2)$ and $(1-\sigma_1) \cdot \sigma_2$ respectively. These four possibilities induce four different changes in beliefs and corresponding reactions. For (T,L) the updated expectations become $(q_1(5), q_2(5)) = (5/9, 5/9)$ and both players choose (T,L) for the future plays leading to the equilibrium with learning $(q_1^\infty, q_2^\infty) = (1, 1)$. Similarly, observation of (B,R) leads to $(q_1^\infty, q_2^\infty) = (0, 0)$. Finally, for (T,R) and updated beliefs $(q_1(5), q_2(5)) = (5/9, 4/9)$, successive behaviour and updating will lead back to $(1/2, 1/2)$ in step six. A similar sequence of learning results for the case where (B,L) is chosen in round 4. As a result, one has after four plays,

- convergence to the pure strategy equilibrium (T,L) with probability $\sigma_1 \cdot \sigma_2$,
- convergence to the pure strategy equilibrium (B,R) with probability $(1-\sigma_1) \cdot (1-\sigma_2)$,
- convergence to the mixed strategy equilibrium in a finite number of steps with probability $(\sigma_1 + \sigma_2 - 2\sigma_1\sigma_2)$.

Since the mixed strategy equilibrium beliefs will be reached in a finite number of steps each time, one can pick a subsequence of rounds in which this indifference situation obtains. In each of these rounds there is a $\sigma_1 \cdot \sigma_2$ chance that learning will lead to the pure strategy equilibrium (T,L) and a $(1-\sigma_1) \cdot (1-\sigma_2)$ chance that updating leads to (B,R). Hence, for any $(\sigma_1, \sigma_2) \in (0, 1)^2$,

$$\begin{aligned} \text{Prob}((T,L)) &= (\sigma_1 \cdot \sigma_2) \cdot \left[\sum_{k=0}^{\infty} (\sigma_1 + \sigma_2 - 2\sigma_1\sigma_2)^k \right] \\ &= (\sigma_1 \cdot \sigma_2) \cdot [1 - \sigma_1 - \sigma_2 + 2\sigma_1\sigma_2]^{-1} \end{aligned}$$

is the probability of reaching the pure strategy equilibrium (T,L) and

$$\text{Prob}((B,R)) = (1-\sigma_1) \cdot (1-\sigma_2) \cdot [1 - \sigma_1 - \sigma_2 + 2\sigma_1\sigma_2]^{-1}$$

is the probability of reaching (B,R). ■

Obviously, the mixed strategy equilibrium is unstable with respect to learning if it is reached in finite time. As the example shows, one of the two pure strategy equilibria will result with probability one. A slight modification of this example will illustrate that the mixed strategy equilibrium may be approached in the limit only.

Example 4.2:

Consider the same payoff matrix as in the previous example, but let the initial beliefs be given by $(\alpha_1, \beta_1) = (0.2, 0.8)$ and $(\alpha_2, \beta_2) = (0.8, 0.2)$. The following table gives the development of beliefs and the associated strategy choices for the first five rounds.

t	$q_1(t)$	$q_2(t)$	$s_1(t)$	$s_2(t)$
1	0.2	0.8	T	R
2	(1.2)/2	(0.8)/2	B	L
3	(1.2)/3	(1.8)/3	T	R
4	(2.2)/4	(1.8)/4	B	L
5	(2.2)/5	(2.8)/5	T	R

It is obvious, that updating will make $q_1(t)$ converge to 0.5 for both players but will never actually reach this belief. Notice that in this case the relative frequency of the pure strategy choices converges to 0.5, the mixed strategy equilibrium. In this sense both players get more and more convinced that the opponent is playing the mixed strategy 0.5, even if they actually choose a pure strategy in each period.

As the previous examples indicate, the following cases may occur:

- (a) Either the equilibrium with learning is one of the three possible Nash equilibria with certainty, or
- (b) the equilibrium with learning is the pure strategy equilibrium (1,1) with probability $p(\sigma_1, \sigma_2)$ and the pure strategy equilibrium (0,0) with probability $(1-p(\sigma_1, \sigma_2))$.

Note that $p(\cdot)$ is a function of the randomisation which the players use in case of indifference.

It is possible, therefore, to partition the space of all prior distributions IR_+^4 into the following subsets:

$$\begin{aligned}
 Q_1 &= \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in IR_+^4 \mid (q_1^\infty, q_2^\infty) = (1,1) \text{ with probability } 1\} \\
 Q_0 &= \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in IR_+^4 \mid (q_1^\infty, q_2^\infty) = (0,0) \text{ with probability } 1\} \\
 Q_m &= \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in IR_+^4 \mid (q_1^\infty, q_2^\infty) = (\bar{m}_T, \bar{m}_L) \text{ with probability } 1\} \\
 Q_{10} &= \{(\alpha_1, \beta_1, \alpha_2, \beta_2) \in IR_+^4 \mid (q_1^\infty, q_2^\infty) = (1,1) \text{ with probability } p(\cdot) \\
 &\quad \text{and } (q_1^\infty, q_2^\infty) = (1,1) \text{ with probability } 1 - p(\cdot)\}.
 \end{aligned}$$

Given any probability distribution on the space of parameters for the prior distributions IR_+^4 , one can determine the probability of each Nash equilibrium. Let μ be any probability measure on IR_+^4 , then

$$\begin{aligned}
 \text{Prob} \{ (q_1^\infty, q_2^\infty) = (1,1) \} &= \mu(Q_1) + \mu(Q_{10}) \cdot p(\sigma_1, \sigma_2), \\
 \text{Prob} \{ (q_1^\infty, q_2^\infty) = (0,0) \} &= \mu(Q_0) + \mu(Q_{10}) \cdot (1 - p(\sigma_1, \sigma_2)), \\
 \text{Prob} \{ (q_1^\infty, q_2^\infty) = (\bar{m}_T, \bar{m}_L) \} &= \mu(Q_m). \quad \blacksquare
 \end{aligned}$$

Obviously, these probabilities sum to one. Though it may not be possible to predict with certainty which Nash equilibrium will be "learned" for arbitrary prior beliefs, it is at least possible to predict that for given prior beliefs, one or the other equilibrium is more likely to arise. Since one can indicate the probability of a certain equilibrium arising, one can

evaluate the expected outcome of a game with multiple equilibria. This is particularly useful in multi-stage games where the optimal choice of a variable in the earlier stages requires knowledge of the result of a game in a later stage.

An additional advantage of this approach to equilibrium selection is the possibility to give a meaning to statements of the following type:

"If player 1 is more optimistic about player 2 playing L, then the equilibrium should be (T,L)."

We conclude this section with an example.

Example 4.3:

Consider the same coordination game as in example 4.1. with $(\sigma_1, \sigma_2) = (0.5, 0.5)$. Assume that α_1 and β_1 are integers for both players and that $(\alpha_1, \beta_1, \alpha_2, \beta_2) \leq (2, 2, 2, 2)$ holds. The following table gives the probability of the Nash equilibrium (1,1) being learned. Note that Q_m is empty in this case.

(α_1, β_1)	(α_2, β_2)	(1,1)	(1,2)	(2,1)	(2,2)
(1,1)		0.5	0	1	0.5
(1,2)		0	0	0.5	0
(2,1)		1	0.5	1	1
(2,2)		0.5	0	1	0.5

Assume further that each of the 16 possible beliefs occurs with the same probability. The probability of the pure strategy equilibrium (1,1) equals the sum of the entries of this table divided by 16, i.e. 0.5. Hence, each of the two pure strategy equilibria will be reached with probability 1/2. ■

5. Counterexamples for alternative distributions

In this section a few examples will demonstrate that prior distributions other than beta distributions will affect the convergence results in particular by constraining the set of possible prior beliefs. Since Bayesian

learning as described in section 3.1 never leads to expectations outside the support of the distribution f_1^0 , one can stop the player from ever learning the behaviour of his opponent by arbitrarily restricting the support of the prior distribution.

Example 5.1:

Consider the game of example 3.1 once again and assume that player 1 believes that it is equally likely that his opponent plays L either with probability 0.25 or with probability 0.33. Similarly, player 2 assumes with probability 0.5 that player 1 chooses T with probability 0.75 and with probability 0.5 that he plays T with probability 0.67. The following table shows the first 10 steps of adjustment.

t	$s_1(t)$	$q_1(t)$	$s_2(t)$	$q_2(t)$
1	B	0.71	T	0.29
2	B	0.70	T	0.30
3	B	0.70	T	0.30
4	B	0.69	T	0.31
5	B	0.69	T	0.31
6	B	0.69	T	0.31
7	B	0.68	T	0.32
8	B	0.68	T	0.32
9	B	0.68	T	0.32
10	B	0.68	T	0.32

As the table shows, both players adjust their prior expectations in response to the observed strategy choice of their opponent: player 1 upwards and player 2 downwards. Since the expected probability of player 1 playing T can never fall below 0.67 and the expected probability of player 2 choosing L can never exceed 0.33, learning leads to beliefs $(q_1^\infty, q_2^\infty) = (0.67, 0.33)$. These beliefs however will never sustain a Nash equilibrium.

Figure 5.1 shows the best response correspondences of the two players, the support of the possible belief distribution (the square), and the limiting belief (⊙).

FIGURE 5.1

Example 5.1 illustrates that constraints on prior beliefs may prevent players from learning to play Nash equilibrium strategies, this need not be the case as the following modification of example 5.1 shows.

Example 5.2:

Consider the same coordination game as in the previous example. The following assumptions on prior beliefs are supposed to hold:

- (i) player 1 believes that player 2 plays mixed strategy $m_L = 0.25$ with probability 0.3, $m_L = 0.33$ with probability 0.3, and $m_L = 0.60$ with probability 0.4;
- (ii) player 2 believes that player 1 plays mixed strategy $m_T = 0.75$ with probability 0.4, $m_T = 0.67$ with probability 0.3, and $m_T = 0.40$ with probability 0.3.

This leads to the following learning process:

t	$s_1(t)$	$q_1(t)$	$s_2(t)$	$q_2(t)$
1	B	0.62	T	0.41
2	B	0.56	T	0.47
3	T	0.51	T	0.52
4	T	0.55	T	0.56
5	T	0.59	T	0.58
6	T	0.63	T	0.59
7	T	0.66	T	0.59
8	T	0.69	T	0.60
9	T	0.70	T	0.60
10	T	0.71	T	0.60

Obviously, beliefs converge in this case to $(q_1^\infty, q_2^\infty) = (0.75, 0.60)$. Note that this does not represent an equilibrium with learning as defined in section 3.2. Nevertheless, one observes each player choosing a Nash equilibrium strategy. Thus, constraints on prior beliefs need not prevent players from converging to Nash equilibrium play even if they rule out learning of the true behaviour. Figure 5.2 depicts this example.

FIGURE 5.2

6. Concluding remarks

The concept of learning and convergence to Nash equilibrium as outlined in this paper is far from satisfying so far. A number of technical questions remain open. To name but a few:

- What assumptions on the form of prior distributions are sufficient to guarantee existence of an equilibrium with learning?
- To what extent can the results obtained in the 2x2 matrix game be extended to games with more than two players and more general strategy spaces?
- How is the speed of adjustment towards Nash equilibrium play related to the structure of the underlying game?

A partial answer to the second question can be found in papers by Shapley (1964) and Rosenmüller (1971) on so-called fictitious play. Shapley (1964) shows in an example that the iterative process of learning induced by the beta distribution family can lead to limit cycles which are different from the unique mixed strategy equilibrium of the respective game if players have three strategies available. The recent papers by Jordan (1990a, 1990b) indicate, however, that general convergence results are possible.

But from a conceptional point of view as well the analysis of the paper raises a number of questions:

- Is Bayesian learning as described in this paper an adequate concept of learning in games?
- For what class of games is such a learning approach to equilibrium selection adequate?

In this sense, it seems as if this paper rather raises questions for further research than answering them.

Nevertheless, given the impasse reached in the analysis of games with multiple Nash equilibria and the need to find an answer in applied work to the question which equilibrium will come about, this new approach seems to be worth pursuing. In addition, for games with incomplete information about

an opponent's payoff, it seems intuitively reasonable to take past experience about this agents behaviour into account for predictions of his future behaviour (compare Schelling (1980), p. 103). In particular, it should be noted that the approach proposed in this paper does not preclude integration of prior information about an opponent into beliefs if it is available.

APPENDIX

Proof of proposition 2.1:

To avoid indices for players, let $(p, q) \in [0, 1]^2$ define mixed strategies for player 1 and 2 respectively. It is a straightforward calculation to show that the expected payoffs can be written as

$$\begin{aligned} u_1(p, q) &= (q \cdot b_1 + (1-q) \cdot d_1) + [\Omega_1 \cdot q - \theta_1] \cdot p \text{ and} \\ u_1(p, q) &= (p \cdot c_2 + (1-p) \cdot d_2) + [\Omega_2 \cdot p - \theta_2] \cdot q. \end{aligned}$$

Obviously, the best reaction correspondences of both players are

$$r_1(q) = \begin{cases} (1) & \text{for } [\Omega_1 \cdot q - \theta_1] > 0 \\ [0, 1] & \text{for } [\Omega_1 \cdot q - \theta_1] = 0 \\ (0) & \text{for } [\Omega_1 \cdot q - \theta_1] < 0 \end{cases}$$

and

$$r_2(p) = \begin{cases} (1) & \text{for } [\Omega_2 \cdot p - \theta_2] > 0 \\ [0, 1] & \text{for } [\Omega_2 \cdot p - \theta_2] = 0 \\ (0) & \text{for } [\Omega_2 \cdot p - \theta_2] < 0 \end{cases}$$

Note that for $\Omega_i \neq 0$, $i = 1, 2$, one can write the conditions for the best responses as $\Omega_1 \cdot [q - \theta_1/\Omega_1]$ and $\Omega_2 \cdot [p - \theta_2/\Omega_2]$ respectively. Therefore for $\Omega_i \neq 0$, there is a mixed strategy equilibrium if and only if $\theta_i/\Omega_i \in (0, 1)$ for $i = 1, 2$ holds. Furthermore, the existence of a mixed strategy equilibrium implies that θ_i and Ω_i must have the same sign.

Case (i): Without loss of generality assume $\theta_1 > 0$ and $\theta_2 < 0$, then $\Omega_1 - \theta_1 > 0$ and $\Omega_2 - \theta_2 < 0$ follows. Hence $\theta_1/\Omega_1 \in (0, 1)$ for both players and $p^* = \theta_2/\Omega_2$ and $q^* = \theta_1/\Omega_1$ form a mixed strategy equilibrium. On the other hand, $p^* = 1$ requires $q \geq \theta_1/\Omega_1 > 0$, but $\theta_2/\Omega_2 \in (0, 1)$ and $\Omega_2 < 0$ imply $q^* = 0$, an obvious contradiction. Similarly, one contradicts the hypothesis $p^* = 0$.

Case (ii): Here θ_1 and θ_2 must be of the same sign. Suppose that both are positive. Clearly, in this case too, $\theta_i/\Omega_i \in (0, 1)$ for both players and $p^* = \theta_2/\Omega_2$ and $q^* = \theta_1/\Omega_1$ form a mixed strategy equilibrium. If $p^* = 1$

then $q \geq \theta_1/\Omega_1 > 0$ must hold, but $p^* = 1$ implies $q^* = 1$. Hence, $(p, q) = (1, 1)$ is a Nash equilibrium. Similar reasoning yields that $(q, p) = (0, 0)$ is an equilibrium too. Let both θ_1 be negative, then $p^* = 1$ requires $q \leq \theta_1/\Omega_1 < 1$, but $p^* = 1$ implies $q^* = 0$. Hence, $(p, q) = (1, 0)$ is a Nash equilibrium in pure strategies. Similarly one obtains $(p, q) = (0, 1)$ as a pure strategy Nash equilibrium.

Case (iii): Without loss of generality assume that $\theta_1 \cdot (\Omega_1 - \theta_1) < 0$ holds. Then player 1 has a dominant strategy, namely $p^* = 1$ if $\theta_1 < 0$ holds and $p^* = 0$ if $\theta_1 > 0$ holds. For $\theta_1 < 0$ and $p^* = 1$, $q^* = 1$ follows if $\Omega_2 - \theta_2 > 0$ holds, and $q^* = 0$ if $\Omega_2 - \theta_2 < 0$ holds. Thus, there is a pure strategy equilibrium if $\Omega_2 \neq \theta_2$ holds. Similar reasoning shows that for $\theta_1 > 0$ there is a pure strategy equilibrium whenever $\theta_2 \neq 0$ holds.

Case (iv): In this case at least one player has $\theta_1 \cdot (\Omega_1 - \theta_1) = 0$. Without loss of generality, let $\theta_1 \cdot (\Omega_1 - \theta_1) = 0$ hold. For $\Omega_1 > \theta_1 = 0$, one can quickly check that

$(p, q) \in [0, \theta_2/\Omega_2] \times \{0\} \cup \{1\} \times \{1\}$	for $(\Omega_2 - \theta_2) > 0$ and $\theta_2 > 0$,
$(p, q) \in [\theta_2/\Omega_2, 1] \times \{0\}$	for $(\Omega_2 - \theta_2) < 0$ and $\theta_2 < 0$,
$(p, q) \in [0, 1] \times \{0\} \cup \{1\} \times [0, 1]$	for $\theta_2 = \Omega_2 \geq 0$,
$(p, q) \in \{1\} \times [0, 1]$	for $\theta_2 = \Omega_2 < 0$,
$(p, q) \in \{0\} \times \{0\} \cup \{1\} \times \{1\}$	for $\Omega_2 > \theta_2 = 0$,
$(p, q) \in [0, 1] \times \{0\} \cup \{1\} \times \{1\}$	for $\Omega_2 < \theta_2 = 0$,
$(p, q) \in [0, 1] \times \{0\}$	for $(\Omega_2 - \theta_2) < 0$ and $\theta_2 > 0$,

is the equilibrium correspondence. By symmetry, the correspondence is of the same form for the other possible cases, $\Omega_1 < \theta_1 = 0$, $\Omega_1 = \theta_1 > 0$, and $\Omega_1 = \theta_1 < 0$. For $\Omega_1 = \theta_1 = 0$, the equilibrium correspondence coincides with the best response correspondence of player 2. Clearly, this correspondence has the claimed properties. The argument for $\theta_2 \cdot (\Omega_2 - \theta_2) = 0$ is identical. ■

REFERENCES

- Aumann, R.J. (1987). "Correlated Equilibrium as an Expression of Bayesian Rationality". *Econometrica* 55, 1-18
- Aumann, R.J. (1976). "Agreeing to Disagree". *The Annals of Statistics* 4, 1236-1239
- Aumann, R.J. (1974). "Subjectivity and Correlation in Randomized Strategies". *Journal of Mathematical Economics* 1, 67-96
- Aumann, R.J., Maschler, M. (1972). "Some Thoughts on the Minimax Principle". *Management Science* 18 (Part 2), P54-P63
- Brown, G.W. (1951). "Iterative Solution of Games with Fictitious Play". In: Koopmans, T. (ed.). *Activity Analysis of Production and Allocation*. Cowles Commission Monographs 13, New York
- Canning, D. (1989a). "Convergence to Equilibrium in a Sequence of Games with Learning". Unpublished paper. London School of Economics
- Canning, D. (1989b). "Social Equilibrium". Unpublished paper. London School of Economics.
- Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W. (1988). "Selection Criteria in Coordination Games: Some Experimental Results". Unpublished paper, Hoover Institution, Stanford, College of Business Administration, University of Iowa
- Cournot, A. (1838). *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Calmann-Lévy 1974 (reprint)
- Crawford, V.P. (1989). "On Interpreting the Results of Van Huyck, Battalio, and Beil's Coordination Experiments". Unpublished paper. Department of Economics, University of California, San Diego
- Crawford, V.P. (1988). "Learning and Mixed-Strategy Equilibria in Evolutionary Games". Unpublished paper. Department of Economics, University of California, San Diego
- Crawford, V.P. (1985). "Learning Behavior and Mixed-Strategy Nash Equilibria". *Journal of Economic Behavior and Organization* 6, 69-78
- Crawford, V.P., Haller, H. (1990). "Learning How to Cooperate: Optimal Play in Repeated Coordination Games". *Econometrica* 58, 571-595
- Cyert, R.M., DeGroot, M.H. (1987). *Bayesian Analysis and Uncertainty in Economic Theory*. Rowman & Littlefield, Totowa
- DeGroot, M.H. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York, London
- Diamond, D.W., Dybvig, P.H. (1983). "Bank Runs, Deposit Insurance, and Liquidity". *Journal of Political Economy* 91, 401-419

Eichberger, J. (1990). "Nash Equilibrium and Learning". In: Borland, J., Creedy, J., Eichberger, J. (eds.). *Recent Developments in Game Theory*. Edward Elgar Publishing Ltd. 1991

Eichberger, J., Milne, F. (1990). "Bank Runs and Capital Adequacy". Unpublished paper, University of Melbourne

Fernandez, R., Glazer, J. (1989). "Striking for a Bargain between Two Completely Informed Agents". Unpublished paper, Boston University, forthcoming in *American Economic Review*

Fudenberg, D., Kreps, D.M. (1988). "A Theory of Learning, Experimentation, and Equilibrium in Games". Unpublished paper. Department of Economics, MIT and Graduate School of Business, Stanford

Haller, H. (1986). "Non-Cooperative Bargaining of $N \geq 3$ Players". *Economics Letters* 22, 11-13

Haller, H., Holden, St. (1990). "A Letter to the Editor on Wage Bargaining". *Journal of Economic Theory* 52, 232-236

Harsanyi, J.C. (1975). "The Tracing Procedure". *International Journal of Game Theory* 4, 61-94

Harsanyi, J.C. (1974). "Oddness of the Number of Equilibrium Points: A New Proof". *International Journal of Game Theory* 3, 235-250

Harsanyi, J.C. (1973). "Games with Randomly Disturbed Payoffs". *International Journal of Game Theory* 2, 1-23

Harsanyi, J.C. (1967). "Games with Incomplete Information Played by Bayesian Players". Parts I, II, III, *Management Science* 14, 159-182, 320-334, 486-502

Harsanyi, J.C., Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge

Hines, W.G.S. (1987). "Evolutionary Stable Strategies: A Review of Basic Theory". *Theoretical Population Biology* 31, 195-272

Jordan, J.S. (1990a). "Bayesian Learning in Normal Form Games". Unpublished paper. University of Minnesota

Jordan J.S. (1990b). "The Exponential Convergence of Bayesian Learning in Normal Form Games". Discussion Paper No. 259, Center for Economic Research, Department of Economics, University of Minnesota

Kalai, E., Lehrer, E. (1990). "Rational Learning Leads to Nash Equilibrium". Discussion Paper No. 858, Department of Managerial Economics and Decision Sciences, Northwestern University

Ledyard, J.O. (1986). "The Scope of the Hypothesis of Bayesian Equilibrium". *Journal of Economic Theory* 39, 59-82

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press

- Milgrom, P., Roberts, J. (1990). "Adaptive and Sophisticated Learning in Repeated Normal Form Games". Unpublished paper
- Moulin, H. (1986). *Game Theory for the Social Sciences*. New York University Press
- Rapoport, A., Guyer, M.J., Gordon, D.G. (1976). *The 2x2 Game*. Ann Arbor
- Robinson, J. (1951). "An Iterative Method of Solving a Game". *Annals of Mathematics* 54, 296-301
- Rosenmüller, J. (1971). "Über Periodizitätseigenschaften spieltheoretischer Lernprozesse". *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 17, 259-308
- Rosenthal, R.W., Landau, H.J. (1979). "A Game-Theoretic Analysis of Bargaining with Reputations". *Journal of Mathematical Psychology* 20, 233-255
- Rubinstein, A. (1982). "Perfect Equilibrium in a Bargaining Model". *Econometrica*, 50, 97-109
- Schelling, T.C. (1980). *The Strategy of Conflict*. Cambridge
- Shapley, L.S. (1964). "Some Topics in Two-Person-Games". In: Dresher, M. et al. (eds.). "Advances in Game Theory", *Annals of Mathematical Studies* 52, 1-28
- Shefrin, H.M. (1981). "Games with Self-Generating Distributions". *Review of Economic Studies* 48, 511-519
- Tan, T.C.-C., Da Costa Werlang, S.R. (1988). "The Bayesian Foundations of Solution Concepts of Games". *Journal of Economic Theory* 45, 370-391
- Taylor, P.T., Jonker, L.B. (1978). "Evolutionary Stable Strategies and Game Dynamics". *Mathematical Biosciences* 40, 145-156
- Tirole, J. (1989). *The Theory of Industrial Organization*. MIT Press, Cambridge
- Van Damme, E. (1987). *Stability and Perfection of Nash Equilibria*. Heidelberg, New York
- Van Huyck, J.B., Battalio, R.C., Beil, R.O. (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure". *American Economic Review* 80, 234-248

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
8943	A. Roell	Dual Capacity Trading and the Quality of the Market
8944	C. Hsiao	Identification and Estimation of Dichotomous Latent Variables Models Using Panel Data
8945	R.P. Gilles	Equilibrium in a Pure Exchange Economy with an Arbitrary Communication Structure
8946	W.B. MacLeod and J.M. Malcomson	Efficient Specific Investments, Incomplete Contracts, and the Role of Market Alternatives
8947	A. van Soest and A. Kapteyn	The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution
8948	P. Kooreman and B. Melenberg	Maximum Score Estimation in the Ordered Response Model
8949	C. Dang	The D_3 -Triangulation for Simplicial Deformation Algorithms for Computing Solutions of Nonlinear Equations
8950	M. Cripps	Dealer Behaviour and Price Volatility in Asset Markets
8951	T. Wansbeek and A. Kapteyn	Simple Estimators for Dynamic Panel Data Models with Errors in Variables
8952	Y. Dai, G. van der Laan, D. Talman and Y. Yamamoto	A Simplicial Algorithm for the Nonlinear Stationary Point Problem on an Unbounded Polyhedron
8953	F. van der Ploeg	Risk Aversion, Intertemporal Substitution and Consumption: The CARA-LQ Problem
8954	A. Kapteyn, S. van de Geer, H. van de Stadt and T. Wansbeek	Interdependent Preferences: An Econometric Analysis
8955	L. Zou	Ownership Structure and Efficiency: An Incentive Mechanism Approach
8956	P. Kooreman and A. Kapteyn	On the Empirical Implementation of Some Game Theoretic Models of Household Labor Supply
8957	E. van Damme	Signaling and Forward Induction in a Market Entry Context
9001	A. van Soest, P. Kooreman and A. Kapteyn	Coherency and Regularity of Demand Systems with Equality and Inequality Constraints

No.	Author(s)	Title
9002	J.R. Magnus and B. Pesaran	Forecasting, Misspecification and Unit Roots: The Case of AR(1) Versus ARMA(1,1)
9003	J. Driffill and C. Schultz	Wage Setting and Stabilization Policy in a Game with Renegotiation
9004	M. McAleer, M.H. Pesaran and A. Bera	Alternative Approaches to Testing Non-Nested Models with Autocorrelated Disturbances: An Application to Models of U.S. Unemployment
9005	Th. ten Raa and M.F.J. Steel	A Stochastic Analysis of an Input-Output Model: Comment
9006	M. McAleer and C.R. McKenzie	Keynesian and New Classical Models of Unemployment Revisited
9007	J. Osiewalski and M.F.J. Steel	Semi-Conjugate Prior Densities in Multi- variate t Regression Models
9008	G.W. Imbens	Duration Models with Time-Varying Coefficients
9009	G.W. Imbens	An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling
9010	P. Deschamps	Expectations and Intertemporal Separability in an Empirical Model of Consumption and Investment under Uncertainty
9011	W. Güth and E. van Damme	Gorby Games - A Game Theoretic Analysis of Disarmament Campaigns and the Defense Efficiency-Hypothesis
9012	A. Horsley and A. Wrobel	The Existence of an Equilibrium Density for Marginal Cost Prices, and the Solution to the Shifting-Peak Problem
9013	A. Horsley and A. Wrobel	The Closedness of the Free-Disposal Hull of a Production Set
9014	A. Horsley and A. Wrobel	The Continuity of the Equilibrium Price Density: The Case of Symmetric Joint Costs, and a Solution to the Shifting-Pattern Problem
9015	A. van den Elzen, G. van der Laan and D. Talman	An Adjustment Process for an Exchange Economy with Linear Production Technologies
9016	P. Deschamps	On Fractional Demand Systems and Budget Share Positivity
9017	B.J. Christensen and N.M. Kiefer	The Exact Likelihood Function for an Empirical Job Search Model

No.	Author(s)	Title
9018	M. Verbeek and Th. Nijman	Testing for Selectivity Bias in Panel Data Models
9019	J.R. Magnus and B. Pesaran	Evaluation of Moments of Ratios of Quadratic Forms in Normal Variables and Related Statistics
9020	A. Robson	Status, the Distribution of Wealth, Social and Private Attitudes to Risk
9021	J.R. Magnus and B. Pesaran	Evaluation of Moments of Quadratic Forms in Normal Variables
9022	K. Kamiya and D. Talman	Linear Stationary Point Problems
9023	W. Emons	Good Times, Bad Times, and Vertical Upstream Integration
9024	C. Dang	The D_2 -Triangulation for Simplicial Homotopy Algorithms for Computing Solutions of Nonlinear Equations
9025	K. Kamiya and D. Talman	Variable Dimension Simplicial Algorithm for Balanced Games
9026	P. Skott	Efficiency Wages, Mark-Up Pricing and Effective Demand
9027	C. Dang and D. Talman	The D_1 -Triangulation in Simplicial Variable Dimension Algorithms for Computing Solutions of Nonlinear Equations
9028	J. Bai, A.J. Jakeman and M. McAleer	Discrimination Between Nested Two- and Three-Parameter Distributions: An Application to Models of Air Pollution
9029	Th. van de Klundert	Crowding out and the Wealth of Nations
9030	Th. van de Klundert and R. Gradus	Optimal Government Debt under Distortionary Taxation
9031	A. Weber	The Credibility of Monetary Target Announcements: An Empirical Evaluation
9032	J. Osiewalski and M. Steel	Robust Bayesian Inference in Elliptical Regression Models
9033	C. R. Wichers	The Linear-Algebraic Structure of Least Squares
9034	C. de Vries	On the Relation between GARCH and Stable Processes
9035	M.R. Baye, D.W. Jansen and Q. Li	Aggregation and the "Random Objective" Justification for Disturbances in Complete Demand Systems

No.	Author(s)	Title
9036	J. Driffill	The Term Structure of Interest Rates: Structural Stability and Macroeconomic Policy Changes in the UK
9037	F. van der Ploeg	Budgetary Aspects of Economic and Monetary Integration in Europe
9038	A. Robson	Existence of Nash Equilibrium in Mixed Strategies for Games where Payoffs Need not Be Continuous in Pure Strategies
9039	A. Robson	An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games
9040	M.R. Baye, G. Tian and J. Zhou	The Existence of Pure-Strategy Nash Equilibrium in Games with Payoffs that are not Quasiconcave
9041	M. Burnovsky and I. Zang	"Costless" Indirect Regulation of Monopolies with Substantial Entry Cost
9042	P.J. Deschamps	Joint Tests for Regularity and Autocorrelation in Allocation Systems
9043	S. Chib, J. Osiewalski and M. Steel	Posterior Inference on the Degrees of Freedom Parameter in Multivariate-t Regression Models
9044	H.A. Keuzenkamp	The Probability Approach in Economic Methodology: On the Relation between Haavelmo's Legacy and the Methodology of Economics
9045	I.M. Bomze and E.E.C. van Damme	A Dynamical Characterization of Evolutionarily Stable States
9046	E. van Damme	On Dominance Solvable Games and Equilibrium Selection Theories
9047	J. Driffill	Changes in Regime and the Term Structure: A Note
9048	A.J.J. Talman	General Equilibrium Programming
9049	H.A. Keuzenkamp and F. van der Ploeg	Saving, Investment, Government Finance and the Current Account: The Dutch Experience
9050	C. Dang and A.J.J. Talman	The D_1 -Triangulation in Simplicial Variable Dimension Algorithms on the Unit Simplex for Computing Fixed Points
9051	M. Baye, D. Kovenock and C. de Vries	The All-Pay Auction with Complete Information
9052	H. Carlsson and E. van Damme	Global Games and Equilibrium Selection
9053	M. Baye and D. Kovenock	How to Sell a Pickup Truck: "Beat-or-Pay" Advertisements as Facilitating Devices

No.	Author(s)	Title
9054	Th. van de Klundert	The Ultimate Consequences of the New Growth Theory; An Introduction to the Views of M. Fitzgerald Scott
9055	P. Kooreman	Nonparametric Bounds on the Regression Coefficients when an Explanatory Variable is Categorized
9056	R. Bartels and D.G. Fiebig	Integrating Direct Metering and Conditional Demand Analysis for Estimating End-Use Loads
9057	M.R. Veall and K.F. Zimmermann	Evaluating Pseudo- R^2 's for Binary Probit Models
9058	R. Bartels and D.G. Fiebig	More on the Grouped Heteroskedasticity Model
9059	F. van der Ploeg	Channels of International Policy Transmission
9060	H. Bester	The Role of Collateral in a Model of Debt Renegotiation
9061	F. van der Ploeg	Macroeconomic Policy Coordination during the Various Phases of Economic and Monetary Integration in Europe
9062	E. Bennett and E. van Damme	Demand Commitment Bargaining: - The Case of Apex Games
9063	S. Chib, J. Osiewalski and M. Steel	Regression Models under Competing Covariance Matrices: A Bayesian Perspective
9064	M. Verbeek and Th. Nijman	Can Cohort Data Be Treated as Genuine Panel Data?
9065	F. van der Ploeg and A. de Zeeuw	International Aspects of Pollution Control
9066	F.C. Drost and Th. E. Nijman	Temporal Aggregation of GARCH Processes
9067	Y. Dai and D. Talman	Linear Stationary Point Problems on Unbounded Polyhedra
9068	Th. Nijman and R. Beetsma	Empirical Tests of a Simple Pricing Model for Sugar Futures
9069	F. van der Ploeg	Short-Sighted Politicians and Erosion of Government Assets
9070	E. van Damme	Fair Division under Asymmetric Information
9071	J. Eichberger, H. Haller and F. Milne	Naive Bayesian Learning in 2 x 2 Matrix Games

P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS

Bibliotheek K. U. Brabant

17 000 01117527 1