Preliminary draft

Comments are welcome

AN EVOLUTIONARY INTERPRETATION OF MIXED-STRATEGY EQUILIBRIA*

Jörg Oechssler**

Department of Economics

Columbia University

April 9, 1994

Abstract

One of the more convincing interpretations of mixed strategy equilibria describes a mixed equilibrium as a steady state in a large population in which all players use pure strategies but the population as a whole mimics a mixed strategy. To be complete, however, this interpretation requires a good story about how the population arrives at the appropriate distribution over pure strategies. In this paper I attempt to give an explanation based on an evolutionary, stochastic learning process. Convergence properties of these processes have been studied extensively but almost exclusively for the case of convergence to pure Nash equilibria. Here I study the conditions under which an evolutionary process converges to population mixed-strategy equilibria. I find that not all mixed equilibria can be justified as the result of the evolutionary learning process even if the equilibrium is unique. For symmetric 2x2 and 3x3 games I give necessary and sufficient conditions for convergence and for n×n games I give a sufficient condition. For cases in which the conditions are not satisfied counterexamples are given, in which the process enters a limit cycle.

---

* I am indebted to Susan Elmes for numerous discussions and very helpful comments.

** Department of Economics, Columbia University, New York, NY 10027, USA.

(212) 749-1578, email: jo33@columbia.edu

I. INTRODUCTION

Although the concept of a mixed strategy is one of the most important concepts in game theory, it is also one of the more controversial topics. Widely used for its convenient property of convexifying the strategy space, it is generally acknowledged to have a relatively weak behavioral basis. As Ariel Rubinstein (1988, p. 9) states rather drastically: "...the naive interpretation of mixed strategies, as an action which is conditional on the outcome of a lottery executed by the player before the game, is intuitively ridiculous."

Uneasiness about mixed strategy equilibria arises from the fact that although a player is indifferent between all of the pure strategies in the support of his mixed strategy, he has to randomize over these pure strategies in such a way that his opponents are exactly indifferent between their strategies. As Rubinstein (1988, p.10) points out this assumed behavior becomes even more problematic if there are small costs associated with using a randomization device.

This is not to say that players never consciously randomize. Most people understand that in rock-scissors-paper it would be foolish to become predictable. Thus in order to avoid being outguessed players may actually randomize. This argument, of course, applies only if the game is played repeatedly and players choices are observable. But what would be the incentive for a player to randomize if the game is played in a large population with anonymous matching and with no hope of observing the past behavior of any particular opponent?

There are two approaches in the literature that interpret mixed strategy equilibria not as states in which players randomize in each round of play but rather as steady state in which a mixed strategy is obtained on *average* (either over time or in a cross section) while all players actually use pure best replies in every given period.[1] The first of these approaches is Harsanyi's (1973) purification idea that states that a player is influenced by small, unmodelled perturbations to his payoffs which are not observable by his opponents. In every repetition of the game a player chooses the unique, pure best reply in the corresponding game of incomplete information depending on the realization of the perturbation. Considering the long run average of choices his opponents are lead to believe that the player actually randomizes between the pure strategies. The main problem with this approach is that the source of the perturbations remains unexplained, which is dissatisfying since if perturbations are payoff relevant, as Harsanyi assumes, they should be explicitly included in the model (Rubinstein, 1988).

The second approach considers a mixed strategy equilibrium as a steady state in a large population (see, e.g., Rosenthal, 1979). A steady state is a *population mixed-strategy equilibrium* if each pure strategy in the support of the mixed equilibrium strategy is being played by the

[1] See Osborne and Rubinstein (forthcoming) for a more detailed discussion of this.

appropriate proportion of the population to give an outside observer - or each individual player - the impression that the population as a whole is playing a mixed strategy even though each individual player uses a pure strategy. But this interpretation of a mixed equilibrium begs another question. How does the population arrive at a population mixed-strategy equilibrium or, expressed differently, how can players be coordinated such that the right proportions are achieved even though each individual player has no reason to prefer a particular pure strategy?

The current paper seeks to answer this question by employing a stochastic evolutionary learning mechanism. Such models have been used extensively in the recent literature (see, for example, Kandori, Mailath and Rob, 1993 or Young, 1993) but almost without exception they were used to study convergence of the learning process to *pure* (and usually, strict) Nash equilibria.[2] In this paper I characterize the conditions under which an evolutionary learning process converges to a *mixed* strategy equilibrium.

The general set up of the model is the following. Players from a large but finite population are repeatedly matched in pairs to play a symmetric normal form game. After each period players observe the percentages with which each action was chosen. Players are *myopic* in the sense that they seek to maximize the payoff of the current period without taking into account the effects their actions might have on the future development of the system. Furthermore, there is *inertia* in the system as players are assumed to be able to adjust their strategies only infrequently. These assumptions define what is sometimes called a *best reply process* or *best response dynamic* (see, for example, Kandori and Rob 1993, and Bergin and Lipman 1993).

I will consider two types of best reply processes which differ in the way players are assumed to calculate their expected payoffs. One version, which I call the *large population case* assumes that the population is sufficiently large that players do not bother to exclude themselves from the strategy profile before calculating the probabilities of their opponents' strategies. To some extent this seems to be in line with the general bounded rationality flavor of the model. Also, I do not assume that players necessarily know the exact population size. All they observe are the proportions of players taking each action. Strictly speaking players are, therefore, not able to perform the calculation of subtracting themselves out of the strategy profile.

Alternatively, I consider a best reply process for the *small population case,* in which players are assumed to know the exact population size and exclude themselves from the profile before calculating their expected payoff. For 2x2 and 3x3 games I report convergence results for

[2] Young (1993) shows that games that are weakly acyclic converge to a strict Nash equilibrium. Kandori and Rob (1993) give a convergence result for pure coordination games and for games with strategic complementarities. A selection of other papers in this area include Canning (1992), Ellison (1993), Foster and Young (1990), and Nöldeke and Samuelson (1993).

both types of best reply processes. Results for general n×n games are only given for the large population case.

The first important result of this paper is that not all population mixed-strategy equilibria can be justified on the ground that they are the outcome of the best reply process. For both, the large and the small population case there are games for which the best reply process can never converge to the mixed equilibrium even though the equilibrium is assumed to be the unique symmetric equilibrium. Instead, the process cycles forever around the equilibrium.

In other words there are two classes of mixed strategy equilibria. The equilibria in the one class can be justified as the outcome of evolutionary learning since the process converges to an equilibrium from any initial state in finite time with probability one. On the other hand there is a class of mixed strategy equilibria that cannot justified in this way. What makes them different? Roughly speaking, the equilibria in the latter category are unstable in a sense that if the system is off the equilibrium by just one player, this player never has an incentive to return the system to equilibrium by switching to the under-represented strategy.

For 2x2 and 3x3 games I present necessary and sufficient conditions for (global) convergence to the equilibrium with both types of best reply processes. Clearly, a necessary condition for every converging process is that there exists a *last step* to the equilibrium. In other words there must exist a population strategy profile such that given the dynamics of the best reply process it is possible to move from that profile to the equilibrium. Surprisingly, it turns out that the existence of any such last step is also the *sufficient* condition for convergence in the 3x3/large population case.

The situation is somewhat different in the 3x3/small population case. In contrast to the large population case not all mixed equilibria are absorbing states of the best reply process in the small population case. Some mixed strategy equilibria are unstable in the sense that if a player excludes himself from the equilibrium profile, his old strategy is not a best reply against the remaining profile. For this case I show that a necessary and sufficient condition is that the equilibrium is absorbing, a condition which is stronger than the necessary and sufficient condition for the large population case.

Finally, I give a sufficient condition for convergence in symmetric n×n games. This condition, which is called *evolutionary stability with respect to pure strategies (ESPS)*, implies a restriction on the payoff structure of the game such that from every strategy profile the dynamics are directed toward the equilibrium.

There is a variety of studies in the literature that try to justify the use of Nash equilibria based on an adaptive or evolutive story rather than on introspection by perfectly rational individuals. These adaptive approaches generally assume that individuals play the game repeatedly while ignoring possible strategic links between repetitions. It is then analyzed whether or not relatively simple adjustment rules lead to Nash equilibria. Although some early progress had been made, in particular in the *fictitious play* literature (Brown, 1951, and Robinson, 1951),[3] research in this direction all but ended after Shapley showed that such adjustment processes do not necessarily converge.

Furthermore, as Crawford (1985) points out, the question whether such adjustment processes converge to *mixed* strategy equilibria has largely be ignored. Crawford (1985) himself gives a very pessimistic answer. He finds that a process, in which players adjust the weights of their mixed strategies according to the relative success of each pure strategy in the past, is locally unstable at the mixed equilibrium for almost all games.

More recently, several papers reconsider the question of learning and mixed Nash equilibria. Fudenberg and Kreps (1993) and Jordan (1993) point out problems with the fictitious play interpretation of mixed Nash equilibria. In particular, in models of fictitious play players do not randomize but rather switch from playing one pure strategy to playing another in cycles of ever-increasing length. Further, both papers, as well as Young (1993), give an example of a version of the Battle of Sexes game in which fictitious play (or adaptive play, as Young calls a modified, stochastic version) leads to a situation in which players' actions are perfectly correlated. Players miscoordinate every period even though the relative frequency with which they use each pure strategy corresponds to the weights in the mixed strategy equilibrium.

Fudenberg and Kreps (1993) also define a new behavioral rule that leads to convergence of intended behavior rather then just to convergence of empirical frequencies. But they find that "... while convergence to mixed behavior is possible, it is hard to see why it should occur." (p. 323) so that again they have to resort to Harsanyi's purification idea to justify why players who are indifferent would randomize with the frequencies required by the mixed Nash equilibrium. In light of all these problems the population mixed-strategy interpretation may be an elegant escape.

The next section introduces the details of the model followed by a short account of the convergence properties of 2x2 games. In Section III a necessary and sufficient condition for convergence for the 3x3/large population case is given. The same is done in section IV for the small population case. And finally section V contains the proof that ESPS is a sufficient condition for convergence of $n \times n$ games.

---

[3] Originally, fictitious play was not interpreted as a justification for Nash equilibria but rather as an algorithm to compute them.

II. THE MODEL

The basic game considered here is a symmetric n×n normal form game with pure strategies $S = \{S_1, ..., S_n\}$. If a player chooses strategy S_i against strategy S_j of his opponent, he receives a payoff of $\Pi(S_i,S_j)$. This stage game is played repeatedly by a large population of N players such that in every period each player is randomly matched with one of his opponents.[4] The *state* of the system at time t can be summarized by a vector $\mu^t = \{\mu^t_1, ..., \mu^t_n\}$ that gives the proportion of the population using each strategy. The *state space* $M = \{\mu \in \Delta^{n-1} \mid \mu_i = \#S_i / N\}$ is a finite set of grid points in the n - 1 dimensional simplex.

Consistent with the evolutionary nature of this model the assumptions about players' information sets are less demanding than usual in game theory. Players are supposed to know only the current state μ and their own payoff matrix (even though the game is symmetric and all payoff matrices are identical this fact is not necessarily known to the players). I assume that a players in a large population who receives the information that the currents state is μ will expect a payoff next period from playing strategy S_i of

$$\Pi(S_i,\mu) = \sum_{j=1}^{n} \mu_j \Pi(S_i,S_j). \quad (1)$$

Two assumptions are implicit in this formulation. The first is that players expect the system to have strong inertia, that is, that the state tomorrow is almost the same as the state today. This assumption is justified if the population is large and few players are able to change their strategies each period, which is the case for the stochastic adjustment process assumed below. The second assumption is that the population is so large that a player does not bother to exclude himself from the profile before calculating the expected payoff. Note that since players are not assumed to know the population size N, they cannot exactly carry out such a calculation anyway.

On the other hand if population size is small and if N is assumed to be known, a player can exclude himself from the profile and will then expect a payoff from playing S_i when he was using S_k in the current period of

$$\Pi(S_i,\mu_{-k}) = \sum_{j \neq k} \frac{\mu_j N}{N-1} \Pi(S_i,S_j) + \frac{\mu_j N - 1}{N-1} \Pi(S_i,S_k), \quad (2)$$

where μ_{-k} denotes strategy profile μ with one players excluded from strategy S_k. Most results in this paper will be given for the large population case but results for 2x2 and 3x3 games are presented for the small population case as well.

---

[4] If N is odd, the odd-man-out is not matched and receives and an arbitrary payoff of π. The expected payoff is then equal to the expected payoff conditional on being matched plus (1 - 1/N) π. All results are unchanged for this case so I will henceforth assume that N is even.

Given the players' expectations about next period's payoffs the set of pure strategy *best replies* is

$$B(\mu) = \{ S_i \in S \mid S_i \in argmax_j \, \Pi(S_j,\mu) \}$$

for the large population case and

$$B(\mu_{-k}) = \{ S_i \in S \mid S_i \in argmax_j \, \Pi(S_j,\mu_{-k}) \}$$

for the small population case.

One feature that one would want every plausible learning process to satisfy is that strategies that have done well in the past should increase in frequency whereas strategies that haven't should decrease. I will follow Samuelson (1991) and Kandori and Rob (1993) in formalizing this general idea by assuming that players are myopic optimizers who - infrequently - have the opportunity to adjust their strategies to the current environment. Infrequent adjustment seems to be a reasonable assumption in situations with adjustment costs that vary stochastically over time. For example, many people buy a new car only when their old car breaks down. Taking the break-down time of cars to be stochastic the (opportunity) cost of switching to a different model is lower at some infrequent points in time.

Infrequent adjustment is not a necessary assumption for this model but it gives some plausibility to the assumption of myopic behavior. If adjustment is infrequent players are almost correct in assuming that the strategy profile tomorrow looks like the profile today. If in addition the future is discounted heavily, then seemingly myopic behavior could be based on rational decisions.

Assumption (best reply process): Every period each player receives the opportunity to adjust his strategy with probability $\theta > 0$. If the opportunity arises, a player chooses a best reply against the current strategy profile μ. If there are several best replies, a player may choose any of them with positive probability unless he is already playing a best reply in which case he remains at his current strategy.

Probability θ may depend on the state and on the strategy the player is currently using, $\theta(\mu,S_i)$. For example, it may be plausible to specify that players with very bad strategies can switch more frequently. The results in this paper are unaffected by this modification as long as all $\theta(\mu,S_i)$ are strictly positive.[5]

The above assumption gives rise to a stationary Markov chain on the state space M. I will say that the process has converged to an *absorbing state* in period τ if $\mu^t = \mu^{t+1}$ for all periods $t \geq \tau$. It is obvious that a state is absorbing if and only if all players use best replies.

[5] See also Kandori and Rob (1993) for this observation.

To better characterize the convergence properties of the best reply process it is useful to define a *best reply region* $R(S_i)$ which is the subset of the state space M for which strategy S_i is a best reply.

$$R(S_i) = \{ \mu \in M \mid S_i \in B(\mu) \}$$

Two properties can be immediately stated. First, the best reply regions are convex sets and second, they cover the entire state space, that is, $\bigcup_{i=1\ldots n} R(S_i) = M$. Let $\mathbf{R}(S_i)$ denote the set on which S_i is the *unique* best reply, that is $\mathbf{R}(S_i) = R(S_j)^c \cap R(S_k)^c$, with "c" denoting the complement of a set in M. I will call $\mathbf{R}(S_i)$ the "interior" of $R(S_i)$ for short.

Below I will frequently employ a diagram of the state space for the case of 3x3 games to illustrate general properties of the best reply process. I will use it here to show the difference between the best reply process for the large and the small population case. In figure 1 the corners correspond to states in which the entire populations plays the same pure strategy and the edge opposite a corner correspond to states in which no one plays this particular strategy.

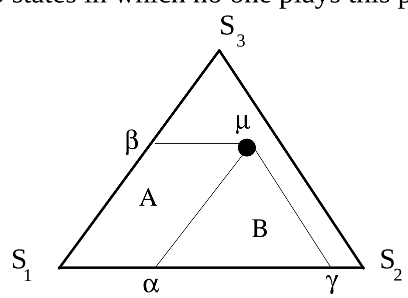

Figure 1

Suppose S_1 is the unique best reply against μ, that is $\mu \in \mathbf{R}(S_1)$. Then any grid point in area A can be reached by the best reply process for the large population case. Which point the process actually moves to depends on which and how many players get the opportunity to change their strategies. For example, if all current S_3-players get the chance to switch but no one else, the process moves to α. If all S_2 get the chance to switch, the process moves to β. Current S_1-players, of course, would not change their strategy even if they had the chance since they are already playing a best reply.

In contrast, in the small population case it is possible that while $\mu \in \mathbf{R}(S_1)$, μ_{-3} is in another best reply region, e.g. $\mu_{-3} \in \mathbf{R}(S_2)$ because a current S_3-player after excluding himself from the profile faces a profile with a lower percentage of S_3-players. In that case, all grid point in A or in B can be reached. For example, if all S_3-players get the opportunity to move, the process moves to γ.

Definition 1: A *population mixed-strategy equilibrium (PME)* $\mu^* = (\mu^*_1,..., \mu^*_n)$ is a distribution of players over strategies such that the proportions μ_i^* equal the weights each strategy receives in a mixed Nash equilibrium.

Clearly, a PME can only exist if N can be partitioned in a way so as to produce the weights of the mixed Nash equilibrium, which excludes games with irrational payoffs from consideration. Given the right N the question is, under what conditions, if any, will a best reply process converge to a PME? In the large population case, that is, with payoffs as specified in (1), each Nash equilibrium, whether pure or mixed, is an absorbing state of the best reply process. However, some mixed equilibria are highly unstable and the best reply process cannot end up there unless it started there.

I will start with the relatively simple case of 2x2 games. If there are two pure symmetric equilibria, there is also a mixed equilibrium μ^*. But the best reply process for the large population case can never converge to μ^* since for all $\mu > \mu^*$ strategy S_1 is the unique best reply and for all $\mu < \mu^*$ strategy S_2 is the unique best reply. Thus, the process always converges away from μ^* to one of the pure strategies. On the other hand if a game has a unique equilibrium it is easy to see that the process will converge to this equilibrium even if it is mixed since the dynamics from all states are directed toward this unique absorbing state. Due to the finiteness of the state space, the mixed equilibrium must be reached eventually.

The situation is similar for the small population case, i.e. with payoffs specified as in (2), with the exception that in the game with two pure and one mixed equilibria the latter is not even an absorbing state. To see why, note that a S_1-player after excluding himself from the mixed equilibrium profile μ^* is not playing a best reply against μ^*_{-1} and would like to switch to S_2 instead. Hence, μ^* is not an absorbing state.

III. 3x3 GAMES: THE LARGE POPULATION CASE

There are several results in the literature concerning the convergence of best reply processes to pure Nash equilibria. For example, Young (1993) shows that weakly acyclic games converge to a (strict) Nash equilibrium. Kandori and Rob (1993) show the same for pure coordination games and games with strategic complementarities. An interesting question is therefore whether the process converges in games that fail to have a pure, symmetric Nash equilibrium. In the remainder of the paper I will restrict myself to games with a unique, mixed equilibrium. The main theorem of this section gives the necessary and sufficient condition for convergence in the case of games with completely mixed Nash equilibria. But before turning to that I will deal with the easier case of games with non-completely mixed Nash equilibria, which is similar to the 2x2 case in the sense that all such games converge.

Proposition 1: Let μ^* *be a mixed but not completely mixed Nash equilibrium. Then the best reply process for the large population case converges always to* μ^*.

Proof: Assume WLOG that $\mu^*_3 = 0$. To show that the process always converges to $\mu^* = (\mu^*_1, 1-\mu^*, 0)$ I will construct a sequence of possible transitions from an arbitrary initial state μ to one of the pure strategy states and from there to μ^*. This shows that all states other than the equilibrium are transient and, hence, the process must converge to μ^* eventually.

If $\mu \in \mathbf{R}(S_i)$, some i, S_i can be reached in one step. If not, an intermediate step in the interior of one of the best reply regions can be reached and from there a pure strategy. To see that μ^* can be reached from any pure strategy consider the *best reply graph* of a game with a unique mixed equilibrium. In the best reply graph pure strategies are connected by a directed path, e.g. $S_i \to S_j$ indicating that $S_j \in B(S_i)$ and, hence, the population could move from a state in which everyone plays S_i to a state in which everyone plays S_j. Two types of best reply graphs are possible for games with a unique, mixed equilibrium. Either there is a cycle $S_1 \to S_2 \to S_3 \to S_1$, in which case μ^* can be reached from all pure strategies, or at least one strategy is a best reply to both other strategies. Note that S_3 cannot be a best reply to both S_1 and S_2. If it were, then due to convexity of the best reply regions all states $(1-\mu, \mu, 0)$, $\mu \in [0,1]$, would be contained in $R(S_3)$. But this implies that the payoffs of both S_1 and S_3 are flat on all states $(1-\mu, \mu, 0)$, $\mu \in [0,1]$ since at μ^* S_1 and S_3 are best replies. Hence, the equilibrium would not be unique.

Suppose WLOG that $S_1 \notin R(S_3)$ which implies that $S_1 \in R(S_2)$ since no pure strategy equilibria exist. The two remaining possible best reply graphs must include the following directed paths (there may or may not be some additional paths in each graph) $S_1 \to S_2 \leftrightarrow S_3$ and $S_3 \to S_1 \leftrightarrow S_2$. In the latter case μ^* can obviously be reached from all pure strategies. In the former case a state $\mu'' = (0, \mu'', 1-\mu'') \in \mathbf{R}(S_1)$ can be reached from both S_2 and S_3. μ'' must exist since otherwise there would be a second equilibrium on the edge connecting S_2 and S_3. Summarizing, the process can move from S_2 or S_3 to S_1 via μ'' and then from S_1 to μ^*.■

The following theorem gives a necessary and sufficient condition for the convergence of the best reply process in a large population to the population mixed equilibrium for the case of 3x3 games with a unique completely mixed equilibrium μ^*.

Theorem 1: For symmetric 3x3 games with a unique, completely mixed Nash equilibrium μ^* *the best reply process in a large population converges to the equilibrium if and only if there exists some* $\mu \in M$ *and some i with the property that* $\mu_i < \mu_i^*$, $\mu_j > \mu_j^*$, $j \neq i$, *and* $\{S_i\} = B(\mu)$.

Proof: To prove that the existence of μ is a necessary condition for convergence to the equilibrium note that every converging process must have a last step. Relative to the equilibrium profile μ^*, the penultimate state μ' of any converging sequence can only have one of the following three formats:

(a) Two strategies are over-represented relative to μ^* and one is under-represented, that is,

$$\exists!\ S_i \text{ such that } \mu_i < \mu_i^* \text{ and } \mu_j > \mu_j^*,\ j \neq i. \tag{3}$$

In order to move from some μ' satisfying (3) to μ^*, S_i must be the unique best reply against μ', that is

$$\exists\ \mu' \text{ satisfying (3) s.t. } \mu' \in \mathbf{R}(S_i). \tag{4}$$

Note that (4) is just the condition of the theorem.

(b) Two strategies are under-represented and one is over-represented, that is,

$$\exists!\ S_i \text{ such that } \mu_i > \mu_i^* \text{ and } \mu_j < \mu_j^*,\ j \neq i. \tag{5}$$

To move from μ' satisfying (5) to μ^* both of the under-represented strategies must be best replies against μ' and the over-represented strategy must not be a best reply, that is,

$$\exists\ \mu' \text{ satisfying (5) s.t. } \mu' \in R(S_j) \cap R(S_k). \tag{6}$$

(c) One strategy is over-represented and one is under-represented, that is,

$$\exists\ i,j,k \text{ such that } \mu_i < \mu_i^*,\ \mu_j > \mu_j^* \text{ and } \mu_k = \mu_k^* \tag{7}$$

In this last case the equilibrium can only be reached if S_i is a best reply and S_j is not:

$$\exists\ \mu' \text{ satisfying (7) s.t. } \mu' \in R(S_i) \cap R(S_j)^c \tag{8}$$

I will now show that whenever condition (6) or (8) is satisfied, condition (4) is satisfied as well. Hence, (4) is a necessary condition for convergence to μ^*. For notational convenience let M be divided into regions as shown in figure 2.

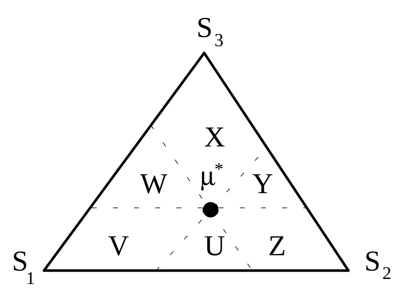

Figure 2

For example, $W = \{ \mu \in M \mid \mu_1 \geq \mu_1^* \text{ and } \mu_3 \geq \mu_3^* \}$. Condition (4) can now be written as

$$(W \cap \mathbf{R}(S_2)) \cup (Y \cap \mathbf{R}(S_1)) \cup (U \cap \mathbf{R}(S_3)) \neq \varnothing \tag{9}$$

and (6) becomes

$$(R(S_2) \cap R(S_3) \cap V) \cup (R(S_1) \cap R(S_2) \cap X) \cup (R(S_1) \cap R(S_3) \cap Z) \neq \varnothing \tag{10}$$

To show that (10) implies (9) assume that (10) is satisfied and, WLOG,

$$\left(R(S_2) \cap R(S_3) \cap V\right) \neq \varnothing.$$

$\mathbf{R}(S_2) \cup \mathbf{R}(S_3)$ cannot be contained in V due to the convexity of the best reply regions (otherwise $R(S_1)$ would contain W and U and, hence, would not be convex). Thus,

$$\left(\mathbf{R}(S_2) \cup \mathbf{R}(S_3)\right) \cap (W \cup U) \neq \varnothing.$$

This can be satisfied either because $(\mathbf{R}(S_2) \cap W) \cup (\mathbf{R}(S_3) \cap U) \neq \varnothing$, in which case (9) is satisfied, or because $(\mathbf{R}(S_2) \cap U) \cup (\mathbf{R}(S_3) \cap W) \neq \varnothing$. In the latter case, $Y \subset R(S_1)$ because no pure strategy equilibrium exists and hence $S_3 \notin R(S_3)$ and $S_2 \notin R(S_2)$. in either case condition (9) is satisfied whenever (10) is satisfied.

Condition (8) can be written as

$$(R(S_1) \cap R(S_2)^c \cap Y \cap Z) \cup (R(S_2) \cap R(S_1)^c \cap W \cap V) \cup \ldots \neq \varnothing.$$

Assume WLOG that $(R(S_1) \cap R(S_2)^c \cap Y \cap Z) \neq \varnothing$. Either $Y \cap \mathbf{R}(S_i) \neq \varnothing$, in which case (9) is satisfied or $Y \cap Z$ forms the boundary between $R(S_1)$ and $R(S_3)$ (see figure 3). Since $S_1 \notin R(S_1)$ and $S_3 \notin R(S_3)$, it follows that $W \subset \mathbf{R}(S_2)$, which implies that (9) is satisfied. This concludes the "only if" part of the proof.

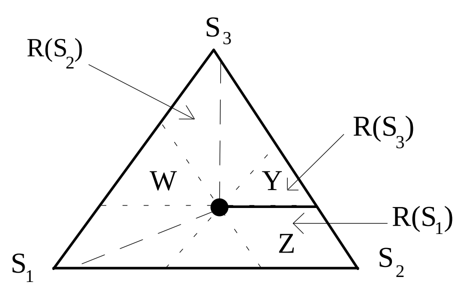

Figure 3

I will now show that the existence of a μ satisfying condition (4) is also a *sufficient* condition for convergence. Given (4) is satisfied it is possible to construct a sequence of possible transitions from any initial state μ' to the equilibrium, which implies that all states other than the equilibrium are transient. Therefore, the process must converge to the equilibrium eventually.

To demonstrate that the equilibrium μ^* can be reached from any arbitrary initial state μ' I will first prove the claim that from any μ' each pure strategy state can be reached. Take any $\mu' \in M$, $\mu' \neq \mu^*$. If $\mu' \in \mathbf{R}(S_i)$, some i, then the pure strategy state S_i can be reached. If not, then an intermediate state in the interior of some best reply region can be reached and from there the process can mover to a pure strategy.

If the best reply graph of the game has a cycle involving all three strategies, e.g. $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_1$, where again $S_i \rightarrow S_j$ means that $S_j \in B(S_i)$, then all pure strategies can be reached starting from any of them. If no such cycle exists, one strategy, say S_3, is not a best reply to any other pure strategy. Since $S_i \notin B(S_i)$, $\forall i$, this implies that $S_1 \leftrightarrow S_2$, that is, S_1 can be reached from S_2 and vice versa. In this case there must exist some $(\alpha, 1-\alpha, 0) \in \mathbf{R}(S_3)$, which can be reached from both S_1 and S_2 since otherwise there would be a second equilibrium besides μ^*. Thus S_3 can be reached from S_1 or S_2 via the state $(\alpha, 1-\alpha, 0)$, which proves the claim.

It remains to be shown that the equilibrium can be reached from a pure strategy state. Assume that some μ satisfying condition (4) exists, that is,

$$\exists\, \mu \in \mathbf{R}(S_i) \text{ s.t. } \mu_i < \mu_i^* \text{ and } \mu_j > \mu_j^*,\ j \neq i.$$

For notational convenience let, WLOG, i = 1,

$$\exists\, \mu \in \mathbf{R}(S_1) \text{ s.t. } \mu_2 < \mu_1^*,\ \mu_3 > \mu_3^* \text{ and } \mu_2 > \mu_2^*.$$

Taking the projection from μ^* through μ onto the boundary of M we get a point $\bar{\mu}$ which must be an element of $\mathbf{R}(S_1)$ as well due to the convexity of the best reply regions.

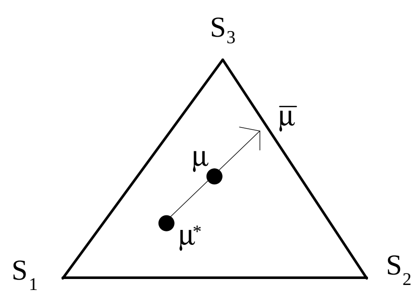

Figure 4

(a) Consider first the case in which $S_3 \in B(S_2)$ or $S_2 \in B(S_3)$.

In this case $\bar{\mu}$ can easily be reached from S_2 or S_3, respectively. In the next step the process can move from $\bar{\mu}$ to the equilibrium since $\bar{\mu} \in \mathbf{R}(S_1)$. Furthermore, since $\bar{\mu}$ is in the interior of $R(S_1)$ there exists a neighborhood of $\bar{\mu}$ from which the equilibrium can be reached as well. Thus, it is not required that the population size N is such that $\bar{\mu}$ is exactly one of the grid points of the state space.

(b) If neither $S_3 \in B(S_2)$ nor $S_2 \in B(S_3)$, strategy S_1's best reply region must look as in figure 5. In this case $\bar{\mu}$ cannot be reached from any of the pure strategies and an alternative path to the equilibrium has to be constructed.

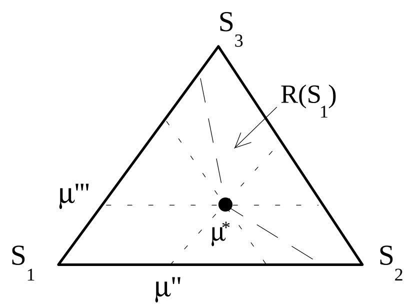

Figure 5

For this purpose consider the point $\mu'' = (1-\mu_2^*, \mu_2^*, 0)$ in figure 5 and note that μ'' can be reached from S_2. Since $S_3 \in R(S_1)$, $\mu'' \notin R(S_1)$ due to convexity of the best reply regions.

If $\mu'' \in R(S_3)$, the equilibrium can be reached from μ''. If not, $\mu'' \in R(S_2)$ and the best reply region for S_3 must be to the right of μ''. The latter statement follows because $R(S_1)$ and $R(S_2)$ cannot have a common point on the baseline of the simplex as such a point would be a second equilibrium. Hence, $\mu''' = (1-\mu_3^*, 0, \mu_3^*)$ must be in the best reply region of S_2 (it cannot be in $R(S_1)$ due to convexity). Therefore, the equilibrium can be reached from μ''' and thus - via S_3 - from any arbitrary initial state, which concludes the proof∎

For an example of how the path to the equilibrium is constructed consider the following game.

	S_1	S_2	S_3
S_1	0	-2	3
S_2	1	0	0
S_3	-2	3	0

The unique, completetely mixed equilibrium is $\mu^* = (1/3, 1/3, 1/3)$ and the best reply regions are as in figure 6.

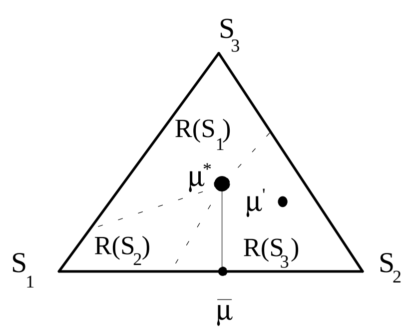

Figure 6

Suppose the initial state is $\mu' = (1/6, 1/2, 1/3)$ and that there exists some $\bar{\mu} \in \mathbf{R}(S_3)$. Since μ' is in the best reply region of strategy S_3, the state in which everyone plays S_3 can be reached if all players get the opportunity to switch their strategies. From there the process can move to S_1. If now $\bar{\mu}_2$ percent of the players get the opportunity to move, $\bar{\mu}$ can be reached. Finally, from $\bar{\mu}$ the process can move to the equilibrium.

IV. 3x3 GAMES: THE SMALL POPULATION CASE

In this section I will consider 3x3 games in the case of a "small" population which should be taken to mean that the exact population size is known and that a player finds it worthwhile to exclude himself from the population before calculating an optimal reply. It does not mean, however, that the population so small that the assumption of anonymous matching becomes senseless. For technical reasons it is sufficient to assume that $N > 1/\mu_i^*$, for all $\mu^*_i > 0$, but one should imagine a larger population than that. One justification for the myopic behavior of players is that a player does not think that his behavior can influence future play by his opponents. Clearly, in a population of four players this would be a bit unrealistic.

As in the last section I will first deal with games that have a unique mixed but not completely mixed equilibrium.

Proposition 2: Let μ^* be a mixed but not completely mixed Nash equilibrium. Then the best reply process for the small population case converges if and only if μ^* is absorbing.

The proof is essentially the same as the proof of proposition 1. There is no significant difference between the small or the large population case in terms of *getting to* the equilibrium because the intermediate steps are pure strategies states in which it does not matter whether a players excludes himself or not. The only difference is that in the small population case not all mixed Nash equilibria are absorbing states of the best reply process. And clearly the process cannot converge to a non-absorbing state.

For an example of a not completely mixed equilibrium that is non-absorbing consider the following game.

	S_1	S_2	S_3
S_1	3	2	3
S_2	4	0	4
S_3	1.5	4	0

The unique symmetric equilibrium is $\mu^* = (2/3, 1/3, 0)$. Figure 7 shows the best reply regions of the game.

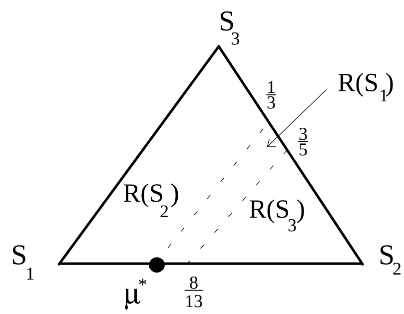

Figure 7

Given a population of 6 (by assumption the minimum population size is $N > 1/\mu_2 = 3$) a player who currently plays S_1 after excluding himself faces a strategy profile of $\mu^*_{-1} = (3/5, 2/5, 0)$. This profile, however, is already in the best reply region of S_3 since $3/5 < 8/13$. The S_1-player would switch to S_3 and, hence, the equilibrium cannot be an absorbing state.

Theorem 2: For symmetric 3x3 games with a unique, completely mixed Nash equilibrium μ^* *the best reply process in the small population case converges to the equilibrium if and only if* μ^* *is an absorbing state.*

Proof: It is obvious that the condition of μ^* being absorbing is necessary for convergence to a population mixed equilibrium. Note that this was of no concern in the large population case as there every equilibrium is absorbing. When players exclude themselves from the profile it is no longer true that every mixed equilibrium is absorbing, though, and clearly the process cannot converge to μ^* if μ^* is not absorbing.

Thus, I have to show that the condition of μ^* being absorbing is also sufficient for convergence. The equilibrium μ^* is absorbing if and only if $S_i \in B(\mu_{-i})$ for all i. Due to the convexity of the best reply regions $S_1 \in B(\mu_{-1})$ together with $S_1 \in B(\mu)$ implies that $\mu' \in B(\mu)$, where $\mu' = \left(0, \frac{\mu_2^*}{1-\mu_1^*}, \frac{\mu_3^*}{1-\mu_1^*}\right)$. Likewise $S_3 \in B(\mu_{-3})$ and $S_3 \in B(\mu)$ imply that $\mu'' \in B(\mu)$, where $\mu'' = \left(\frac{\mu_1^*}{1-\mu_3^*}, \frac{\mu_2^*}{1-\mu_3^*}, 0\right)$. Consider now point α in figure 8, where $\alpha = (0, \mu_2^*, 1-\mu_2^*)$.

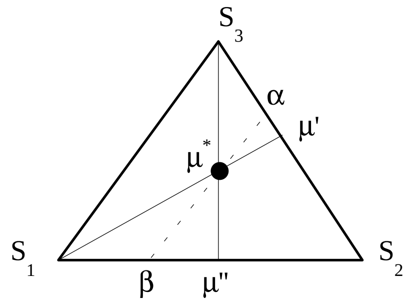

Figure 8

I claim that if $\alpha \in R(S_1)$, then the equilibrium can be reached from α. For this to be possible current S_3-players must want to switch to S_1. At α the proportion of S_3-players faced by a current S_3-player (after excluding himself) is $\frac{N(1-\mu_2^*)-1}{N-1}$. I will show that this proportion is larger than μ'_3, which implies that all point between α and μ' are contained in $R(S_1)$ and, hence, $\alpha \in \mathbf{R}_3(S_1)$.

$$\frac{N(1-\mu_2^*)-1}{N-1} > \mu_3' = \frac{\mu_3^*}{1-\mu_1^*} \Leftrightarrow$$
$$N\left((1-\mu_2^*)(1-\mu_1^*)-\mu_3^*\right) > 1-\mu_1^*-\mu_3^* \Leftrightarrow$$
$$\mu_1^* N(1-\mu_1^*-\mu_3^*) > 1-\mu_1^*-\mu_3^* \Leftrightarrow$$
$$\mu_1^* N > 1$$

The last inequality is satisfied by assumption. Thus μ^* can be reached from α if $\alpha \in R(S_1)$.

If $\alpha \notin R(S_1)$, then $\alpha \in \mathbf{R}(S_2)$. By convexity β cannot be in $R(S_2)$ and therefore $\beta \in R(S_3)$. Again, by the same argument as above, since $\mu'' \in R(S_3)$, $\beta \in R_1(S_3)$ and the equilibrium can be reached from β.

It remains to prove that from any initial state α and β can be reached. As in the proof of theorem 1 some pure strategy can be reached easily and once one pure strategy is reached all others can be reached as well. Again, if there is a cycle $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_1$, α and β can be reached from all pure strategies. If there is no such cycle, there is a strategy that is a best reply to both other strategies. Let S_2 be this strategy. This assures that there is a connection on the best reply path between S_1 and S_2 and between S_2 and S_3 and, hence, α and β can be reached.■

For an example of a unique, completely mixed equilibrium that is not an absorbing state of the best reply process for the small population case consider the following game.

	S_1	S_2	S_3
S_1	0	x	1
S_2	1	0	x
S_3	x	1	0

The equilibrium is (1/3, 1/3, 1/3). Note that for all x < 0 the equilibrium is not absorbing since $S_i \notin B(\mu^*_{-i})$. For example, if N = 9, a S_1-player faces a profile of (2/8, 3/8, 3/8) against which S_3 is the unique best reply. The process would always move away from the equilibrium. In contrast, the necessary and sufficient condition for convergence in the large population case is satisfied for all x > -1.

V. n×n GAMES

Generalizing the results from section III to all symmetric games n×n games was not (yet) completely successful. While I am able to give a sufficient condition for convergence to the population mixed equilibrium in the large population case, this sufficient condition is probably too strong. I suspect that a much simpler necessary and sufficient similar to the one given in theorem 1 exists but I was unable to prove this (but neither could I find a counterexample).

I will consider symmetric n×n games which either give rise to a unique, completely mixed equilibrium or can be reduced by iterative elimination of strictly dominated strategies to a game with a completely mixed equilibrium. That is, I will consider games with a mixed equilibrium whose carrier equals the set of rationalizable pure strategies. Let $T = \{T_1, ..., T_k\}$ denote the set of rationalizable strategies.

Since the best reply process gives rise to a stationary Markov chain on the finite state space M, M can be partitioned into transient and recurrent states. With time, either the process reaches an absorbing state or it enters a recurrent class (ergodic set) of states. Thus, the process may fail to converge only if there exists a non-singleton ergodic set $E \subseteq M$.

The aim is therefore to find a condition that guarantees that all ergodic sets are singletons. Note first that one can restrict attention to states in which only rationalizable strategies are used since all states in which strictly dominated strategies are used must be transient (if all players using a strictly dominated strategy get the opportunity to switch strategies, it is impossible for the process to return to a state in which such a dominated strategy is played by anyone). Therefore, if there exists a non-singleton ergodic set E, its *carrier* $K(E) = \{S_i \in S \mid \mu_i > 0 \text{ for some } \mu \in E\}$ cannot include strategies that are iteratively dominated, that is $K(E) \subseteq T$.

I will now introduce a condition that is sufficient for convergence to the unique mixed Nash equilibrium.[6] The intuitive idea is to impose a restriction on the payoff matrix such that in any arbitrary state the adjustment dynamics are directed towards the equilibrium. Let

$$O(\mu) \equiv \{ T_j \in T \mid \mu_j - \mu^*_i \geq \mu_i - \mu^*_i, \forall i = 1...k \}$$

[6] This condition was first introduced in an earlier paper of mine (Oechssler, 1993).

be the set of strategies in T that are the most over-represented in state μ relative to the equilibrium profile μ^*.

Definition 2: A game is *evolutionary stable with respect to pure strategies* (ESPS) if
$\forall\ \mu \in \Delta T, \mu \neq \mu^*, \exists\ T_i \in O(\mu)$ and $T_j \notin O(\mu)$ s.t. $T_i \notin B(\mu)$ and $T_j \in B(\mu)$.

A simpler, but overly restrictive, stability condition would be to require that none of the over-represented strategies can be a best reply. This way it would be easy to guarantee that the process converges to an equilibrium. ESPS requires only that *at least one* of the strategies that are over-represented *the most* in μ relative to the equilibrium profile is not a best reply to μ; and that at least one of the best replies does not belong to the strategies that are over-represented the most. In particular, ESPS implies that if just one player deviates from the equilibrium profile, his new strategy would not be a best reply against the new strategy profile and, hence, ESPS implies the weaker condition of theorem 1 for 3x3 games.

Theorem 3: The best reply process for the large population case converges to the unique mixed equilibrium for all symmetric $n \times n$ games satisfying ESPS.

As in the theorem 1 the strategy for the proof is to demonstrate that from any initial state a pure strategy state can be reached. This is done in the following lemma. Subsequently, it is shown that there exists a possible sequence of transitions from a pure strategy to the equilibrium. Hence, all states other than μ^* are transient and the process must eventually converge to the equilibrium.

Lemma: If $E \subseteq M$ is a non-singleton ergodic set, then there exists a $\mu' \in E$ such that $\mu_i = 1$ and $\mu_j = 0, \forall\ j \neq i$.

Proof: Again let $K(\mu)$ be the carrier of μ: $K(\mu) = \{ S_i \in S \mid \mu_i > 0 \text{ for some i} \}$. Take any $\mu^0 \in E$ in which different players use different strategies. Since μ^0 is not an absorbing state, not all strategies in the carrier of μ^0 can be best replies to μ^0. If there exists a strategy $S_j \in B(\mu^0)$, with $S_j \notin K(\mu^0)$, then the desired μ' can be reached in one step when all players simultaneously switch to S_j, which happens with positive probability. If no such S_j exists, the process moves with positive probability to some

$$\mu^1 \in \{ \mu \in E \mid \mu_i \geq \mu^0_i \text{ if } S_i \in B(\mu^0) \text{ and } \mu_i = 0 \text{ if } S_i \notin B(\mu^0)\}.$$

Note that the carrier of μ^1 must be a *strict* subset of the carrier of μ^0: $K(\mu^1) \subset K(\mu^0)$. Again, if there exists a strategy $S_j \in B(\mu^1)$, with $S_j \notin K(\mu^1)$, then μ' can be reached. If not, then with positive probability the process moves to some

$$\mu^2 \in \{ \mu \in E \mid \mu_i \geq \mu^1_i \text{ if } S_i \in B(\mu^1) \text{ and } \mu_i = 0 \text{ if } S_i \notin B(\mu^1)\},$$

where $K(\mu^2) \subset K(\mu^1)$. Repeating this procedure, we must eventually arrive at the desired μ' since the number of strategies is finite.■

Proof of theorem 3: The strategy for proving the theorem is to show that all ergodic sets are singletons, that is, they are absorbing states themselves. The supposition of an ergodic set being a non-singleton will be shown to yield a contradiction.

Assume to the contrary that there exists an ergodic set E that is not a singleton. It suffices to consider the reduced game consisting of the strategies in T since $K(E) \subseteq T$. By assumption there exists a symmetric equilibrium μ^* satisfying ESPS that is completely mixed on T. I will show by construction that the process converges from any $\mu \in E$ to the equilibrium μ^*, which is an absorbing state, contradicting the assumption of E being an ergodic set.

As a starting point take some $\mu \in E$ such that $\mu_i = 1$, which exists by the lemma and can be reached from any initial state in E since E is ergodic. WLOG let strategy T_i be called T_1, giving a starting profile $\mu^0 = (1,0,0,...,0)$.

Since $T_1 \notin B(\mu^0)$, all players would like to switch to some other strategy which is a best reply, say T_2. With positive probability exactly $(1-\mu_1^*)$ percent of players have the opportunity to do so while the remaining μ_1^* percent have to retain T_1. The resulting profile is $\mu^1 = (\mu_1^*, 1-\mu_1^*, 0, 0,...,0)$.

If $\mu^1 = \mu^*$, an absorbing state is reached. If not, $T_2 \notin B(\mu^1)$ because μ^* is assumed to satisfy ESPS: since $\mu^1_2 = 1-\mu_1^* > \mu_2^*$, $\{T_2\} = O(\mu^1)$, which implies that $T_2 \notin B(\mu^1)$.

The remaining possibilities are:

[1] $\{T_1\} = B(\mu^1)$, i.e. T_1 is the unique best reply. In this case players will switch from T_2 to T_1 and increase μ_1 until a profile $\overline{\mu}^1$ is reached, such that $T_1, T_2 \notin B(\overline{\mu}^1)$. Both, T_1 and T_2 are over-represented in $\overline{\mu}^1$. Note that $T_1, T_2 \in B(\overline{\mu}^1)$ is not possible because $\overline{\mu}^1$ would be a second equilibrium. Also, $\overline{\mu}^1_2 \geq \mu_2^*$ because if $\overline{\mu}^1_2$ became smaller than μ_2^* at some point before reaching $\overline{\mu}^1$, T_1 would be the only over-represented strategy and, hence, by ESPS could not be a best reply. From $\overline{\mu}^1$ we can proceed with the construction.

[2] $\{T_1\} \neq B(\mu^1) \rightarrow$ proceed with the construction.

In either case [1] or [2] there must be some other strategy, call it T_3, that is a best reply against $\overline{\mu}^1$ or μ^1, respectively. There exists a strictly positive probability that $1-\mu_1^*-\mu_2^*$ percent of the players switch to T_3 while the others stay at T_1 and T_2 respectively, yielding the new profile $\mu^2 = (\mu_1^*, \mu_2^*, 1-\mu_1^*-\mu_2^*, 0, 0,...,0)$.

If $\mu^2 = \mu^*$, an absorbing state is reached. If not, $T_3 \notin B(\mu^2)$ due to ESPS. By the same logic as above, either

[1] $\{T_1, T_2\} \not\supset B(\mu^2) \rightarrow$ proceed.

[2] $\{T_1, T_2\} \supseteq B(\mu^2)$, i.e., either T_1 or T_2 or both are the only best replies.

[2.1] Consider first the case $\{T_1, T_2\} = B(\mu^2)$. With positive probability the process moves to a state $\mu' = (\mu_1', \mu_2', \mu_3', 0, 0, ..., 0)$, where

$$\mu_1^{'} - \mu_1^{*} = \mu_2^{'} - \mu_2^{*} = \mu_3^{'} - \mu_3^{*}. \qquad (11)$$

Since $O(\mu') = \{T_1, T_2, T_3\}$, by ESPS there exists a $T_4 \in B(\mu')$ and at least one $T_i \in O(\mu')$, s.t. $T_i \notin B(\mu')$. Three cases can occur:

[2.11] $T_i \notin B(\mu')$, $\forall i = 1,2,3 \rightarrow$ proceed.

[2.12] Only one strategy remains a best reply. Assume, without loss of generality, that $T_1, T_2 \notin B(\mu')$. Due to the upper hemi-continuity of the best reply correspondence,

$$\exists\ \mu'' \equiv (\mu_1' - \varepsilon, \mu_2' - \varepsilon, \mu_3', 2\varepsilon, 0,.., 0), \text{ s.t. } T_1, T_2 \notin B(\mu'').$$

Since $O(\mu'') = \{T_3\}$, ESPS implies that $T_3 \notin B(\mu'')$. Furthermore, there exists a $T_4' \in B(\mu'')$. If $T_4 \in B(\mu'')$, let $T_4' = T_4$. Otherwise, everyone currently playing T_4 switches to $T_4' \rightarrow$ proceed.

[2.13] Two strategies remain a best reply. This case can be handled by applying [2.12] twice.

[2.2] The remaining cases are $\{T_1\} = B(\mu^2)$ or $\{T_2\} = B(\mu^2)$. Assume, WLOG, $\{T_1\} = B(\mu^2)$. With positive probability the process moves to $\mu''' = (\mu_1''', \mu_2^*, \mu_3''', 0, 0, ..., 0)$, where the μ_i''' are defined as in equation (11). By ESPS either T_1 or T_3 is not a best reply to μ'''. If both are not a best reply against $\mu''' \rightarrow$ proceed. If one of them remains a best reply, assume, WLOG, $T_1 \notin B(\mu''')$ and $T_3 \notin B(\mu''')$.

[2.21] $T_2 \in B(\mu''')$. If $T_2 \in B(\mu''')$, $\hat{\mu} = (\mu_1''' - \varepsilon', \mu_2^*, \mu_3''' + \varepsilon', 0, ..,0)$ can be reached where $\hat{\mu}$ is such that $T_1 \notin B(\hat{\mu})$ and $T_2 \in B(\hat{\mu})$. Since $O(\hat{\mu}) = \{T_3\}$, by ESPS, $T_3 \notin B(\hat{\mu})$. Hence, μ' can be reached and one can proceed as in [2.1].

[2.22] $T_2 \notin B(\mu''')$. In this case ESPS implies that there exists a $T_4 \in B(\mu''')$. [2.11] and [2.12] can then be applied analogously.

For all the above cases I have shown that a state can be reached in which none of the strategies in $\{T_1, T_2, T_3\}$ is a best reply. The construction can be continued from this state, since there is a positive probability that $1-\mu_1^*-\mu_2^*-\mu_3^*$ percent of players switch to T_4 while all others stay put, yielding $\mu^3 = (\mu_1^*, \mu_2^*, \mu_3^*, 1- \Sigma\, \mu_i^*, 0, 0,...,0)$.

Continuing in this fashion it is clear that an absorbing state must be reached starting from μ^0 in a finite number of steps with some probability $q > 0$. Since $\mu^0 \in E$, it will be visited infinitely often if E is an ergodic set as assumed. Hence, the probability of not reaching an absorbing state from μ^0 after u visits to μ^0, $(1-q)^u$, approaches zero as $u \rightarrow \infty$. Therefore, E must be a singleton.∎

The following example might be helpful to understand the logic of the proof. Consider the 4x4 game in figure 9.[7]

[7] Since the game is symmetric, only player I's payoffs are shown.

	A	B	C	D
A	0	0	10	0
B	10	0	0	0
C	0	10	0	0
D	4	4	4	-2

Figure 9

The unique, completely mixed equilibrium of the game is $\mu^* = (1/4,1/4,1/4,1/4)$. It can easily be checked that the game satisfies ESPS. The following table gives the sequence of transitions.

state μ^t	$B(\mu^t)$	step in proof
(1, 0, 0, 0)	B	
(1/4, 3/4, 0, 0)	C	
(1/4, 1/4, 1/2, 0)	A	[2.2]
(3/8, 1/4, 3/8, 0)	A,B	[2.21]
(3/8, 1/4+ε, 3/8−ε, 0)	B	
(1/3, 1/3, 1/3, 0)	D	[2.11]
(1/4, 1/4, 1/4, 1/4)	A,B,C,D	

Starting from an arbitrary pure strategy state, say $\mu^0 = (1,0,0,0)$, the process follows the steps outlined in the proof. Since B is the unique best reply against μ^0, the proportion of B-players must increase. With positive probability exactly 3/4 of players get the opportunity to switch and the new profile becomes (1/4, 3/4, 0, 0). Against this profile C is the best reply and the transition to the state (1/4, 1/4, 1/2, 0) is possible. This situation corresponds to stage [2.2] in the proof since against (1/4, 1/4, 1/2, 0) strategy A, is the best reply, that is, a strategy that is already at its 'right' proportion. Thus, as in [2.2] a state $\mu''' = (3/8, 1/4, 3/8, 0)$ is the next step. Against μ''' C is not a best reply, but A and B are, which corresponds to case [2.21] in the proof. If just one player switches from C to B, B becomes the unique best reply. Hence, as in [2.11] a state $\mu' = (1/3, 1/3 ,1/3, 0)$ can be reached. Finally, since D is the unique best reply to μ', the equilibrium can be reached from there.

As mentioned before the ESPS condition is not a necessary condition and can most likely be weakened.. The following example shows that there are 4x4 games (any $n\times x$ game could be constructed in a similar way) that converge with probability one even if they do not satisfy ESPS.

	A	B	C	D
A	0	0	-1	1
B	1	0	0	-1
C	-1	1	0	0
D	0	-1	1	0

The unique symmetric equilibrium is $\mu^* = (1/4, 1/4, 1/4, 1/4)$. To see that the game violates ESPS consider the profile $\mu' = (0, 1/2, 1/2, 0)$ for which both B and C are over-represented the most, i.e., $\{B, C\} = O(\mu')$. The unique best reply against μ' is C, i.e., $\{C\} = B(\mu')$. Since there is no strategy outside of $O(\mu)$ that is a best reply, ESPS is violated.

However, it is easy to see that the best reply process for this game converges from all initial states to μ^*. From any initial state a pure strategy state can be reached either directly or indirectly through an intermediate step. Since the best reply graph connects all pure strategies with the cycle $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$, one can start WLOG at the state (1, 0, 0, 0) against which B is the best reply. Thus, the process can move to (1/4, 3/4, 0, 0) and C becomes the best reply, which makes it possible to reach (1/4, 1/4, 1/2, 0). Against this state both B and D are best replies and, hence, the equilibrium can be reached when 1/4 of the population switches from C to D.

REFERENCES

BERGIN, J. and LIPMAN, B. (1993), "Evolution with State-Dependent Mutations" (mimeo: Queens University).

BROWN, G. (1951), "Iterative Solutions of Games by Fictitious Play", in Koopman, T., *Activity Analysis of Production and Allocation* (New York: Wiley).

CANNING, D. (1992), "Evolution of Language Conventions in Common Interest Signaling Games" (mimeo: Columbia University).

CRAWFORD, V. (1985), "Learning Behavior and Mixed-Strategy Nash Equilibria", *Journal of Economic Behavior and Organization*, **6**, 69-78.

ELLISON, G. (1993), "Learning, Local Interaction, and Coordination", *Econometrica*, **61**, 1047-1071.

FUDENBERG, D. and KREPS, D. (1993), "Learning Mixed Equilibria", *Games and Economic Behavior*, **5**, 320-367.

FOSTER, D. and YOUNG, P. (1990), "Stochastic Evolutionary Game Dynamics", *Theoretical Population Biology*, **38**, 219-232.

HARSANYI, J. (1973), "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed Strategy Equilibrium Points", *International Journal of Game Theory,* **2**, 1-23.

JORDAN, J. (1993), "Three Problems in Learning Mixed-Strategy Nash-Equilibria", *Games and Economics Behavior*, **5**, 368-386.

KANDORI, M., MAILATH, G. and ROB, R. (1993), "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica,* **61**, 29-56.

KANDORI, M. and ROB, R. (1993), "Evolution of Equilibria in the Long Run: A General Theory and Applications" (revised version, mimeo Princeton University).

NÖLDEKE, G. and SAMUELSON, L. (1993), "An Evolutionary Analysis of Backward and Forward Induction", *Games and Economic Behavior,* **5**, 425-454.

OECHSSLER, J. (1993), "Competition among Conventions", (mimeo: Columbia University).

OSBORNE, M. and RUBINSTEIN, A. (forthcoming), *A Course in Game Theory* (Cambridge, Mass.: MIT Press).

ROBINSON, J. (1951), "An Iterative Method of Solving a Game", *Annals of Mathematics*, **54**, 296-301.

ROSENTHAL, R. (1979), "Sequences of Games with Varying Opponents", *Econometrica*,**47**, 1353-1366.

RUBINSTEIN, A. (1988), "Comments on the Interpretation of Game Theory", STICERD-London School of Economics Working Paper No. TE/88/181.

SAMUELSON, L. (1991), "How to tremble if You Must" (mimeo, University of Wisconsin).

SHAPLEY, L. (1964), "Some Topics in Two-Person Games", in Dresher, M., Shapley, L. and Tucker, A. (eds.), *Advances in Game Theory* (Princeton: Princeton University Press).

YOUNG, P. (1993), "The Evolution of Conventions", *Econometrica,* **61**, 57-84.