Basins of Attraction and Equilibrium Selection Under Different Learning Rules

Russell Golman and Scott E Page*

May 22, 2009

*This work was funded by an NSF-IGERT grant and a U.S. Air Force MURI grant.

Abstract

A deterministic learning model applied to a game with multiple equilibria produces distinct basins of attraction for those equilibria. In symmetric two-by-two games, basins of attraction are invariant to a wide range of learning rules including best response dynamics, replicator dynamics, and fictitious play. In this paper, we construct a class of three-by-three symmetric games for which the overlap in the basins of attraction under best response learning and replicator dynamics is arbitrarily small. We then derive necessary and sufficient conditions on payoffs for these two learning rules to create basins of attraction with vanishing overlap. The necessary condition requires that pure, uniformly evolutionarily stable strategies are almost never initial best responses. The existence of parasitic or misleading actions allows subtle differences in the learning rules to accumulate.

KEYWORDS: Adjustment dynamics, attainability, basins of attraction, best response dynamics, coordination game, equilibrium selection, evolutionary game, learning, replicator dynamics.

JEL classification code: C73

1 Background

The existence of an equilibrium in a game is insufficient proof of its plausibility as an outcome. We must also describe a process through which players can achieve it. The distinction between the existence of an equilibrium and its attainability, and the necessity of the latter, rests at the foundations of game theory. In Nash's 1951 thesis, he proposed an adjustment dynamic built on a mass action model to support the convergence to an equilibrium (Weibull, 1996). The Nash adjustment dynamic relies on self interested behavior to move a population of players toward equilibrium. Unfortunately, it fails to achieve equilibria for many games. For this reason, game theorists building on Nash's original work focused instead on fictitious play, a learning rule in which players successively choose a pure strategy which is optimal against the cumulated history of the opponent's plays (Brown, 1951). More recent research by economists, psychologists, and theoretical biologists has produced a variety of adjustment dynamics, many of which fall into two broad categories: *belief based learning* and *reinforcement based learning*.[1] In the former, players take actions based on their beliefs of the actions of others. In the latter, players mimic actions that have been successful in the past (see Fudenberg and Levine, 1998; Camerer, 2003; Swinkels, 1993).

In this paper, we focus on two learning dynamics /adjustment processes: *continuous time best response dynamics* (Gilboa and Matsui, 1991) and *replicator dynamics*

[1]These categories also go by the terms *epistemic learning* and *behavioral learning* respectively (Walliser, 1998).

(Taylor and Jonker, 1978) and explore the extent to which they can differ in their basins of attraction for symmetric games with strict equilibria. For any two-by-two symmetric game, these two learning rules produce identical basins of attraction. We show that by adding a single action, we can produce a game in which these two learning rules create basins of attraction that have arbitrarily small overlap. In other words, best response dynamics lead to a different equilibrium than replicator dynamics almost always. Within our class of three-by-three games, the equilibrium found by the replicator dynamics is a *uniformly evolutionarily stable strategy*, but it is almost never the initial best response. The condition that pure, uniformly evolutionarily stable strategies satisfy this *never an initial best response property* proves to be a necessary requirement for the two learning rules to share vanishing overlap in their basins of attraction for strict equilibria. These results extend to classes of dynamics that generalize the best response and replicator rules.

To show how these rules can produce such different outcomes, we must first describe how best response and replicator dynamics model a population of adapting agents in the aggregate. In general, we assume players are randomly matched from large population pools. Best response dynamics are a form of belief-based learning – players' action choices depend on their beliefs about the actions of other players. In continuous time best response dynamics, a population of players moves toward a best response to the current state of the opposing population. Fictitious play relies on belief-based learning in discrete time. In each period, the rule assigns new beliefs based on the average play of the opponent. Actions are chosen rationally given those beliefs. The best response dynamics can be thought of as the extension of fictitious play to continuous time (Hofbauer and Sigmund, 2003).[2]

In contrast to best response dynamics, replicator dynamics are a form of reinforcement learning – actions spread based on their past success (Erev and Roth, 1998).[3] Replicator dynamics have ecological foundations: payoffs are analogous to fitness, and fitter actions are more apt to survive and grow. Note that actions initially not present in the population can never be tried with replicator dynamics.

In this paper, we consider symmetric matrix games. The learning dynamics thus operate in a single, large, well-mixed population. In this setting, the continuous time best response dynamics and replicator dynamics can be derived as the expected behavior of agents with stochastic protocols for switching their actions (Sandholm, 2009). In the best response dynamics, some infinitesimal proportion of the agents are always switching their action to match the current best response. The resulting flows are piecewise linear. In the replicator dynamics, agents copy better performing members of the population (Schlag, 1998). Players do not rely on beliefs about the actions of others. They need only know the payoffs of actions they encounter. Learning by imitation at the agent level thus leads to reinforcement learning at the

[2]The connection between fictitious play and best response dynamics requires the view that in fictitious play, a new agent enters the population each round with an action that is fixed forever. The state variable must then take on an interpretation as the opponent's population mixed strategy.

[3]The aforementioned Nash learning rule, or what is now called the Brown - von Neumann - Nash (BNN) dynamics also can be interpreted as a form of reinforcement learning (Brown and von Neumann, 1950; Skyrms, 1990).

population level.

Belief-based learning rules, such as best response, and reinforcement learning rules, such as replicator dynamics can be combined in a single learning rule called experience-weighted attraction (EWA) learning (Camerer and Ho, 1999). EWA can be made to fit either model exactly or to create a hybrid model that balances beliefs about future plays against past history of success. In experimental tests across a variety of games, belief-based learning, reinforcement learning, and EWA learning all predict behavior with reasonable accuracy. EWA outperforms the two pure models, though this is partly due to the fact that it has more free parameters.

The extant theoretical and empirical literature suggests that often these distinct learning rules make similar predictions about rates of change of actions and that for many games, they select identical equilibria. We know, for example, that any strict pure Nash Equilibrium will be dynamically stable under nearly all learning dynamics and that interior evolutionarily stable strategies are globally stable for both replicator dynamics (Hofbauer et al., 1979) and best response dynamics (Hofbauer, 1995; Hofbauer, 2000). Hopkins (1999) shows that stability properties of equilibria are robust across many learning dynamics, and, most relevant for our purposes, that best response dynamics and replicator dynamics usually have the same asymptotic properties. Best response dynamics and replicator dynamics are both myopic adjustment dynamics – they both flow towards higher immediate payoffs (Swinkels, 1993). Feltovich (2000) finds that belief-based learning and reinforcement learning generate qualitatively similar patterns of behavior, as does Salmon (2001), whose analytic survey concludes that only subtle differences exist across the various learning rules in extant experiments. Thus, advocates of each learning rule can point to substantial empirical support.

Our finding, that the choice of learning rule has an enormous effect on the choice of equilibrium, points to the importance of determining how people actually learn. And while the experimental work just mentioned has found this a difficult prospect, our class of games offers an opportunity to distinguish between different types of learning. Experiments on our games would have to find one, or possibly both, of the learning rules to be inconsistent with observed behavior.

Our results may at first seem to contradict the existing current literature. We want to make clear that they do not. First, many experiments consider two-by-two games. And as we review here, the two learning rules generate identical basins of attraction for two-by-two symmetric matrix games. The learning rules differ only in the time that they take to reach those equilibria. Second, our analysis focuses on *basins of attraction*, i.e. we ask which equilibrium is reached given an initial point. Most of the existing theorems consider *stability*, i.e. whether an equilibrium is stable to perturbations. Proving that an equilibrium is stable does not imply anything about the size of its basin of attraction. An equilibrium with a basin of attraction of measure epsilon can be stable. Thus, results that strict equilibria are stable for both replicator dynamics and best response dynamics do not imply that the two dynamics generate similar basins of attraction.

Conditions on payoff matrices that imply that best response dynamics, replicator dynamics, and Nash dynamics all produce similar stability properties need not place much restriction on basins of attraction, unless the stability is global. Conditions for global stability of each dynamic, for example if the mean payoff function is strictly concave (Hofbauer and Sigmund, 2003), imply identical basins of attraction. However, such conditions also imply a unique stable equilibrium.[4] One branch of the learning literature does consider games in which stability depends on the learning dynamic (Kojima, 2006) as well as games with distinct basins of attraction for different learning rules (Hauert et al., 2004). Those models rely on nonlinear payoff structures. Here, we consider matrix games with linear payoffs.

Of course, in a symmetric rock-paper-scissors game or an asymmetric matching pennies game, best response dynamics converges to the mixed equilibrium while replicator dynamics cycles. In these games, the mixed equilibrium is attracting under best response dynamics, but is only neutrally stable under replicator dynamics. Rock-paper-scissors is a knife edge case, where a slight change in payoffs could make the equilibrium stable under replicator dynamics, but matching pennies illustrates the inability of replicator dynamics to attain a mixed equilibrium in any asymmetric game. Our focus here is different. We analyze symmetric games with strict equilibria. The equilibria are asymptotically stable under both dynamics. We identify divergent

[4]Similar logic applies to repelling equilibria: if the mean payoff function is strictly convex, then a possible interior Nash Equilibrium must be repelling for each dynamic. Hofbauer and Sigmund's theorem (2003) follows from earlier work with each dynamic (Hofbauer and Sigmund, 1998; Hofbauer, 2000; Hopkins, 1999).

behavior of the learning rules, not because one rule fails to attain an equilibrium, but because the two rules select different equilibria.

To prove our results, we consider each possible initial distribution over actions and then characterize how the various learning rules specify the path of future distributions. In the games we consider, these continuous flows attain equilibria. Thus, the equilibrium selected can be thought of as a function of the initial population distribution of actions and the learning rule.

Our result that the choice of learning rule can determine the equilibrium selected can be interpreted through the lens of the equilibrium refinement literature (Harsanyi and Selten, 1988; Govindan and Wilson, 2005; Samuelson, 1997; Basov, 2004). In games with multiple strict Nash Equilibria, dynamical models with persistent randomness select long run, stochastically stable equilibria, which generalize the notion of risk dominance from two-by-two games (Foster and Young, 1990; Kandori et al., 1993). The stochastically stable equilibrium in a 3-by-3 game can vary with the learning dynamic (Ellison, 2000). Long run stochastic stability depends on the relative sizes of basins of attraction, given the underlying deterministic dynamic. Thus, even though we deal with deterministic dynamics only, our result complements the literature on stochastic stability by further supporting the conclusion that long run equilibria can be sensitive to how players learn. Our findings establish that the importance of learning style in equilibrium selection does not strictly rely on the presence of shocks that shift the population from one equilibrium to another.

The remainder of the paper is organized as follows. In the next section, we define the learning rules and show how they generate similar behavior in a simple three-by-three coordination game. Then, we present our main results, which show that belief-based learning and reinforcement learning can be very different. In Section 4, we introduce generalizations of best response and replicator dynamics and extend our results to these classes of dynamics. We conclude with a discussion of the relevance of the attainability of equilibria.

2 The Learning Rules

In a population game, the state space for a given population X is the unit simplex $\triangle$. A point $\mathbf{x} \in \triangle$ denotes the fraction of the population playing each action and is thus called a population mixed strategy. A learning rule for population X operates on the state space $\triangle$ by specifying for any given payoff structure a dynamical system $\dot{\mathbf{x}} = \mathbf{V}_{\pi}(\mathbf{x}, t)$ such that $\triangle$ is forward invariant, i.e., trajectories stay within the simplex. We interpret the learning dynamic as tracking the changes in the proportions of agents choosing the various actions.

We first introduce our learning rules in the context of a two-player game with large populations X and Y of randomly matched agents with n and m actions respectively. Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ be the population mixed strategy vectors. The component x_i (or y_i) is the fraction of population X (or Y) choosing action i. We will refer to the fraction of population X (or Y) choosing an action other than i as x_{-i} (or y_{-i}). Denote by π_i^{μ} the payoff a player in population μ gets from action

i. Of course, payoffs are a function of the opposing population mixed strategy, but we omit the function's argument for ease of notation, writing π_i^X in place of $\pi_i^X(\mathbf{y})$. Denote the vector of these payoffs by $\vec{\pi}^\mu = (\pi_1^\mu, \ldots, \pi_n^\mu)$.

The continuous time replicator dynamics can be written as

$$\begin{aligned} \dot{x}_i &= x_i(\pi_i^X - \bar{\pi}^X) \\ \dot{y}_i &= y_i(\pi_i^Y - \bar{\pi}^Y) \end{aligned}$$

where $\bar{\pi}^\mu$ is the average payoff in population μ. Specifically, $\bar{\pi}^X = \mathbf{x} \cdot \vec{\pi}^X$ and $\bar{\pi}^Y = \mathbf{x} \cdot \vec{\pi}^Y$.

Let BR($\mathbf{y}$) be the set of best replies to $\mathbf{y}$ (for a player in population X),

$$\mathrm{BR}(\mathbf{y}) = \operatorname*{arg\,max}_{\mathbf{v} \in \triangle^{n-1}} \mathbf{v} \cdot \vec{\pi}^X.$$

Similarly, the set of best replies to $\mathbf{x}$ is:

$$\mathrm{BR}(\mathbf{x}) = \operatorname*{arg\,max}_{\mathbf{v} \in \triangle^{m-1}} \mathbf{v} \cdot \vec{\pi}^Y.$$

Continuous time best response dynamics can be written as

$$\dot{\mathbf{x}} \in \mathrm{BR}(\mathbf{y}) - \mathbf{x} \qquad \dot{\mathbf{y}} \in \mathrm{BR}(\mathbf{x}) - \mathbf{y}.$$

The discrete fictitious play learning rule can be written as

$$\mathbf{x}(t+1) = \frac{t\mathbf{x}(t) + \mathbf{b}(t)}{t+1}$$

where $\mathbf{x}(t)$ is the vector of frequencies each action has been played through period t and $\mathbf{b}(t)$ is a best response to the opponent's history at this point. Fictitious play closely approximates continuous time best response dynamics. To avoid repetition, we focus on the best response dynamics. Results for best response hold for fictitious play as well.

This paper focuses on symmetric matrix games. In these games, both players have the same set of available actions and payoffs are linear. We can define the learning rules in the context of a single, well-mixed population, suitable for a symmetric game. The replicator dynamics are

$$\dot{x}_i = x_i(\pi_i - \bar{\pi}).$$

The superscripts can be dropped because in the single population setting there is no ambiguity in referring to the payoff to the average payoff $\bar{\pi}$ or the payoff to action i, π_i. The best response dynamics are

$$\dot{\mathbf{x}} \in \mathrm{BR}(\mathbf{x}) - \mathbf{x}.$$

2.1 An Example

To show how to apply these learning rules, we begin with an example of a simple three-by-three coordination game. In this game, the various learning rules generate similar basins of attraction. We borrow this game from Haruvy and Stahl (1999; 2000) who used it to study learning dynamics and equilibrium selection in experiments with human subjects. The payoff matrix for the Haruvy-Stahl game is written as follows:

$$\begin{pmatrix} 60 & 60 & 30 \\ 30 & 70 & 20 \\ 70 & 25 & 35 \end{pmatrix}.$$

The entry in row i and column j gives the payoff to a player who chooses action i and whose opponent chooses action j. This game has two strict pure Nash Equilibria: $(0, 1, 0)$ and $(0, 0, 1)$ as well as a mixed Nash Equilibrium at $(0, \frac{1}{4}, \frac{3}{4})$. It can be shown for both best response dynamics and replicator dynamics that the two pure equilibria are stable and that the mixed equilibrium is unstable.

Given that this game has three possible actions, we can write any distribution of actions in the two dimensional simplex $\triangle^2$. To locate the basins of attraction of each equilibrium, we must first identify those regions of the simplex $\triangle^2$ in which each action is a best response. This is accomplished by finding the lines where each pair of actions performs equally well. Let π_i be the payoff from action i. We find $\pi_1 = \pi_2$ when $4x_2 + 2x_3 = 3$, $\pi_2 = \pi_3$ when $17x_2 + 5x_3 = 8$, and $\pi_1 = \pi_3$ when $9x_2 + x_3 = 2$. These three lines determine the best response regions shown in Figure 1.

We can use Figure 1 to describe the equilibrium chosen under best response dynamics. Regions A, B, and C all lie the basin of attraction of action 3, while region D is in the basin of action 2. Note that the boundary of the basins of attraction under best response dynamics is a straight line.

In Figure 2, we characterize the basins of attraction for replicator dynamics. The boundary separating the basins of attraction here becomes a curve from the point

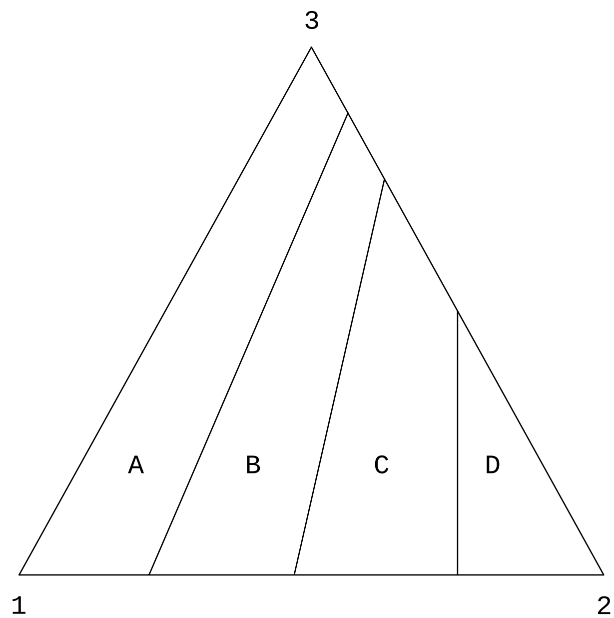

Figure 1: Best response regions. In region A, action 3 is the best response. In regions B and C, action 1 is the best response, but in B $\pi_3 > \pi_2$, while in C the opposite is true. In region D, action 2 is the best response.

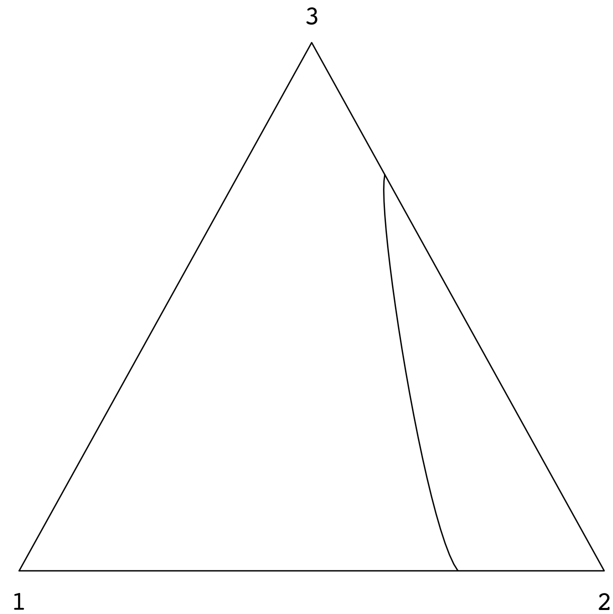

Figure 2: Basins of attraction under replicator dynamics.

$(\frac{1}{4}, \frac{3}{4}, 0)$ to $(0, \frac{1}{4}, \frac{3}{4})$ entirely within region C of Figure 1. Notice that the basins of attraction under best response dynamics and replicator dynamics differ. Best response dynamics creates basins with straight edges. Replicator dynamics creates basins with curved edges. This curvature arises because the second best action can also grow in the population under replicator dynamics. As it grows in proportion, it can become the best response. As a result, the population can slip from one best response basin into another one. Even so, notice that the difference in the two basins of attraction comprises a small sliver of the action space. We show this in Figure 3.

In games such as this, the two dynamics not only select the same equilibrium almost all of the time, but also generate qualitatively similar behavior. If the initial distribution of actions is close to $(0, 1, 0)$, the dynamics flow to that equilibrium point.

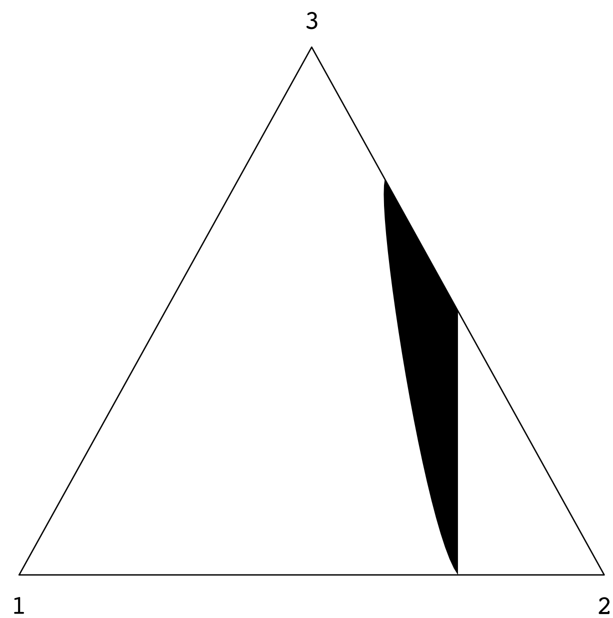

Figure 3: The small difference between best response and replicator dynamics. The shaded area flows to action 2 under replicator dynamics, to action 3 with best response dynamics.

If not, they flow to $(0, 0, 1)$.

In this game, the two learning rules create similar basins of attraction. Intuitively, we might expect only these small differences for all games with three actions, given the similarities of the learning rules. However, as we show in the next section, even with three-by-three games, the sliver can become almost the entire simplex.

3 Results

We now turn to our main results. We first present the well known fact that best response dynamics and replicator dynamics are identical for games with two possible actions. We consider learning dynamics to be identical if the direction of their flows is the same. This allows for differences in the speed of the flow. We then define a class of games with three actions in which the two learning rules generate basins of attraction with vanishing overlap. Within that class of games, an equilibrium action is almost never the initial best response. We show that to be a necessary condition for any symmetric game for which the two learning rules almost always lead to different strict equilibria.

Proposition 1 *For symmetric two-by-two matrix games, best response dynamics and replicator dynamics produce identical dynamics (Fudenberg and Levine, 1998).*

Proof The best response dynamics reduces to

$$\dot{x_i} = x_j$$

$$\dot{x_j} = -x_j$$

when $\pi_i > \pi_j$, and to $\dot{\mathbf{x}} = 0$ when they payoffs are equal. The replicator dynamics reduces to

$$\dot{x_1} = (\pi_1 - \pi_2)x_1x_2$$

$$\dot{x_2} = (\pi_2 - \pi_1)x_1x_2.$$

In both dynamics, the action with the higher payoff increases until the two payoffs become equal or the other action is completely eliminated.

Our first theorem says that there are three-by-three matrix games such that the two learning dynamics lead to different outcomes, for nearly all initial conditions. The claim cannot hold for all initial conditions because of the case where the initial point is a Nash Equilibrium of the game.

Theorem 1 *For any ϵ, there is a three-by-three game such that the fraction of the space of initial conditions from which best response dynamics and replicator dynamics lead to the same outcome is less than ϵ.*

We present a proof by construction. Consider the class of games with payoff matrix

$$\begin{pmatrix} 1 & -N & -N^{-1} \\ 2 - N^3 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1}$$

Lemma 1 *For any $N > 1$, both best response dynamics and replicator dynamics have two stable fixed points at:* $\mathbf{x} = (1, 0, 0)$ *and* $\mathbf{x} = (0, 1, 0)$.

Proof Both configurations are strict Nash Equilibria because both actions are strict best responses to themselves. Thus, a population in which all players take action 1 (resp. 2) would remain fixed. Strict Nash Equilibria are necessarily stable fixed points of both best response and replicator dynamics. The game also has an interior Nash Equilibrium which is unstable under either learning rule. These stable fixed points have to be Nash Equilibria, and no other Nash Equilibria exist.

Note that $(0, 0, 1)$ is not a Nash Equilibrium because action 2 is a best response. While it is a fixed point with respect to replicator dynamics, it cannot be stable.

Given two stable rest points, the eventual choice of one or the other depends on the initial distribution of play. The next result shows that for large N, best response dynamics almost always converge to all players taking action 2. The accompanying Figure 4 shows the flow diagram for the best response dynamics when $N = 5$.

Lemma 2 *For any ϵ, there exists M such that for all $N \geq M$, the basin of attraction of* $(0, 1, 0)$ *given best response dynamics is at least* $1 - \epsilon$ *of the action space.*

Proof First we show any point with $x_1 > \frac{1}{N}$ and $x_2 > \frac{1}{N}$ is in the basin of attraction of $(0,1,0)$, assuming $N > 2$. For such a point, action 3 is initially a best response because $\pi_3 = 0$ whereas $\pi_1 = x_1 - Nx_2 - \frac{1}{N}x_3 < 0$ and $\pi_2 = 2 - N^3 x_1 < 0$. Then, as we show, action 1 never becomes a best response. So, eventually, the dynamic flows toward action 2.

Because actions which are not best responses have the same relative decay rate,

$$\frac{x_1(t)}{x_1(0)} = \frac{x_2(t)}{x_2(0)}$$

for t such that action 3 is still a best response. So $x_1(t) - Nx_2(t) < 0$ for all t because it holds for $t = 0$. Action 3 dominates action 1. Action 3 is not a Nash Equilibrium, so eventually another action must become the best response, and the only candidate is action 2. Once x_1 falls to $\frac{2}{N^3}$, action 2 dominates forever.

Thus, by choosing N large enough, the basin of attraction of $(0,1,0)$ can be made as large as desired.

The next lemma shows that for large N, replicator dynamics leads to all players taking action 1 for almost any initial condition. Figure 5 shows the replicator dynamics flow pattern when $N = 5$.

Lemma 3 *For any ϵ, there exists M such that for all $N \geq M$, the basin of attraction of $(1,0,0)$ given replicator dynamics is at least $1-\epsilon$ of the action space.*

Proof

$$\dot{x_1} = x_1 \left((x_1 - Nx_2 - \frac{1}{N}x_3)(1 - x_1) - 2x_2 + N^3 x_1 x_2 \right).$$

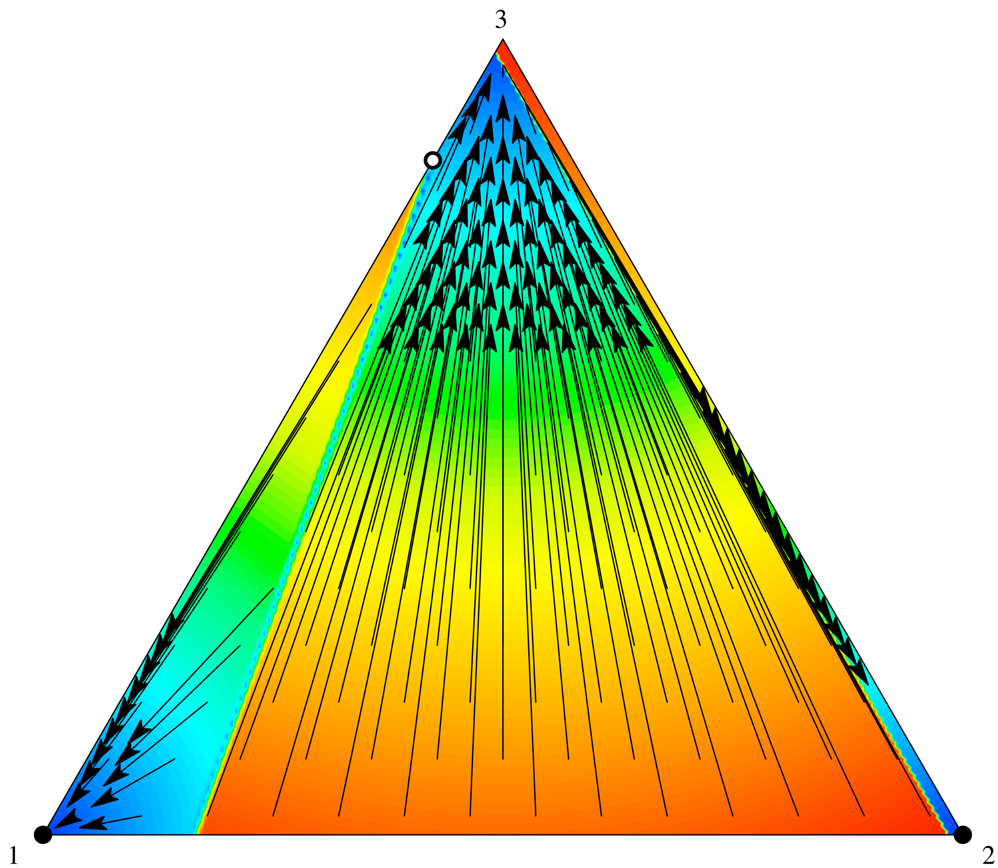

Figure 4: Phase diagram for the best response dynamics in the game used to prove Theorem 1, setting $N = 5$. Black (white) circles are stable (unstable) rest points. Most trajectories initially move towards action 3, but from this corner then flow to action 2. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

If $x_1 > \frac{1}{N}$, then $x_1 - \frac{1}{N}x_3 > 0$. For $N > 2$, $x_1 > \frac{1}{N}$ also implies $-Nx_2(1-x_1) - 2x_2 + N^3x_1x_2 > 0$ because $N^3x_1 > N^2 > N(1-x_1) + 2$.

So, for $N > 2$, if $x_1 > \frac{1}{N}$, then $\dot{x_1} > 0$. This means the replicator dynamics will flow to action 1.

By choosing N large enough, the basin of attraction of $(1, 0, 0)$ can be made as large as desired.

Thus, we have proved Proposition 2, that as N approaches infinity, best response dynamics and replicator dynamics converge to different equilibria.

Proposition 2 *In the limit as $N \to \infty$, the Lebesgue measure of the set of initial starting points for which best response dynamics and replicator dynamics flow to the same equilibrium tends to zero.*

This completes the proof of Theorem 1 above. Notice that in the class of games used in the proof, neither of the equilibrium strategies is an initial best response almost anywhere in the action space when N grows large. We say that these strategies satisfy the *Never an Initial Best Response Property* for such a sequence of games. To formally define this property, we must introduce some notation.

Let m be the Lebesgue measure on the action space. Given a vector of parameter values $\vec{P}$, let $G(\vec{P})$ be a class of games with payoffs that depend on those parameters. Let $\mathrm{BR}^{-1}(\mathbf{s})$ be the set of points $\mathbf{x}$ for which strategy $\mathbf{s}$ is a best response.

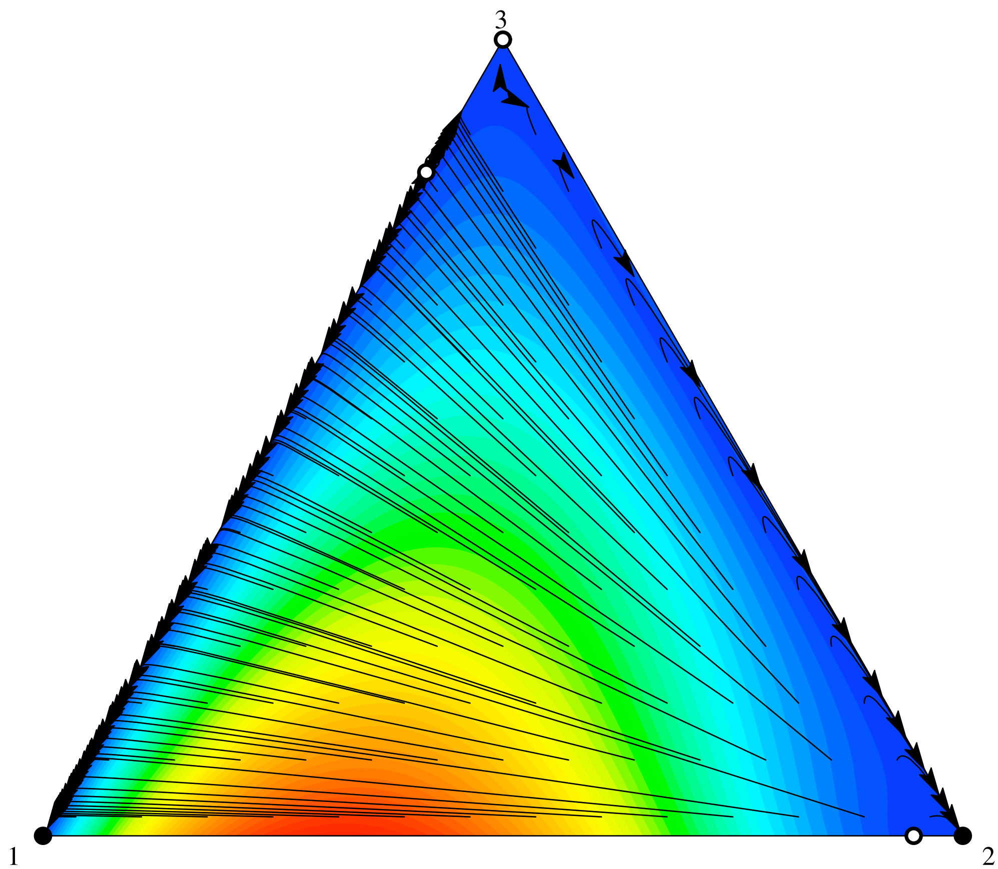

Figure 5: Phase diagram for the replicator dynamics in the game used to prove Theorem 1, setting $N = 5$. Most trajectories flow away from action 2 and then towards action 1. Figure made by the Dynamo program (Sandholm and Dokumaci, 2007).

Definition Strategy $\mathbf{s}$ satisfies the *Never an Initial Best Response Property* at $\vec{\hat{P}}$ if

$$\lim_{\vec{P}\to\vec{\hat{P}}} m\left(\mathrm{BR}^{-1}(\mathbf{s})\right) = 0.$$

Our next result makes use of the Never an Initial Best Response Property in establishing a necessary condition for there to be vanishing overlap in the basins of attraction created by best response dynamics and replicator dynamics. Before presenting this result, we need to lay down some more groundwork.

Recall that a strict equilibrium of a game is one in which each player's strategy is a strict best response to that equilibrium. We now extend the definition of a strict equilibrium to the limit of a sequence of games. Note that only pure Nash Equilibria can be strict.

Definition An equilibrium $\mathbf{s}$ is strict in the limit as $\vec{P} \to \vec{\hat{P}}$ if for all i such that $s_i > 0$,

$$\lim_{\vec{P}\to\vec{\hat{P}}} f(\vec{P})\left(\pi_i(\mathbf{s}) - \pi_j(\mathbf{s})\right) > 0 \text{ for all } j \neq i \text{ and some } f(\vec{P}) > 0. \tag{2}$$

Condition (2) is equivalent to the following condition: for all $\vec{P} \neq \vec{\hat{P}}$ in some neighborhood of $\vec{\hat{P}}$,

$$(\pi_i(\mathbf{s}) - \pi_j(\mathbf{s})) > 0 \text{ for all } j \neq i.$$

Strict equilibrium actions are also evolutionarily stable strategies (ESS), as we can see from Maynard Smith's original (1974) definition. An equilibrium $\mathbf{s}$ is an ESS if for all $\mathbf{s}' \neq \mathbf{s}$,

$$\mathbf{s} \cdot \vec{\pi}(\mathbf{s}) \geq \mathbf{s}' \cdot \vec{\pi}(\mathbf{s}),$$

with equality implying

$$\mathbf{s} \cdot \vec{\pi}(\mathbf{s}') > \mathbf{s}' \cdot \vec{\pi}(\mathbf{s}').$$

We can think of an ESS as an equilibrium satisfying an evolutionary stability condition that says that once it is fixed in the population, it will do better than any invading strategy as long as this invader is rare. Thomas (1985) reformulates this definition to allow for payoff functions that might be nonlinear.

Definition An equilibrium $\mathbf{s}$ is an ESS if for all $\mathbf{s}' \neq \mathbf{s}$ in some neighborhood of $\mathbf{s}$,

$$\mathbf{s} \cdot \vec{\pi}(\mathbf{s}') > \mathbf{s}' \cdot \vec{\pi}(\mathbf{s}').$$

We would like to extend this definition of an ESS to the limit of a sequence of games, but there are two ways to do this, depending on whether a different neighborhood of $\mathbf{s}$ may be chosen for each game in the sequence or a single neighborhood of $\mathbf{s}$ is chosen for the entire sequence. We are interested in the latter concept, which is a stronger condition, and we call it a *uniformly evolutionarily stable strategy.*

Definition An equilibrium $\mathbf{s}$ is a uniformly ESS in the limit as $\vec{P} \to \vec{\hat{P}}$ if there is a punctured neighborhood $\dot{U}(\mathbf{s})$ of $\mathbf{s}$ (i.e., a neighborhood from which the point $\mathbf{s}$ is removed) such that for all $\mathbf{s}' \in \dot{U}(\mathbf{s})$ and all $\vec{P} \neq \vec{\hat{P}}$ in some neighborhood of $\vec{\hat{P}}$,

$$\mathbf{s} \cdot \vec{\pi}(\mathbf{s}') > \mathbf{s}' \cdot \vec{\pi}(\mathbf{s}').$$

Note that if equilibrium $\mathbf{s}$ is strict in the limit as $\vec{P} \to \vec{\hat{P}}$, this implies that for all $\vec{P} \neq \vec{\hat{P}}$ in some neighborhood of $\vec{\hat{P}}$, the state $\mathbf{s}$ is an ESS, but it does *not* imply that $\mathbf{s}$ is a *uniformly ESS* in this limit.

An example of a uniformly ESS can be found in the class of games used to prove Theorem 1, with payoff matrix given by (1). In the limit as $N \to \infty$, the equilibrium strategy $(1, 0, 0)$ is a uniformly ESS, but the equilibrium strategy $(0, 1, 0)$ is not.

Our next results will make use of some additional notation. Given a learning rule $\mathcal{R}$ and an equilibrium action a of the game $G(\vec{P})$, let $B(\mathcal{R}, a, \vec{P})$ denote the basin of attraction of $(x_a = 1, x_{-a} = 0)$. Let $\mathbf{R}$ denote the replicator dynamics and $\mathbf{B}$ the best response dynamics.

In Theorem 2 below and the associated Corollary 1, we show that requiring pure, uniformly ESS to satisfy the *Never an Initial Best Response Property* is necessary if best response dynamics and replicator dynamics are to have basins of attraction with vanishing overlap. In the examples put forth here, this necessary condition entails the existence of either a parasitic action – an action that feeds off other actions but cannot survive on its own – or a misleading action – an action that looks good initially but eventually becomes less attractive as the population evolves.

Theorem 2 *Suppose for some action s,*

$$\lim_{\vec{P} \to \vec{\hat{P}}} m\left(B(\mathbf{R}, s, \vec{P}) \cap B(\mathbf{B}, s, \vec{P})\right) = 0.$$

Then, if $(x_s = 1, x_{-s} = 0)$ is a uniformly ESS, it satisfies the Never an Initial Best Response Property at $\vec{\hat{P}}$.[5]

[5] If we were to suppose that best response dynamics and replicator dynamics share vanishing overlap in their basins of attraction for an interior equilibrium, we could immediately conclude that this equilibrium is not a uniformly ESS. Interior ESS are, as already mentioned, globally asymptotically stable for both replicator and best response dynamics.

Proof We will denote the equilibrium point $(x_s = 1, x_{-s} = 0)$ by $\mathbf{s}$. Suppose that $\mathbf{s}$ is a uniformly ESS such that $m\left(\mathrm{BR}^{-1}(\mathbf{s})\right)$ remains strictly positive in the limit $\vec{P} \to \vec{\hat{P}}$. We will identify a nonvanishing region inside the basins of attraction of $\mathbf{s}$ for both replicator dynamics and best response dynamics.

Let $U(\mathbf{s})$ be a neighborhood of $\mathbf{s}$ such that $\dot{U}(\mathbf{s}) = U(\mathbf{s})\backslash\{\mathbf{s}\}$ satisfies the condition for $\mathbf{s}$ to be a uniformly ESS. Let $\nu = \sup_{\mathbf{x} \notin U(\mathbf{s})} x_s$. Define the neighborhood $W(\mathbf{s}) \subseteq U(\mathbf{s})$ of all points satisfying $x_s > \nu$. We have constructed $W(\mathbf{s})$ such that $\mathbf{x} \in \dot{W}(\mathbf{s})$ implies that $\dot{x}_s > 0$ under the replicator dynamics (because by the ESS condition, action s has better than average payoff here) and in turn, $\dot{x}_s > 0$ implies that $\mathbf{x}$ remains in $W(\mathbf{s})$.

We now observe that $\mathrm{BR}^{-1}(\mathbf{s})$ is a convex set because of the linearity of payoffs. Additionally, since $\mathbf{s}$ is a pure Nash Equilibrium, $\mathbf{s} \in \mathrm{BR}^{-1}(\mathbf{s})$. Thus, $\mathrm{BR}^{-1}(\mathbf{s})$ and $W(\mathbf{s})$ have positive intersection. By the fact that $W(\mathbf{s})$ is independent of $\vec{P}$ and our hypothesis that $\mathrm{BR}^{-1}(\mathbf{s})$ is nonvanishing, we conclude that $m\left(W(\mathbf{s}) \cap \mathrm{BR}^{-1}(\mathbf{s})\right)$ remains strictly positive in the limit $\vec{P} \to \vec{\hat{P}}$. Note that by the ESS condition and the linearity of payoffs, we can rule out the possibility that there are multiple best responses anywhere in the interior of $\mathrm{BR}^{-1}(\mathbf{s})$. For points $\mathbf{x}$ in the interior of $W(\mathbf{s}) \cap \mathrm{BR}^{-1}(\mathbf{s})$, best response dynamics flows to $\mathbf{s}$ because $\mathrm{BR}(\mathbf{x}) = \{\mathbf{s}\}$ and replicator dynamics flows to $\mathbf{s}$ because $\mathbf{x} \in W(\mathbf{s})$.

Theorem 2 leads directly to the following corollary.

Corollary 1 *Suppose*

$$\lim_{\vec{P}\to\vec{\hat{P}}} \sum_s m\left(B(\mathbf{R}, s, \vec{P}) \cap B(\mathbf{B}, s, \vec{P})\right) = 0.$$

Then every pure, uniformly ESS satisfies the Never an Initial Best Response Property at $\vec{\hat{P}}$.

Corollary 1 provides a necessary condition for non-overlapping basins. We can also derive several different sets of conditions that are sufficient to generate vanishing overlap in the basins of attraction of strict equilibria with best response and replicator dynamics. We present one such set of sufficient conditions for a symmetric three-by-three game here. Observe that the conditions we present are satisfied by the class of games used in the proof of Theorem 1.

To describe these conditions, we introduce some new notation and some simplifying assumptions. Let π_{ij} be the payoff to action i against action j, which by definition depends on the parameters $\vec{P}$. Since both dynamics are invariant under the transformations $\pi_{ij} \to \pi_{ij} + c$ for all i and fixed j and $\pi_{ij} \to k\pi_{ij}$ for all i, j with $k > 0$, we can set $\pi_{3j} = 0$ for all j and $|\pi_{11}| \in \{0, 1\}$. Also without loss of generality we can renumber the three actions so that $(x_1 = 1, x_{-1} = 0)$ denotes the equilibrium attained by replicator dynamics and $(x_2 = 1, x_{-2} = 0)$ the equilibrium attained by best response dynamics. Because these equilibria are strict in the limit as $\vec{P} \to \vec{\hat{P}}$, we have that for $j \in \{1, 2\}$, $i \neq j$, $\lim_{\vec{P}\to\vec{\hat{P}}} f_{jji}(\vec{P})(\pi_{jj} - \pi_{ij}) > 0$ for some functions $f_{jji} > 0$. And, by our choice of which equilibrium is to be found by each dynamic, we also have $\lim_{\vec{P}\to\vec{\hat{P}}} f_{321}(\vec{P})(\pi_{23} - \pi_{13}) > 0$ for some function $f_{321} > 0$.

Theorem 3

$$\lim_{\vec{P}\to\hat{\vec{P}}}\sum_{i=1}^{2} m\left(B(\mathbf{R},i,\vec{P})\cap B(\mathbf{B},i,\vec{P})\right)=0$$

if: i) $\pi_{23}>0$*; ii)* $\pi_{13}\leq 0$ *and* $\lim_{\vec{P}\to\hat{\vec{P}}}\pi_{13}=0$*;*[6] *iii)* $\lim_{\vec{P}\to\hat{\vec{P}}}\pi_{12}=-\infty$*; iv)* $\lim_{\vec{P}\to\hat{\vec{P}}}\frac{\pi_{21}}{\pi_{12}}=\infty$*; v)* $\lim_{\vec{P}\to\hat{\vec{P}}}\frac{\pi_{21}}{\pi_{22}}=-\infty$*; and vi)* $\lim_{\vec{P}\to\hat{\vec{P}}}\frac{\pi_{21}}{\pi_{23}}=-\infty$*.*

The proof relies on two lemmas, one for each learning dynamic.

Lemma 4 *As* $\vec{P}$ *approaches* $\hat{\vec{P}}$*, the fraction of the action space inside* $B(\mathbf{B},2,\vec{P})$ *approaches* 1*.*

Proof We first show that actions 1 and 2 satisfy the *Never an Initial Best Response Property* at $\hat{\vec{P}}$, that action 3 is initially a best response in all but an arbitrarily small part of the action space when $\vec{P}$ nears $\hat{\vec{P}}$. By the normalization condition, $\pi_3=0$. Therefore, it suffices to show $\pi_1<0$ and $\pi_2<0$.

1. $\pi_{\mathbf{2}}<\mathbf{0}$. Assuming $x_1>0$, $\pi_2=x_1\left(\pi_{21}+\frac{x_2}{x_1}\pi_{22}+\frac{x_3}{x_1}\pi_{23}\right)$. Condition (v) implies π_{21} dominates $\frac{x_2}{x_1}\pi_{22}$. Condition (vi) implies π_{21} dominates $\frac{x_3}{x_1}\pi_{23}$. And π_{21} is negative. So, for $\vec{P}$ near $\hat{\vec{P}}$, $\pi_2<0$.

2. $\pi_{\mathbf{1}}<\mathbf{0}$. Assuming $x_2>0$, $\pi_1=x_2\left(\pi_{12}+\frac{x_1}{x_2}\pi_{11}+\frac{x_3}{x_2}\pi_{13}\right)$. The normalization conditions imply $\pi_{11}=1$. Condition (iii) states that π_{12} approaches $-\infty$ while condition (ii) states that $\pi_{13}\leq 0$. So, for $\vec{P}$ near $\hat{\vec{P}}$, $\pi_1<0$.

Thus, for any point in the interior of the action space, $\vec{P}$ can be chosen such that action 3 is initially a best response.

[6] Another set of sufficient conditions might allow $\pi_{13}>0$, but would then require additional conditions to ensure that the best response dynamics avoids selecting $(1,0,0)$.

Now we show that under best response dynamics, action 3 dominates action 1 along the path towards $(0,0,1)$. Under best response dynamics, actions which are not best responses have the same relative decay rates. So $\frac{x_1}{x_2}$ remains constant along the path towards $(0,0,1)$. So π_1 remains negative along this path. By condition (i), action 3 is not a best response to itself. Eventually action 2 becomes the best response.

As the dynamic then moves toward $(0,1,0)$, π_1 remains negative because the π_{12} term becomes even more significant relative to the others. Action 1 never becomes the best response, so the best response dynamics lead to $(0,1,0)$.

Lemma 5 *As $\vec{P}$ approaches $\vec{\hat{P}}$, the fraction of the action space inside $B(\mathbf{R},1,\vec{P})$ approaches* 1.

Proof Under the replicator dynamics,

$$\dot{x_1} = x_1\left(\pi_{11}x_1(x_2+x_3) + \pi_{12}x_2(x_2+x_3) + \pi_{13}x_3(x_2+x_3) - \pi_{21}x_1x_2 - \pi_{22}x_2^2 - \pi_{23}x_3x_2\right).$$

Consider initial points that satisfy $x_1 > -\pi_{13}$ and $x_2 > 0$. Recalling that $\pi_{11} = 1$, this gives

$$\pi_{11}x_1(x_2+x_3) + \pi_{13}x_3(x_2+x_3) > 0. \tag{3}$$

By conditions (iv), (v), and (vi), $|\pi_{21}|$ grows faster than $|\pi_{12}|$, π_{22}, and π_{23} as $\vec{P}$ nears $\vec{\hat{P}}$. Consequently, the term with π_{21} dominates the other remaining terms in the expansion of $\dot{x_1}$. So, for $\vec{P}$ near $\vec{\hat{P}}$,

$$\pi_{12}x_2(x_2+x_3) - \pi_{21}x_1x_2 - \pi_{22}x_2^2 - \pi_{23}x_3x_2 > 0. \tag{4}$$

Thus, initially $\dot{x_1} > 0$. Moreover, by choosing $\vec{P}$ such that $\pi_{21} < \frac{1}{x_1(0)}(\pi_{12} - \pi_{22} - \pi_{23})$, we can be sure equation (4) holds as x_1 increases. As x_1 increases, it remains above $-\pi_{13}$, so equation (3) continues to hold as well. Thus, $\dot{x_1} > 0$ at all times.

It remains to show that the fraction of the action space satisfying $x_1 > -\pi_{13}$ and $x_2 > 0$ approaches 1 as $\vec{P}$ approaches $\hat{\vec{P}}$. This follows from (ii), which states that $\lim_{\vec{P} \to \hat{\vec{P}}} \pi_{13} = 0$. This implies that a point $\mathbf{x}$ need only satisfy $x_1 > 0$ and $x_2 > 0$ to be in $B(\mathbf{R}, 1, \vec{P})$ for some $\vec{P}$ near $\hat{\vec{P}}$.

We have thus described a set of six conditions which generate vanishing overlap in basins of attraction with best response dynamics and replicator dynamics in a class of games with only three actions.

Admittedly, none of the games within this class may be likely to arise in the real world. However, if we widen our scope and allow for more strategies, we can find games that map more tightly to real world phenomena and exhibit this same behavior. Consider the following symmetric matrix game with four actions, selected from a class of generalized stag hunt games (Golman and Page, 2008):

$$\begin{pmatrix} 2 & 2 & 2 & 2 \\ 1 & N+1 & 1 & 1 \\ 0 & 0 & 0 & N^2 \\ 0 & 0 & -N^2 & 0 \end{pmatrix}.$$

In this game, the first action is a safe, self interested action like hunting hare. The second action represents an attempt to cooperate, to hunt a stag, for example.

The third action is predatory toward the fourth action, which can be thought of as a failed attempt at cooperation. This fourth action fails to protect itself from the predator, fails to accrue benefits from coordination, and fails to guarantee itself a positive payoff. Clearly, a rational player would not choose it, and it is not played in equilibrium. Nevertheless, introducing predation into the stag hunt enhances the strategic context. This game captures a choice between the security of self-reliance, the productivity of cooperation, or the temptation of exploiting those agents who haven't yet learned what not to do. As we now show, when N goes to infinity, best response dynamics flow to an equilibrium in which all players choose action 1, but replicator dynamics flow to an equilibrium in which all players choose action 2.

Proposition 3 *In the four-by-four game above, as $N \to \infty$, the Lebesgue measure of the set of initial starting points for which best response dynamics and replicator dynamics flow to the same equilibrium tends to zero.*

Once again, the proof relies on three lemmas, one to identify the stable equilibria and two to describe the behavior of the learning rules.

Lemma 6 *Both best response dynamics and replicator dynamics have two stable fixed points:* $\mathbf{x} = (1, 0, 0, 0)$ *and* $\mathbf{x} = (0, 1, 0, 0)$.

Proof Here again, both configurations are strict Nash Equilibria because each of action 1 and 2 is a strict best response to itself. The only other Nash Equilibrium, $\mathbf{x} = \left(\frac{N-1}{N}, \frac{1}{N}, 0, 0\right)$, is clearly unstable given either dynamics. Note that action 4 is strictly dominated, and if we apply iterated elimination of strictly dominated actions, action 3 becomes strictly dominated once action 4 is eliminated.

The next lemma shows that for large N, best response dynamics leads to action 1 starting from almost any initial condition.

Lemma 7 *For any ϵ, there exists M such that for all $N \geq M$, the basin of attraction of $(1, 0, 0, 0)$ given best response dynamics is at least $1 - \epsilon$ of the action space.*

Proof First we show any point with $x_4 > \frac{2}{N}$ is in the basin of attraction of $(1, 0, 0, 0)$, assuming $N > 2$. For such a point, action 3 is initially a best response because $\pi_3 > 2N$ whereas $\pi_1 = 2$, $\pi_2 < 1 + N$, and $\pi_4 < 0$. Then, as we show, action 1 becomes a best response before action 2. Once it becomes a best response, it remains one forever, because its payoff is constant, while the payoffs to actions 2 and 3 are decreasing. So, once action 1 becomes a best response, the dynamic flows toward it thereafter.

Now we show that action 1 does become the best response before action 2. We define

$$\alpha(t) = \frac{x_1(t)}{x_1(0)} = \frac{x_2(t)}{x_2(0)} = \frac{x_4(t)}{x_4(0)}$$

for t such that action 3 is still a best response. The latter equalities hold because actions which are not best responses have the same relative decay rate. Note that $\alpha(t)$ is a strictly decreasing function. Now

$$\pi_1 = \pi_3 \text{ when } \alpha = \frac{2}{N^2\,(x_4(0))}.$$

But

$$\pi_2 < \pi_3 \text{ if } \alpha > \frac{1}{N\left(Nx_4(0) - x_2(0)\right)}.$$

Action 1 eventually becomes the best response because

$$\frac{2}{N^2\left(x_4(0)\right)} > \frac{1}{N\left(Nx_4(0) - x_2(0)\right)},$$

as long as $Nx_4(0) > 2x_2(0)$. This condition holds if $x_4(0) > \frac{2}{N}$.

Thus, by choosing N large enough, the basin of attraction of $(1, 0, 0, 0)$ can be made as big as desired.

Unlike best response dynamics, for large N, replicator dynamics leads to almost all players taking action 2 for almost any initial condition.

Lemma 8 *For any ϵ, there exists M such that for all $N \geq M$, the basin of attraction of $(0, 1, 0, 0)$ given replicator dynamics is at least $1 - \epsilon$ of the action space.*

Proof We now have $\dot{x_2} = x_2\left((1 + Nx_2)(1 - x_2) - 2x_1\right)$. So $\dot{x_2} \geq 0$ if $x_2 \geq \frac{1}{N}$. By choosing N large enough, the basin of attraction of $(0, 1, 0, 0)$ can be made as big as desired.

This completes the proof of Proposition 3. In this class of games, replicator dynamics flows to the equilibrium with the higher payoff, whereas in the class of games used in the proof of Theorem 1, the best response dynamics flows to the optimal equilibrium. Neither learning dynamic can find the optimal equilibrium in all classes of games because a different set of normalization conditions can change which equilibrium is optimal.

4 Broader Classes of Dynamics

We introduce two new classes of adjustment dynamics: one-sided payoff positive dynamics, which generalize the replicator dynamics, and threshold dynamics, a generalization of the best response dynamics. We then extend our results from the previous section to describe vanishing overlap in the basins of attraction of a one-sided payoff positive dynamic and a threshold dynamic.

As the name suggests, our *one-sided payoff positive dynamics* are closely related to the commonly known payoff positive dynamics (Weibull, 1995). Payoff positive dynamics assume that actions with above average payoffs have positive relative growth rates and actions with below average payoffs have negative relative growth rates.[7] The one-sided class of dynamics still captures the property that actions with above average payoffs grow in the population, but does not address what happens to actions with below average payoffs. Thus, the class of one-sided payoff positive dynamics includes all the payoff positive dynamics. They in turn contain the replicator dynamics, which prescribe a relative growth rate proportional to the difference between action's payoff and population mean payoff. Neither class of dynamics specifies precise rates of growth the way replicator does, making them both quite general.

Definition A *one-sided payoff positive dynamic* is one that satisfies the following condition:

$$\dot{x}_i > 0 \text{ if } \pi_i > \bar{\pi} \text{ and } x_i > 0 \tag{5}$$

[7] These dynamics are also termed sign-preserving (Nachbar, 1990).

as well as the requirements that Nash Equilibria are rest points and that if $\mathbf{z}$ is the limit of an interior orbit for $t \to \infty$, then $\mathbf{z}$ is a Nash Equilibrium[8].

These basic requirements are satisfied by most adjustment dynamics as part of a folk theorem (see Cressman, 2003). The second requirement holds whenever relative growth rates are Lipschitz continuous, for example. The other statements of the folk theorem, namely that stable rest points are Nash Equilibria and that strict equilibria are asymptotically stable, can be shown to hold from our definition of one-sided payoff positive dynamics.

Lipschitz continuous one-sided payoff positive dynamics fit into the even broader class of weakly payoff positive dynamics, which assume that some action with above average payoff will have positive relative growth whenever there is such an action. The distinction is that weak payoff positivity does not guarantee growth for *all* the above average actions in the population.

In contrast to one-sided payoff positive dynamics, in which agents seek actions with above average payoffs, we can conceive of a learning rule in which agents switch actions when their payoffs are at or below the median. But, there is no need to hold the 50th percentile payoff in such special regard as the threshold for switching. We define *threshold dynamics* by the property that agents switch away from actions with

[8]Theorem 5 would hold without these requirements, but with the possibility that the one-sided payoff positive dynamics have measure zero basins of attraction for all strict equilibria. We want to focus on the case that the one-sided payoff positive dynamic selects a different equilibria than the threshold dynamic.

payoffs at or below the K^{th} percentile as long as there is something better to switch to. We do not restrict ourselves to a particular threshold by setting a value for K. Instead, we allow K to vary over time within a range that is bounded below by some $\hat{K} > 0$. We sometimes omit writing the input t when dealing with a fixed instant in time and use K to mean the function $K(t)$. But $\hat{K}$ is always a constant. We want agents to be as averse as possible to actions whose payoffs fall below the threshold, but recognizing that the speed of the dynamics can be adjusted by an overall scaling factor, we only require that such an action has a relative rate of decline as quick as any in the population.

Definition Consider any $K(t) \geq \hat{K}$ where $\hat{K} > 0$. A *threshold dynamic* is one that satisfies the following condition:

If at time t

$$\sum_{\mu:\, \pi_\mu < \pi_i} x_\mu < K \text{ and for some } l,\ \pi_l > \pi_i, \tag{6}$$

then when $x_i > 0$,

$$\frac{\dot{x}_i}{x_i} \leq \frac{\dot{x}_j}{x_j} \text{ for all actions } j \text{ such that } x_j > 0 \tag{7}$$

and when $x_i = 0$, $\dot{x}_i = 0$.

Note that if two actions both have payoffs below the K^{th} percentile, they must have the same relative rate of decline. In addition, it is always the case that the action with the worst payoff declines in the population. On the other hand, there is no guarantee that the best response grows in the population unless other actions are sufficiently rare.

Definition A threshold dynamic is *properly scaled* if the speed of the dynamic has a lower bound $v(\mathbf{x})$ such that $v(\mathbf{x}) \geq \kappa d$ where d represents the distance to the nearest equilibrium point and κ is some constant of proportionality.

As previously mentioned, the speed of a dynamic can be adjusted by an overall scaling factor. Usually we do not care about speed at all because the scaling of the time parameter does not have physical significance. In this case, we must assume a threshold dynamic is properly scaled in order to ensure that it does not slow to a halt at an arbitrary point in the strategy space. In this paper, we consider only properly scaled threshold dynamics.

Let us briefly consider a few particular constant functions we might use for $K(t)$ in the threshold dynamics. If we do want the median payoff to be the threshold, we can choose $K(t) = .5$ for all t. Then all actions with payoffs equal to or below the median will have the same relative rate of decline, and actions with payoffs above the median will do no worse. That is, an action with a higher payoff may still decline at the same relative rate as the former or may grow very quickly; the threshold dynamics allow for either. This example suggests that when $K(t)$ is small, the dynamics allow for quite a bit of freedom in the center of the strategy space.

Alternatively, we could set the threshold below which agents switch actions to be the 100th percentile payoff at all times, $K(t) = 1$. Obviously, every action has a payoff at or below the 100th percentile payoff, so inequality (7) applies to every action present in the population that is not a best response. An agent already playing a best response cannot find something better to switch to, and thus best responses are not subject to this inequality. Every action that is subject to inequality (7) has the same relative rate of decline, and if we set this rate to be -1, we then obtain the best response dynamics. Thus, threshold dynamics are a generalization of the best response dynamics.

We now show that requiring pure, uniformly ESS to satisfy the *Never an Initial Best Response Property* is necessary if a one-sided payoff positive dynamic and a threshold dynamic are to have basins with vanishing overlap, just as Theorem 2 and Corollary 1 showed it is for best response and replicator dynamics. Let **OSPP** denote any one-sided payoff positive dynamic and **TD** any threshold dynamic.

Theorem 4 *Suppose for some action s,*

$$\lim_{\vec{P} \to \hat{\vec{P}}} m\left(B(\mathbf{OSPP}, s, \vec{P}) \cap B(\mathbf{TD}, s, \vec{P})\right) = 0.$$

Then, if $(x_s = 1, x_{-s} = 0)$ is a uniformly ESS, it satisfies the Never an Initial Best Response Property at $\hat{\vec{P}}$.

Proof The proof here mirrors the one for Theorem 2. We construct the neighborhood $W(\mathbf{s})$ in the same way, but with the additional condition that $x_s > 1 - \hat{K}$. We need only show that for $\mathbf{x} \in \mathrm{int}\big(W(\mathbf{s}) \cap \mathrm{BR}^{-1}(\mathbf{s})\big)$, both classes of dynamics flow to $\mathbf{s}$. Under one-sided payoff positive dynamics, $\dot{x}_s > 0$ for $\mathbf{x} \in W(\mathbf{s})$ because action s has an above average payoff, and such a flow cannot leave $W(\mathbf{s})$. Under threshold dynamics, when $\mathbf{x} \in \mathrm{int}\big(W(\mathbf{s}) \cap \mathrm{BR}^{-1}(\mathbf{s})\big)$, inequality (7) applies to all actions other than s because they have payoffs below the $\hat{K}^{\text{th}}$ percentile. All other actions must have the same negative growth rate, so $\dot{\mathbf{x}} = \kappa(\mathbf{s} - \mathbf{x})$ for some positive constant κ.

Corollary 2 *Suppose*

$$\lim_{\vec{P} \to \vec{\hat{P}}} \sum_s m\left(B(\mathbf{OSPP}, s, \vec{P}) \cap B(\mathbf{TD}, s, \vec{P})\right) = 0.$$

Then every pure, uniformly ESS satisfies the Never an Initial Best Response Property at $\vec{\hat{P}}$*.*

We also find that the same set of conditions used in Theorem 3 is sufficient for a one-sided payoff positive dynamic and a threshold dynamic to share vanishing overlap in their basins. Recall that the setting for this theorem is a symmetric three-by-three game with two strict equilibria.

Without loss of generality we choose $(x_1 = 1, x_{-1} = 0)$ to be the equilibrium attained by the one-sided payoff positive dynamic and $(x_2 = 1, x_{-2} = 0)$ the equilibrium attained by the threshold dynamic. Because these equilibria are strict in the limit as $\vec{P} \to \vec{\hat{P}}$, we have that for $j \in \{1, 2\}$, $i \neq j$, $\lim_{\vec{P} \to \vec{\hat{P}}} f_{jji}(\vec{P})(\pi_{jj} - \pi_{ij}) > 0$ for some functions $f_{jji} > 0$. And, by our choice of which equilibrium is to be found by each dynamic, we also have $\lim_{\vec{P} \to \vec{\hat{P}}} f_{321}(\vec{P})(\pi_{23} - \pi_{13}) > 0$ for some function $f_{321} > 0$. Once again, we set $\pi_{3j} = 0$ for all j and $\pi_{11} = 1$, but no longer are our dynamics necessarily invariant under positive affine transformations of the payoffs. If the dynamics happen to retain this invariance, then this still amounts to a choice of payoff normalization. However, in general, we are making an additional assumption about payoffs here.

Theorem 5

$$\lim_{\vec{P}\to\vec{\hat{P}}} \sum_{i=1}^{2} m\left(B(\mathbf{OSPP}, i, \vec{P}) \cap B(\mathbf{TD}, i, \vec{P})\right) = 0$$

if: i) $\pi_{23} > 0$*; ii)* $\pi_{13} \leq 0$ *and* $\lim_{\vec{P}\to\vec{\hat{P}}} \pi_{13} = 0$*; iii)* $\lim_{\vec{P}\to\vec{\hat{P}}} \pi_{12} = -\infty$*; iv)* $\lim_{\vec{P}\to\vec{\hat{P}}} \frac{\pi_{21}}{\pi_{12}} = \infty$*; v)* $\lim_{\vec{P}\to\vec{\hat{P}}} \frac{\pi_{21}}{\pi_{22}} = -\infty$*; and vi)* $\lim_{\vec{P}\to\vec{\hat{P}}} \frac{\pi_{21}}{\pi_{23}} = -\infty$.

Again, we break up the proof into two lemmas, one for each learning dynamic.

Lemma 9 *As* $\vec{P}$ *approaches* $\vec{\hat{P}}$*, the fraction of the action space inside* $B(\mathbf{TD}, 2, \vec{P})$ *approaches* 1.

Proof Consider the threshold dynamics with any value of $\hat{K}$ and any threshold function $K(t) \geq \hat{K}$. We show that if initially $x_2 > 0$, then for $\vec{P}$ near $\vec{\hat{P}}$, $\pi_1 < \pi_3$ at all times. As $\pi_3 = 0$ by the normalization condition, this amounts to showing $\pi_1 < 0$ forever. We know $\pi_1 = \pi_{11}x_1 + \pi_{12}x_2 + \pi_{13}x_3$. Recall that $\pi_{11} = 1$. Condition (ii) states that $\pi_{13} \leq 0$. So $\pi_1 < 0$ as long as

$$x_1 + \pi_{12}x_2 < 0. \tag{8}$$

Consider first the case that $x_2(0) \geq \hat{K}$. Take $\vec{P}$ near enough $\vec{\hat{P}}$ that $\pi_{12} < -\frac{1}{\hat{K}}$. Condition (iii) makes this possible. As long as $x_2 \geq \hat{K}$, equation (8) holds and we still have $\pi_1 < 0$.

In the case that $x_2(0) < \hat{K}$, condition (iii) allows us to take $\vec{P}$ near enough $\vec{\hat{P}}$ that $\pi_{12} < -\frac{1}{x_2(0)}$. This guarantees that $\pi_1 < 0$ initially.

If ever $x_2 < K$, then π_1 is below the K^{th} percentile and $\frac{\dot{x_1}}{x_1} \leq \frac{\dot{x_2}}{x_2}$, so equation (8) continues to hold. Still $\pi_1 < 0$.

In fact, the only way to avoid $x_2 < K$ at some time would involve $\pi_2 > 0$ pretty quickly. But, if indeed $x_2 < K$ at some time, then the decline in x_1 also would lead to $\pi_2 > 0$ eventually. So, one way or another, action 2 becomes the best response and x_1 has to decline. When $x_1 < \hat{K} \leq K$, the K^{th} percentile payoff is either 0 or π_2, and when additionally $\pi_2 > 0$, then only x_2 can grow. From then on, the dynamic moves straight toward $(0, 1, 0)$.

Lemma 10 *As $\vec{P}$ approaches $\vec{\hat{P}}$, the fraction of the action space inside $B(\mathbf{OSPP}, 1, \vec{P})$ approaches* 1*.*

Proof The proof of Lemma 5, which applied replicator dynamics to this game, carries over here, applying to all one-sided payoff positive dynamics with only trivial changes. We no longer have an exact formula for $\dot{x_1}$, but the argument that it is always positive still applies because it was based on the fact that $\pi_1 > \bar{\pi}$ at all times for almost all initial points. The definition of a one-sided payoff positive dynamic requires that the limit of an interior orbit is a Nash Equilibrium, and the only one that can be approached with $\dot{x_1} > 0$ is $(1, 0, 0)$.

We have thus extended our finding of vanishing overlap in basins of attraction for strict equilibria to entire classes of dynamics.

5 Discussion

In this paper, we have shown that it is possible to construct three-by-three symmetric games in which two common learning rules – replicator dynamics, and best response dynamics – have vanishing overlap in their basins of attraction. That so few actions are required is surprising, making the game we have constructed of significant pedagogical value. Our more general results describe necessary and sufficient conditions for vanishing overlap. The necessary condition – that for any game in which the learning rules attain distinct strict equilibria from almost any starting point the initial best response cannot be a uniformly ESS - has an intuitive explanation. The initial incentives must be misleading. They should point the agents away from equilibria and in some other directions. In doing so, these initial incentives allow for even small differences in the dynamics to take root and drive the two learning rules to distinct equilibria.

We also derived a set of sufficient conditions for the basins of attraction of two stable equilibria under best response learning and replicator dynamics to have almost no overlap. Other sufficient conditions could also be constructed. What appears invariant to the construction is that some payoffs must grow arbitrarily large.

Our focus on basins of attraction differentiates this paper from previous studies that consider stability. Nash was aware that the existence of an equilibrium is not sufficient proof that it will arise. Nor is proof of its local stability. We also need to show how to attain an equilibrium from an arbitrary initial point (Binmore and Samuelson, 1999). And, as we have just shown, the dynamics of how people learn can determine whether or not a particular equilibrium is attained. Richer models of individual and firm behavior can also support diverse choices of equilibria (Allen, Strathern, and Baldwin, 2007). Here, we emphasized the minimal conditions necessary for the learning rule to matter.

We also focused on the extreme case of no overlap. That said, our general findings about the necessity of misleading actions and the nature of our sufficient conditions should help us to identify games in which learning rules might matter. In particular, the idea that temporary best responses create opportunity for differences in learning rules to accumulate would seem to have wide applicability. It provides logical foundations for the intuition that learning rules matter more in more complex environments.

In conclusion, we might add that games in which best response dynamics and replicator dynamics make such different equilibrium predictions would seem to lend themselves to experiments. These games would allow experimenters to distinguish among learning rules more decisively than games in which the learning rules converge to the same equilibrium.

References

[1] Allen P.M, Strathern, M., Baldwin, J. S., 2007. Complexity and the limits to learning. *J. Evol. Econ.* 17, 401-431.

[2] Basov, S., 2004. Equilibrium selection in coordination games: Why do dominated strategies matter? *Econ. Bull.* 3, 1-3.

[3] Binmore, K., Samuelson, L., 1999. Evolutionary Drift and Equilibrium Selection. *Rev. Econ. Stud.* 66, 363-394.

[4] Brown, G.W.. 1951. Iterative Solutions of Games by Fictitious Play. In: Koopmans, T.C. (Ed.), *Activity Analysis of Production and Allocation.* Wiley, New York.

[5] Brown, G.W., von Neumann, J., 1950. Solutions of games by differential equations. *Ann. Math. Stud.* 24, 73-79.

[6] Camerer, C., 2003. *Behavioral Game Theory: Experiments on Strategic Interaction.* Princeton University Press, Princeton.

[7] Camerer, C., Ho, T.H., 1999. Experience Weighted Attraction Learning in Normal Form Games. *Econometrica* 67, 827-874.

[8] Cressman, R., 2003. *Evolutionary Dynamics and Extensive Form Games.* MIT Press, Cambridge.

[9] Ellison, G. 2000. Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution. *Rev. Econ. Stud.* 67 (1), 17-45.

[10] Erev, I., Roth, A.E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Amer. Econ. Rev.* 88 (4), 848-881.

[11] Feltovich, N., 2000. Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games. *Econometrica* 68, 605-641.

[12] Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. *Theoretical Population Biology* 38, 219-232.

[13] Fudenberg, D., Levine, D.K., 1998. *The Theory of Learning in Games.* MIT Press, Cambridge.

[14] Gilboa, I., Matsui, A., 1991. Social Stability and Equilibrium. *Econometrica* 59, 859-867.

[15] Golman, R., Page, S., 2008. Individual and Cultural Learning in Stag Hunt Games With Multiple Actions.

[16] Govindan, S., Wilson, R., 2005. Refinements of Nash Equilibrium. In: Durlauf, S., Blume, L. (Eds.), *The New Palgrave Dictionary of Economics* 2. Palgrave Macmillan, London.

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

[18] Haruvy, E., Stahl, D.O., 1999. Empirical Tests of Equilibrium Selection based on Player Heterogeneity. mimeo, University of Texas at Austin.

[19] Haruvy, E., Stahl, D.O., 2000. Robust Initial Conditions for Learning Dynamics. mimeo, University of Texas at Austin.

[20] Hauert, C., Haiden, N., Sigmund, K., 2004. The Dynamics of Public Goods. *Discrete and Continuous Dynamical Systems - Series B* 4, 575-587.

[21] Hofbauer, J., 1995. Stability for the best response dynamics. mimeo, University College London.

[22] Hofbauer, J., 2000. From Nash and Brown to Maynard Smith: equilibria, dynamics and ESS, *Selection* 1, 81-88.

[23] Hofbauer, J., Schuster, P., Sigmund, K., 1979. A note on evolutionary stable strategies and game dynamics. *J. Theor. Biol.* 81, 609-612.

[24] Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, U.K.

[25] Hofbauer, J., Sigmund, K., 2003. Evolutionary Game Dynamics. *Bull. Am. Math. Soc.* 40, 479-519.

[26] Hopkins, E., 1999. A Note on Best Response Dynamics. *Games Econ. Behav.* 29, 138-150.

[27] Kandori, M., Mailath, G., Rob, R., 1993. Learning, mutation and long run equilibria in games. *Econometrica* 61, 29-56.

[28] Kojima, F., 2006. Stability and Instability of the Unbeatable Strategy in Dynamic Processes. *Int. J. Econ. Theory* 2, 41-54.

[29] Maynard Smith, J., 1974. The Theory of Games and the Evolution of Animal Conflicts. *J. Theor. Biol.* 47, 209-221.

[30] Nachbar, J. H., 1990. "Evolutionary" Selection Dynamics in Games: Convergence and Limit Properties. *Int. J. Game Theory* 19, 59-89.

[31] Salmon, T., 2001. An Evaluation of Econometric Models of Adaptive Learning. *Econometrica* 69, 1597-1628.

[32] Samuelson, L., 1997. *Evolutionary Games and Equilibrium Selection*. MIT Press, Cambridge.

[33] Sandholm, W. H., 2009. Pairwise Comparison Dynamics and Evolutionary Foundations for Nash Equilibrium, working paper, University of Wisconsin. http://www.ssc.wisc.edu/˜whs.

[34] Sandholm, W.H., Dokumaci, E., 2007. *Dynamo: Phase Diagrams for Evolutionary Dynamics*. Software suite. http://www.ssc.wisc.edu/˜whs/dynamo.

[35] Schlag, K. H., 1998. Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-armed Bandits. *J. Econ. Theory* 78, 130-156.

[36] Skyrms, B., 1990. *The Dynamics of Rational Deliberation*. Harvard University Press, Cambridge.

[37] Swinkels, J. M., 1993. Adjustment dynamics and rational play in games. *Games Econ. Behav.* 5, 455-484.

[38] Taylor, P.D., Jonker, L., 1978. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* 40, 145-156.

[39] Thomas, B., 1985. On Evolutionarily Stable Sets. *J. Math. Biol.* 22, 105-115.

[40] Vaughan, R. N. Equilibrium Selection: Basins of Attraction vs. Potential Maxima. mimeo, University College London.

[41] Walliser, B., 1998. A spectrum of equilibration processes in game theory. *J. Evol. Econ.* 8, 67-87.

[42] Weibull, J.W., 1995. *Evolutionary Game Theory.* MIT Press, Cambridge.

[43] Weibull, J.W., 1996. Contribution to Nobel Seminar: The Work of John Nash in Game Theory. *J. Econ. Theory* 69, 153-185.