Adjustment Dynamics and Equilibrium Selection: Why How Players Learn Matters

Russell Golman[a] and Scott E Page[b]

May 14, 2007

[a] To whom correspondence should be addressed. Department of Applied and Interdisciplinary Mathematics, University of Michigan, 2082 East Hall, 530 Church Street, Ann Arbor, MI 48109, USA. Telephone: 734 764-7436. Fax: 734 763-0937. E-mail: rgolman@umich.edu

[b] Center for the Study of Complex Systems, University of Michigan, 321A West Hall, 1085 S. University Ave., Ann Arbor, MI 48109, USA. E-mail: spage@umich.edu

Abstract

A deterministic learning model applied to a game with multiple equilibria produces distinct basins of attraction for those equilibria. In this paper, we investigate the possibility that the basins of attraction for best response learning and replicator dynamics have vanishing overlap, and thus that the learning models select different equilibria with probability one. In two-by-two games, we show that basins of attraction are invariant to a wide range of learning rules including best response dynamics, replicator dynamics, and fictitious play. However, we find that we can construct three-by-three symmetric games for which the overlap in the basins of attraction is arbitrarily small. We also show that for any game in which these learning models have basins of attraction with vanishing overlap, then with probability one, the initial best response is not an equilibrium.

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).

Fictitious play relies on discrete updating. In each period, the rule assigns players new beliefs based on the average play of their opponents. Players then choose actions rationally given those beliefs. More recent research by economists, psychologists, and theoretical biologists has produced a variety of other adjustment dynamics, many of which fall into two broad categories: *belief based learning* and *reinforcement based learning*. 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; Weibull, 1995).

In this paper, we focus on two learning dynamics /adjustment processes: *continu-*

ous time best response dynamics (Gilboa and Matsui, 1991) and *replicator dynamics* (Taylor and Jonker, 1978) and explore the extent to which they can differ in their basins of attraction. We show first that for any two-by-two symmetric game, these two learning rules produce identical basins of attraction. We then 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 actions must almost never be the initial best response. This *never an initial best response property* proves to be a necessary condition for the two learning rules to have vanishing overlap in their basins of attraction.

To show how these learning rules can produce such different outcomes, we must first describe the dynamics they create in the space of actions. 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. The resulting flows are piecewise linear. The best response dynamics can be thought of as the extension of fictitious play to continuous time (Hofbauer and Sigmund, 2003).

In contrast to best response dynamics, replicator dynamics are a form of reinforcement learning – action choices arise from the success of actions in previous rounds

(Erev and Roth, 1998).[1] Replicator dynamic have ecological foundations: more successful actions, i.e. more fit actions, spread in the population, while less successful ones do not. In replicator dynamics, players do not rely on beliefs about the actions of others. They need only know the payoffs of each action and its share of the population. Note that actions initially not present in the population can never be tried with replicator dynamics.

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 replica-

[1] 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).

tor 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. Similarly, 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 results – that the choice of learning rule has an enormous effect on the choice of equilibrium - at first seem to contradict the existing current literature. Thus, we want to make clear that they do not. First, many experiments consider two-by-two games. And as we show here, the two learning rules generate identical basins of attraction for two-by-two 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.[2] 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 symmetric games with linear payoffs.

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; Kandori et al.,

[2]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).

1993; Basov, 2004). Relatedly, one could link refinements with learning dynamics by choosing the equilibrium that maximizes a potential function. This selection criterion also depends on the type of learning rule because the maximum of the best response potential function may occur at a different equilibrium from the maximum of the replicator potential function (Vaughan, n.d.).

The remainder of this 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. We conclude with a discussion of the relevance of the attainability of equilibria.

2 The Learning Rules

We first define three learning rules: fictitious play, best response dynamics, and replicator dynamics. The fictitious play learning rule can be written as follows:

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

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. Continuous time best response dynamics can be written as

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

where $\mathbf{x}$ and $\mathbf{y}$ are the mixed strategy vectors for the two populations and $\mathrm{BR}(\mathbf{y})$ is the set of best replies to $\mathbf{y}$. These dynamics would arise in the large population limit with random matching.

We consider here symmetric matrix games. In these games, both players have the same set of available actions and payoffs are linear. The learning dynamics thus operate in a single, large, well-mixed population. The best response dynamics can then be written as

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

The continuous time replicator dynamics for this class of games can be written as follows:

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

where π_i is the payoff to action i and $\bar{\pi}$ is the average payoff.

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 S_3. To locate the basins of attraction of each equilibrium, we must first identify those regions of the simplex S_3 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

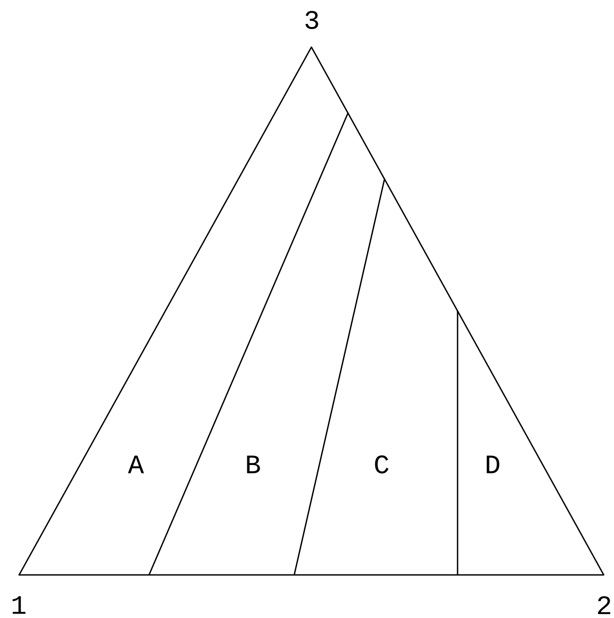

Figure 1: Best response regions. In region A, action 3 is the best response. In regions B and C, action 1 is the best response, but in B $\pi_3 > \pi_2$, while in C the opposite is true. In region D, action 2 is the best response.

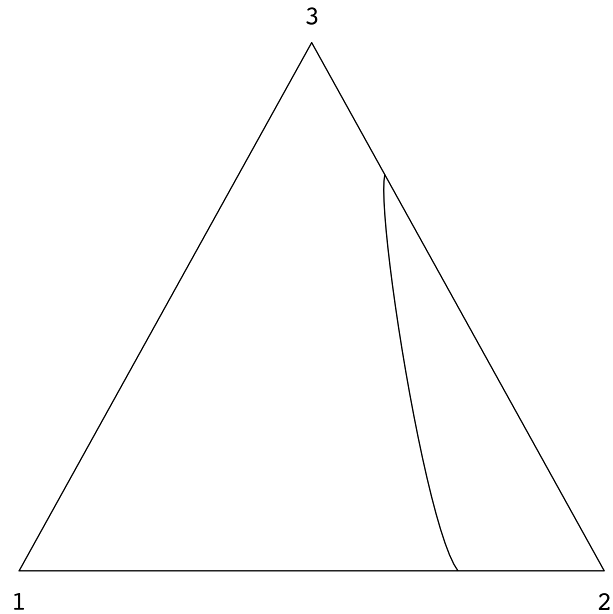

Figure 2: Basins of attraction under replicator dynamics.

boundary separating the basins of attraction here becomes a curve from the point $(\frac{1}{4}, \frac{3}{4}, 0)$ 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

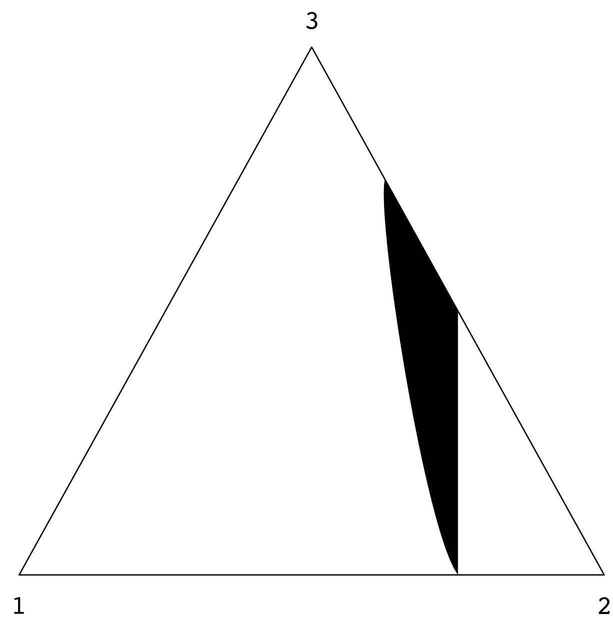

Figure 3: The small difference between best response and replicator dynamics. The shaded area flows to action 2 under replicator dynamics, to action 3 with best response dynamics.

distribution of actions is close to $(0, 1, 0)$, the dynamics flow to that equilibrium point. 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 prove 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.

Theorem 1 *For symmetric two-by-two matrix games, best response dynamics and replicator dynamics produce identical dynamics.*

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. ∎

The next 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 2 *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

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

Lemma 1 *We first show that 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.

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.

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

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

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 $-Nx_2(1 - x_1) - 2x_2 + N^3 x_1 x_2 > 0$ because $N^3 x_1 > N^2 > N(1 - x_1) + 2$.

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 1, that as N approaches infinity, best response dynamics and replicator dynamics converge to different equilibria.

Proposition 1 *In the limit as $N \to \infty$, best response dynamics and replicator dynamics flow to different equilibria except on a set of measure zero.*

This completes the proof of Theorem 2 above. Notice that in the class of games used in the proof, neither of the equilibrium actions is an initial best response almost anywhere in the action space when N is large. We call this the *Never an Initial Best Response Property.* 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 $BR^{-1}(a)$ be the set of points $\vec{x}$ for which action a is a best response.

Action a satisfies the *Never an Initial Best Response Property* at $\vec{\hat{P}}$ if

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

In our next theorem, we show that the Never an Initial Best Response Property is necessary for there to be vanishing overlap in the basins of attraction of a set of strict equilibria with best response dynamics and replicator dynamics.

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.

Definition An equilibrium $\vec{s}$ is strict in the limit as $\vec{P} \to \vec{\hat{P}}$ if for all i,

$$\lim_{\vec{P}\to\vec{\hat{P}}} f(\vec{P})\,(\pi_i(s_i, s_{-i}) - \pi_i(s', s_{-i})) > 0$$

for all s' and some $f(\vec{P}) > 0$.

Note that only pure Nash Equilibria can be strict.

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 $(a, a)\ |\mathcal{R}, G(\vec{P})$. Let $\mathbf{R}$ denote the replicator dynamics and $\mathbf{B}$ the best response dynamics. Let $\mathcal{S}$ be a set of equilibrium actions such that the equilibria are strict in the limit as $\vec{P} \to \vec{\hat{P}}$.

Theorem 3 *Suppose*

$$\lim_{\vec{P}\to\vec{\hat{P}}} \sum_{a\in\mathcal{S}} m\left(B(\mathbf{R}, a, \vec{P}) \cap B(\mathbf{B}, a, \vec{P})\right) = 0.$$

Then for all $a \in \mathcal{S}$, a satisfies the Never an Initial Best Response Property at $\vec{\hat{P}}$.

Proof Given that equilibrium (a, a) is strict in the limit as $\vec{P} \to \vec{\hat{P}}$, for $\vec{P}$ near $\vec{\hat{P}}$ the indifference curves relating payoffs to a and all other actions a' do not include the point $(x_a = 1, x_{-a} = 0)$, and $BR^{-1}(a)$ includes $(x_a = 1, x_{-a} = 0)$ and is bounded by these indifference curves. Suppose that $m\,(BR^{-1}(a))$ remains strictly positive in the limit $\vec{P} \to \vec{\hat{P}}$. It follows that the indifference curves do not hit $(x_a = 1, x_{-a} = 0)$ in this limit. Thus, for all $\vec{P}$ in a neighborhood of $\vec{\hat{P}}$, there exists $\epsilon > 0$, such that all points with $x_a > 1 - \epsilon$ are in $BR^{-1}(a)$. For such points, $\dot{x_a} > 0$ with either best response or replicator dynamics, because for both dynamics a best response spreads in the population. Thus, in the limit $\vec{P} \rightarrow \vec{\hat{P}}$, there exists a nonvanishing region with $x_a > 1-\epsilon$ that is in the basin of attraction of $(x_a = 1, x_{-a} = 0)$ for both best response and replicator dynamics. $\blacksquare$

The *Never an Initial Best Response Property* 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 with best response and replicator dynamics. We present one such set of sufficient conditions for a symmetric three-by-three game here. We leave to the reader the construction of other sufficient conditions. Observe that the conditions we present are satisfied by the class of games used in the proof of Theorem 2.

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 $\pi_{ij} \rightarrow \pi_{ij} + c$ for all i and fixed j and $\pi_{ij} \rightarrow k\pi_{ij}$ 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. So, for $j \in \{1, 2\}$, $i \neq j$, $\lim_{\vec{P} \rightarrow \vec{\hat{P}}} f_{jji}(\vec{P})(\pi_{jj} - \pi_{ij}) > 0$ for some functions $f_{jji} > 0$. And we also have $\lim_{\vec{P} \rightarrow \vec{\hat{P}}} f_{321}(\vec{P})(\pi_{23} - \pi_{13}) > 0$ for some function $f_{321} > 0$.

Theorem 4

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

if: i) $\pi_{23} > 0$*; ii)* $\pi_{13} \leq 0$ *and* $\lim_{\vec{P}\to\vec{\hat{P}}} \pi_{13} = 0$*;*[3] *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)* $\lim_{\vec{P}\to\vec{\hat{P}}} \frac{\pi_{21}}{\pi_{23}} = -\infty$.

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

Lemma 4 *As* $\vec{P}$ *approaches* $\vec{\hat{P}}$*, 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 $\vec{\hat{P}}$, that action 3 is initially a best response in all but an arbitrarily small part of the action space when $\vec{P}$ nears $\vec{\hat{P}}$. 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 $\vec{\hat{P}}$, $\pi_2 < 0$.

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 $\vec{\hat{P}}$, $\pi_1 < 0$.

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.

[3] 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,

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

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

$$\pi_{11}x_1(x_2+x_3) + \pi_{13}x_3(x_2+x_3) > 0. \quad (1)$$

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 $\vec{\hat{P}}$,

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

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 (2) holds as x_1 increases. As x_1 increases, it remains above $-\pi_{13}$, so Equation (1) 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 $\vec{x}$ need only satisfy $x_1 > 0$ and $x_2 > 0$ to be in $B(\mathbf{R}, 1, \vec{P})$ for some $\vec{P}$ near $\hat{\vec{P}}$. $\blacksquare$

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 games modeling collective action problems (Golman and Page, 2007):

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

In this game, the first action is a safe, self interested action. The second action represents an attempt to solve a collective action problem. The third action is predatory toward the fourth action, which can be thought of as a failed attempt to solve the collective action problem. In this game, as 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 2 *In the four-by-four game above, as $N \to \infty$, best response dynamics and replicator dynamics flow to different equilibria except on a set of measure 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

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

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 $\beta(t)$ is a strictly decreasing function. Now

$$\pi_1 = \pi_3 \text{ when } \beta = \frac{2}{N^2\left(x_4(0)\right)}.$$

But

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

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 2. 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 2, 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 (Golman, 2007).

4 Conclusion

In a class of three-by-three symmetric games, we have shown that how players learn can influence equilibrium selection. We have provided 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. We have also shown that if in any game the learning rules attain distinct strict equilibria from almost any starting point, then almost everywhere the initial best response cannot be an equilibrium action. Games in which best response dynamics and replicator dynamics make such different equilibrium predictions would seem to lend themselves to experiments to determine whether people actually learn according to one of these rules.

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.

References

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

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

[3] 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.

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

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

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

[7] 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.

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

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

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

[11] Golman, R., 2007. PhD diss. University of Michigan, under preparation.

[12] Golman, R., Page, S., 2007. Individual and Cultural Learning in Collective Action Games. Unpublished.

[13] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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