Individual and Cultural Learning in Stag Hunt Games With Multiple Actions

Russell Golman and Scott E Page

May 22, 2009

Abstract

We compare *cultural learning* and *individualistic, belief-based learning* in a class of generalized stag hunt games. Agents can choose from among multiple potentially cooperative actions or can take a secure, self interested action. We assume that a proportion of the cooperative actions prove effective, while others can be undermined by a predatory action. For this class of games, the set of stable equilibria is identical under the two learning rules. However, we show that the basins of attraction for the efficient equilibria are much larger for cultural learning. We further show that as the stakes grow arbitrarily large, cultural learning always locates an efficient equilibrium while belief-based learning never does.

KEYWORDS: Belief-based learning; cooperation; coordination game; cultural learning; stag hunt.

JEL classification codes: C73; D79; H41

1 Introduction

Game theorists motivate the prediction of equilibrium outcomes by assuming either that rational agents select equilibrium strategies or that agents learn and settle into an equilibrium. When a game possesses multiple equilibria, the assumption of rational agents requires the introduction of refinement criteria to select from among the equilibria. With learning agents, the explicit model of behavior determines a basin of attraction for each equilibrium. The size of a basin can vary depending on the learning rule.[1] Thus, learning can be seen as a selector of equilibria.[2]

A natural practical question to ask is whether differences in learning rules can lead to qualitatively different outcomes. In particular, does one type of learning rule make cooperative behavior more likely to emerge in equilibrium than another? To get at that question, in this paper, we explore the extent to which the type of learning rule influences equilibrium selection in a class of generalized stag hunt games. We compare two canonical learning models: *cultural learning* and *individual belief-based learning.* We find that the former more often settles into a cooperative equilibrium. This does

---

[1] The stability of an equilibrium is usually invariant for broad classes of learning rules. Even locally asymptotically stable equilibria, however, may have small basins (Epstein, 2003). In extreme cases a basin of attraction may be arbitrarily small for one learning rule and arbitrarily large for another (Golman and Page, 2008).

[2] Basins of attraction have been the focus of a related literature that considers learning models with persistent randomness and selects stochastically stable equilibria (Foster and Young, 1990; Kandori et al., 1993; Young, 1993; Kandori and Rob, 1995). These dynamics favor risk dominant solutions.

not imply that cultural learning is better for all games (see Golman and Page 2008), but does suggest that for games that involve coordination on a cooperative action, cultural learning may be a preferred learning rule.

The original stag hunt game traces back to 1773, when Rousseau proposed the story of a stag hunt to represent a choice in which the benefits of cooperation conflict with the security of acting alone. In the story, two individuals must each choose to hunt a stag or to hunt a hare. Hunting stags can only be successful with cooperation, while hunting a hare does not require the other player's help. The catch is that the stag offers both hunters a lot more meat than the hare. Thus, the stag hunt obliges a choice between productivity and security. Skyrms (2001) argues that the stag hunt captures the incentives present in choices whether to adopt or modify the social contract.

Rousseau's stag hunt has been modeled as a two-by-two game with two strict pure Nash Equilibria: an efficient one in which both hunt stag and an inefficient one in which both hunt hare. In playing a stag hunt, agents try to figure out which equilibrium action to choose. Equilibrium selection arguments can be invoked in favor of either action. While hunting the stag is payoff dominant, hunting the hare is risk dominant (Harsanyi and Selten, 1988).

The stag hunt game, along with the prisoner's dilemma, is often invoked as a framework with which to study collective action problems (Medina, 2007). In a prisoner's dilemma or a public goods game, incentives lead to everybody defecting. In contrast, in a stag hung game, players have an incentive to cooperate provided enough of the other players do so as well. Stag hunt differs from the prisoners' dilemma in that achieving cooperation does not require higher order strategies such as tit for tat (Axelrod, 1984), trigger mechanisms (Abreu et al., 1990), or norm based strategies (Bendor and Swistak, 1977), all of which allow defectors to be punished outside of the context of the original game. Higher order strategies that produce cooperation in the prisoner's dilemma, in effect, transform that game into a stag hunt game, where the choice to defect corresponds to the inefficient, but safer action (Skyrms, 2001). Therefore, many of the results of this paper can be interpreted through the lens of the repeated prisoner's dilemma.

The choice between stag and hare simplifies a more complex reality in which agents might choose between multiple stags and a hare. By that we mean a society or a community would often have more than a single cooperative action to pursue. So, here, we extend the canonical two-by-two stag hunt game to allow for more actions. This increase in actions also provides sufficient space for learning rules to matter, whereas in a two-by-two game, how agents learn has no effect. As in the canonical stag hunt game, the models we consider include an insulated *self interested* action that does not require agents to coordinate. It generates only a modest payoff, but it is safe. We differ from the canonical model in that we allow for multiple potentially cooperative actions.

In our basic model, any cooperative action would be optimal if the population coordinated on it. We then expand our model to consider the possibility that some of these actions would be *effective*, e.g., successful if followed, but that others, termed *naive*, would be undermined by a *predatory* action.[3] Thus, we add predation (Conlisk, 2001) to the stag hunt. The predatory action can be seen as a form of defection. It robs those agents who undertake naive actions. This framework captures situations in which a community may have several potential "stag" actions of which some would prove fruitful and others would not. To illustrate, a community may have several mechanisms to share water, but some of these may prove exploitable by predatory actions.

Within this class of models, we find that cultural learning more often locates the efficient cooperative equilibria than does individual belief-based learning. To be precise, we show that cultural learning dominates individual learning: given any starting point for which individual learning results in cooperation, so does cultural learning.

Within the expanded framework that includes naive and predatory strategies, we highlight three additional results. First, contrary to intuition, we find that the effect of the learning rule becomes amplified as the stakes increase. As the stakes grow infinitely large, cultural learning converges to full cooperation, always locating an effective action, while belief-based learning converges to a zero probability of finding an effective action. Thus, ramping up incentives makes the type of learning rule more, not less important. Second, we find that as the number of potentially cooperative actions increases, so does the probability of finding the self interested action. This

---

[3] Though the naive actions are not played in equilibrium, such dominated strategies play a crucial role in equilibrium selection in coordination games (Ochs, 1995; Basov, 2004).

coordination failure aligns with basic intuition that lots of options makes coordinating on a single one more difficult. Third, and counter to intuition, given cultural learning we find that changing one action from naive to effective can move some initial points from the basin of attraction of a cooperative equilibrium into the basin of attraction of the safe equilibrium.

2 Individual and Cultural Learning

Given the variety of potential learning rules, we feel it necessary to motivate our decision to compare *individual, belief-based learning* to *cultural evolutionary learning* (Camerer 2003, Fudenberg and Levine 1999). These two learning rules differ in how they characterize behavior. Belief-based learning is prospective and individualistic; cultural evolutionary learning is retrospective and social. By considering these two extreme forms of learning we investigate the possibility that the type of learning rule might matter for equilibrium selection.

We model individual learning using a simple best response learning rule (Gilboa and Matsui, 1991; Hofbauer and Sigmund, 2003). Elaborated models of individual learning, such as logit learning and quantal response learning, include noise terms and individual errors. The extra degree of freedom introduced with this error term implies that they can fit experimental data better than the simpler best response dynamic. Nevertheless, we stick here with best response learning owing to its foundational nature and analytic tractability.

Following convention, we use replicator dynamics (Taylor and Jonker, 1978) to capture cultural learning in a population of players (Henrich and Boyd, 2002). Replicator dynamics can be seen as capturing situations in which agents compare payoffs with each other and copy better performing agents.[4] Cultural learning is less *greedy* than best response dynamics. Agents don't move only in the direction of the best action but towards all actions that have above average payoffs.

For the sake of analytic tractability, we consider continuous time dynamics.[5] These would arise in the limit of a large, well-mixed population. In this limit, both dynamics can be derived as the expected behavior of agents with stochastic protocols for switching their actions (Sandholm, 2009). A simple revision protocol in which agents occasionally switch to the best response to the current population state generates the best response dynamics. Imitative revision protocols, such as *imitation driven by dissatisfaction* (Björnerstedt and Weibull, 1996) or *pairwise proportional imitation* (Schlag, 1998), lead to the replicator dynamics.

The differences between belief-based learning rules and cultural evolutionary learning have been the subject of substantial theoretical, experimental, and empirical investigation. For the most part, the theoretical literature focuses on how the rules operate and, in particular, on the stability of equilibria under the two types of rules. That literature shows that in many games both rules produce the same set of stable

---

[4] Cultural learning can be performance-based if it depends on payoffs or conformist if it depends on popularity. Here, we consider performance-based cultural learning.

[5] Showing that our main results hold with discrete dynamics as well is a straightforward, though somewhat involved, exercise.

equilibria (Hopkins, 1999; Hofbauer et al., 1979; Hofbauer, 2000). Counterexamples rely on knife edge assumptions.

Though the stable equilibria may be the same across learning rules, the paths to those equilibria differ. Experimental and empirical literatures attempt to flesh out which rule people apply in practice. As the two rules differ in their informational and cognitive requirements, we should expect each rule to be better suited to some environments than the other. Cultural learning requires knowledge of the success of others. Given that information, a cultural learning rule doesn't require much cognitive effort: agents need only copy someone doing better than they are. Best response learning, on the other hand, does not require any information about the success of others – other than the payoffs to the game – but it does require knowledge of the full distribution of actions and calculation of the payoff from each possible action. Thus, we shouldn't expect to see best responses unless people understand the game.

In two-by-two games, we might therefore expect best response learning to better predict behavior. In fact, Cheung and Friedman (1998) find greater support for belief-based learning than for replicator dynamics. However, the performance of these and other learning rules is often so similar as to be almost indistinguishable (Feltovich, 2000; Salmon, 2001). What differences that do exist between the behavior predicted by these rules and the data can often be explained by considering a hybrid model that includes both belief-based and reinforcement learning (Camerer and Ho, 1999).[6]

---

[6] Reinforcement learning can also give rise to the replicator dynamics (Börgers and Sarin, 1997).

Learning rules have also been studied in the field. Evidence from the real world generally tilts towards cultural learning. Henrich (2001), in surveying evidence on the adoption of innovations, finds S-shaped adoption curves to be prevalent. Cultural learning, which relies on imitation, produces S-shaped curves. Individual learning does not. A hybrid learning model would also produce S-shaped adoption curves. Therefore, the empirical evidence should not lead us to declare cultural learning the winner so much as it tells us that people do take into account how others act.[7]

Our interests here tend less toward the empirical question of what people do and more in the direction of the theoretical question of what would happen if people were to follow one rule at the expense of the other. We find that cultural learning locates an efficient equilibrium more often than does best response learning. This could imply that societies that have a more collective orientation might be better equipped to coordinate their efforts and cooperate in the face of strategic uncertainty. That said, cultural learning is not the only mechanism through which a society can achieve efficient, coordinated cooperation. Many other mechanisms have been shown sufficient, including focal points (Schelling, 1960) and, in the context of an evolutionary dynamic, cheap talk (Skyrms, 2004). Alternatively, if groups or bands of people engage in battles with neighboring groups, then group selection could produce altruistic strategies that ensure cooperation (Gintis et al., 2003).

In what follows, we introduce our model of a stag hunt game with multiple co-

---

[7] *Prestige bias*, the coupling of prestige to success, is further evidence that imitation is a part of how people learn (Henrich and Gil-White, 2001).

operative actions, provide an example, and prove some general results. Then, in Section 4, we modify our model to allow some of the cooperative actions to fail and be undermined by a predatory action. In Section 5, we present the bulk of our results, analyzing this model. Section 7 concludes with a discussion of the types of learning we consider.

3 The Basic Model

In our basic model, we assume a *self interested* action that offers agents a risk-free return. We also assume n *cooperative* actions. Each offers the agent a reward that is assuredly positive and increases with the number of others taking the same action. In the canonical example, hunting rabbits would be the self interested action. Hunting stags would be a cooperative action, as would participating in a collective whale hunt. In these cases, hunting cooperatively greatly increases the chances of a hearty meal, while hunting alone for a stag or a whale gives the player a small (but nonzero) chance at finding something to eat. (Perhaps while being faithful to the larger cause, the hunter can still bag a small animal or fish on the side.) Taking a cooperative action has *positive spillovers* in that it helps others who join in. Thus, taking a cooperative action becomes the better choice (only) when a significant percentage of other agents also take this action.

Formally, our game consists of $n+1$ actions, with the action set $\mathcal{A} = \{1, \cdots, n, S\}$. We denote the set of cooperative actions $\mathcal{C} = \{1, \cdots, n\}$. We consider a single unit-mass population of agents. The state space is therefore the n-dimensional unit simplex $\triangle^n$, with a point $\mathbf{x} \in \triangle^n$ denoting the fraction of the population choosing each action. The vector $\mathbf{x} = (x_1, \ldots, x_n, x_S)$ is called the population mixed strategy.

Payoffs in a population game are a function of the population state $\mathbf{x}$. We normalize payoff magnitudes by attributing a cooperative action taken in isolation a payoff of 1, assuming it's the same for any cooperative action, and we let the parameter $\beta > 0$ capture the relative benefit of coordinating on a cooperative action. Larger β imply greater benefits from achieving cooperation. The cooperative actions vary in their efficiency according to a family of parameters, θ_i for each $i \in \mathcal{C}$. The maximum reward, if everybody coordinates on effective action i, is an additional $\theta_i\beta$. Finally, the insulated, self interested action S receives a payoff of $c > 1$ regardless of the actions of the other agents. The payoffs can therefore be written as follows:

$$\begin{aligned} \pi(i, \mathbf{x}) &= 1 + \theta_i \beta x_i \text{ for } i \in \mathcal{C} \\ \pi(S, \mathbf{x}) &= c > 1. \end{aligned}$$

We make a technical assumption, (A1) $\theta_i\beta + 1 > c$ for all $i \in \mathcal{C}$, to create the proper ordering over payoffs. This guarantees that successfully coordinating on a cooperative action yields a higher payoff than taking the self interested action. Thus, we have a strict pure Nash Equilibrium at every action in the game.

Learning rules operate on a 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. The learning dynamic produces changes in the proportions of agents playing the various actions. In what follows we compare two learning rules: *best response dynamics* and *replicator dynamics.* In best response dynamics, some infinitesimal proportion of the agents are switching their action to match the current best response. In replicator dynamics, agents are comparing payoffs and learning from the success of others.

Best Response Dynamics *(Individual Learning)* $\dot{\mathbf{x}} \in \mathrm{BR}(\mathbf{x}) - \mathbf{x}$ where $\mathrm{BR}(\mathbf{x})$ is the set of best replies to $\mathbf{x}$.

Replicator Dynamics *(Cultural Learning)* $\dot{x}_i = x_i(\pi_i - \bar{\pi})$ where π_i is the payoff to action i and $\bar{\pi}$ is the average payoff.

3.1 An Example

A simple example with $n = 2$ demonstrates how these learning rules perform. In this example, we take $c = 3$ and $\theta_1\beta = \theta_2\beta = 5$. This game has three pure strategy equilibria, one for each action. Each equilibrium is strict and is therefore asymptotically stable for both learning rules.

We represent a distribution of actions as a point in the two dimensional simplex $\triangle^2$. To locate the basins of attraction under best response dynamics, we identify the regions of the simplex $\triangle^2$ in which each action is a best response. These regions are defined by the lines where each pair of actions earns the same payoff. We find $\pi_1 = \pi_2$ when $x_1 = x_2$, $\pi_1 = \pi_S$ when $x_1 = \frac{2}{5}$, and $\pi_2 = \pi_S$ when $x_2 = \frac{2}{5}$. Because the payoff to a cooperative action increases as the action spreads, the equilibrium chosen

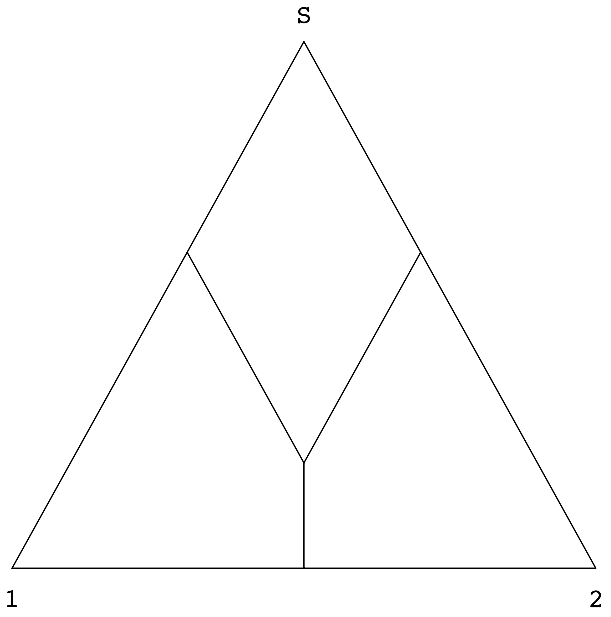

Figure 1: Basins of attraction under best response dynamics

under best response dynamics consists of the action that is initially a best response. This is a feature of our basic model, but it will not be true of our modified model. It means that the best response regions are the basins of attraction of the pure equilibria under best response dynamics. They are shown in Figure 1. The corresponding flow diagram for the best response dynamics is shown in Figure 2.

Figure 3 contains the flow diagram for the replicator dynamics, and then, in Figure 4, we characterize the basins of attraction for replicator dynamics. Here, the boundary separating the basins of attraction includes curves, not only lines. This curvature arises because under replicator dynamics a cooperative action can grow in the population even if the self interested action is the best response. As it grows, the cooperative action becomes the best response. As a result, the population can slip

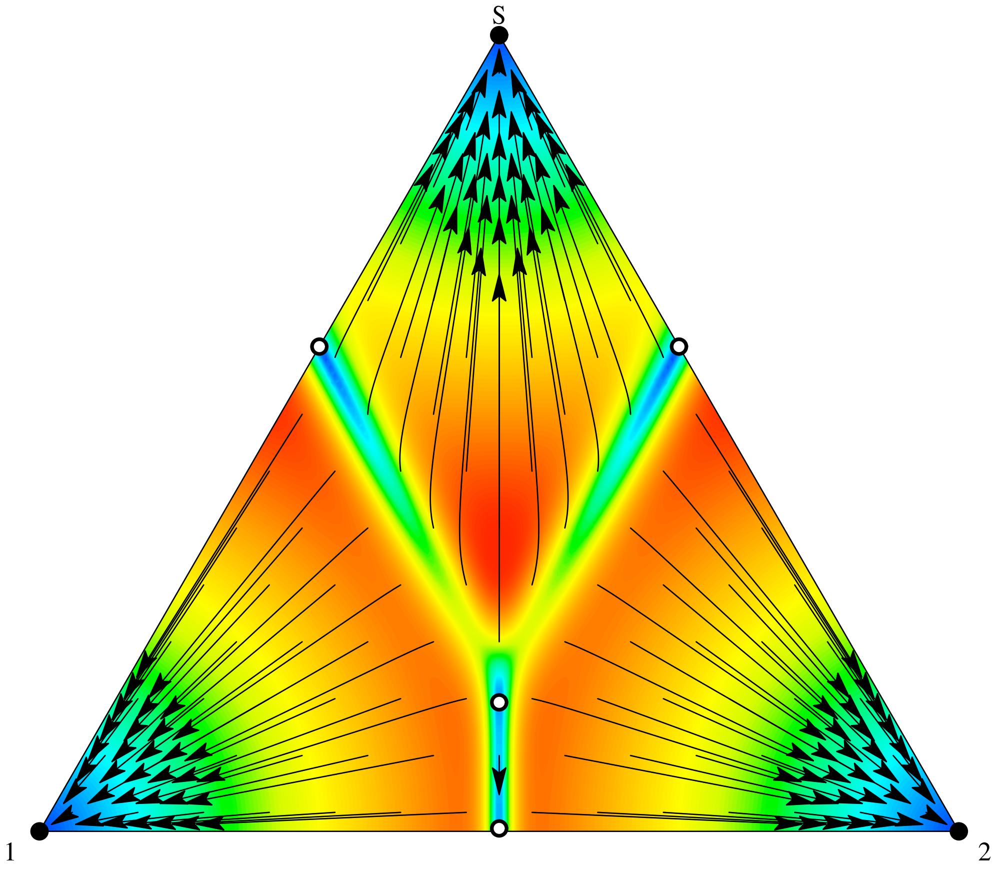

Figure 2: Phase diagram for the best response dynamics. Black (white) circles are stable (unstable) rest points. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

from the self interested action's best response region into a cooperative action's best response region. Thus, cooperation is more likely to arise under replicator dynamics.

3.2 General Results

We now show that the intuition developed in the example holds more generally. That is, the replicator dynamics is more likely to lead to cooperation. In what follows, we assume an initial distribution of actions with full support, bounded density, and no mass points on the boundary of the strategy space, and then apply the two learning rules. We first show for both learning rules that if a cooperative action ever has the best payoff, it remains best forever and thus persists in the resulting equilibrium.

Lemma 1 *For both best response dynamics and replicator dynamics, if for some* $i \in \mathcal{C}$, $\pi_i > \pi_j$ *for all* $j \neq i \in \mathcal{A}$ *at some time, it remains so at all later times.*

Proof For best response dynamics, the result is straightforward. Only a best response grows in the population. If a cooperative action is a best response, it becomes more widespread, and consequently its payoff increases. Meanwhile, other actions become less common, so the payoffs to the other cooperative actions decrease.

For replicator dynamics, suppose $\pi_i \geq \pi_j$ for all $j \in \mathcal{A}$. For $j \in \mathcal{C}$,

$$\dot{\pi_j} = \theta_j \beta \dot{x_j} = \theta_j \beta x_j (\pi_j - \bar{\pi}) = (\pi_j - 1)(\pi_j - \bar{\pi}). \tag{1}$$

So $\dot{\pi_i} \geq \dot{\pi_j}$ for all $j \in \mathcal{C}$, and $\dot{\pi_i} > 0$. The payoff to the cooperative action which is the best response increases faster than the payoff to other cooperative actions. The

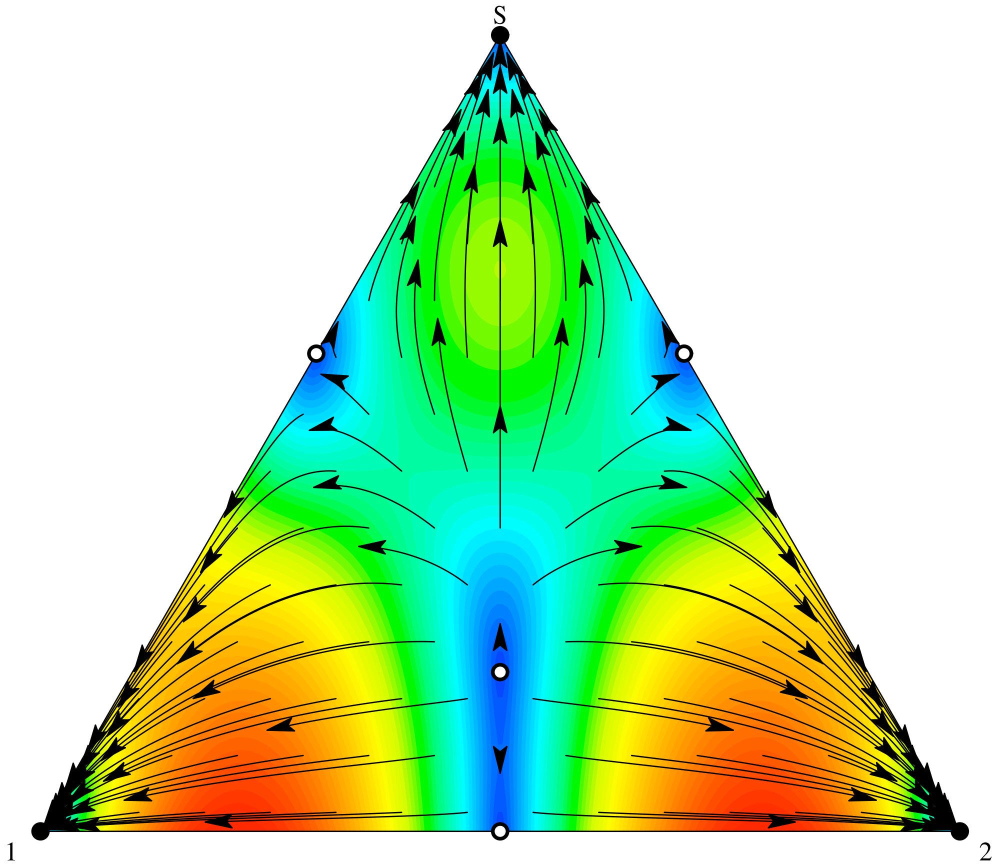

Figure 3: Phase diagram for the replicator dynamics. Black (white) circles are stable (unstable) rest points. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

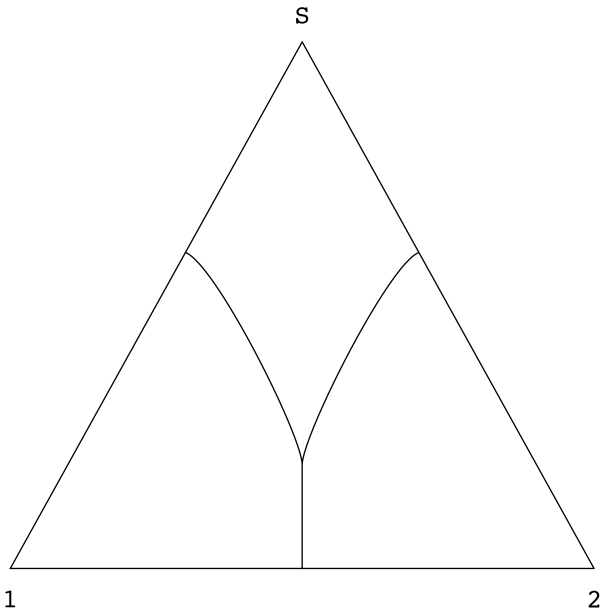

Figure 4: Basins of attraction under replicator dynamics

self interested action has a constant payoff, so it cannot become the best response. Cooperative action i must remain a best response forever. ∎

Our first theorem identifies the basins of attraction under best response dynamics. In an abuse of notation, we refer to the equilibrium with everyone choosing cooperative action $i \in \mathcal{C}$, $(x_i = 1, x_{-i} = 0)$ as *cooperative equilibrium* $i \in \mathcal{C}$. Similarly, we use the term *self interested equilibrium* to stand for the equilibrium with everyone choosing self interested action, $(x_S = 1, x_{-S} = 0)$. For clarity in our presentation, we define some new parameters that help us compare the payoffs of a cooperative action and the self interested action. Let $T_i = \frac{c-1}{\theta_i \beta}$ be the threshold frequency for cooperative action $i \in \mathcal{C}$ to be better than the self interested action. That is, $\pi_i \geq \pi_S$ if and only if $x_i \geq T_i$, with equality in one following from equality in the other.

Theorem 1 *Given best response dynamics, a point* $\mathbf{x}$ *is in the basin of attraction of the self interested equilibrium if and only if for all* $j \in \mathcal{C}$, $x_j < T_j$.

A point $\mathbf{x}$ *is in the basin of attraction of cooperative equilibrium* $i \in \mathcal{C}$ *if and only if* $x_i > T_i$ *and* $x_i > \frac{\theta_j}{\theta_i} x_j$ *for all* $j \neq i \in \mathcal{C}$.

Proof The inequalities given in the theorem define the best response regions. Lemma 1 tells us that the best response regions of cooperative actions are contained in the basins of attraction of cooperative equilibria. Similarly under best response dynamics, if the self interested action is a best response, cooperative actions decline in the population and cooperative payoffs decrease. The self interested action thus remains a best response, and the dynamic leads to the self interested equilibrium. ∎

The next theorem states that the basins of attraction of the cooperative equilibria under best response dynamics are proper subsets of these basins under replicator dynamics.

Theorem 2 *Assume* $n \geq 2$. *The basin of attraction of any cooperative equilibrium* $i \in \mathcal{C}$ *under best response dynamics is a proper subset of the basin of attraction of this equilibrium under replicator dynamics.*

Proof Lemma 1 and Theorem 1 together imply that the basin of attraction of a cooperative equilibrium under best response dynamics is contained in the basin under replicator dynamics because the former consists only of points for which the cooperative action is the best response. It remains to show that the basin of attraction of any cooperative equilibrium $i \in \mathcal{C}$ under replicator dynamics includes some points for which the self interested action is the best response. This is done in the appendix.

▮

Theorem 2 tells us that when there are multiple cooperative actions, replicator dynamics is more likely to select one than best response dynamics. (When there is just one cooperative action, the learning rules behave alike, as they do in all two-by-two games (Golman and Page, 2008).) Seeing that cultural learning outperforms belief-based learning in this environment, we next consider the question of how much better it can be.

4 A Model with a Predatory Strategy

We now modify our model by introducing a predatory strategy and making some of the cooperative strategies susceptible to it. We maintain the n potentially cooperative actions, but now we assume that only k of these cooperative actions are *effective*. These actions have positive spillovers and are immune from predation. The others we call *naive* because they are susceptible to attack from a *predatory* action. In the context of the literal stag hunt, consider a scenario in which players have the option of domesticating animals and sharing the burden of looking after them. This is a potentially cooperative action, but in the case that nobody can be trusted to exercise costly vigilance while watching over somebody else's animals, it is naive. The animals may be stolen.

The predatory action can be thought of as a strong form of defecting. An agent who takes the predatory action benefits at the expense of those attempting naive actions. Neither the predatory action nor the self interested action aids the cooperative efforts of other players, but unlike the predatory action, the self interested one does not undermine the success of anybody who is trying to cooperate. An example clarifies this distinction. Hunting rabbits instead of stags indirectly lowers the payoffs to the stag hunters by denying them another participant. But, a player who steals the aforementioned domesticated animals is taking a *predatory* action. This directly harms those taking taking naive actions and creates no additional surplus. The predatory action will often be an initial best response, but it is not an equilibrium. The nature of these incentives proves important, as we shall show.

This modified game consists of $n+2$ actions: a *predatory* action, a *self interested* action, k *effective* actions, and $n-k$ *naive* actions. We assume $n > k \geq 1$ ensuring that both the set of effective actions and the set of naive actions are nonempty.

The **action set** $\mathcal{A} = \{1, \cdots, n, P, S\}$ *with partition* $f : \{1, \cdots, n\} \to \{E, N\}$ *where* $\mathcal{E} = \{i | f(i) = E\}$ *denotes the* **effective** *actions and* $\mathcal{N} = \{i | f(i) = N\}$ *denotes the* **naive** *actions.*

Note that $k = |\mathcal{E}|$. Effective actions $i \in \mathcal{E}$ have the same payoff as in the basic model when all cooperative actions were effective. Naive actions $i \in \mathcal{N}$ differ in their payoff structure. They each get a negative payoff per agent playing the predatory action. The parameter γ measures the value that could be lost to or gained by predation. If the population were to tend towards 100% predatory action, the payoff deducted from each naive action would be γ. The payoffs lost by naive actions are gained by agents taking the predatory action. This predatory action P gets nothing from everybody else, so its payoff scales with the prevalence of naive actions, approaching γ as the population tends toward 100% naive actions. The payoffs from the various actions can be written as follows:

$$\begin{aligned}
\pi(i, \mathbf{x}) &= 1 + \theta_i \beta x_i \ \text{ for } i \in \mathcal{E} \\
\pi(i, \mathbf{x}) &= 1 + \theta_i \beta x_i - \gamma x_P \ \text{ for } i \in \mathcal{N} \\
\pi(P, \mathbf{x}) &= \gamma \sum_{i \in \mathcal{N}} x_i \\
\pi(S, \mathbf{x}) &= c > 1.
\end{aligned}$$

Assumption A1 from the basic model still holds for all $i \in \mathcal{E}$, but not necessarily for $i \in \mathcal{N}$. We now have (A1′) $\theta_i \beta + 1 > c$ for all $i \in \mathcal{E}$. We make two additional technical assumptions, (A2) $\gamma > c$ and (A3) $\gamma > \max_{i \in \mathcal{N}} \{\theta_i\} \beta \frac{c}{c-1}$. A2 guarantees that in a population full of naive actions, the predatory action has a higher payoff then the self interested action. A3 guarantees that a naive action is never best, i.e., a predator can steal even more than the positive spillovers generated by a naive action.

5 Results

We now turn to our main results: how the equilibrium attained depends strongly on the learning rule. We begin with an obvious theorem about the set of pure strategy equilibria to this class of games.

Theorem 3 *Given assumptions* $A1'$ *and* $A3$ *and* $c > 1$*, this game has* $k + 1$ *pure strategy equilibria: one in which all players take the self interested action and* k *equilibria in which they all take the same effective action.*

We maintain our abuse of notation in referring to one of these k *effective equilibria* as effective equilibrium $i \in \mathcal{E}$. We now show that Lemma 1 still applies to the effective cooperative actions under best response dynamics, and can be broadened to include the self interested action as well. However, we will need to slightly modify the lemma for the replicator dynamics. For the best response dynamics, we can say that if an effective action or the self interested action ever has the best payoff, it remains best forever. This is not the case for the predatory action. (By assumption A3, a naive action is never best.)

Lemma 2 *For best response dynamics, if for some* $i \in \mathcal{E} \cup \{S\}$, $\pi_i > \pi_j$ *for all* $j \neq i \in \mathcal{A}$ *at some time, it remains so at all later times.*

Proof Note that naive actions are dominated by a mixed strategy that plays the self interested action with probability $\frac{1}{c}$ and the predatory action with probability $\frac{c-1}{c}$. So, naive actions never have the highest payoff and are always decreasing. This means the payoff to the predatory action is always decreasing. So, exempting the naive actions, an effective action has the only increasing payoff when it is the strict best response, and the self interested action has the only nondecreasing payoff when it is the strict best response. ∎

Our next result identifies a sufficient condition for replicator dynamics to yield effective cooperation. We do not have a similar sufficient condition for fixation of the self interested action. This lemma is similar to Lemma 1 in that it applies to an effective action that at some time is better than the self interested action and has the most positive spillovers, but we do not require it to be better than the predatory action. Of particular importance is the fact that this condition is independent of the predation parameter γ.

Lemma 3 *Under the replicator dynamics, if for some* $i \in \mathcal{E}$*,*

1. $\pi_i > \pi_S$ *and*
2. $x_i > \frac{\theta_j}{\theta_i} x_j$ *for all* $j \neq i \in \mathcal{E} \cup \mathcal{N}$

at some time, then conditions 1 and 2 remain in effect at all later times and the replicator dynamics leads to the equilibrium $(x_i = 1, x_{-i} = 0)$*.*

Proof A piece of the average payoff, $x_P \pi_P + \sum_{j \in \mathcal{N}} x_j \pi_j$, partially cancels, leaving only $\sum_{j \in \mathcal{N}} (1 + \theta_j \beta x_j) x_j$. Thus, the average payoff $\bar{\pi} = x_S \pi_S + \sum_{j \in \mathcal{E} \cup \mathcal{N}} (1 + \theta_j \beta x_j) x_j$. Now, conditions 1 and 2 together imply that $\pi_i > \pi_j$ for all $j \in \mathcal{A} \setminus \{i, P\}$, and in turn,

$$\pi_i - \bar{\pi} > 0. \tag{2}$$

Condition 2 alone implies that

$$\pi_i - \bar{\pi} > \pi_j - \bar{\pi} \text{ for all } j \neq i \in \mathcal{E} \cup \mathcal{N}. \tag{3}$$

Inequality (3) gives action i the highest relative growth rate in $\mathcal{E} \cup \mathcal{N}$, ensuring that condition 2 continues to hold. Inequality (2) means that action i does indeed have positive growth, maintaining condition 1 and leading to the equilibrium $(x_i = 1, x_{-i} = 0)$. ▮

We now derive the basins of attraction under best response dynamics. We will make use of a new parameter $r(\mathbf{x})$ that denotes the ratio of the predatory payoff to the self interested payoff. (It depends on the prevalence of naive actions.) We have $r(\mathbf{x}) = \frac{1}{c}\gamma \sum_{j \in \mathcal{N}} x_j$. Also, recall that $T_i = \frac{c-1}{\theta_i \beta}$, now for $i \in \mathcal{E}$, is the threshold frequency for effective action i to be better than the self interested action.

Theorem 4 *Given best response dynamics, a point* $\mathbf{x}$ *is in the basin of attraction of the self interested equilibrium if and only if for all* $j \in \mathcal{E}$,

$$x_j < \max\left[T_j,\, r(\mathbf{x})T_j\right]. \tag{4}$$

A point $\mathbf{x}$ *is in the basin of attraction of effective equilibrium* $i \in \mathcal{E}$ *if and only if*

$$x_i > \max\left[T_i,\, r(\mathbf{x})T_i,\, \frac{\theta_j}{\theta_i} x_j\right] \text{ for all } j \neq i \in \mathcal{E}.^{8} \tag{5}$$

Proof See appendix.

The condition $x_i > \frac{\theta_j}{\theta_i} x_j$ for all $j \neq i \in \mathcal{E}$ that is necessary for best response dynamics to obtain effective equilibrium $i \in \mathcal{E}$ just says that action i is the best of all

---

[8] In the zero probability event that an equality holds exactly and neither of the conditions applies, the best response dynamics finds a mixed equilibrium.

the effective actions. It determines which effective equilibrium is attained, but could be dropped from a claim asserting just that the equilibrium attained is of the form $(x_S = 0, \sum_{i \in \mathcal{E}} x_i = 1)$.

Theorem 4 establishes a strict condition for best response dynamics to achieve effective cooperation. The condition that an effective action does better than the self interested action and all other effective actions proves to be necessary but not on its own sufficient for best response dynamics to attain it as an equilibrium. On the other hand, the condition that the self interested action initially does better than any effective action, is sufficient (and not even necessary) for the best response dynamics to lead to universal self interested action.

The next claim states that cultural learning more often achieves effective cooperation: *whenever best response dynamics attains an effective equilibrium, replicator dynamics attains it as well.* The converse will not be true. Thus, cultural learning is strictly preferred to individual learning.

Theorem 5 *The basin of attraction of any effective equilibrium $i \in \mathcal{E}$ under best response dynamics is contained in the basin of attraction of this equilibrium under replicator dynamics.*

Proof Using Lemma 3 and Theorem 4, we show that the necessary condition for the best response dynamics to be in the basin of $(x_i = 1, x_{-i} = 0)$ for some $i \in \mathcal{E}$ is sufficient conditions for the replicator dynamics to be in the basin of this equilibrium. The condition $x_i > T_i$ in Theorem 4 is equivalent to condition 1 in Lemma 3. The requirement $x_i > r(\mathbf{x})T_i$ in the theorem along with assumption A3 implies that $x_i > \max_{j\in\mathcal{N}} \frac{\{\theta_j\}}{\theta_i} \sum_{l\in\mathcal{N}} x_l$, by plugging the A3 inequality in for the γ hiding in $r(\mathbf{x})$. This easily gives us $x_i > \frac{\theta_j}{\theta_i} x_j$ for all $j \in \mathcal{N}$. Theorem 4 explicitly fills in for all $j \neq i \in \mathcal{E}$ and thus satisfies condition 2 in Lemma 3. ▮

These results characterize fully the basins of attraction under best response dynamics, but do not do so for replicator dynamics. Because the replicator dynamics are nonlinear, we cannot solve for the basins exactly. In Section 6, we make a simplifying assumption that allows us to proceed without an exact description of the replicator dynamics' basins. We would like to make analysis of the replicator dynamics more tractable without losing generality from our results for the best response dynamics. To see how this is possible, let us refer back to Theorem 4, which characterizes the best response dynamics' basins of attraction. Observe that Theorem 4 makes no reference to the naive actions' efficiency parameters. As long as the $\{\theta_j\}_{j\in\mathcal{N}}$ obey assumption A3, they have no effect on the best response dynamics. These parameters do influence the replicator dynamics, but it appears that the behavior of the replicator dynamics is qualitatively similar for a range of possible values (see appendix).

6 Comparative Statics

We now consider the case that naive actions fail to produce positive spillovers for others taking the same action. In what follows, we assume $\theta_j = 0$ for all $j \in \mathcal{N}$. This gives each of the naive actions the same payoff,

$$\pi(i, \mathbf{x}) = 1 - \gamma x_P \text{ for } i \in \mathcal{N}.$$

While we may still refer to naive actions as potentially cooperative, in this framework they are not actually cooperative in the sense that they do not create positive externalities. As we have indicated, results for the best response dynamics in this section would still hold for nonzero $\{\theta_j\}_{j \in \mathcal{N}}$, but with our assumption analysis of the replicator dynamics is more tractable.

6.1 Changing the Incentives to Coordinate and Predate

The next two results consider the effects of increasing the stakes by ramping up both the incentives to coordinate on an effective action and the benefits of predation. The result in the first claim aligns with the intuition that increasing incentives increases the probability of an efficient outcome. It states that as β, the benefit from coordinating on an effective action, grows large, replicator dynamics almost always achieves coordinated, effective cooperation. However, the cooperative action taken may not be the most efficient one. Since the result follows from an application of Lemma 3, it holds regardless of whether the incentives to predate are large as well.

The second claim states that under best response dynamics, as the benefits to predation grow, the basin of attraction of the self interested action expands to the entire space. This holds even if β goes to infinity as well, so long as the benefits of predation are sufficiently larger.

Together, these two results imply that as the stakes rise, we need not worry about the initial distribution of strategies. As long as an effective action is played occasionally, a population using replicator dynamics will learn to coordinate effective cooperation. And as long as naive actions cannot be ruled out entirely, if the incentive to predate rises faster than the incentive to cooperate, then a population using best response dynamics will learn to take safe, self interested action.

Theorem 6 *As $\beta \to \infty$, the basins of attraction of the equilibria featuring only effective actions ($x_S = 0, \sum_{i\in\mathcal{E}} x_i = 1$) approach the entire strategy space under replicator dynamics. This holds even if $\gamma \to \infty$ as well.*

Proof Let $m = \arg\max_j(\theta_j x_j)$ be the set of effective actions with highest payoff. By the logic that proved Lemma 1 and the fact that equation (1) still holds for our modified model, this set of best effective actions remains constant over time. If for $i \in m$, $x_i > T_i$, then Lemma 3 applies in the case that $|m| = 1$ and a straightforward extension of it applies when multiple effective actions tie for the highest payoff. In short, the replicator dynamics flow to an equilibrium that satisfies ($\sum_{i\in m} x_i = 1, x_j = 0 : j \notin m$). As $\beta \to \infty$, every $T_i \to 0$ and the set of points satisfying $x_i > T_i$ for $i \in m$ approaches the entire strategy space. ∎

Theorem 7 *Under best response dynamics the basin of attraction of the self interested equilibrium monotonically increases in γ.*

Proof As γ increases, the condition $x_j < T_j \frac{\gamma}{c} \sum_{l\in\mathcal{N}} x_l$ for all $j \in \mathcal{E}$ is satisfied for more initial points. By Theorem 4, if this condition is met, the best response dynamics flow to $(x_S = 1, x_{-S} = 0)$. ∎

Corollary 1 *As* $\gamma \to \infty$*, the basin of attraction of* $(x_S = 1, x_{-S} = 0)$ *approaches the entire strategy space under best response dynamics. This conclusion holds even if* $\beta \to \infty$ *as well, as long as* $\frac{\gamma}{\beta} \to \infty$*.*

Proof Points with $\sum_{i \in \mathcal{N}} x_i > \frac{c\theta_j \beta}{(c-1)\gamma}$ (and thus, $r(\mathbf{x})T_j > 1$) for all $j \in \mathcal{E}$ satisfy inequality (4) and are in the basin of attraction of $(x_S = 1, x_{-S} = 0)$ given best response dynamics. So, as $\gamma \to \infty$ faster than β, the basin of attraction of $(x_S = 1, x_{-S} = 0)$ approaches the entire strategy space. ∎

In the case where β and $\frac{\gamma}{\beta}$ approach infinity, Theorem 6 and Corollary 1 taken together show that best response dynamics and replicator dynamics predict entirely different outcomes. Replicator dynamics *always* achieves effective cooperation. Best response dynamics *never* does.

6.2 Changing the Number of Effective Actions

Up to this point, we have taken the partitioning function f as exogenous and examined the effects of varying payoff parameters on the basins of attraction. In practice, stag hunt games may vary in their number of effective actions: some communities may have a choice of many collective goods to produce; others may have few. To see the effect of having more potentially cooperative actions, we now treat the number of actions as a parameter. We find that when the self interested action is sufficiently attractive and the number of potentially cooperative actions grows large, both learning dynamics settle on the self interested action. Note that the condition we place on the payoff to the self interested action is only sufficient and could be weakened.

Theorem 8 *Assume* $c > 3$. *As* $n \to \infty$, *the fraction of the strategy space in the basin of attraction of the self interested equilibrium grows to* 1 *under both best response and replicator dynamics.*

Proof See appendix.

Now, we fix the payoff parameters and the number of potentially cooperative actions and examine the effect of changing which of these actions are effective. That is, we compare basins of attraction across different partition functions. Given partition f_i, let $B(f_i)$ be the union of the basins of attraction of the equilibria featuring only effective actions ($x_S = 0, \sum_{j \in \mathcal{E}_i} x_j = 1$).

Definition Partition f_2 *effectively contains* f_1 if $\mathcal{E}_1 \subset \mathcal{E}_2$ and for all $j \in \mathcal{E}_1$, θ_j is the same for both partitions.

If f_2 effectively contains f_1, then $k_2 > k_1$. That is, more of the n potentially cooperative actions are effective with partition f_2 than with f_1. We might expect that making more of the potentially cooperative actions effective would improve the probability of locating such an equilibrium. And, indeed, the basins of attraction of these good equilibria do grow under best response dynamics as more actions are made effective, as the next claim states.

Theorem 9 *If f_2 effectively contains f_1, then under best response dynamics, any initial point that results in coordination on an effective action under partition f_1 also does under partition f_2, i.e., $B(f_1) \subset B(f_2)$.*

Proof From Theorem 4 we know that under best response dynamics a point is in $B(f_i)$ if and only if for some $l \in \mathcal{E}_i$, (i) $x_l > T_l$ and (ii) $x_l > T_l \frac{\gamma}{c} \sum_{j \in \mathcal{N}_i} x_j$. If these inequalities are satisfied for a given x_l under partition f_1, they must still be satisfied for x_l under f_2 because (i) is unchanged and (ii) is weaker because $\mathcal{N}_2 \subset \mathcal{N}_1$. Moreover, there are additional actions in $\mathcal{E}_2$ for which these inequalities may be satisfied. So $B(f_1) \subset B(f_2)$. ▮

Surprisingly, a similar result does not hold for replicator dynamics. Increasing the number of effective actions creates a *crowding effect*. It can raise the average payoff in the population and therefore prevent any effective action from growing in the population.

Theorem 10 *If f_2 effectively contains f_1, then under replicator dynamics there can exist initial points that lead to coordination on an effective action under partition f_1 that do not under partition f_2, i.e., $B(f_1) \backslash B(f_2)$ need not be empty.*

Proof See appendix.

In interpreting these last two claims, we must keep in mind our earlier result that the basin of attraction for effective actions is always larger under replicator dynamics. Thus, when we change one cooperative action from naive to effective, any initial points that we move out of the basin of attraction of an effective action and into the basin of attraction of the self interested action under replicator dynamics must remain in the basin of attraction of the self interested action under best response dynamics as well.

7 Discussion

In this paper, we have analyzed a class of generalized stag hunt games and shown that cultural learning dominates individual learning: any initial condition for which individual, belief-based learning achieves effective cooperation necessarily leads to effective cooperation under replicator dynamics too. Moreover, we have shown that as the stakes grow large, cultural learning, as captured by replicator dynamics, almost always achieves coordinated, effective cooperation. In contrast, individualistic, belief-based learning captured by best response dynamics almost never does.

These dichotomous limiting results are clearly a product of our specification. Nevertheless, the core intuition holds generally: cultural learning allows moderately successful attempts at coordination to gain a foothold and potentially grow in the population, while best response learning only rewards the best action, which rarely is a nascent attempt at cooperation.[9] It is the greediness of the best response dynamics that distinguishes its behavior from the replicator dynamics. In general, we would

---

[9] Given that as a general rule, small changes in payoffs have correspondingly small effects on the basins of attraction, we can expect that slight variations in the payoff structure of our model have only mild effects on our results.

expect the greediness of a learning dynamic to be the key factor in determining outcomes here, even more so than whether the underlying revision protocol is imitative or direct.[10]

Our results reveal the importance of learning rules for equilibrium selection. We show that ramping up incentives may not be sufficient to produce a desired outcome, as it may make equilibrium selection even more dependent on the learning rule. Also, our finding that increasing the number of effective cooperative actions can move some initial conditions into the basin of an inefficient equilibrium demonstrates the costs of abundance. More effective actions can produce a crowding effect.

The primary result – that cultural learning performs better here – suggests the possibility that it might evolve. Stag hunt games are common, and societies that use cultural learning should have greater success in them. Group selection, either cultural or genetic, could promote cultural learning over belief-based learning. This hypothesis requires further study into the mechanism through which societies adopt new learning styles (see Henrich and McElreath, 2003). Evolution is not the only force. Learning may well include transference of behaviors learned in one context to other contexts (Bednar and Page 2007).

Recent work has found that behavior in a common experimental setting varies widely across cultures and that some of that variation can be explained by features of those cultures (Henrich et al., 2001). For example, cultures that engage in collective

---

[10] Sandholm (2009) defines a direct revision protocol as one in which a revising agent's choice of an alternative action to consider is without regard to its popularity.

enterprises, like whale hunting, appear more likely to share. These findings suggest that distinct cultures may differ in their learning styles. Relatedly, a substantial body of survey and case study research shows that cultures vary in their levels of individualism and collectivism (Inglehart, 1997). In more collectivist societies, people may have richer social networks giving them better knowledge of the actions of others and the payoffs of those actions. This suggests a possible link between collectivism and cultural learning. Comparably, in individualistic societies, people may be less informed about others and more concerned with acting optimally. These characteristics would point to best response learning. Seeing that, in our framework, cultural learning induces cooperation, which could reinforce underlying collectivist sentiment, it is conceivable that cooperation, collectivism, and cultural learning form a positive feedback loop that could amplify initial differences between societies.

While there is logic to the idea that more collectivist societies might have an easier time mustering cooperation, levels of collectivism and individualism do not independently determine how societies learn. Best response learning requires greater cognitive effort than cultural learning. The members of a collectivist society with strong attachment to rationality could indeed use best response learning. And members of an individualistic society might turn to cultural learning when a game becomes too complicated to think through. Surprisingly, additional cognitive effort could be counterproductive for the society if it means belief-based learning prevails over cultural learning in contexts such as ours.

Our findings also have relevance for institutional design. How agents learn may not seem to be a choice variable, and yet, each learning style requires information about the payoffs or actions of others. That information can be made more or less available through institutional choices. Additionally, when societal differences in learning processes result in significant differences in outcomes, we can try to design the games created by economic and political institutions with respect for the prevailing learning style so that the society attains better outcomes.

We conclude by reiterating the observation that the type of learning rule in use matters most when a model admits multiple equilibria. Except in those cases where learning rules do not attain an equilibrium (Hofbauer and Swinkels, 1996), single equilibrium models leave little room for how people learn, or for culture more generally, to have any effect. Thus, if we want to understand the implications of variation in learning rules, we need to consider games with more than two strategies.

Appendix A

Completing the Proof of Theorem 2.

We now identify points for which the self interested action is the best response, but for which the replicator dynamics leads to the cooperative equilibrium ($x_i = 1, x_{-i} = 0$). Consider points of the form $x_i = \frac{c-1-\epsilon}{\theta_i \beta}$, $x_l = \frac{c-1}{2\theta_l \beta}$ for some other $l \in \mathcal{C}$, and $x_j < \frac{\theta_i}{\theta_j} x_i$ for all $j \neq i \in \mathcal{C}$. Assume ϵ is small. Such points have been chosen so that $\pi_i = c - \epsilon$ and $\pi_l = \frac{c+1}{2}$. The self interested action is the best response at this point, but cooperative action i is very close and better than all other cooperative actions.

Plugging in for π_i and parts of $\bar{\pi}$, we have

$$\begin{aligned}\dot{x_i} &= x_i\left((c-\epsilon)(1-x_i) - cx_S - \sum_{j\neq i\in\mathcal{C}} \pi_j x_j\right)\\ &= x_i\left(-\epsilon(1-x_i) + \sum_{j\neq i\in\mathcal{C}} (c-\pi_j)\, x_j\right).\end{aligned}$$

The second step here used the fact that $\sum_{j\in\mathcal{A}} x_j = 1$. Dropping some positive terms from the right hand side, we get the inequality $\dot{x_i} > x_i\left((c-\pi_l)\, x_l - \epsilon\right)$. Plugging in for π_l and simplifying, we have

$$\dot{x_i} > x_i\left(\frac{c-1}{2}x_l - \epsilon\right). \tag{6}$$

As long as ϵ is small, the right hand side of inequality (6) is positive and x_i increases past $\frac{c-1}{\theta_i\beta}$ right away. This makes cooperative action i a best response, and then Lemma 1 applies and we know the dynamics leads to the cooperative equilibrium $(x_i = 1, x_{-i} = 0)$. ▮

Proof of Theorem 4.

For the purposes of the proof, it is helpful to write inequality (4) as

$$x_j < T_j \tag{7}$$

or

$$x_j < r(\mathbf{x})T_j, \tag{8}$$

and inequality (5) as the following set of three conditions:

i) $x_i > T_i$;

ii) $x_i > r(\mathbf{x})T_i$; and

iii) $x_i > \frac{\theta_j}{\theta_i} x_j$ for all $j \neq i \in \mathcal{E}$.

We point out that if for some $i \in \mathcal{E}$, $x_i > T_i$ and $x_i < r(\mathbf{x})T_i$, then $x_i < \frac{1}{\theta_i \beta}\left(\gamma \sum_{l \in \mathcal{N}} x_l - 1\right)$. (Actually, if equality holds in one but not both of the conditions, we still obtain the desired inequality.) This fact follows just from rearranging terms, using a fair bit of basic algebra. It means that if we are relying on inequality (8) to establish that a point is in the basin of the self interested equilibrium (i.e., when inequality (7) fails and some effective action is initially better than the self interested action), then inequality (8) ensures that the predatory action is initially better than this effective action. Alternatively, if we are establishing that a point is in the basin of attraction of an effective equilibrium, we can say that in order for an effective action to be initially better than the self interested action and the predatory action, condition (ii) must hold.

Now, assume best response dynamics. We will invoke Lemma 2, which tells us that if either the self interested action or an effective action is initially a best response or becomes one, it remains a best response forever, so the best response dynamics flow towards the equilibrium featuring this action.

For effective action i to initially be the best response, conditions (i) and (iii) obviously must hold and the need for condition (ii) is described in the proof's first paragraph. For the self interested action S to initially be the best response, inequality (7) is clearly necessary.

If neither the self interested action nor any of the effective actions are initially a best response, then the predatory action P must initially be the best response. In this case, the equilibrium attained depends on which action next becomes the best response. So, let us now consider points for which the predatory action P is the best response. If for all $j \in \mathcal{E}$, inequality (7) holds and $x_j < T_j$, then the self interested action S has a higher payoff than any of the effective actions. As the dynamic moves toward P, the payoffs to the predatory and effective actions decrease, so eventually S becomes the best response. Alternatively, suppose for some $i \in \mathcal{E}$, condition (i) holds and $x_i > T_i$. Then we define $\alpha(t) = \frac{x_j(t)}{x_j(0)}$ for $j \neq P$ and t such that action P is still a best response. This definition is independent of j because actions which are not best responses have the same relative decay rate. Note that $\alpha(t)$ is a strictly decreasing function. Now either

$$\pi_S = \pi_P \text{ when } \alpha = \frac{c}{\gamma \sum_{l \in \mathcal{N}} x_l(0)} \tag{9}$$

or for some $i \in \mathcal{E}$,

$$\pi_i = \pi_P \text{ when } \alpha = \frac{1}{\left(\gamma \sum_{l \in \mathcal{N}} x_l(0) - \theta_i \beta x_i(0)\right)}, \tag{10}$$

whichever happens first. Equation (9) follows from $\pi_P = \gamma \sum_{l \in \mathcal{N}} \alpha(t) x_l(0)$. Equation (10) depends on this as well as on $\pi_i = 1 + \theta_i \beta \alpha(t) x_i(0)$ for $i \in \mathcal{E}$. If inequality (8) applies, i.e., if for all $j \in \mathcal{E}$, $x_j(0) < \frac{(c-1)\gamma}{c\theta_j\beta} \sum_{l \in \mathcal{N}} x_l(0)$, then rearranging terms produces

$$\frac{c}{\gamma \sum_{l \in \mathcal{N}} x_l(0)} > \frac{1}{\left(\gamma \sum_{l \in \mathcal{N}} x_l(0) - \theta_j \beta x_j(0)\right)},$$

and this means action S eventually becomes the best response. On the other hand, if for some $i \in \mathcal{E}$, $x_i(0) > \frac{(c-1)\gamma}{c\theta_i\beta}\sum_{l\in\mathcal{N}} x_l(0)$ and $x_i(0) > \frac{\theta_j}{\theta_i}x_j(0)$ for all $j \neq i \in \mathcal{E}$, conditions (ii) and (iii) respectively, then action i always has the highest payoff of all the effective actions and becomes the best response before the self interested action does. ▮

Proof of Theorem 8.

Let $M\{*\}$ denote the fraction of the strategy space satisfying $\{*\}$. As $n \to \infty$, $M\{x_j < T_j \text{ for all } j \in \mathcal{E}\}$ approaches 1. By Theorem 4, all points satisfying this condition are in the basin of attraction of $(x_S = 1, x_{-S} = 0)$ with the best response dynamics.

Assume replicator dynamics. A sufficient condition to be in this basin of attraction is at some time t,

$$cx_S(t) > 1 + \theta_j\beta x_j(t) \text{ for all } j \in \mathcal{E}, \tag{11}$$

as this ensures that $\pi_j < \bar{\pi}$ then and at all future times. An alternative condition is that

$$\bar{\pi} - \sum_{i\in\mathcal{N}} x_i \geq 2\pi_j - 1 \text{ for all } j \in \mathcal{E} \tag{12}$$

at some time. Inequality (11) is self enforcing because it ensures that x_S increases while π_j decreases for all $j \in \mathcal{E}$. To see that inequality (12) is self enforcing is slightly more involved. It too ensures that π_j decreases for all $j \in \mathcal{E}$, but now we must take a time derivative of $\bar{\pi} - \sum_{i \in \mathcal{N}} x_i$ and show that it is positive. We get

$$\begin{aligned} \dot{\bar{\pi}} - \sum_{i \in \mathcal{N}} \dot{x}_i &= \sum_{l \in \mathcal{A}} \dot{x}_l \pi_l + x_l \dot{\pi}_l - \sum_{i \in \mathcal{N}} \dot{x}_i \\ &= \sum_{l \in \mathcal{A}} \dot{x}_l (\pi_l - \bar{\pi}) + x_l \dot{\pi}_l - \sum_{i \in \mathcal{N}} \dot{x}_i . \end{aligned}$$

The last step here uses $\sum_{l \in \mathcal{A}} \dot{x}_l = 0$. We can write $\dot{x}_l(\pi_l - \bar{\pi})$ as $x_l(\pi_l - \bar{\pi})^2$. For $l \in \mathcal{E}$, $x_l \dot{\pi}_l = x_l(\pi_l - 1)(\pi_l - \bar{\pi})$, and $x_l(\pi_l - \bar{\pi})^2 > x_l(\pi_l - 1)(\bar{\pi} - \pi_l)$ by inequality (12) itself. So

$$\sum_{l \in \mathcal{E}} \dot{x}_l (\pi_l - \bar{\pi}) + x_l \dot{\pi}_l > 0. \tag{13}$$

We can plug in the payoffs for the predatory and the naive actions and as a shortcut use $\sum_{l \in \{P\} \cup \mathcal{N}} x_l \pi_l = \sum_{i \in \mathcal{N}} x_i$, taking the time derivative of both sides, to write

$$\sum_{l \in \{P\} \cup \mathcal{N}} \dot{x}_l (\pi_l - \bar{\pi}) + x_l \dot{\pi}_l = \sum_{l \in \{P\} \cup \mathcal{N}} \dot{x}_l (-\bar{\pi}) + \sum_{i \in \mathcal{N}} \dot{x}_i .$$

We find that $\sum_{l \in \{P\} \cup \mathcal{N}} \dot{x}_l \leq 0$ as long as $\bar{\pi} \geq 1$ because $\sum_{l \in \mathcal{N}} \dot{x}_l = \sum_{l \in \mathcal{N}} x_l \left(1 - \gamma x_P - \bar{\pi}\right)$ and $\dot{x_P} = x_P \left(\sum_{l \in \mathcal{N}} \gamma x_l - \bar{\pi}\right)$ imply that

$$\sum_{l \in \{P\} \cup \mathcal{N}} \dot{x}_l = \sum_{l \in \mathcal{N}} x_l (1 - \bar{\pi}) + x_P(-\bar{\pi}).$$

And we know $\bar{\pi} \geq 1$ from inequality (12). So

$$\sum_{l \in \{P\} \cup \mathcal{N}} \dot{x}_l (\pi_l - \bar{\pi}) + x_l \dot{\pi}_l - \sum_{i \in \mathcal{N}} \dot{x}_i > 0. \tag{14}$$

Finally, $x_S(\pi_S - \bar{\pi})^2$ is clearly positive and $\dot{\pi_S} = 0$, so

$$\dot{x}_S (\pi_S - \bar{\pi}) + x_S \dot{\pi}_S > 0. \tag{15}$$

Thus, piecing together inequalities (13), (14) and (15), we get $\dot{\bar{\pi}} - \sum_{i \in \mathcal{N}} \dot{x}_i > 0$.

Let j be a best effective action. We can place an upper bound on the rate at which action j spreads, $\dot{x_j} = x_j(1 + x_j\theta_j\beta - \bar{\pi}) < x_j(1 + x_j\theta_j\beta)$. This bound has the form of a logistic differential equation. The solution is then bounded by the logistic function,

$$x_j(t) \leq \frac{x_j(0)}{(1 + \theta_j\beta x_j(0))e^{-t} - \theta_j\beta x_j(0)}. \tag{16}$$

Because inequality (12) is sufficient for the replicator dynamics to flow to $(x_S = 1, x_{-S} = 0)$, we consider an assumption that $\bar{\pi} < 2\pi_j - 1 + \sum_{i\in\mathcal{N}} x_i \leq 2\pi_j$. This allows us to place a lower bound on the rate at which action S spreads, $\dot{x_S} = x_S(c - \bar{\pi}) > x_S(c - 2 - 2\theta_j\beta x_j)$. Then, plugging in inequality (16) for $x_j(t)$ and integrating,

$$x_S(t) \geq x_S(0)e^{(c-2)t}(1 + \theta_j\beta x_j(0)(1 - e^t))^2.$$

Applying this to our first sufficient condition, inequality (11), it is sufficient to show

$$cx_S(0)e^{(c-2)t}(1 + \theta_j\beta x_j(0)(1 - e^t))^2 > 1 + \frac{\theta_j\beta x_j(0)}{(1 + \theta_j\beta x_j(0))e^{-t} - \theta_j\beta x_j(0)}$$

or equivalently,

$$cx_S(0)e^{(c-2)t}(1 + \theta_j\beta x_j(0)(1 - e^t))^3 > 1 + \theta_j\beta x_j(0). \tag{17}$$

This last step requires some algebra. The left hand side of (17) is maximized at

$$t = \ln\left(\frac{1 + \theta_j\beta x_j(0)}{\theta_j\beta x_j(0)\frac{c+1}{c-2}}\right).$$

Plugging in for t in (17), the sufficient condition becomes

$$cx_S(0)\left(\frac{(1 + \theta_j\beta x_j(0))(c-2)}{\theta_j\beta x_j(0)(c+1)}\right)^{c-2}\left(1 + \theta_j\beta x_j(0)\left(1 - \frac{(1 + \theta_j\beta x_j(0))(c-2)}{\theta_j\beta x_j(0)(c+1)}\right)\right)^3$$
$$> 1 + \theta_j\beta x_j(0).$$

As $n \to \infty$, $x_j(0)$ becomes small, so we keep only terms of lowest order in $x_j(0)$. This simplifies our sufficient condition to

$$c x_S(0) \left(\frac{c-2}{\theta_j \beta x_j(0)(c+1)} \right)^{c-2} \left(\frac{3}{c+1} \right)^3 > 1.$$

It remains only to show that this condition is met almost everywhere when n is large.

Our sufficient condition holds if

$$x_j(0) \le \frac{1}{n^{\frac{c}{3(c-2)}}} \text{ and } x_S(0) > \frac{1}{c} \left(\frac{\theta_j \beta (c+1)}{c-2} \right)^{c-2} \left(\frac{c+1}{3} \right)^3 \frac{1}{n^{\frac{c}{3}}}. \qquad (18)$$

Because these two inequalities are positively correlated, [11]

$$M\{\text{Constraint (18)}\} \ge$$
$$M\left\{ x_j(0) \le \frac{1}{n^{\frac{c}{3(c-2)}}} \right\} \cdot M\left\{ x_S(0) > \frac{1}{c} \left(\frac{\theta_j \beta (c+1)}{c-2} \right)^{c-2} \left(\frac{c+1}{3} \right)^3 \frac{1}{n^{\frac{c}{3}}} \right\}.$$

Consider the first of these inequalities. We have

$$\begin{aligned} M\left\{ x_j(0) > \frac{1}{n^{\frac{c}{3(c-2)}}} \right\} &\le M\left\{ x_i(0) > \frac{1}{n^{\frac{c}{3(c-2)}}} \text{ for some } i \in \mathcal{E} \cup \mathcal{N} \right\} \\ &\le n \cdot M\left\{ x_1(0) > \frac{1}{n^{\frac{c}{3(c-2)}}} \right\} \\ &= n \left(1 - \frac{1}{n^{\frac{c}{3(c-2)}}} \right)^{n+1}. \end{aligned}$$

Here, and again in equation (19), we evaluate the fraction of the strategy space satisfying a given inequality simply by integrating over the strategy space. Now, $\frac{c}{3(c-2)} < 1$ because we assumed $c > 3$, so

$$\lim_{n \to \infty} n \left(1 - \frac{1}{n^{\frac{c}{3(c-2)}}} \right)^{n+1} = 0.$$

---

[11] Recall that $M\{*\}$ denotes the fraction of the strategy space satisfying $\{*\}$.

Thus,

$$\lim_{n\to\infty} M\left\{ x_j(0) \leq \frac{1}{n^{\frac{c}{3(c-2)}}} \right\} = 1.$$

Now consider the second inequality. We have

$$M\left\{ x_S(0) > \frac{1}{c}\left(\frac{\theta_j\beta(c+1)}{c-2}\right)^{c-2}\left(\frac{c+1}{3}\right)^3 \frac{1}{n^{\frac{c}{3}}} \right\}$$
$$= \left[1 - \frac{1}{c}\left(\frac{\theta_j\beta(c+1)}{c-2}\right)^{c-2}\left(\frac{c+1}{3}\right)^3 \frac{1}{n^{\frac{c}{3}}}\right]^{n+1}. \quad (19)$$

This approaches 1 as $n \to \infty$ because $\frac{c}{3} > 1$. Thus,

$$\lim_{n\to\infty} M\left\{ c x_S(0)\left(\frac{c-2}{\theta_j\beta x_j(0)(c+1)}\right)^{c-2}\left(\frac{3}{c+1}\right)^3 > 1 \right\} = 1.$$

The fraction of the strategy space satisfying a condition that puts it in the basin of attraction of $(x_S = 1, x_{-S} = 0)$ approaches 1. ■

Proof of Theorem 10.

We construct a specific counterexample for the case $n = 3$ that can be extended to a more general case. Let $\mathcal{E}_1 = \{1\}$, $\mathcal{E}_2 = \{1, 2\}$, $\theta_1 = 1$, $\theta_2 = 1$ under partition f_2 (whereas, of course, $\theta_2 = 0$ under partition f_1), and $\beta > 2c$. Then any point which satisfies

$$x_1 = x_2 = \frac{c-1}{\beta} - \epsilon,\ x_S = \frac{c-\beta\epsilon}{c}\left(1 + 2\epsilon - 2\frac{c-1}{\beta}\right)$$

for small enough ϵ will be in $B(f_1)$ but not in $B(f_2)$.

Consider partition f_1. Recall that $\dot{x}_1 = x_1(\pi_1 - \bar{\pi})$. By construction $\pi_1 = (c - \beta\epsilon)$ and still $\pi_S = c$. Plugging in and simplifying, we get the average payoff

$$\bar{\pi} = (c - \beta\epsilon)\left(1 - \frac{c-1}{\beta} + \epsilon\right) + \frac{c-1}{\beta} - \epsilon + x_3.$$

We combine terms and find that at our initial point, $\pi_1 - \bar{\pi} = \beta \left(\frac{c-1}{\beta} - \epsilon \right)^2 - x_3$. Therefore, initially, $\dot{x_1} = \beta \left(\frac{c-1}{\beta} - \epsilon \right)^3 - \left(\frac{c-1}{\beta} - \epsilon \right) x_3$. From the fact that $\sum_{j \in \mathcal{A}} x_j = 1$, we know our initial point satisfies $x_3 \leq \frac{\beta\epsilon}{c} \left(1 + 2\epsilon - 2\frac{c-1}{\beta} \right)$. This gives us a minimum initial value for $\dot{x_1}$,

$$\dot{x_1} \geq \beta \left(\frac{c-1}{\beta} - \epsilon \right)^3 - \left(\frac{c-1}{\beta} - \epsilon \right) \frac{\beta\epsilon}{c} \left(1 + 2\epsilon - 2\frac{c-1}{\beta} \right). \tag{20}$$

Observe that the right hand side of (20) has a positive leading order term with no ϵ dependence. As ϵ is small, x_1 soon grows larger than $\frac{c-1}{\beta}$. By Lemma 3, the point must be in $B(f_1)$.

Now consider partition f_2. The average payoff is larger with this partition. In fact, $\pi_1 = \pi_2 = c - \beta\epsilon$, and it turns out $\bar{\pi} = c - \beta\epsilon + x_3$ at our initial point. This means that initially $\pi_1 = \pi_2 = \bar{\pi} - x_3$. We will now see that the state of

$$\pi_1 = \pi_2 \leq \bar{\pi} - x_3 \tag{21}$$

must persist forever because it is self enforcing. Note that $x_1 = x_2$ and $\pi_1 = \pi_2$ at all times by the symmetry of their initial conditions. We can plug in $\bar{\pi} - x_3 = cx_S + 2\pi_i x_i$ with $i \in \{1, 2\}$ and then rewrite equation (21) as

$$\pi_i (1 - 2x_i) \leq cx_S. \tag{22}$$

We now compare time derivatives of both sides of inequality (22) and show $c\dot{x_S} > \dot{\pi_i} (1 - 2x_i) + \pi_i (-2\dot{x_i})$. In particular, $\dot{\pi_i} < 0$ because of inequality (21) itself, and clearly $1 - 2x_i > 0$. It remains to show $c\dot{x_S} + 2\pi_i \dot{x_i} \geq 0$. We have $c > \pi_i > 0$, so it is sufficient to show $\dot{x_S} + 2\dot{x_i} \geq 0$. And, recognizing that $\sum_{j \in \mathcal{A}} \dot{x_j} = 0$, it is fine to show $\dot{x_3} + \dot{x_P} \leq 0$. In the proof of Theorem 8, we show this will be negative as long as $\bar{\pi} > 1$. We know $\bar{\pi} > cx_S$, and we know x_S has been increasing because $c > \bar{\pi}$. (To check this last inequality, just examine the formula for the average payoff and recall that we have already argued that $c > \pi_i$ in our persistent state.) Finally, we obtain

$$\begin{aligned} cx_S(0) &= (c - \beta\epsilon)\left(1 + 2\epsilon - 2\frac{c-1}{\beta}\right) \\ &= c - \frac{2c}{\beta}(c-1) + O(\epsilon) \; > 1, \end{aligned}$$

using in the last step the facts that ϵ is small and $2c < \beta$. Because the average payoff always remains above the payoff to either of the effective actions, the effective actions become rarer, and it follows that the initial point is not in $B(f_2)$. ∎

Appendix B

Figures 5 and 6 show sample trajectories of the two learning dynamics for a game with just one effective action and one naive action. The images have a three-dimensional perspective because the strategy space for a game with four actions is the three-dimensional simplex $\triangle^3$. The parameter values ($c = 2$, $\gamma = 100$, $\beta = 10$, $\theta_E = \theta_N = 1$)[12] are inspired by our results in Section 6.1. As we can see in Figure 5, the best response dynamics often approach the predatory action at first, before flowing directly to the self interested action; although not shown, some points near the effective action would flow to that equilibrium as well. In Figure 6, we see that the replicator dynamics often curves towards the effective action as the naive action decays; again,

---

[12] We denote the single effective action by E and the naive action by N.

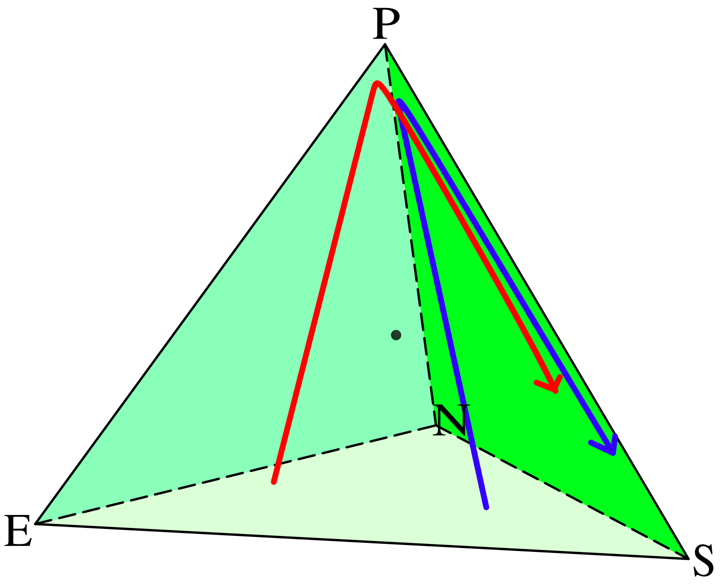

Figure 5: Sample trajectories for the best response dynamics in a game with just one effective action and one naive action, setting $c = 2$, $\gamma = 100$, $\beta = 10$, $\theta_E = \theta_N = 1$. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

some trajectories (not shown) in the neighborhood of the self interested action would reach that equilibrium, too. We show in Figures 7 and 8 sample trajectories of the replicator dynamics while varying θ_N and retaining the other parameter values used in Figure 6.

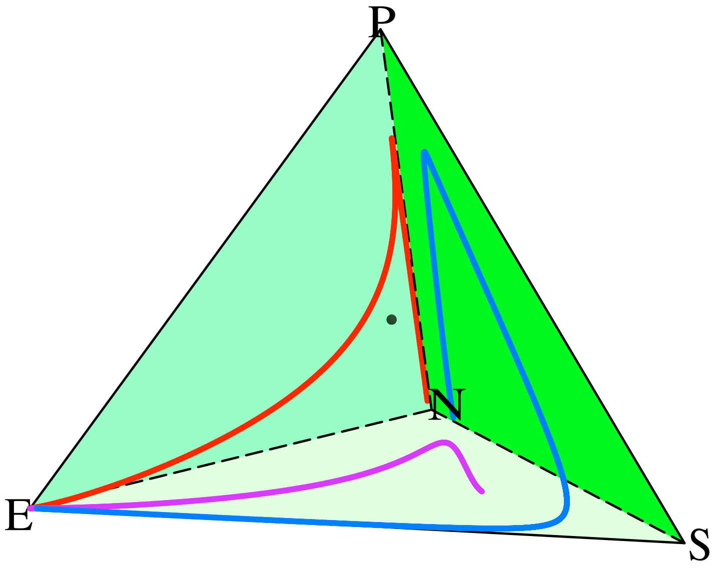

Figure 6: Sample trajectories for the replicator dynamics in a game with just one effective action and one naive action, setting $c = 2$, $\gamma = 100$, $\beta = 10$, $\theta_E = \theta_N = 1$. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

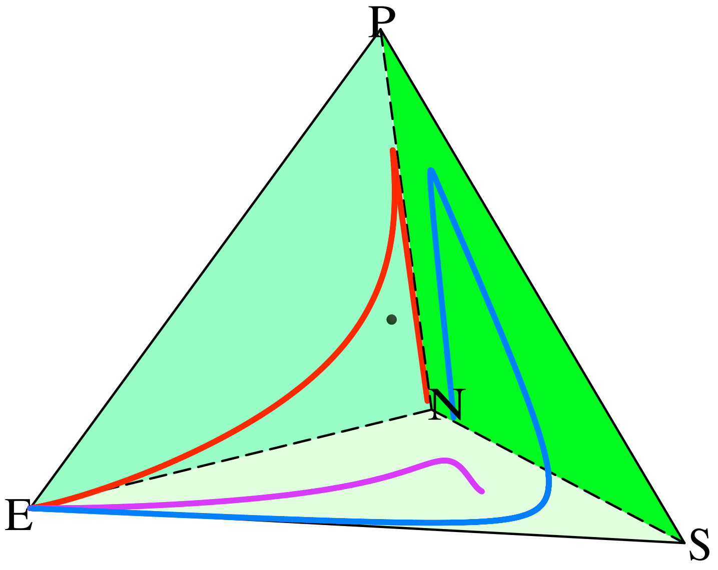

Figure 7: Sample trajectories for the replicator dynamics, setting $c = 2$, $\gamma = 100$, $\beta = 10$, $\theta_E = 1$, $\theta_N = 0$. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

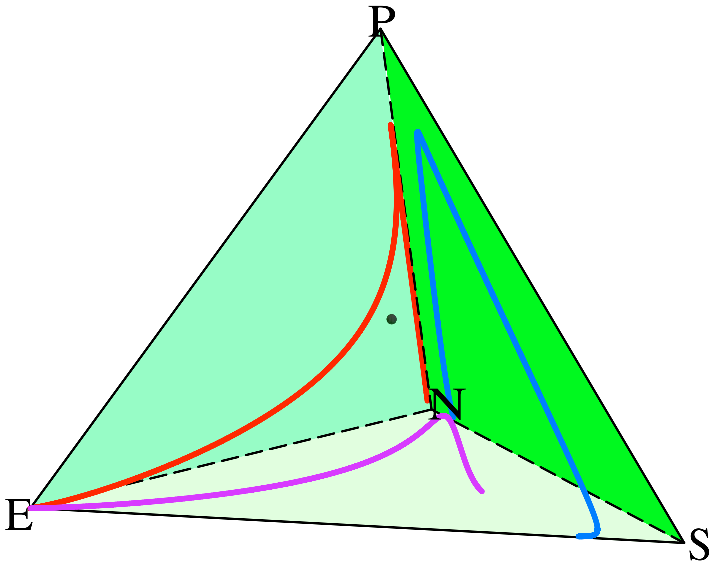

Figure 8: Sample trajectories for the replicator dynamics, setting $c = 2$, $\gamma = 100$, $\beta = 10$, $\theta_E = 1$, $\theta_N = 2$. Figure made by the game dynamics simulation program Dynamo (Sandholm and Dokumaci, 2007).

References

[1] Abreu, D. , Pearce, D., Stacchetti, E., 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring. *Econometrica* 58, 1041-1063.

[2] Axelrod, R., 1984. *The Evolution of Cooperation.* New York: Basic Books.

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

[4] Bednar, J., Page, S.E., 2007. Can Game(S) Theory Explain Culture? *Rationality and Society* 19(1) 65-97. forthcoming.

[5] Bendor, J., Swistak, P., 1997. The Evolutionary Stability of Cooperation. *American Political Science Review* 91(2), 290-307.

[6] Björnerstedt, J., Weibull, J.W., 1996. Nash Equilibrium and Evolution by Imitation. *The Rational Foundations of Economic Behavior*. Arrow, K. J., et. al., 155-181. New York: St. Martins Press.

[7] Börgers, T., Sarin, R., 1997. Learning Through Reinforcement and Replicator Dynamics. *Journal of Economic Theory* 77, 1-14.

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

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

[10] Cheung, Y.-W, Friedman, D., 1998. A Comparison of Learning and Replicator Dynamics Using Experimental Data. *Journal of Economic Behavior and Organization* 35 (3), 263-280.

[11] Conlisk, J., 2001. Costly Predation and the Distribution of Competence. *The American Economic Review* 91 (3), 475-484.

[12] Epstein, J., 2003. *Generative Social Science.* Princeton: Princeton University Press.

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

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

[15] Fudenberg, D., Levine, D., 1999. *The Theory of Learning in Games.* Cambridge: MIT Press.

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

[17] Gintis, H., Bowles, S., Boyd, R., Fehr, E., 2003. Explaining Altruistic Behavior in Humans. *Evolution and Human Behavior* 24, 153-172.

[18] Golman, R., Page, S., 2008. Basins of Attraction and Equilibrium Selection Under Different Learning Rules. *Journal of Evolutionary Economics*, forthcoming.

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

[20] Henrich, J. 2001. Cultural Transmission and the Diffusion of Innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change and much of sociocultural evolution. *American Anthropologist*, 103: 992-1013.

[21] Henrich, J., Boyd, R., 2002. On Modeling Cognition and Culture: Why Cultural Evolution Does Not Require Replication of Representations. *Journal of Cognition and Culture* 2(2), 87-112.

[22] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., 2001. In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review* 91, 73-78.

[23] Henrich, J., Gil-White, F., 2001. The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission. *Evolution and Human Behavior* 22, 165-196.

[24] Henrich, J., McElreath, R., 2003. The Evolution of Cultural Evolution. *Evolutionary Anthropology* 12, 123-135.

[25] Hofbauer, J., 2000. From Nash and Brown to Maynard Smith: Equilibria, Dynamics and ESS. *Selection* 1, 81-88.

[26] Hofbauer, J., Schuster, P., Sigmund, K., 1979. A Note on Evolutionary Stable Strategies and Game Dynamics. *Journal of Theoretical Biology* 81, 609-612.

[27] Hofbauer, J., Sigmund, K., 2003. Evolutionary Game Dynamics. *Bulletin of the American Mathematical Society* 40, 479-519.

[28] Hofbauer, J., Swinkels, J.M., 1996. A Universal Shapley Example. Unpublished manuscript, University of Vienna and Northwestern University.

[29] Hopkins, E., 1999. A Note on Best Response Dynamics. *Games and Economic Behavior* 29, 138-150.

[30] Inglehart, R., 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies.* Princeton: Princeton University Press.

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

[32] Kandori, M., Rob, R., 1995. Evolution of equilibria in the long run: a general theory and applications. *Journal of Economic Theory* 65, 383-414.

[33] Medina, L., 2007. *A Unified Theory of Collective Action and Social Change.* Ann Arbor: University of Michigan Press.

[34] Ochs, J., 1995. Coordination Problems. *Handbook of Experimental Economics.* Kagel, J., Roth, A. Princeton: Princeton University Press.

[35] Rousseau, J., 1773. *A Discourse on Inequality.*

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

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

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

[39] Schelling, T., 1960. *The Strategy of Conflict*. Harvard University Press.

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

[41] Skyrms, B., 2001. The Stag Hunt. *Proceedings and Addresses of the American Philosophical Association* 75(2), 31-41.

[42] Skyrms, B., 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.

[43] Taylor, P.D., Jonker, L., 1978. Evolutionarily Stable Strategies and Game Dynamics. *Mathematical Biosciences* 40, 145-156.

[44] Young, H.P., 1993. The evolution of conventions. *Econometrica* 61, 57-84.