

Survival of Dominated Strategies under Evolutionary Dynamics*

Josef Hofbauer[†] and William H. Sandholm[‡]

April 6, 2007

Abstract

We show that any evolutionary dynamic that satisfies three mild requirements—*continuity*, *positive correlation*, and *innovation*—does not eliminate strictly dominated strategies in all games. Likewise, we demonstrate that existing elimination results for evolutionary dynamics are not robust to small changes in the specifications of the dynamics.

1. Introduction

One fundamental issue in evolutionary game theory concerns the relationship between its predictions and those provided by traditional, rationality-based solution concepts. Indeed, much of the early interest in the theory among economists is due to its ability to justify traditional equilibrium predictions as consequences of myopic decisions made by simple agents.

Some of the best known results in this vein link the rest points of an evolutionary dynamic with the Nash equilibria of the game being played. There are many dynamics whose rest points include all Nash equilibria, and one can construct classes of dynamics whose rest points exactly match the Nash equilibria of the underlying game.¹

*We thank Larry Samuelson and many seminar audiences for helpful comments. Financial support from NSF Grants SES-0092145 and SES-0617753, the Centre for Economic Learning and Social Evolution (ELSE) at University College London, and the University of Vienna is gratefully acknowledged.

[†]Department of Mathematics, University of Vienna, Nordbergstrasse 15, A-1090 Vienna, Austria. e-mail: josef.hofbauer@univie.ac.at; website: <http://homepage.univie.ac.at/josef.hofbauer>.

[‡]Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706, USA. e-mail: whs@ssc.wisc.edu; website: <http://www.ssc.wisc.edu/~whs>.

¹For results of the latter sort, see Brown and von Neumann (1950), Smith (1984), and Sandholm (2005, 2006a).

In order to improve upon these results, one might instead seek a connection between dynamically stable sets and Nash equilibria. Such a connection would provide a strong defense of the Nash prediction, as agents who began play in some disequilibrium state could be expected to find their way to Nash equilibrium. Unfortunately, results of this kind cannot be proved. Hofbauer and Swinkels (1996) and Hart and Mas-Colell (2003) show that no reasonable evolutionary dynamic converges to Nash equilibrium in all games: there are some games in which cycling or more complicated limit behavior far from any Nash equilibrium is the only plausible long run prediction.

These negative results lead us to consider a more modest question. Rather than seek evolutionary support for equilibrium play, we instead turn our attention to a more basic rationality requirement: namely, the avoidance of strategies that are strictly dominated by a pure strategy.

Research on this question to date has led to a number of positive results. Two of the canonical evolutionary dynamics are known to eliminate strictly dominated strategies, at least from most initial conditions. Akin (1980) shows that starting from any interior population state, the replicator dynamic (Taylor and Jonker (1978)) eliminates strategies that are strictly dominated by a pure strategy. Samuelson and Zhang (1992), building on work of Nachbar (1990), extend this result to a broad class of evolutionary dynamics driven by imitation: namely, dynamics under which strategies' percentage growth rates are ordered by their payoffs.² Elimination results are also available for dynamics based on traditional choice criteria: the best response dynamic (Gilboa and Matsui (1991)) eliminates strictly dominated strategies by construction, as under this dynamic revising agents always switch to optimal strategies.

Since the elimination of strategies strictly dominated by a pure strategy is the mildest requirement employed in standard game-theoretic analyses, it may seem unsurprising that two basic evolutionary dynamics obey this dictum. Nevertheless, there is no a priori reason to expect dominated strategies to be eliminated. Evolutionary dynamics are built upon the notion that agents switch to strategies whose current payoffs are reasonably good. But even if a strategy is dominated, it can have reasonably good payoffs at many population states. Put differently, domination is a "global" property, depending on payoffs at all states, while decision making in evolutionary models is "local", depending only on the payoffs available at present. By this logic, there is no reason to expect evolutionary dynamics to eliminate dominated strategies as a general rule.

In this paper, we argue that evolutionary support for the elimination of dominated

²Samuelson and Zhang (1992) and Hofbauer and Weibull (1996) also introduce classes of imitative dynamics under which strategies strictly dominated by a mixed strategy are eliminated.

strategies is much more tenuous than the results noted above suggest. In particular, we prove that all evolutionary dynamics satisfying three mild conditions—continuity, positive correlation, and innovation—fail to eliminate strictly dominated strategies in all games. Dynamics satisfying these conditions include not only well-known dynamics from the evolutionary literature, but also slight modifications of the dynamics under which elimination is known to occur. In effect, this paper shows that the dynamics known to eliminate strictly dominated strategies are the only ones one should expect to do so, and that even these elimination results are knife-edge cases.

An important predecessor of this study is the work of Berger and Hofbauer (2006), who present a game in which a strictly dominated strategy survives under the Brown-von Neumann-Nash (BNN) dynamic (Brown and von Neumann (1950)). We begin the present study by showing how Berger and Hofbauer’s (2006) analysis can be extended to a variety of other dynamics, including the pairwise difference (PD) dynamic (Smith (1984)) and certain generalizations of the BNN and PD dynamics (Hofbauer (2000), Sandholm (2005, 2006a)). While this analysis is relatively simple, it is not general, as it depends on the functional forms of the dynamics at issue. Since in practice it is difficult to know exactly how agents will update their choices over time, a more compelling elimination result would require only minimal structure.

Our main theorem provides just such a result. Rather than specifying functional forms for the evolutionary dynamics under consideration, the theorem allows for any dynamic satisfying three mild conditions. The first, *continuity*, asks that a dynamic change continuously as a function of the payoff vector and the population state. The second, *positive correlation*, is a weak monotonicity condition: it demands that away from equilibrium, the correlation between strategies’ payoffs and growth rates always be positive. The final condition, *innovation*, is a requirement that has force only at non-Nash boundary states: if at such a state some unused strategy is a best response, the growth rate of this strategy must be positive. Both innovation and positive correlation rule out the replicator dynamic and the other purely imitative dynamics noted above; however, they allow arbitrarily close approximations of these dynamics, under which agents usually imitate successful opponents, but occasionally select new strategies directly.

To prove the main theorem, we construct a four-strategy game in which one strategy that is strictly dominated by another pure strategy. We show that under any dynamic satisfying our three conditions, the strictly dominated strategy survives along solutions starting from most initial conditions.

Section 2 introduces population games and evolutionary dynamics. Section 3 establishes the survival result for the PD dynamic and other related dynamics. Section 4

presents the main result. Section 5 discusses key ingredients of the analysis, possible alternative proofs, and the nonrobustness of existing elimination results. Section 6 concludes. Auxilliary results and proofs omitted from the text are provided in the Appendix.

2. The Model

2.1 Population Games

We consider games played by a single unit mass population of agents. All agents choose from the finite set of strategies $S = \{1, \dots, n\}$. The set of population states is therefore the simplex $X = \{x \in \mathbf{R}_+^n : \sum_{i \in S} x_i = 1\}$, where x_i is the proportion of agents choosing strategy $i \in S$. The standard basis vector $e_i \in \mathbf{R}^n$ represents the state at which all agents choose strategy i .

If we take the set of strategies as fixed, we can identify a game with a Lipschitz continuous *payoff function* $F : X \rightarrow \mathbf{R}^n$, which assigns each population state $x \in X$ a vector of payoffs $F(x) \in \mathbf{R}^n$. The component $F_i : X \rightarrow \mathbf{R}$ represents the payoffs to strategy i alone. We also let $\bar{F}(x) = \sum_{i \in S} x_i F_i(x)$ denote the population's average payoff, and let $B^F(x) = \operatorname{argmax}_{y \in X} y'F(x)$ denote the set of (mixed) best responses at population state x .

The simplest examples of population games are generated by random matching in symmetric normal form games. An n -strategy symmetric normal form game is defined by a payoff matrix $A \in \mathbf{R}^{n \times n}$. A_{ij} is the payoff a player obtains when he chooses strategy i and his opponent chooses strategy j ; this payoff does not depend on whether the player in question is called player 1 or player 2. When agents are randomly matched to play this game, the (expected) payoff to strategy i at population state x is $F_i(x) = \sum_{j \in S} A_{ij}x_j$; hence, the population game associated with A is the linear game $F(x) = Ax$.

While random matching generates population games with linear payoffs, many population games that arise in applications have payoffs that are nonlinear in the population state—see Section 5.1. Games with nonlinear payoff functions will play a leading role in the analysis to come.

2.2 Evolutionary Dynamics

An evolutionary dynamic assigns each population game F an ordinary differential equation $\dot{x} = V^F(x)$ on the simplex X . One simple and general way of defining an evolutionary dynamic is via a growth rate function $g : \mathbf{R}^n \times X \rightarrow \mathbf{R}^n$; here $g_i(\pi, x)$ represents the (absolute) growth rate of strategy i as a function of the current payoff vector $\pi \in \mathbf{R}^n$

and the current population state $x \in X$. Our notation suppresses the dependence of g on the number of strategies n .

To ensure that the simplex is forward invariant under the induced differential equations, the function g must satisfy

$$g_i(\pi, x) \geq 0 \text{ whenever } x_i = 0, \text{ and } \sum_{i \in S} g_i(\pi, x) = 0.$$

In words: strategies that are currently unused cannot become less common, and the sum of all strategies' growth rates must equal zero. A growth rate function g satisfying these conditions defines an evolutionary dynamic as follows:

$$\dot{x}_i = V_i^F(x) = g_i(F(x), x).$$

One can also build evolutionary dynamics from a more structured model that not only provides explicit microfoundations for the dynamics, but also is inclusive enough to encompass all dynamics considered in the literature.³ In this model, the growth rate function g is replaced by a *revision protocol* $\rho : \mathbf{R}^n \times X \rightarrow \mathbf{R}_+^{n \times n}$, which describes the process through which individual agents make decisions. As time passes, agents are chosen at random from the population and granted opportunities to switch strategies. When an i player receives such an opportunity, he switches to strategy j with probability proportional to the conditional switch rate $\rho_{ij}(\pi, x)$. Aggregate behavior in the game F is then described by the differential equation

$$\dot{x}_i = V_i^F(x) = \sum_{j \in S} x_j \rho_{ji}(F(x), x) - x_i \sum_{j \in S} \rho_{ij}(F(x), x).$$

The first term above captures the inflow of agents into strategy i from other strategies, while the second term captures the outflow of agents from strategy i to other strategies.

Table I presents four canonical examples of evolutionary dynamics, along with revision protocols that generate them. For further discussion of these dynamics, see Sandholm (2006a,b).

3. Survival under the Pairwise Difference Dynamic

Berger and Hofbauer (2006) prove that strictly dominated strategies can survive under the BNN dynamic of Brown and von Neumann (1950). To prepare for our main result,

³For explicit accounts of microfoundations, see Benaïm and Weibull (2003) and Sandholm (2003).

| Revision protocol | Evolutionary dynamic | Name | Origin |
|---------------------------------|---|------------------------------|------------------------------|
| $\rho_{ij} = x_j[F_j - F_i]_+$ | $\dot{x}_i = x_i(F_i(x) - \bar{F}(x))$ | replicator | Taylor and Jonker (1978) |
| $\rho_{ij} = B_j^F(x)$ | $\dot{x} \in B^F(x) - x$ | best response | Gilboa and Matsui (1991) |
| $\rho_{ij} = [F_j - \bar{F}]_+$ | $\dot{x}_i = [F_i(x) - \bar{F}(x)]_+$ $-x_i \sum_{j \in S} [F_j(x) - \bar{F}(x)]_+$ | Brown-von Neumann-Nash (BNN) | Brown and von Neumann (1950) |
| $\rho_{ij} = [F_j - F_i]_+$ | $\dot{x}_i = \sum_{j \in S} x_j [F_i(x) - F_j(x)]_+$ $-x_i \sum_{j \in S} [F_j(x) - F_i(x)]_+$ | pairwise difference (PD) | Smith (1984) |

Table I

we adapt Berger and Hofbauer’s (2006) arguments to prove a survival result for the pairwise difference (PD) dynamic of Smith (1984). We also show how these analyses can be extended to other dynamics with similar functional forms.

We begin with our result for the PD dynamic.

Theorem 3.1. *There is a population game F in which under the PD dynamic, along solutions from most initial conditions, there is a strictly dominated strategy played by a fraction of the population that is bounded away from 0 and that exceeds $\frac{1}{6}$ infinitely often as time approaches infinity.*

The proof of Theorem 3.1 proceeds in three steps. First, we show that the PD dynamic converges to a limit cycle in the bad Rock-Paper-Scissors game (Figure 1). Second, we introduce a new strategy, Twin, which duplicates the strategy Scissors, and show that in the resulting four-strategy game, solutions to the PD dynamic from almost all initial conditions converge to a cycling attractor; this attractor sits on the plane where Scissors and Twin are played by equal numbers of agents, and has regions where both Scissors and Twin are played by more than one sixth of the population (Figure 2). Third, we uniformly reduce the payoff of the new strategy by ε , creating a “feeble twin”, and use a continuity argument to show that the attractor persists (Figure 3). Since the feeble twin is a strictly dominated strategy, this last step completes the proof of the theorem.

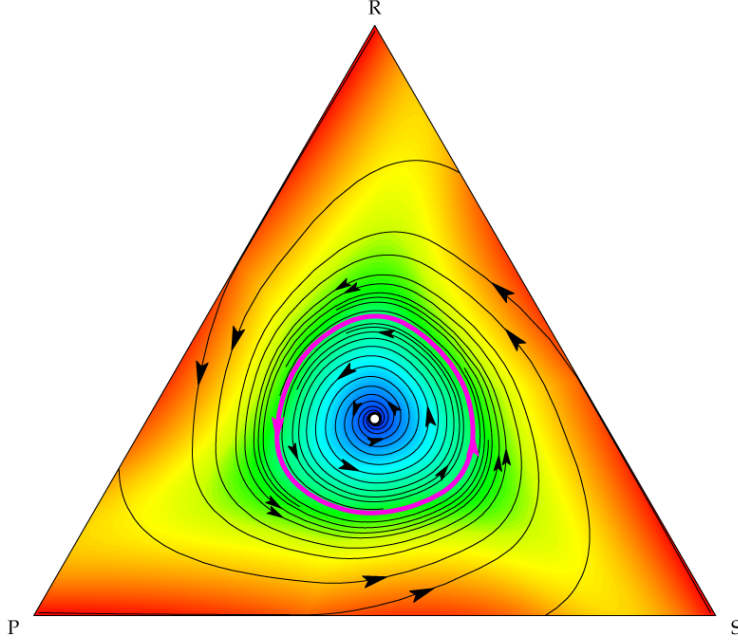


Figure 1: The PD dynamic in bad RPS.

Proof. To begin, we introduce the bad Rock-Paper-Scissors game:

$$G(x) = Ax = \begin{pmatrix} 0 & -b & a \\ a & 0 & -b \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \text{where } b > a > 0.$$

(Since $b > a$, the cost of losing a match exceeds the benefit of winning a match.) For any choice of $b > a > 0$, the unique Nash equilibrium of this game is $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. One can show that x^* is unstable under the PD dynamic. (We do not use this fact directly in our analysis, but it can be demonstrated by slightly modifying the proof of Lemma B.2.)

Next, following Berger and Hofbauer (2006), we introduce a four-strategy game F , which we obtain from bad RPS by introducing an “identical twin” of Scissors.

$$(1) \quad F(x) = \tilde{A}x = \begin{pmatrix} 0 & -b & a & a \\ a & 0 & -b & -b \\ -b & a & 0 & 0 \\ -b & a & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

The set of Nash equilibria of F is the line segment $NE = \{x^* \in X : x^* = (\frac{1}{3}, \frac{1}{3}, c, \frac{1}{3} - c)\}$.

Since Scissors and Twin always earn the same payoffs ($F_3(x) \equiv F_4(x)$), we can derive a

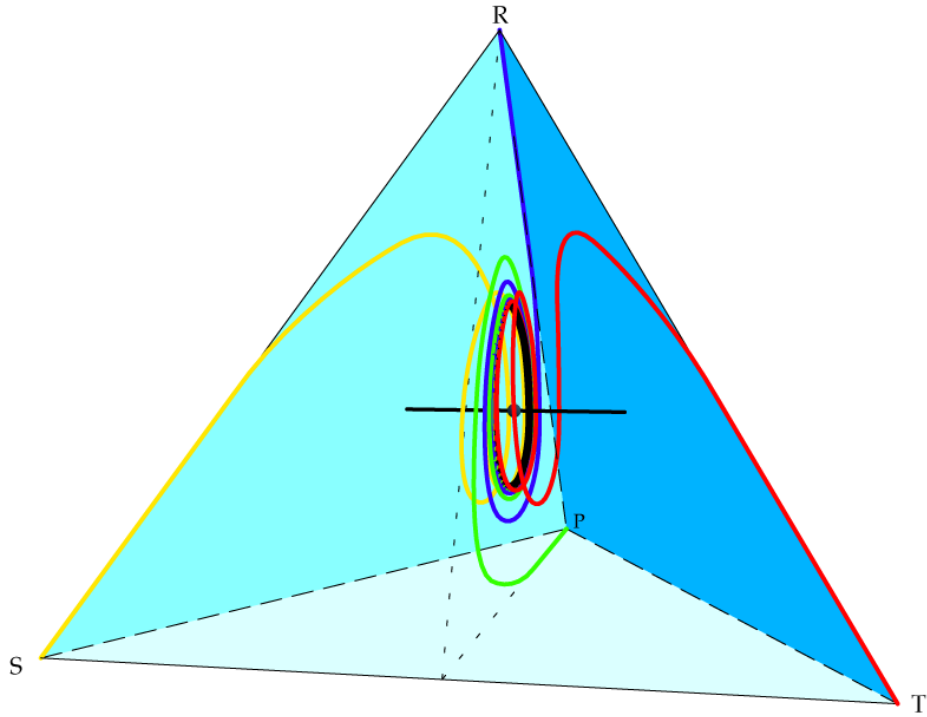


Figure 2: The PD dynamic in “bad RPS with a twin”.

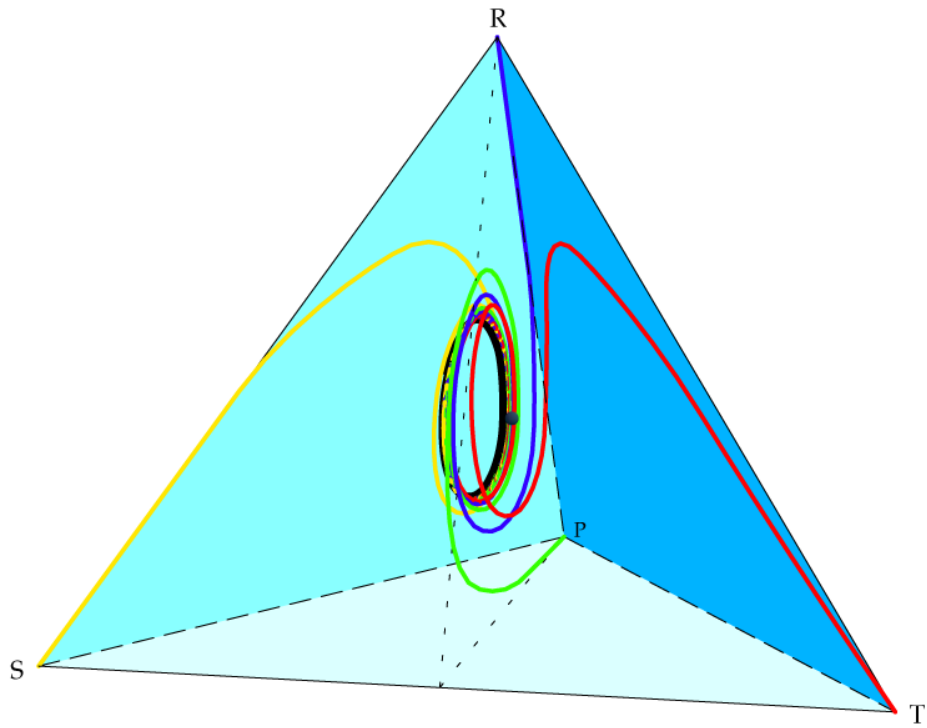


Figure 3: The PD dynamic in “bad RPS with a feeble twin”.

simple expression for the rate of change of the difference between their utilization levels:

$$\begin{aligned}
(2) \quad \dot{x}_3 - \dot{x}_4 &= \left(\sum_{j \in S} x_j [F_3(x) - F_j(x)]_+ - x_3 \sum_{j \in S} [F_j(x) - F_3(x)]_+ \right) \\
&\quad - \left(\sum_{j \in S} x_j [F_4(x) - F_j(x)]_+ - x_4 \sum_{j \in S} [F_j(x) - F_4(x)]_+ \right) \\
&= -(x_3 - x_4) \sum_{j \in S} [F_j(x) - F_4(x)]_+.
\end{aligned}$$

Thus, except at Nash equilibrium states, the dynamic moves toward the plane $P = \{x \in X : x_3 = x_4\}$ on which the identical twins receive equal weight.

Next, we characterize the behavior of the PD dynamic in a neighborhood of the set of Nash equilibria.

Lemma 3.2. *The set NE is a repeller under the PD dynamic: that is, there is a neighborhood U of NE such that all trajectories starting in U – NE leave U and never return.*

The proof of this lemma requires a long computation. We establish the lemma as an immediate corollary of a more general result, Lemma B.2 in Appendix B.

Together, equation (2) and Lemma 3.2 imply that all solution trajectories other than those starting in NE converge to an attractor \mathcal{A} . \mathcal{A} is compact (see Appendix A), contained in the invariant plane P , and disjoint from the set NE. Since the Nash equilibrium on plane P is state $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6})$, and since the attractor $\mathcal{A} \subset P$ circumnavigates x^* , there are portions of \mathcal{A} where more than one sixth of the population plays Twin.

Finally, we modify the game F by making Twin “feeble”: in other words, by uniformly reducing its payoff by ε :

$$F_\varepsilon(x) = \tilde{A}_\varepsilon x = \begin{pmatrix} 0 & -b & a & a \\ a & 0 & -b & -b \\ -b & a & 0 & 0 \\ -b - \varepsilon & a - \varepsilon & -\varepsilon & -\varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

If $\varepsilon > 0$, strategy 4 is strictly dominated by strategy 3.

As increasing ε from 0 continuously changes the game from F to F_ε , doing so also continuously changes the dynamic from V^F to V^{F_ε} . It thus follows from results on continuation of attractors (Theorem A.1 in Appendix A) that for small ε , the attractor \mathcal{A} of V^F continues to an attractor \mathcal{A}_ε that is contained in a neighborhood of \mathcal{A} , and that the basin of attraction of \mathcal{A}_ε contains all points outside of a thin tube around the set NE.

On the attractor \mathcal{A} , the speed of rotation under V^F around the segment NE is bounded away from 0. Therefore, by continuity, the attractor \mathcal{A}_ε of V^{F_ε} must encircle NE , and so must contain states at which x_4 , the weight on the strictly dominated strategy Twin, is more than $\frac{1}{6}$. By the same logic, solutions of V^{F_ε} that converge to \mathcal{A}_ε have ω -limit sets that encircle NE and contain states with $x_4 > \frac{1}{6}$. In conclusion, we have shown that most solutions of V^{F_ε} converge to the attractor \mathcal{A}_ε , a set on which x_4 is bounded away from 0, and that these solutions satisfy $x_4 > \frac{1}{6}$ infinitely often in the long run. This completes the proof of Theorem 3.1. ■

It is worth noting that the number $\frac{1}{6}$, the bound that the weight on the dominated strategy perpetually exceeds, is not as large as possible. By replacing A , a cyclically symmetric version of bad Rock-Paper-Scissors, with an asymmetric version of this game, we can move the unstable Nash equilibrium from $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to a state where the fraction of the population choosing Scissors is as close to 1 as desired (see Hofbauer and Sigmund (1998, p. 80)). Then repeating the rest of the proof above, we find that the bound of $\frac{1}{6}$ in the statement of Theorem 3.1 can be replaced by any number less than $\frac{1}{2}$.

The analysis of Berger and Hofbauer (2006) and our modification of this analysis make explicit use of the functional forms of the BNN and PD dynamics. This occurs first in the demonstration that the set of Nash equilibria of “bad RPS with a twin” is a repeller. The Lyapunov functions used to accomplish this depend on the two dynamics’ functional forms; even worse, there are evolutionary dynamics for which the equilibrium of bad RPS is attracting instead of repelling. Functional forms are also important in showing that almost all solutions to the two dynamics lead to the plane on which the identical twins receive equal weights. For arbitrary dynamics, particularly ones that do not respect the symmetry of the game, convergence to this plane is not guaranteed. To establish our main result, in which nothing is presumed about functional forms, both of these steps from the proof above will need to be replaced by more general arguments.

On the other hand, dynamics whose functional forms are similar enough to those of the BNN and PD dynamics can be handled by the foregoing analysis. Such dynamics, which are introduced by Hofbauer (2000) and Sandholm (2005, 2006a), are addressed in Theorem 3.3. To state this result, we let $\phi : \mathbf{R} \rightarrow \mathbf{R}_+$ be a Lipschitz continuous function satisfying

$$(3) \quad \text{sgn}(\phi(u)) = \text{sgn}([u]_+) \quad \text{and} \quad \frac{d}{du^+} \phi(u) \Big|_{u=0} > 0.$$

Theorem 3.3. *Suppose that V is an evolutionary dynamic based on a revision protocol of form*

$$(i) \quad \rho_{ij} = \phi(F_j - \bar{F}), \quad \text{or}$$

$$(ii) \quad \rho_{ij} = \phi(F_j - F_i),$$

where ϕ satisfies condition (3). Then there is a game F such that under V^F , along solutions from most initial conditions, there is a strictly dominated strategy played by a fraction of the population that is bounded away from 0 and that exceeds $\frac{1}{6}$ infinitely often as time approaches infinity.

The proof of Theorem 3.3 is provided in Appendix B.

4. The Main Theorem

4.1 Statement of the Theorem

The main result of this paper is Theorem 4.1.

Theorem 4.1. *Suppose the evolutionary dynamic V satisfies*

- | | | |
|------|----------------------|--|
| (C) | Continuity | g is Lipschitz continuous. |
| (PC) | Positive correlation | $V^F(x)'F(x) \geq 0$, with equality if and only if $x \in NE(F)$. |
| (IN) | Innovation | If $x \notin NE(F)$, $x_i = 0$, and $e_i \in B^F(x)$, then $V_i^F(x) > 0$. |

Then there is a game F such that under V^F , along solutions from most initial conditions, there is a strictly dominated strategy played by a fraction of the population bounded away from 0.

None of the conditions imposed above seems especially demanding. *Continuity* (C) requires that small changes in aggregate behavior or payoffs not lead to large changes in the law of motion $V^F(x) = g(F(x), x)$. Since discontinuous revision protocols can only be executed by agents with extremely accurate information, this condition seems natural in most contexts where evolutionary models are appropriate. Of course, this condition excludes the best response dynamic from our analysis, but it does not exclude continuous approximations thereof—see Section 5.2.3.

Positive correlation (PC) is a mild payoff monotonicity condition. It requires that whenever the population is not at a Nash equilibrium, there is a positive correlation between strategies' growth rates and payoffs. From a geometric point of view, condition (PC) requires that the directions of motion $V^F(x)$ and the payoff vectors $F(x)$ always form acute angles with one another; this interpretation will be helpful for understanding the constructions to come. Finally, *innovation* (IN) requires that when a non-Nash population state includes an unused optimal strategy, this strategy's growth rate must be strictly positive. In other words, if an unplayed strategy is sufficiently rewarding, some members of the population will discover it and select it.

A few further comments about conditions (PC) and (IN) may help with the interpretation of our results. First, condition (PC) is among the weakest monotonicity conditions proposed in the evolutionary literature.⁴ Thus, our arguments that appeal to this condition are robust, in that they apply to any dynamic that respects the payoffs from the underlying game to some weak extent. Second, since condition (PC) requires a positive correlation between growth rates and payoffs at all population states, it rules out evolutionary dynamics under which the boundary of the state space is repelling due to “mutations” or other forms of noise. Consequently, condition (PC) excludes the possibility that a dominated strategy survives for trivial reasons of this sort.

Third, conditions (PC) and (IN) both rule out dynamics based exclusively on imitation. At the same time, both of these conditions are satisfied by dynamics under which agents usually imitate, but occasionally evaluate strategies in a more direct fashion. See Section 5.2.1 for an explicit discussion of this point.

Before proceeding, we should point out that the conclusion of Theorem 4.1 is weaker than those of Theorems 3.1 and 3.3 in one notable respect: while the previous results ensured that at least $\frac{1}{6}$ of the population would play the dominated strategy infinitely often, Theorem 4.1 only ensures that the strategy is always used by a proportion of the population bounded away from 0. The reason for this weaker conclusion is the absence of any assumption that the dynamic V^F treats different strategies symmetrically. Adding such a symmetry assumption would allow us to recover the stronger conclusion offered by the earlier results. See Section 4.2.4 for further discussion.

4.2 Proof of the Theorem

As we noted earlier, the analyses in Section 3 took advantage of explicit functional forms for the dynamics at issue. Since Theorem 4.1 provides no such structure, its proof will require some new ideas.

Our first task is to construct a replacement for the bad RPS game. More precisely, we seek a three-strategy game in which dynamics satisfying condition (PC) will fail to converge to Nash equilibrium from almost all initial conditions. Our construction relies on the theory of potential games, developed in the normal form context by Monderer and Shapley (1996) and Hofbauer and Sigmund (1998) and in the population game context by Sandholm (2001, 2006c).

⁴Conditions similar to (PC) have been proposed, for example, in Friedman (1991), Swinkels (1993), and Sandholm (2001).

4.2.1 Potential Games

A population game F is a *potential game* if there exists a continuously differentiable function $f : \mathbf{R}_+^n \rightarrow \mathbf{R}$ satisfying

$$\nabla f(x) = F(x) \text{ for all } x \in X.$$

Put differently, each strategy's payoff function must equal the appropriate partial derivative of the *potential function*:

$$\frac{\partial f}{\partial x_i}(x) = F_i(x) \text{ for all } i \in S \text{ and } x \in X.$$

Games satisfying this condition include common interest games and congestion games, among many others. A basic fact about potential games is that reasonable evolutionary dynamics increase potential: if the dynamic V^F satisfies condition (PC), then along each solution trajectory $\{x_t\}$ we have that

$$\frac{d}{dt}f(x_t) = \nabla f(x_t)' \dot{x}_t = F(x_t)' V^F(x_t) \geq 0,$$

with equality only at Nash equilibria. This observation along with standard results from dynamical systems imply that each solution trajectory of V^F converges to a connected set of Nash equilibria—see Sandholm (2001).

As an example, suppose that agents are randomly matched to play the pure coordination game

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The resulting population game, $F^C(x) = Cx = x$, is a potential game; its potential function, $f^C(x) = \frac{1}{2}x'Cx = \frac{1}{2}((x_1)^2 + (x_2)^2 + (x_3)^2)$, is the convex function pictured in Figure 4(i). Solutions to any evolutionary dynamic satisfying condition (PC) ascend this function. Indeed, solutions from almost all initial conditions converge to a vertex of X —that is, to a strict equilibrium of F^C .

The ability to draw the game F^C itself will prove useful in the analysis to come. Notice that F^C is a map from the simplex $X \subset \mathbf{R}^3$ to \mathbf{R}^3 , and so can be viewed as a vector field. Rather than draw F^C as a vector field in \mathbf{R}^3 , we draw a projected version of F^C on the

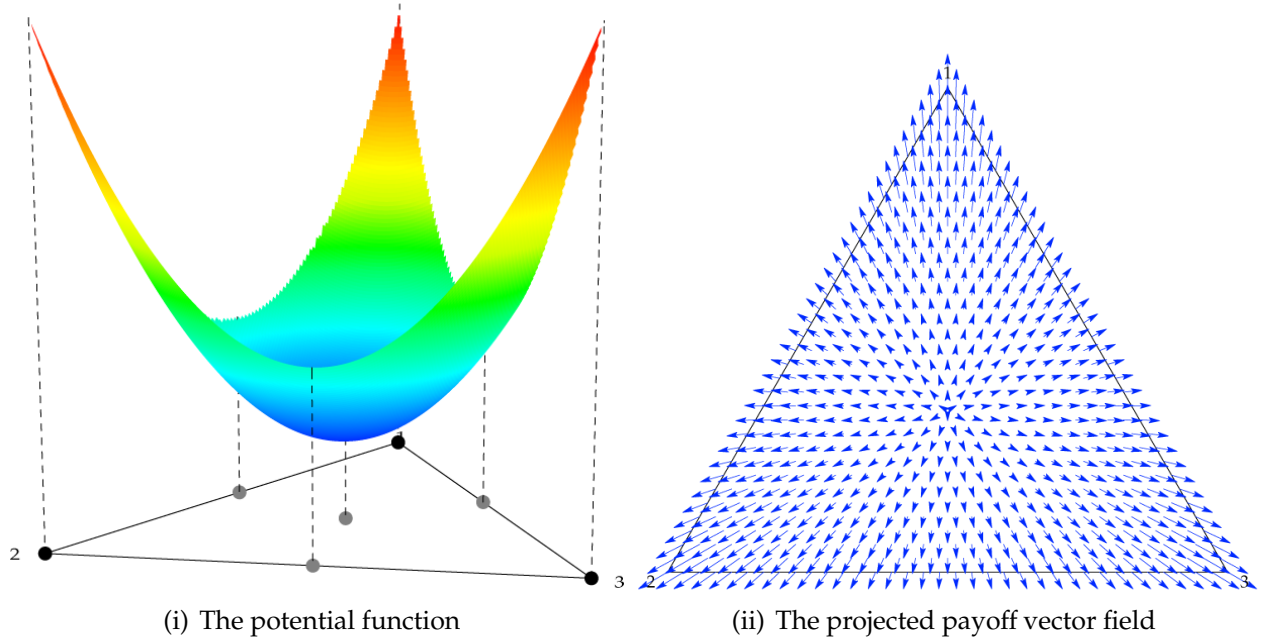


Figure 4: A coordination game.

hyperplane in \mathbf{R}^3 that contains the simplex.⁵ The vectors drawn in Figure 4(ii) represent the directions of maximal increase of the function f^C , and so point outward from the center of the simplex. Dynamics satisfying condition (PC) always travel at acute angles to the vectors in Figure 4(ii), and so tend toward the vertices of X , and solutions from almost all initial conditions converge to a vertex of X

As a second example, suppose that agents are randomly matched to play the anticoordination game C . In Figures 5(i) and 5(ii), we draw the resulting population game $F^{-C}(x) = -Cx = -x$ and its concave potential function $f^{-C}(x) = -\frac{1}{2}x'Cx = -\frac{1}{2}((x_1)^2 + (x_2)^2 + (x_3)^2)$. Both pictures reveal that under any evolutionary dynamic satisfying condition (PC), all solution trajectories converge to the unique Nash equilibrium $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

4.2.2 The Hypnodisk Game

The construction of our replacement for bad RPS is easiest to describe in geometric terms. Begin with the coordination game $F^C(x) = Cx$ pictured in Figure 4(ii). Then draw two circles centered at state $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ with radii $0 < r < R < \frac{1}{\sqrt{6}}$, as shown in Figure 6(i); the second inequality ensures that both circles are contained in the simplex. Finally, twist

⁵More precisely, we draw the vector field ΦF^C , where $\Phi = I - \frac{1}{3}\mathbf{1}\mathbf{1}' \in \mathbf{R}^{3 \times 3}$ is the orthogonal projection of \mathbf{R}^3 onto $TX = \{z \in \mathbf{R}^3 : \sum_{i \in S} z_i = 0\}$, the tangent space of the simplex X . The projection Φ forces the components of $\Phi F^C(x)$ to sum to zero while preserving their differences, so that $\Phi F^C(x)$ preserves all information about incentives contained in payoff vector $F^C(x)$. See Sandholm et al. (2006) for further discussion.

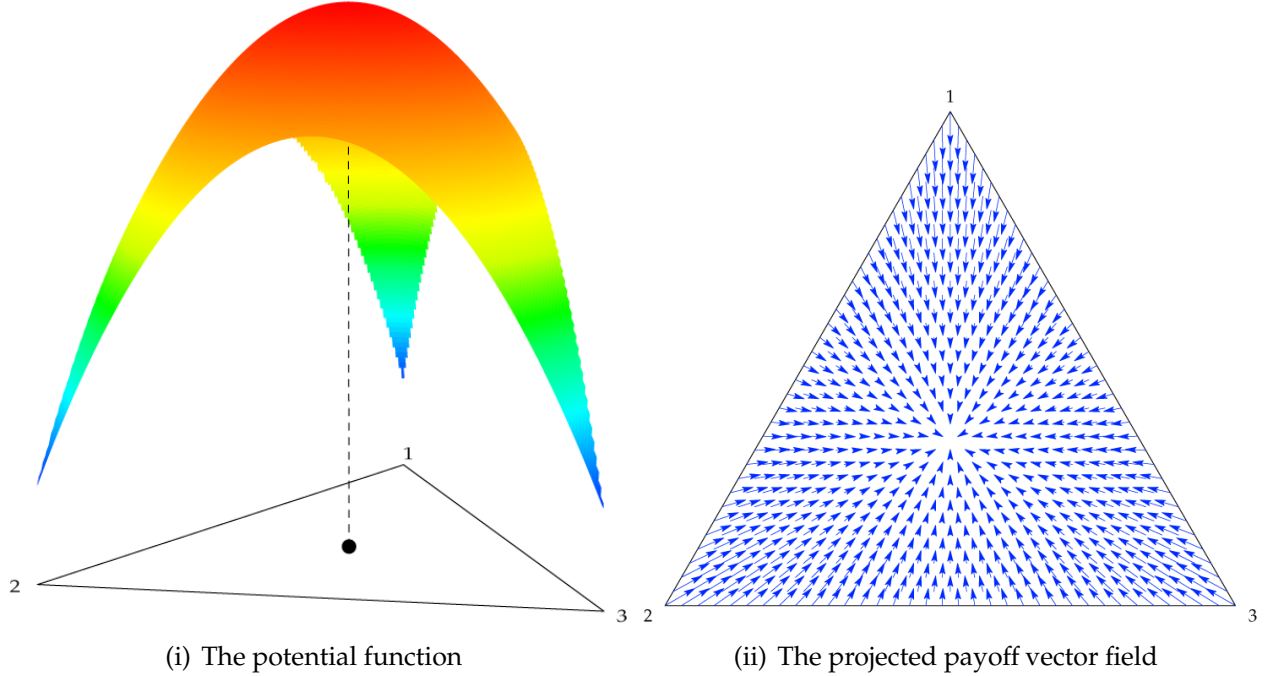


Figure 5: An anticoordination game.

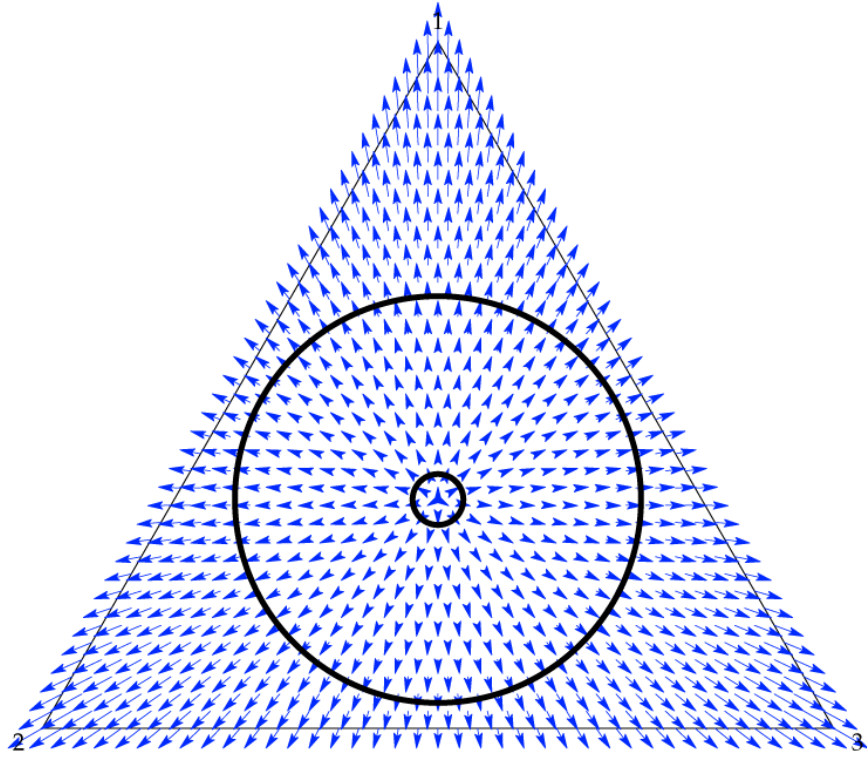
the portion of the vector field lying outside of the inner circle in a clockwise direction, excluding larger and larger circles as the twisting proceeds, so that the outer circle is reached when the total twist is 180° . The resulting vector field is pictured in Figure 6(ii). It is described analytically by

$$H(x) = \cos(\theta(x)) \begin{pmatrix} x_1 - \frac{1}{3} \\ x_2 - \frac{1}{3} \\ x_3 - \frac{1}{3} \end{pmatrix} + \frac{\sqrt{3}}{3} \sin(\theta(x)) \begin{pmatrix} x_2 - x_3 \\ x_3 - x_1 \\ x_1 - x_2 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

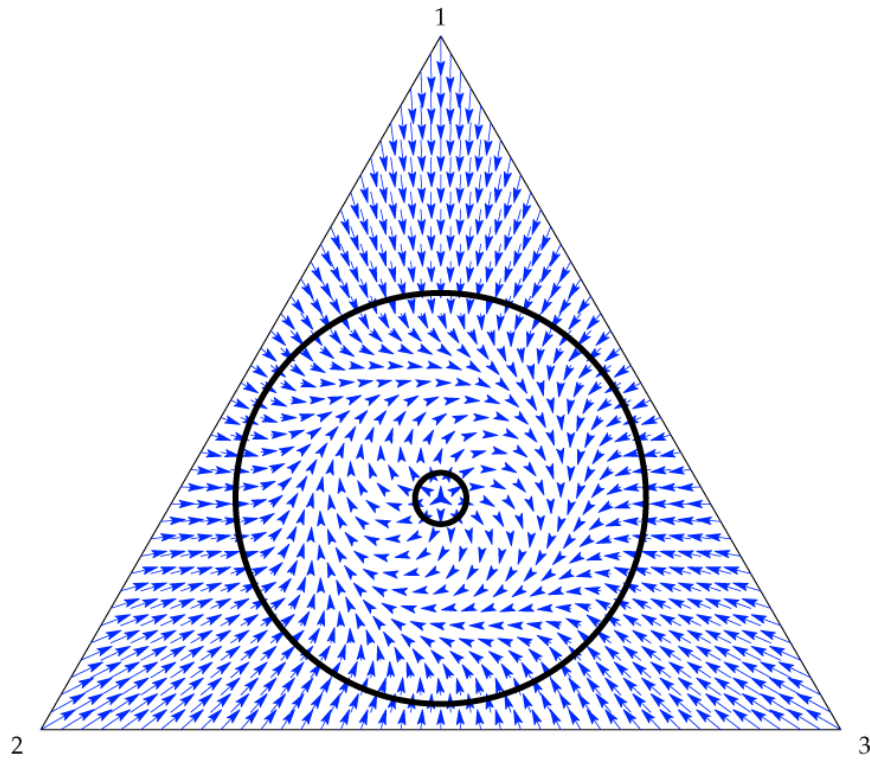
where $\theta(x)$ equals 0 when $|x - x^*| \leq r$, equals π when $|x - x^*| \geq R$, and varies linearly in between. We call the game H the *hypnodisk game*.

What does this construction accomplish? Inside the inner circle, H is identical to the coordination game F^C . Thus, solutions to dynamics satisfying (PC) starting at states in the inner circle besides x^* must leave the inner circle. At states outside the outer circle, the drawing of H is identical to the drawing of the anticoordination game F^{-C} .⁶ Therefore, solutions to dynamics satisfying (PC) that begin outside the outer circle must enter the outer circle. Finally, at each state x in the annulus bounded by the two circles, $H(x)$ is not a componentwise constant vector. Therefore, states in the annulus are not Nash equilibria,

⁶At states x outside the outer circle, $H(x) = -x + \frac{2}{3}\mathbf{1} \neq -x = F^{-C}(x)$. But since $\Phi H(x) = -x + \frac{1}{3}\mathbf{1} = \Phi F^{-C}(x)$ at these states, the pictures of H and F^{-C} , and hence the incentives in the two games, are the same.



(i) Projected payoff vector field for the coordination game



(ii) Projected payoff vector field for the hypnodisk game

Figure 6: Construction of the hypnodisk game.

and so are not rest points of dynamics satisfying (PC). We assemble these observations in the following lemma.

Lemma 4.2. *Suppose that V is an evolutionary dynamic that satisfies conditions (C) and (PC), and let H be the hypnodisk game. Then every solution to V^H other than the stationary solution at x^* enters the annulus with radii r and R and never leaves.*

In fact, since there are no rest points in the annulus, the Poincaré-Bendixson Theorem implies that every nonstationary solution to V^H converges to a limit cycle.

4.2.3 The Twin

Now, let F be the four-strategy game obtained from H by adding a twin: $F_i(x_1, x_2, x_3, x_4) = H_i(x_1, x_2, x_3 + x_4)$ for $i \in \{1, 2, 3\}$, and $F_4(x) = F_3(x)$. The set of Nash equilibria of F is the line segment

$$NE = \left\{ x^* \in X : x_1^* = x_2^* = x_3^* + x_4^* = \frac{1}{3} \right\}.$$

Let

$$I = \left\{ x \in X : (x_1 - \frac{1}{3})^2 + (x_2 - \frac{1}{3})^2 + (x_3 + x_4 - \frac{1}{3})^2 \leq r^2 \right\} \text{ and}$$

$$O = \left\{ x \in X : (x_1 - \frac{1}{3})^2 + (x_2 - \frac{1}{3})^2 + (x_3 + x_4 - \frac{1}{3})^2 \leq R^2 \right\}$$

be concentric cylindrical regions in X surrounding NE , as pictured in Figure 7. By construction, we have that

$$F(x) = \tilde{C}x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

at all $x \in I$. Therefore, solutions to dynamics satisfying (PC) starting in $I - NE$ ascend the potential function $f^{\tilde{C}}(x) = \frac{1}{2}((x_1)^2 + (x_2)^2 + (x_3 + x_4)^2)$ until leaving the set I . At states outside the set O , we have that $F(x) = -\tilde{C}x$, so solutions starting in $X - O$ ascend $f^{-\tilde{C}}(x) = -f^{\tilde{C}}(x)$ until entering O . In summary:

Lemma 4.3. *Suppose that V is an evolutionary dynamic that satisfies conditions (C) and (PC), and let F be the “hypnodisk with a twin” game. Then every solution to V^F other than the stationary solutions at states in NE enter region $D = O - I$ and never leave.*

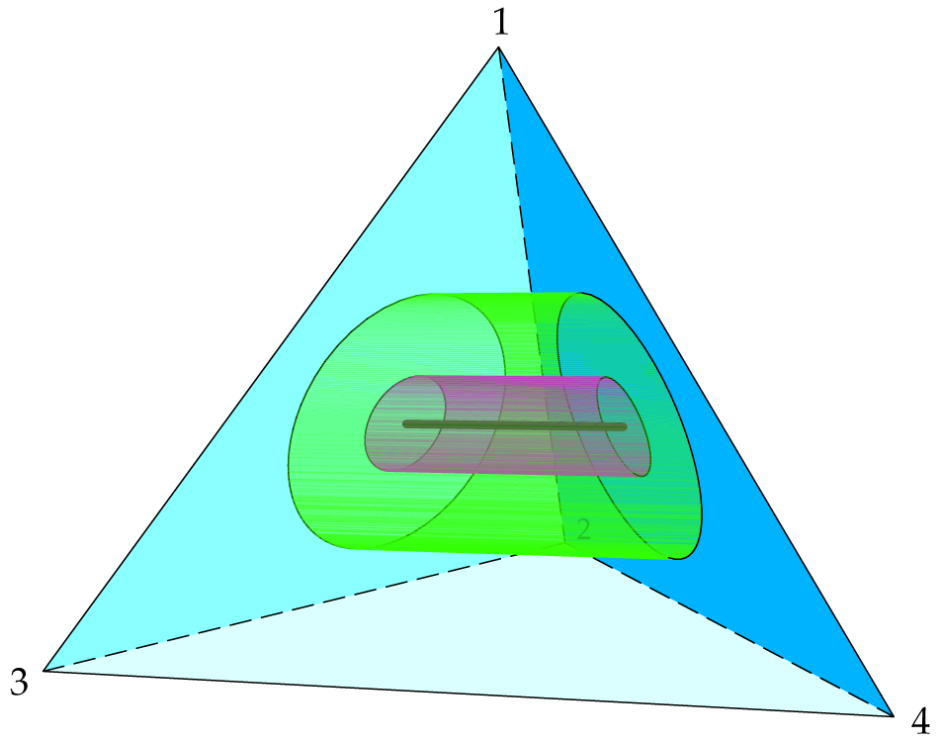


Figure 7: Regions O , I , and $D = O - I$.

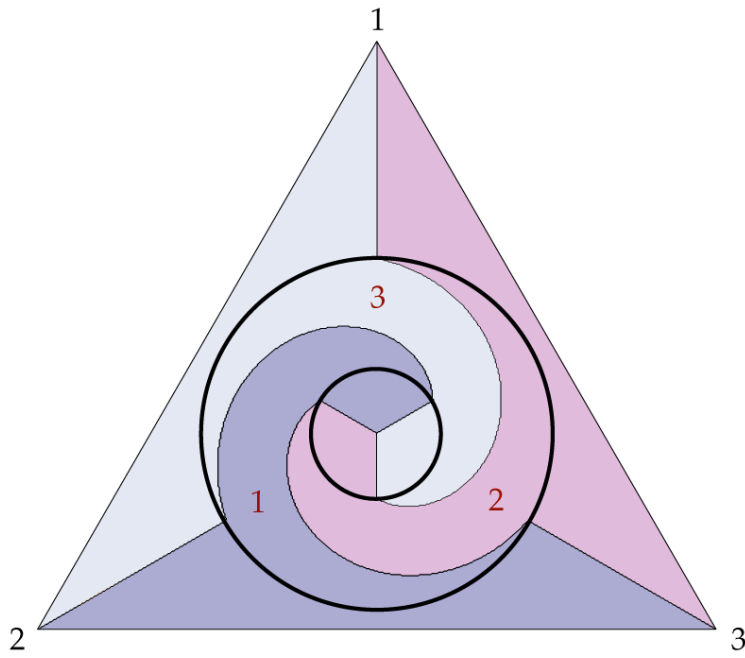


Figure 8: The best response correspondence of the hypnodisk game.

4.2.4 The Feeble Twin

In equation (2) of Section 3, we saw that nonstationary solutions of the PD dynamic equalize the utilization levels of twin strategies. If we presently focus on dynamics that not only satisfy conditions (C), (PC), and (IN), but also treat different strategies symmetrically, equation (2) shows that in the “hypnodisk with a twin” game F , all nonstationary solutions of V^F converge not only to region D , but also to the plane $P = \{x \in X : x_3 = x_4\}$. Continuing with the argument from Section 3 then allows us to conclude that in F_ε , the game obtained from F by turning strategy 4 into a feeble twin (that is, by reducing the payoff to strategy 4 uniformly by $\varepsilon > 0$), the fraction x_4 playing the feeble twin exceeds $\frac{1}{6}$ infinitely often.

Since we seek a result that imposes as little structure on permissible evolutionary dynamics, Theorem 4.1 avoids the assumption that different strategies are treated symmetrically. Since this means that agents may well be biased against choosing the dominated strategy, the conclusion of Theorem 4.1 is necessarily weaker than those of Theorems 3.1 and 3.3. Nevertheless, even without symmetry assumptions, we can prove that dominated strategies persist. To accomplish this, it is enough to show that in game F , most solutions of the dynamic V^F converge to a set on which x_4 is bounded away from 0. If we can do this, then repeating the continuity argument that concluded the proof of Theorem 3.1 shows that in the game F_ε , the dominated strategy 4 survives.

A complete proof that most solutions of V^F converge to a set on which x_4 bounded away from 0 is presented in Appendix C. We summarize the argument here. To begin, it can be shown that all solutions to V^F starting outside a small neighborhood of the segment of Nash equilibria NE converge to an attractor \mathcal{A} , a compact set that is contained in region D and that is an invariant set of the dynamic V^F .

Now suppose by way of contradiction that the attractor \mathcal{A} intersects $Z = \{x \in X : x_4 = 0\}$, the face of X on which Twin is unused. The Lipschitz continuity of the dynamic V^F implies that Z is backward invariant under V^F . Since \mathcal{A} is itself invariant, the fact that \mathcal{A} intersects Z implies the existence of a closed orbit $\gamma \subset \mathcal{A} \cap Z$ that circumnavigates the disk $I \cap Z$. Examining the best response correspondence of the hypnodisk game (Figure 8), we find that such an orbit γ must pass through a region in which strategy 3 is a best response. But since the twin strategy 4 is also a best response in this region, innovation (IN) tells us that solutions passing through this region must reenter the interior of X , contradicting that the attractor \mathcal{A} intersects the face Z .

5. Discussion

5.1 Constructing Games in which Dominated Strategies Survive

Consider an evolutionary dynamic that satisfies condition (PC). A dominated strategy can survive along a given solution trajectory of this dynamic only if that trajectory does not converge: if the trajectory did converge, the limit point would be a rest point, and hence a Nash equilibrium. Therefore, to construct games in which dominated strategies can survive, one should first look for games with poor convergence properties.

The hypnodisk game, the starting point for the proof of the main theorem, is a population game with nonlinear payoff functions. Such games were uncommon in the early literature on evolution in games, which focused on random matching settings. But population games with nonlinear payoffs are more common now, in part because of their appearance in applications. For example, the standard model of driver behavior in a highway network is a congestion game with nonlinear payoff functions.⁷ The nonlinearity reflects the fact that delays on each network link are increasing, convex functions of the number of drivers using the link. For this reason, we do not view the use of a game with nonlinear payoffs as a shortcoming of our analysis. But despite this, it seems worth asking whether our results could be proved within the linear, random matching framework.

In Section 3, where we considered dynamics with prespecified functional forms, we were able to prove survival results within the linear setting. More generally, if we fix an evolutionary dynamic before seeking a population game, finding a linear game that exhibits cycling seems a feasible task. Still, a virtue of our analysis in Section 4 is that it avoids this case-by-case analysis: the hypnodisk game generates cycling under all of the relevant dynamics simultaneously, enabling us to prove survival of dominated strategies under all of these dynamics at once.

Could we have done the same using linear payoffs? Consider the following game due to Hofbauer and Swinkels (1996) (see also (Hofbauer and Sigmund, 1998, Sec 8.6)):

$$F^\delta(x) = A^\delta x = \begin{pmatrix} 0 & 0 & -1 & \delta \\ \delta & 0 & 0 & -1 \\ -1 & \delta & 0 & 0 \\ 0 & -1 & \delta & 0 \end{pmatrix} x.$$

When $\delta = 0$, the game F^0 is a potential game with potential function $f(x) = -(x_1x_3 + x_2x_4)$.

⁷Congestion games with a continuum of agents are studied by Beckmann et al. (1956) and Sandholm (2001); for finite player congestion games, see Rosenthal (1973) and Monderer and Shapley (1996).

It has two components of Nash equilibria: one is a singleton containing the completely mixed equilibrium $x^* = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$; the other is the closed curve γ containing edges $\overline{e_1e_2}$, $\overline{e_2e_3}$, $\overline{e_3e_4}$, and $\overline{e_4e_1}$. The former component is a saddle point of f , and so is unstable under dynamics that satisfy (PC); the latter component is the maximizer set of f , and so attracts most solutions of these dynamics.

If δ is positive but sufficiently small, Theorem A.1 implies that most solutions of dynamics satisfying (PC) lead to an attractor near γ . But once δ is positive, the unique Nash equilibrium of F^δ is the mixed equilibrium x^* . Therefore, the attractor near γ is far from any Nash equilibrium.

If we now introduce a feeble twin, it seems natural to expect that this dominated strategy would survive in the resulting five-strategy game. But in this case, evolutionary dynamics run on a four-dimensional state space. Proving survival results when the dimension of the state space exceeds three is very difficult, even if we fix the dynamic under consideration in advance. This points to another advantage of the hypnodisk game: it allows us to work with dynamics on a three-dimensional state space, where the analysis is still tractable.

5.2 The Fragility of Elimination Results

Results on the elimination of dominated strategies exist in two important cases: imitative dynamics with monotone percentage growth rates, and the best response dynamic. We now explain why these results fail to be robust to small changes in the dynamics.

5.2.1 Input-Output Symmetry and Imitative Dynamics

In Section 2.2, we introduced a model of evolutionary dynamics based on revision protocols $\rho : \mathbf{R}^n \times X \rightarrow \mathbf{R}_+^{n \times n}$. To interpret ρ , we imagine that agents are selected from the population at random and offered opportunities to switch strategies. An i player who receives such an opportunity switches to strategy $j \neq i$ with probability proportional to the conditional switch rate $\rho_{ij} = \rho_{ij}(\pi, x)$. In aggregate, this procedure generates the law of motion

$$\dot{x} = \sum_{j \in S} x_j \rho_{ji} - x_i \sum_{j \in S} \rho_{ij}.$$

Notice that the outflow term in this equation, $-x_i \sum_{j \in S} \rho_{ij}$, is proportional to the utilization of strategy i , while the inflow term, $\sum_{j \in S} x_j \rho_{ji}$, is not. This asymmetry follows directly from the random assignment of revision opportunities: the rate of switches from

strategy i must be proportional to the number of agents currently playing strategy i , but the rate of switches to strategy i is not. It turns out that this asymmetry is crucial to our elimination results. We now argue that if the revision protocol ρ abrogates this asymmetry, elimination results are closer at hand.

Suppose that ρ is an imitative revision protocol: that is, a protocol of the form

$$\rho_{ij}(\pi, x) = x_j r_{ij}(\pi, x).$$

The x_j term reflects the fact that when an agent receives a revision opportunity, he selects an opponent at random, and then decides whether or not to imitate this opponent's strategy. Such protocols lead to aggregate dynamics of the form

$$\begin{aligned} \dot{x}_i &= \sum_{j \in S} x_j \rho_{ji} - x_i \sum_{j \in S} \rho_{ij} \\ &= \sum_{j \in S} x_j x_i r_{ji} - x_i \sum_{j \in S} x_j r_{ij} \\ &= x_i \sum_{j \in S} x_j (r_{ji} - r_{ij}). \end{aligned}$$

In other words, each strategy's growth rate is proportional to its level of utilization:

$$(4) \quad \dot{x}_i = x_i p_i(\pi, x), \quad \text{where } p_i(\pi, x) = \sum_{j \in S} x_j (r_{ji}(\pi, x) - r_{ij}(\pi, x))$$

is the percentage growth rate of strategy i .

To ensure that agents' choices respect incentives, one can assume, following Samuelson and Zhang (1992), that *percentage growth rates are monotone*:

$$p_i(\pi, x) \geq p_j(\pi, x) \text{ if and only if } \pi_i \geq \pi_j.$$

To see the consequences of these assumptions for elimination of dominated strategies, apply the quotient rule to equation (4) to show that

$$(5) \quad \frac{d}{dt} \left(\frac{x_i}{x_j} \right) = \frac{x_i}{x_j} (p_i - p_j).$$

If strategy i strictly dominates strategy j , then the right hand side of equation (5) is positive at all $x \in \text{int}(X)$. Therefore, the dominated strategy j must vanish along every solution in $\text{int}(X)$. This is Samuelson and Zhang's (1992) result.

If instead strategies i and j are twins, the right hand side of equation (5) is always

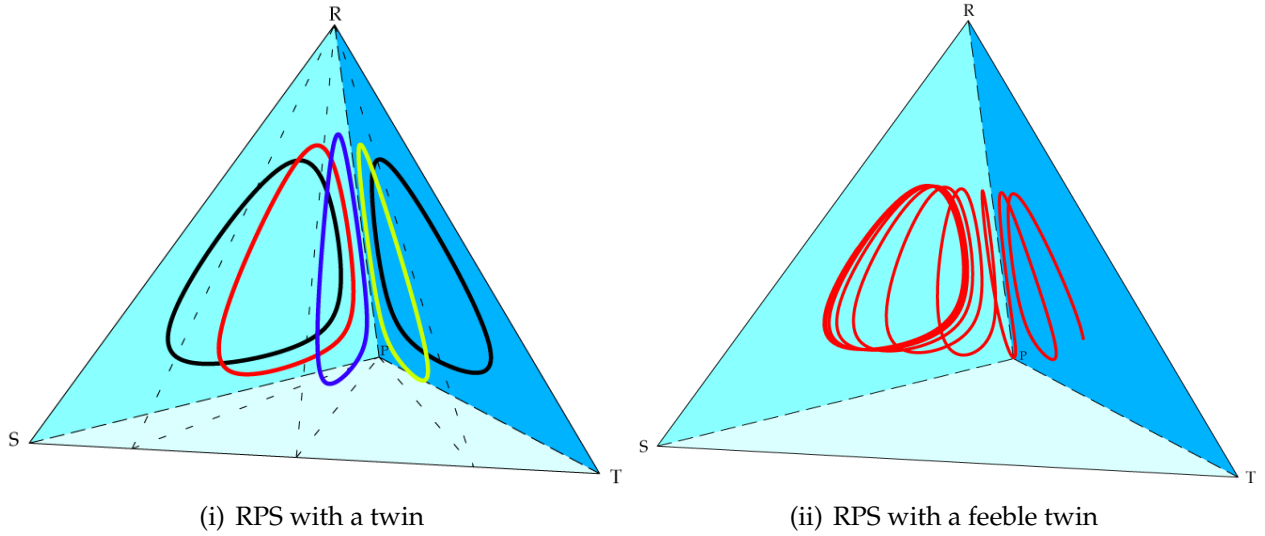


Figure 9: The replicator dynamic in two games.

zero, so the ratio $\frac{x_i}{x_j}$ is constant along every solution trajectory. This observation underlies the fragility of elimination results for imitative dynamics. Consider Figure 9(i), which presents the phase diagram for the replicator dynamic in “(standard) RPS with a twin”. Evidently, the planes on which the ratio $\frac{x_S}{x_T}$ is constant are all invariant sets. If we make the twin feeble by lowering its payoff uniformly by ε , we obtain the dynamics pictured in Figure 9(ii): now the ratio $\frac{x_S}{x_T}$ increases monotonically, and the dominated strategy is eliminated.

The existence of a continuum of invariant hyperplanes in games with identical twins is crucial to this argument. At the same time, dynamics with a continuum of invariant hyperplanes are structurally unstable. If we fix the game but slightly alter the agents’ revision protocol, these invariant sets can collapse, overturning the elimination result.

To make this argument concrete, suppose that instead of always following an imitative revision protocol, agents occasionally use a protocol that requires direct evaluation of payoffs. Such a situation is illustrated in Figure 10(i), which contains the phase diagram for a “bad RPS with a twin” game under a convex combination of the replicator and PD dynamics.⁸ While Figure 9(i) displayed a continuum of invariant hyperplanes, Figure 10(i) shows almost all solution trajectories converging to a limit cycle on the plane where $x_S = x_T$. If we then make the twin feeble, the limit cycle moves slightly to the left, and the dominated strategy survives (Figure 10(ii)).

⁸In particular, we consider the bad RPS game with payoffs 1, 0, and $-\frac{11}{10}$, and the combined dynamic that puts weight $\frac{9}{10}$ on the replicator dynamic and weight $\frac{1}{10}$ on the PD dynamic. This dynamic is generated by the corresponding convex combination of the underlying revision protocols: $\rho_{ij} = \frac{9}{10}x_j[F_j - F_i]_+ + \frac{1}{10}[F_j - F_i]_+$.

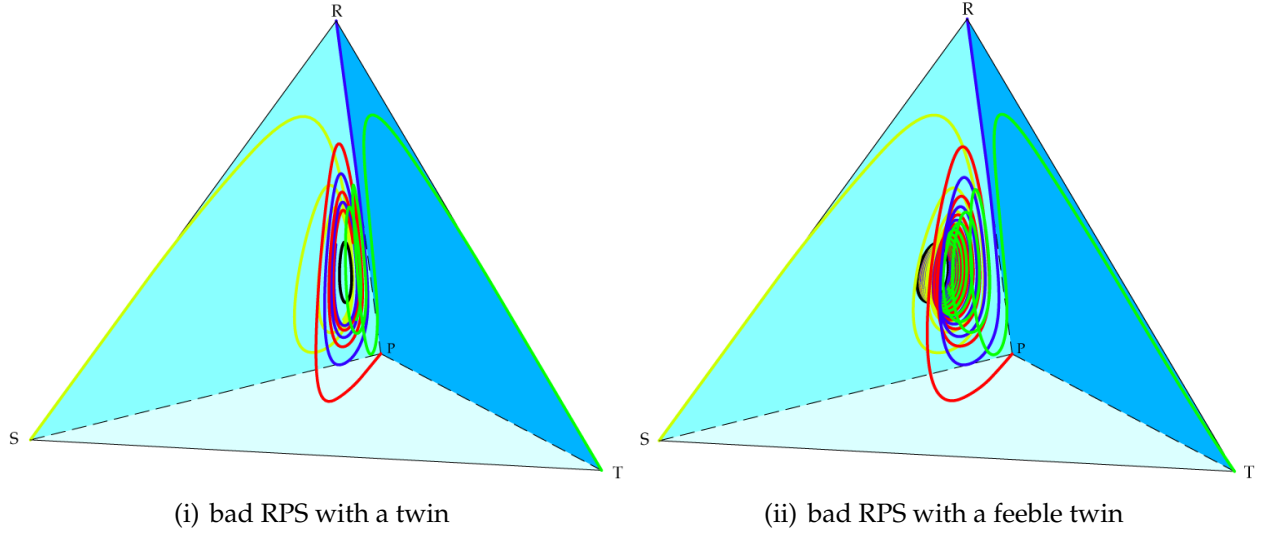


Figure 10: The $\frac{9}{10}$ replicator + $\frac{1}{10}$ PD dynamic in two games.

5.2.2 Input-Output Symmetry and the Projection Dynamic

Imitation is not the only way to create input-output symmetry at interior population states. Instead of supposing that the conditional switch rate from strategy i to strategy j is proportional to x_j , one can suppose instead that it is *inversely* proportional to x_i :

$$\rho_{ij}(\pi, x) = \frac{1}{x_i} q_{ij}(\pi, x).$$

Revision protocols of this form capture the notion that unpopular strategies are more likely to be abandoned. If, for example, we assume that

$$q_{ij}(\pi, x) = \pi_j - \frac{1}{n} \sum_{k \in S} \pi_k,$$

we obtain the following aggregate dynamic on $\text{int}(X)$:

$$\begin{aligned}
 (6) \quad \dot{x}_i &= \sum_{j \in S} x_j \rho_{ji} - x_i \sum_{j \in S} \rho_{ij} \\
 &= \sum_{j \in S} x_j \cdot \frac{1}{x_j} q_{ji} - x_i \sum_{j \in S} \frac{1}{x_i} q_{ij} \\
 &= \sum_{j \in S} (q_{ji} - q_{ij})
 \end{aligned}$$

$$= F_i(x) - \frac{1}{n} \sum_{k \in S} F_k(x).$$

This equation defines the *projection dynamic* on $\text{int}(X)$ (Nagurney and Zhang (1997), Sandholm et al. (2006)). If strategies i and j are twins, the difference $x_i - x_j$ is constant under the dynamic (6) on $\text{int}(X)$. If we then make twin j feeble, this difference must fall over time, suggesting an elimination result. However, equation (6) alone does not define a legitimate evolutionary dynamic on all of X , since it can assign negative growth rates to unused strategies. It is possible to alter how the dynamic is defined on $\text{bd}(X)$ so as to ensure not only forward invariance of X , but other desirable properties as well: for instance, existence and uniqueness of solutions and property (PC). Unfortunately, this discontinuous alteration of the law of motion allows dominated strategies to survive. See Sandholm et al. (2006) for additional details.

5.2.3 The Best Response Dynamic

Dominated strategies are also eliminated under the *best response dynamic*:

$$\dot{x} \in B^F(x) - x,$$

where $B^F(x)$ is the set of mixed best responses to population state x . Like those for imitative dynamics, this elimination result is also fragile. Suppose that strategies i and j are twins. At states where these strategies are not optimal, the best response dynamic reduces the mass on each strategy at a rate proportional to its current level of use. But when these strategies are optimal, multiple directions of motion are possible: the dynamic can move toward vertex e_i , toward vertex e_j , or toward any state in between. If we make twin j feeble, the law of motion changes dramatically: now, the target states in $B^F(x)$ must always lie on the face where j is unused, leading this strategy to be eliminated.

Suppose we replace the best response dynamic with a continuous approximation thereof. This is often done using the *logit dynamic* with a small but positive *noise level* η (Fudenberg and Levine (1998)):

$$(L) \quad \dot{x}_i = \frac{\exp(\eta^{-1}F_i(x))}{\sum_{k \in S} \exp(\eta^{-1}F_k(x))} - x_i$$

This approximation is close at states where one strategy is distinctly better than the rest, but it is necessarily quite poor when the best response is not unique. Indeed, if strategies i and j are twins, then the difference $x_i - x_j$ converges to zero along every solution trajectory. So while the best response dynamic eliminates strictly dominated strategies, one can show

that under the logit dynamic with any small noise level η , in “bad RPS with a feeble twin”, the strictly dominated strategy Twin is played by a significant proportion of the population infinitely often. This claim is formalized in the following analogue of Theorem 3.1, which we state without proof.⁹

Theorem 5.1. *Let F^ε be “bad RPS with a feeble twin”, and let (L) be the logit(η) dynamic for F^ε , where $\eta \in (0, \frac{1}{10})$. For small enough $\varepsilon > 0$, along solutions to (L) from most initial conditions, there is a strictly dominated strategy played by a fraction of the population that exceeds $\frac{1}{6}$ infinitely often as time approaches infinity.*

We should emphasize that the order in which the parameters ε and η are chosen is important here. The theorem fixes the logit noise level η , and states if the payoff disadvantage ε of Twin is small enough, then Twin will be played recurrently by a significant fraction of the population. On the other hand, suppose we fix the payoff disadvantage $\varepsilon > 0$ first. Then if the logit noise level η is made small enough (and hence the Lipschitz coefficient of (L) large enough), the fraction of the population choosing Twin will stay very small.

6. Conclusion

This paper demonstrates that the elimination of dominated strategies, even those that are strictly dominated by a pure strategy, is an uncommon and nonrobust feature of evolutionary dynamics. This discrepancy between traditional and evolutionary predictions can be traced back to qualitative differences in how traditional and evolutionary agents make decisions. Choices that respect dominance are based on global comparisons: to know that one strategy dominates another, one must be able to compare their payoffs at all population states. In contrast, decision making in evolutionary models is local in nature: agents examine the current payoffs of the available strategies, and select strategies whose payoffs are reasonably high at the moment of choice. There is no obvious reason why a choice criterion based on global comparisons should be respected by dynamics built upon local comparisons. This paper can be viewed as an expatiation on this simple idea.

⁹That $\eta < \frac{1}{10}$ is enough to ensure the existence of a cycle in bad RPS—see Sandholm (2007).

Appendix

A. Continuation of Attractors

Let X be a compact metric space and let ϕ be a semi-flow on X ; thus, $\phi : [0, \infty) \times X \rightarrow X$ is a continuous map satisfying $\phi^0(x) = x$ and $\phi^t(\phi^s(x)) = \phi^{t+s}(x)$ for all $s, t \geq 0$ and $x \in X$. A set $A \subset X$ is an *attractor* of ϕ if there is a neighborhood U of A such that $\omega(U) = A$ (see Conley (1978)). Here the ω -limit set of U is defined as $\omega(U) = \bigcap_{t>0} \text{cl}(\phi^{[t, \infty)}(U))$, where for $T \subset \mathbf{R}$ we let $\phi^T(U) = \bigcup_{t \in T} \phi^t(U)$. An attractor is compact and *invariant* ($\phi^t(A) = A$ for all t). Observe that an attractor can strictly contain another attractor.

The *basin* of the attractor is defined as $B(A) = \{x : \omega(x) \subseteq A\}$. For each open set U with $A \subset U \subset \bar{U} \subset B(A)$ we have $\omega(\bar{U}) = A$, see Section II.5.1.A of Conley (1978). Furthermore, if $\phi^t(\bar{U}) \subset U$ holds for some $t > 0$ and for some open set U (which is then called a *trapping region*), then $\omega(U)$ is an attractor, see Section II.5.1.C of Conley (1978).

For a flow $(\phi^t, t \in \mathbf{R})$, the complement of the basin $B(A)$ of the attractor A is called the *dual repeller* of A . For all $x \in B(A) - A$, $\phi^t(x)$ approaches this dual repeller as $t \rightarrow -\infty$.

Consider now a one-parameter family of differential equations $\dot{x} = V_\varepsilon(x)$ in \mathbf{R}^n (with unique solutions $x(t) = \Phi_\varepsilon^t(x(0))$) such that $(\varepsilon, x) \mapsto V_\varepsilon(x)$ is continuous. Then $(\varepsilon, t, x) \mapsto \Phi_\varepsilon^t(x)$ is continuous as well. Suppose that $X \subset \mathbf{R}^n$ is compact and forward invariant under the semi-flows Φ_ε . For $\varepsilon = 0$ we omit the subscript in Φ^t .

The following continuation theorem for attractors is part of the folklore of dynamical systems; compare, e.g., Proposition 8.1 of Smale (1967).

Theorem A.1. *Let A be an attractor for Φ with basin $B(A)$. Then for each small enough $\varepsilon > 0$ there exists an attractor A_ε of Φ_ε with basin $B(A_\varepsilon)$, such that the map $\varepsilon \mapsto A_\varepsilon$ is upper hemicontinuous and the map $\varepsilon \mapsto B(A_\varepsilon)$ is lower hemicontinuous.*

Upper hemicontinuity cannot be replaced by continuity in this result. Consider the family of differential equations $\dot{x} = (\varepsilon + x^2)(1 - x)$ on the real line. The semi-flow Φ corresponding to $\varepsilon = 0$ admits $A = [0, 1]$ as an attractor, but when $\varepsilon > 0$ the unique attractor of Φ_ε is $A_\varepsilon = \{1\}$. This example shows that perturbations can cause attractors to implode; the theorem shows that perturbations cannot cause attractors to explode.

Theorem A.1 is a direct consequence of the following lemma, which is sufficient to prove the results in Sections 3 and 4.

Lemma A.2. *Let A be an attractor for Φ with basin $B(A)$, and let U_1 and U_2 be open sets satisfying $A \subseteq U_1 \subseteq U_2 \subseteq \bar{U}_2 \subseteq B(A)$. Then for each small enough $\varepsilon > 0$ there exists an attractor A_ε of Φ_ε with basin $B(A_\varepsilon)$, such that $A_\varepsilon \subseteq U_1$ and $U_2 \subseteq B(A_\varepsilon)$.*

In this lemma, one can always set $U_1 = \{x : \text{dist}(x, A) < \delta\}$ and $U_2 = \{x \in B(A) : \text{dist}(x, X - B(A)) > \delta\}$ for some small enough $\delta > 0$.

Proof of Lemma A.2. Since A is an attractor and $\omega(\bar{U}_2) = A$, there is a $T > 0$ such that $\Phi^t(\bar{U}_2) \subset U_1$ for $t \geq T$. By the continuous dependence of the flow on the parameter ε and the compactness of $\Phi^T(\bar{U}_2)$, we have that $\Phi_\varepsilon^T(\bar{U}_2) \subset U_1 \subseteq U_2$ for all small enough ε . Thus, U_2 is a trapping region for the semi-flow Φ_ε , and $A_\varepsilon \equiv \omega(U_2)$ is an attractor for Φ_ε . Moreover, $A_\varepsilon \subset U_1$ (since $A_\varepsilon = \Phi_\varepsilon^T(A_\varepsilon) \subseteq \Phi_\varepsilon^T(\bar{U}_2) \subset U_1$) and $U_2 \subset B(A_\varepsilon)$. ■

B. Proofs Omitted from Section 3

Our tasks in this section are to prove Lemma 3.2 and Theorem 3.3. Lemma 3.2 is an immediate consequence of Lemma B.2 below. Theorem 3.3 is proved in the same fashion as Theorem 3.1, with the role of Lemma 3.2 in the proof of Theorem 3.1 being played by Lemma B.1 (for Theorem 3.3(i)) and Lemma B.2 (for Theorem 3.3(ii)). Thus, all that remains is to state and prove these two lemmas.

The two lemmas prove the instability of the set of Nash equilibria $NE = \{x^* \in X : x^* = (\frac{1}{3}, \frac{1}{3}, c, \frac{1}{3} - c)\}$ in the “bad RPST” game $F(x) = \tilde{A}x$ under the dynamics defined in Theorem 3.3. To begin the analysis, recall the Lipschitz continuous function $\phi : \mathbf{R} \rightarrow \mathbf{R}_+$ defined in Theorem 3.3, which satisfies $\text{sgn}(\phi(u)) = \text{sgn}([u]_+)$ and $\frac{d}{du^+} \phi(u)|_{u=0} > 0$. It follows immediately that for any $\bar{d} > 0$, there are positive constants k_1 and k_2 such that

$$(7) \quad k_1 d \leq \phi(d) \leq k_2 d \quad \text{for all } d \in [0, \bar{d}].$$

Choosing $\bar{d} = 2 \max_{i,x} |F_i(x)|$ ensures that bound (7) holds for all values of d relevant to our analysis. If we define

$$\psi(d) = \int_0^d \phi(u) du,$$

then integrating (7) shows that

$$(8) \quad \psi(d) \leq \frac{k_2}{2} d^2 \quad \text{for all } d \in [0, \bar{d}].$$

Lemma B.1. *Let $F(x) = \tilde{A}x$ be the bad RPST game from equation (1), and let V^F be the excess payoff dynamic for this game generated by $\rho_{ij} = \phi(F_j - \bar{F})$. Then the set NE of Nash equilibria of F is a repeller under V^F .*

Proof. Define the *excess payoff* to strategy i by $\hat{F}_i(x) = F_i(x) - \bar{F}(x)$. Then our excess payoff dynamic can be expressed as

$$(EP) \quad \dot{x}_i = \phi(\hat{F}_i(x)) - x_i \sum_{j \in S} \phi(\hat{F}_j(x)).$$

The rest points of (EP) are the Nash equilibria of F ; moreover, if we let

$$\Phi(x) = \sum_{j \in S} \phi(\hat{F}_j(x)),$$

then $\Phi(x) \geq 0$, with equality if and only if x is a Nash equilibrium of F (see Sandholm (2005)).

Consider the Lyapunov function

$$U(x) = \sum_{i \in S} \psi(\hat{F}_i(x)).$$

Theorem 6.1 of Hofbauer (2000) (or Theorem 5.1 of Hofbauer and Sandholm (2006)) shows that $U(x) \geq 0$, with equality holding if and only if x is a Nash equilibrium of F . The proof of this theorem shows that the time derivative of U under the dynamic (EP) can be expressed as

$$(9) \quad \dot{U}(x) = \dot{x}' \tilde{A} \dot{x} - \Phi(x) F(x)' \dot{x}.$$

To prove our lemma, we need to show that $\dot{U}(x) > 0$ whenever $x \notin NE$ and $\text{dist}(x, NE)$ is sufficiently small.

Let $TX = \{z \in \mathbf{R}^n : z' \mathbf{1} = 0\}$, the tangent space of the simplex X , so that $\dot{x} \in TX$, and suppose that $z \in TX$. Then letting $(\zeta_1, \zeta_2, \zeta_3) = (z_1, z_2, z_3 + z_4)$, we have that

$$(10) \quad \begin{aligned} z' \tilde{A} z &= (a - b) (z_1 z_2 + z_2 (z_3 + z_4) + (z_3 + z_4) z_1) \\ &= (a - b) (\zeta_1 \zeta_2 + \zeta_2 \zeta_3 + \zeta_3 \zeta_1) \\ &= \frac{b - a}{2} \left(\left(\sum_{i=1}^3 \zeta_i \right)^2 - 2 \sum_{1 \leq i < j \leq 3} \zeta_i \zeta_j \right) \\ &= \frac{b - a}{2} \sum_{i=1}^3 \zeta_i^2. \\ &= \frac{b - a}{2} \left((z_1)^2 + (z_2)^2 + (z_3 + z_4)^2 \right). \end{aligned}$$

Now if $x \notin NE$, we can write (EP) as

$$(11) \quad \dot{x} = \Phi(x) (\sigma(x) - x)$$

where $\sigma(x) \in X$ is given by $\sigma_i(x) = \phi(\hat{F}_i(x))/\Phi(x)$. Since $x \notin NE$, some strategy i has a below average payoff ($F_i(x) < \bar{F}(x)$), implying that $\sigma_i(x) = 0$, and hence that $\sigma(x) \in \text{bd}(X)$. In fact, since strategies 3 and 4 always earn the same payoff, we have that $\sigma_3(x) = 0$ if and only if $\sigma_4(x) = 0$.

If we now write $y = (x_1, x_2, x_3 + x_4)$ and $\tau(x) = (\sigma_1(x), \sigma_2(x), \sigma_3(x) + \sigma_4(x))$, equation (11) becomes

$$\dot{y} = \Phi(x)(\tau(x) - y).$$

The arguments in the previous paragraph show that $\tau(x)$ is on the boundary of the simplex in \mathbf{R}^3 . Therefore, if we fix a small $\varepsilon > 0$ and assume that $\text{dist}(x, NE) < \varepsilon$, then $|y - (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})| < \varepsilon$, giving us a uniform bound on the distance between $\tau(x)$ and y , and hence a uniform lower bound on $|\dot{y}|$:

$$|\dot{y}| \geq c \Phi(x)$$

for some $c > 0$. By squaring and rewriting in terms of \dot{x} , we obtain

$$(12) \quad \dot{x}_1^2 + \dot{x}_2^2 + (\dot{x}_3 + \dot{x}_4)^2 \geq c^2 \Phi(x)^2.$$

Thus, combining equations (10) and (12) shows that if $\text{dist}(x, NE) < \varepsilon$, then

$$(13) \quad \dot{x}' \tilde{A} \dot{x} \geq \frac{b-a}{2} c^2 \Phi(x)^2.$$

To bound the second term of equation (9), use equation (7) to show that

$$(14) \quad \begin{aligned} \Phi(x) F(x)' \dot{x} &= \Phi(x) (\hat{F}(x) - \bar{F}(x) \mathbf{1})' \dot{x} \\ &= \Phi(x) \hat{F}(x)' \dot{x} && \text{since } \mathbf{1}' \dot{x} = 0 \\ &= \Phi(x) \sum_{i \in S} \hat{F}_i(x) (\phi(\hat{F}_i(x)) - x_i \Phi(x)) \\ &= \Phi(x) \sum_{i \in S} \hat{F}_i(x) \phi(\hat{F}_i(x)) && \text{since } \hat{F}(x)' x = 0 \\ &\geq \Phi(x) k_1 \sum_{i \in S} \hat{F}_i(x)^2 \end{aligned}$$

$$\begin{aligned}
&\geq \Phi(x) \frac{k_1}{n} \left(\sum_{i \in S} \hat{F}_i(x) \right)^2 \\
&\geq \Phi(x) \frac{k_1}{nk_2^2} \left(\sum_{i \in S} \phi(\hat{F}_i(x)) \right)^2 \\
&= \frac{k_1}{nk_2^2} \Phi(x)^3.
\end{aligned}$$

Combining inequalities (13) and (14) with equation (9), we find that for x close enough to NE ,

$$\dot{U}(x) \geq \frac{b-a}{2} c^2 \Phi(x)^2 - \frac{k_1}{nk_2^2} \Phi(x)^3.$$

Since $\Phi(x) \geq 0$, with equality only when $x \in NE$, we conclude that $\dot{U}(x) > 0$ whenever $x \notin NE$ is close enough to NE , and therefore that NE is a repeller under (EP). ■

Lemma B.2. *Let $F(x) = \tilde{A}x$ be the bad RPST game from equation (1), and let V^F be the pairwise comparison dynamic for this game generated by $\rho_{ij} = \phi(F_j - F_i)$. Then the set NE of Nash equilibria of F is a repeller under V^F .*

Proof. The pairwise comparison dynamics considered here are defined by

$$(PWC) \quad \dot{x}_i = \sum_{j \in S} x_j \phi(F_i(x) - F_j(x)) - x_i \sum_{j \in S} \phi(F_j(x) - F_i(x)).$$

Sandholm (2006a) shows that the rest points of (PWC) are the Nash equilibria of F .

Our analysis relies on the following Lyapunov function:

$$\Psi(x) = \sum_{i \in S} \sum_{j \in S} x_i \psi(F_j(x) - F_i(x)).$$

Theorem 5.3 of Hofbauer and Sandholm (2006) (also see Smith (1984)) shows that $\Psi(x) \geq 0$, with equality holding if and only if x is a Nash equilibrium of F . The proof of that theorem shows that the time derivative of Ψ under the dynamic (PWC) can be expressed as

$$\begin{aligned}
(15) \quad \dot{\Psi}(x) &= \dot{x}' \tilde{A} x + \sum_{i \in S} \sum_{j \in S} \left(x_j \phi(F_i(x) - F_j(x)) \sum_{k \in S} \left(\psi(F_k(x) - F_i(x)) - \psi(F_k(x) - F_j(x)) \right) \right) \\
&\equiv T_1(x) + T_2(x).
\end{aligned}$$

Equation (10) tells us that $T_1(x) \geq 0$, with equality when $x \in NE$ (i.e., when $\dot{x} = 0$). Hofbauer and Sandholm (2006) show that $T_2(x) \leq 0$, with equality only when $x \in NE$. To prove the lemma, we must show that $T_1(x) + T_2(x) > 0$ whenever $x \notin NE$ and $\text{dist}(x, NE)$ is sufficiently small.

To begin, observe that since F is linear, we have that

$$(16) \quad [F_j(x) - F_i(x)]_+ \leq c_1 \text{dist}(x, NE).$$

for some $c_1 > 0$. Equations (7), (8), and (16) immediately yield a cubic bound on T_2 :

$$(17) \quad |T_2(x)| \leq c_2 \text{dist}(x, NE)^3$$

for some $c_2 > 0$.

To obtain a lower bound on T_1 , first note that the linearity of F implies that

$$(18) \quad \max_{i \in S} F_i(x) - \min_{j \in S} F_j(x) \geq c_3 \text{dist}(x, NE)$$

for some $c_3 > 0$. If $F_1(x) \geq F_2(x) \geq F_3(x) = F_4(x)$, then equations (7) and (18) imply that

$$\dot{x}_1 = \sum_{j=2}^4 x_j \phi(F_1(x) - F_j(x)) \geq (x_3 + x_4) \phi(F_1(x) - F_3(x)) \geq (x_3 + x_4) c_3 k_1 \text{dist}(x, NE).$$

Similarly, if $F_1(x) \leq F_2(x) \leq F_3(x) = F_4(x)$, then

$$|\dot{x}_1| = x_1 \sum_{j=2}^4 \phi(F_j(x) - F_1(x)) \geq x_1 \phi(F_3(x) - F_1(x)) \geq x_1 c_3 k_1 \text{dist}(x, NE).$$

Obtaining bounds on $|\dot{x}_1|$ and on $|\dot{x}_2|$ for the remaining four cases in like fashion, we find that for some $c_4 > 0$ and some $\varepsilon > 0$, for any x with $\text{dist}(x, NE) \leq \varepsilon$ (and hence $|x_1 - \frac{1}{3}| \leq \varepsilon$, $|x_2 - \frac{1}{3}| \leq \varepsilon$, and $|(x_3 + x_4) - \frac{1}{3}| \leq \varepsilon$), we have that

$$(19) \quad \dot{x}_1^2 + \dot{x}_2^2 + (\dot{x}_3 + \dot{x}_4)^2 \geq \max(\dot{x}_1^2, \dot{x}_2^2) \geq c_4 \text{dist}(x, NE)^2.$$

Thus, equations (10) and (19) together imply that for such x ,

$$(20) \quad T_1 = \dot{x}' \tilde{A} \dot{x} \geq \frac{b-a}{2} c_4 \text{dist}(x, NE)^2.$$

Combining equations (15), (17), and (20), we find that

$$\dot{\Psi}(x) \geq \text{dist}(x, NE)^2 \left(\frac{b-a}{2} c_4 - c_2 \text{dist}(x, NE) \right) \quad \text{whenever } \text{dist}(x, NE) \leq \varepsilon.$$

We therefore conclude that $\dot{\Psi}(x) > 0$ whenever x is sufficiently close to but not in NE , and so that NE is a repellor under (PWC). ■

C. Proofs Details Omitted from Section 4.2.4

Our analysis relies on the notion of attractor-repellor pairs introduced by Conley (1978); see Robinson (1995) for a textbook treatment. Define the flow from the set $U \subseteq X$ under the dynamic V^F by

$$\phi_t(U) = \left\{ \xi \in X : \text{there is a solution } \{x_s\} \text{ to } \dot{x} = V^F(x) \text{ with } x_0 \in U \text{ and } x_t = \xi. \right\}$$

In words, $\phi_t(U)$ contains the time t positions of solutions to V^F whose initial conditions are in U .

Recall that solutions to V^F starting in $I - NE$ ascend the function $f^{\tilde{C}}$ until leaving the set I . It follows that the set NE is a *repellor* under V^F : all backward-time solutions to V^F that begin near NE converge to NE . More precisely, there is a neighborhood U of NE such that

$$\bigcap_{t < 0} \phi_t(\text{cl}(U)) = NE.$$

The *dual attractor* \mathcal{A} of the repellor NE is the forward-time limit of the flow of V^F starting from the complement of $\text{cl}(U)$:

$$\mathcal{A} = \bigcap_{t > 0} \phi_t(X - \text{cl}(U)).$$

Lemma 4.3 tells us that $\mathcal{A} \subset D$. Three other key properties of the attractor \mathcal{A} are noted next.

Lemma C.1. *The attractor \mathcal{A} is nonempty, compact, and forward and backward invariant under V^F .*

These properties of attractors are well known—for proofs, see Conley (1978) or Robinson (1995).

We now show that \mathcal{A} is contained in the interior of the simplex. To do so, we let $Z = \{x \in X : x_4 = 0\}$ be the face of X on which the twin strategy is unused. We prove

Lemma C.2. *The attractor \mathcal{A} and the face Z are disjoint.*

Proof. Recall that $V^F(x) = g(F(x), x)$, where the growth rate function g is Lipschitz continuous (by condition (C)) and satisfies $g_i(\pi, x) \geq 0$ whenever $x_i = 0$. It follows that solutions to V^F that start in $X - Z$ cannot approach Z more than exponentially quickly, and in particular cannot reach Z in finite time. Equivalently, backward solutions to V^F starting from states in Z cannot enter $\text{int}(X)$.

Now suppose by way of contradiction that there exists a state ξ in $\mathcal{A} \cap Z$. Then by Lemma C.1 and the previous paragraph, the entire backward orbit from ξ is also contained in $\mathcal{A} \cap Z$, and hence in $D \cap Z$. Since the latter set contains no rest points (by condition (PC)), the Poincaré-Bendixson Theorem implies that the backward orbit from ξ converges to a closed orbit γ in $D \cap Z$ that circumnavigates $I \cap Z$.

By construction, the annulus $D \cap Z$ can be split into three regions: one in which strategy 1 is the best response, one in which strategy 2 is the best response, and one in which strategy 3 (and hence strategy 4) is a best response. Each of these regions is bounded by a simple closed curve that intersects the inner and outer boundaries of the annulus. Therefore, the closed orbit γ , on which strategy 4 is unused, passes through the region in which strategy 4 is optimal. This contradicts innovation (IN). ■

References

- Akin, E. (1980). Domination or equilibrium. *Mathematical Biosciences*, 50:239–250.
- Beckmann, M., McGuire, C. B., and Winsten, C. B. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven.
- Benaïm, M. and Weibull, J. W. (2003). Deterministic approximation of stochastic evolution in games. *Econometrica*, 71:873–903.
- Berger, U. and Hofbauer, J. (2006). Irrational behavior in the Brown-von Neumann-Nash dynamics. *Games and Economic Behavior*, 56:1–6.
- Brown, G. W. and von Neumann, J. (1950). Solutions of games by differential equations. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games I*, volume 24 of *Annals of Mathematics Studies*, pages 73–79. Princeton University Press, Princeton.
- Conley, C. (1978). *Isolated Invariant Sets and the Morse Index*. American Mathematical Society, Providence, RI.

- Friedman, D. (1991). Evolutionary games in economics. *Econometrica*, 59:637–666.
- Fudenberg, D. and Levine, D. K. (1998). *Theory of Learning in Games*. MIT Press, Cambridge.
- Gilboa, I. and Matsui, A. (1991). Social stability and equilibrium. *Econometrica*, 59:859–867.
- Hart, S. and Mas-Colell, A. (2003). Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, 93:1830–1836.
- Hofbauer, J. (2000). From Nash and Brown to Maynard Smith: Equilibria, dynamics, and ESS. *Selection*, 1:81–88.
- Hofbauer, J. and Sandholm, W. H. (2006). Stable Games. Unpublished manuscript, University of Vienna and University of Wisconsin.
- Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Hofbauer, J. and Swinkels, J. M. (1996). A universal Shapley example. Unpublished manuscript, University of Vienna and Northwestern University.
- Hofbauer, J. and Weibull, J. W. (1996). Evolutionary selection against dominated strategies. *Journal of Economic Theory*, 71:558–573.
- Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.
- Nachbar, J. H. (1990). ‘Evolutionary’ selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory*, 19:59–89.
- Nagurney, A. and Zhang, D. (1997). Projected dynamical systems in the formulation, stability analysis, and computation of fixed demand traffic network equilibria. *Transportation Science*, 31:147–158.
- Robinson, C. (1995). *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC Press, Boca Raton, FL.
- Rosenthal, R. W. (1973). A class of games possessing pure strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67.
- Samuelson, L. and Zhang, J. (1992). Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57:363–391.
- Sandholm, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory*, 97:81–108.
- Sandholm, W. H. (2003). Evolution and equilibrium under inexact information. *Games and Economic Behavior*, 44:343–378.

- Sandholm, W. H. (2005). Excess payoff dynamics and other well-behaved evolutionary dynamics. *Journal of Economic Theory*, 124:149–170.
- Sandholm, W. H. (2006a). Pairwise comparison dynamics and evolutionary foundations for Nash equilibrium. Unpublished manuscript, University of Wisconsin.
- Sandholm, W. H. (2006b). Population games and evolutionary dynamics. Unpublished manuscript, University of Wisconsin.
- Sandholm, W. H. (2006c). Potential functions for population games. Unpublished manuscript, University of Wisconsin.
- Sandholm, W. H. (2007). Evolution in Bayesian games II: Stability of purified equilibria. *Journal of Economic Theory*, forthcoming.
- Sandholm, W. H., Dokumaci, E., and Lahkar, R. (2006). The projection dynamic, the replicator dynamic, and the geometry of population games. Unpublished manuscript, University of Wisconsin.
- Smale, S. (1967). Differentiable dynamical systems. *Bulletin of the American Mathematical Society*, 73:747–817.
- Smith, M. J. (1984). The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov. *Transportation Science*, 18:245–252.
- Swinkels, J. M. (1993). Adjustment dynamics and rational play in games. *Games and Economic Behavior*, 5:455–484.
- Taylor, P. D. and Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.