

Article

Pairwise Comparison Dynamics and Evolutionary Foundations for Nash Equilibrium

William H. Sandholm

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>

Received: 15 October 2009; in revised form: 26 November 2009 / Accepted: 26 November 2009 / Published: 1 December 2009

Abstract: We introduce a class of evolutionary game dynamics — *pairwise comparison dynamics* — under which revising agents choose a candidate strategy at random, switching to it with positive probability if and only if its payoff is higher than the agent’s current strategy. We prove that all such dynamics satisfy *Nash stationarity*: the set of rest points of these dynamics is always identical to the set of Nash equilibria of the underlying game. We also show how one can modify the replicator dynamic and other imitative dynamics to ensure Nash stationarity without increasing the informational demands placed on the agents. These results provide an interpretation of Nash equilibrium that relies on large numbers arguments and weak requirements on payoff observations rather than on strong equilibrium knowledge assumptions.

Keywords: evolutionary game theory; dynamical systems; Nash equilibrium

1. Introduction

Nash equilibrium is the traditional basis for predicting behavior in noncooperative games, and is the starting point for analysis in most applications of game theory. Nevertheless, the program of justifying the Nash prediction by way of more primitive behavioral assumptions remains incomplete. Many researchers have explored the epistemic foundations of Nash equilibrium. But the conditions they propose to ensure Nash play are quite stringent, and seem too demanding to be appropriate in most applications.¹

¹ See Dekel and Gul [1] and the references therein.

Rather than focusing on epistemic issues, one can instead take the traditional view of Nash equilibrium as a necessary condition for stationary behavior among rational agents. But even this more modest stance is founded on the assumption of equilibrium knowledge: namely, that each player correctly anticipates how his opponents will act. In many settings, particularly those with large numbers of agents, this assumption seems too strong.

This paper uses techniques from evolutionary game theory to provide an interpretation of Nash equilibrium for large population settings. We study an explicitly dynamic model in which agents stochastically and myopically update their choices in response to their current strategic environment. By appealing to a suitable form of the law of large numbers, we derive deterministic dynamics that describe the evolution of the agents' aggregate behavior. We prove that the stationary states of these deterministic dynamics are identical to the Nash equilibria of the underlying game. We thereby connect the traditional game-theoretic notion of equilibrium behavior with the usual notion of stasis from dynamical systems. In so doing, we forgo strong equilibrium knowledge assumptions in favor of large numbers arguments and weak assumptions about agents' observations of current payoff opportunities.

In our model, agents from a large population recurrently receive opportunities to switch strategies. Upon receiving an opportunity, an agent decides what to do next by applying a *revision protocol*. A revision protocol is a map from currently available payoffs and current aggregate behavior to *conditional switch rates*, which are the rates at which agents who receive revision opportunities switch from one strategy to another. A population game and a revision protocol together define a stochastic evolutionary process. By applying an appropriate formulation of the law of large numbers, one can show that the behavior of this stochastic process is well-approximated by the solutions to a deterministic *mean dynamic*, which is defined by the stochastic process's expected motion.² Our present analysis takes this law of large numbers for granted and studies the behavior of the deterministic system.

To obtain our desired interpretation of Nash equilibrium, we would like to construct dynamics satisfying *Nash stationarity*: the stationary states of the dynamics should always coincide with the Nash equilibria of the game at hand. At the same time, we would like to derive the dynamics from revision protocols that make limited informational demands on the agents who use them.

None of the usual dynamics from the evolutionary literature achieves both of these goals. For instance, the replicator dynamic (Taylor and Jonker [5]) and other dynamics based on imitation (Björnerstedt and Weibull [6], Weibull [7], Hofbauer [8]) can be derived from revision protocols requiring a bare minimum of information: each agent need only be aware of the payoff to his current strategy. But imitative dynamics fail Nash stationarity, as they admit boundary rest points that are not Nash equilibria of the underlying game. For the best response dynamic (Gilboa and Matsui [9]), the situation is reversed: under this dynamic rest points and Nash equilibria are identical, but the protocol that generates the dynamic is discontinuous, requiring agents to know the exact payoffs of all available strategies in order to determine the current best response. Finally, the BNN dynamic (Brown and von Neumann [10]) and related dynamics (Weibull [11], Hofbauer [12], Sandholm [13]) satisfy Nash stationarity, and are based on continuous revision protocols. But these protocols also require agents to know the average payoff in the population—a piece of data that seems hard to obtain if it is not provided by a central source.

²For formal treatments of this approximation result, see Kurtz [2], Benaïm and Weibull [3], and Sandholm [4]; also see Section 2.2. below.

The dynamics we study in this paper are based on continuous revision protocols of an especially simple form. When an agent receives an opportunity to switch strategies, he chooses a candidate strategy at random, and switches to this strategy with positive probability if and only if its payoff is higher than his current strategy's payoff. To implement such a protocol, an agent need only know the payoffs of two strategies: his current strategy, and the randomly chosen candidate strategy. Nevertheless, we prove that the induced aggregate dynamics, which we dub *pairwise comparison dynamics*, satisfy Nash stationarity: their rest points are precisely the Nash equilibria regardless of the game being played.

To obtain the simplest pairwise comparison dynamic, one sets the probability of switching from the current strategy i to the candidate strategy j to be proportional to the difference between these strategies' payoffs. It is quite interesting to note that the resulting evolutionary dynamic is not new: it appears in the transportation science literature in the work of M. J. Smith [14], who uses it to study the stability of equilibrium behavior in highway networks. Our analysis shows that the functional form used by Smith [14] is not essential to obtain his dynamic's desirable properties.

We noted earlier that dynamics based solely on imitation must fail Nash stationarity: under such dynamics, any state at which all agents choose the same strategy is stationary, as no alternative strategies are available for imitation. Because imitation is a common component of human decision processes, it is important to know whether an exact link between imitative behavior and Nash equilibrium can be salvaged. To accomplish this, we introduce *hybrid dynamics*: we assume that rather than always imitating or always choosing candidate strategies at random, agents instead use hybrid revision protocols that require a bit of each. We show that as long as the weight placed on random selection of candidate strategies is strictly positive, the resulting hybrid dynamics satisfy Nash stationarity. In other words, Nash stationarity and imitation are only in conflict if the latter is used exclusively as the basis for decisions.

Section 2. introduces population games and evolutionary dynamics. Section 3. proposes our desiderata for revision protocols and evolutionary dynamics. Section 4. defines pairwise comparison dynamics, and proves that they satisfy Nash stationarity. Section 5. extends this property to a broad range of hybrid dynamics. Section 6. concludes with a discussion of convergence results.

2. The Model

2.1. Population Games

We consider games played by a *society* consisting of one or more *populations* $p \in \mathcal{P} = \{1, \dots, p\}$. Population p contains a continuum of agents of mass m^p who choose pure strategies from the set $S^p = \{1, \dots, n^p\}$. The total number of pure strategies in all populations is $n = \sum_{p \in \mathcal{P}} n^p$. The set of *population states* $X^p = \{x^p \in \mathbf{R}_+^{n^p} : \sum_{i \in S^p} x_i^p = m^p\}$ contains all empirical distributions of strategies for population p , while the set of *social states* $X = \{x = (x^1, \dots, x^p) \in \mathbf{R}_+^n : x^p \in X^p\}$ consists of empirical distributions of strategies for all populations.

When an agent in population p plays strategy $i \in S^p$, his payoff is described by a function $F_i^p : X \rightarrow \mathbf{R}$ of the current social state. $F^p : X \rightarrow \mathbf{R}^{n^p}$ is the vector of payoff functions for strategies in S^p , while $F : X \rightarrow \mathbf{R}^n$ is the vector of all payoff functions. Similar notational conventions are used throughout the paper. However, when we consider games with a single population, we assume that the population's mass is one and omit the redundant superscript p .

As usual, state $x \in X$ is a *Nash equilibrium* of F ($x \in NE(F)$) if each strategy in use at x is a best response to x . Formally, x is a Nash equilibrium if

$$\text{For all } p \in \mathcal{P} \text{ and } i \in S^p, x_i^p > 0 \text{ implies that } i \in \operatorname{argmax}_{j \in S^p} F_j^p(x).$$

The most natural setting for the model of choice introduced next is one in which the payoff F_i^p is deterministic, rather than the expected payoff from a random match.³ This so-called "playing the field" framework is common in applications of population games; it is used, for instance, in models of network congestion, macroeconomic spillovers, and other sorts of multilateral externalities.⁴

2.2. Revision Protocols

To define our model of evolution, we suppose that all agents in a society are equipped with rate R Poisson alarm clocks. A ring of an agent's clock signals an opportunity for this agent to switch strategies. The agent's decisions at such instances are described by a *revision protocol* $\rho^p : \mathbf{R}^{n^p} \times X^p \rightarrow \mathbf{R}_+^{n^p \times n^p}$. (We sometimes refer to the whole collection $\rho = (\rho^1, \dots, \rho^p)$ as a revision protocol when it is convenient to do so.)

The function ρ^p takes a payoff vector $\pi^p \in \mathbf{R}^{n^p}$ and a population state $x^p \in X^p$ as inputs and returns a matrix of *conditional switch rates* $\rho_{ij}^p(\pi^p, x^p) \in [0, R]$ as outputs. If at social state $x \in X$ an agent playing strategy $i \in S^p$ receives a revision opportunity, then with probability $\rho_{ij}^p(F^p(x), x^p)/R$ the agent switches to strategy $j \neq i$, while with probability $1 - \sum_{j \neq i} \rho_{ij}^p(F^p(x), x^p)/R$ the agent continues to play strategy i . For this last probability to be well-defined, the rate R of the agents' Poisson alarm clocks must satisfy

$$R \geq \sum_{j \neq i} \rho_{ij}^p(F^p(x), x^p) \text{ for all } x \in X, i \in S^p, \text{ and } p \in \mathcal{P}.$$

Note that the diagonal component $\rho_{ii}^p(F^p(x), x^p)$ is merely a placeholder, and plays no formal role in the model.

We can distinguish protocols according to the manner in which agents obtain candidate strategies to consider switching to during revision opportunities. We say that a protocol is *imitative* if it can be expressed in the form

$$\rho_{ij}^p(F^p(x), x^p) = \frac{x_j^p}{m^p} r_{ij}^p(F^p(x), x^p) \quad (1)$$

for some function $r_{ij}^p : \mathbf{R}^{n^p} \times X^p \rightarrow \mathbf{R}_+^{n^p \times n^p}$. We can interpret the revision procedure of an agent who follows an imitative protocol in the following way. When the agent receives a revision opportunity, he selects a member of his population at random and observes this member's strategy; thus, the probability that a strategy $j \neq i$ player is observed is x_j^p/m^p . Strategy j becomes the revising agent's candidate

³For work focusing on properties of learning rules in non-strategic environments with random payoffs, see Schlag [15], Börgers et al. [16], and the references therein.

⁴For specific applications, see Beckman et al. [17], Cooper [18], and Sandholm [19,20]. The term "playing the field" is due to mathematical biologist John Maynard Smith, who observes that while the pairwise random matching framework is mathematically appealing, "contests against the field are probably more widespread and important than pairwise contests" in biological applications (Maynard Smith [21], p. 23); his view would seem to apply equally well to economic modeling. Still, our limited information conditions retain their simple interpretations within the random matching framework so long as matches occur sufficiently quickly relative to the rate of strategy revision.

strategy. He switches to this candidate strategy with probability $r_{ij}^p(F^p(x), x^p)/R$. We will see in Section 3. that the replicator dynamic is among the dynamics that can be derived from protocols of this form. Notice that imitative protocols preclude the choice of unused strategies. Some of the special properties of imitative dynamics can be traced to this source.

Under the remaining revision protocols considered in Sections 3. and 4., a strategy's popularity does not immediately influence the probability with which it is chosen by a revising agent. We sometimes use the term *direct protocol* to refer to revision protocols fitting this description.⁵ It will be convenient in this case to define the pre-protocol r^p to be identical to the direct protocol ρ^p itself.

2.3. Information Requirements for Revision Protocols

In principle, a revision protocol can consist of arbitrary functions that map payoff vectors and population states to conditional switch rates. But in the settings where evolutionary models are most relevant, we expect agents' information about the strategic environment to be limited. Our goal is to show that such limitations on agents' knowledge are consistent with the use of traditional solution concepts to predict agents' aggregate behavior.

One basic requirement in this spirit is that revision protocols be *continuous*.

(C) ρ^p is continuous in π^p and x^p .

Continuity ensures that small changes in aggregate behavior, changes which may be difficult for agents to detect, do not lead to large changes in agents' responses. In large population settings, the exact information about payoffs that is needed to use a discontinuous protocol can be difficult to obtain, and the myopic agents considered in evolutionary models are unlikely to make the necessary efforts to do so. For these reasons, we suggest that continuous protocols are preferable to discontinuous ones for describing how myopic agents in large population environments select new strategies.

Revision protocols also vary in terms of the specific pieces of data needed to employ them. A particularly simple protocol might only condition on the payoff to the agent's current strategy. Others might require agents to gather information about the payoffs to other strategies, whether by briefly experimenting with these strategies, or by asking others about their experiences with them. Still other protocols might require data beyond that provided by payoffs alone.

To organize the discussion, we introduce five classes of data requirements for revision protocols. In general we do so by expressing these requirements in terms of the function ρ^p . But since we do not view the act of observing the strategy of a randomly chosen opponent as imposing an informational burden, we express the data requirements for imitative protocols in terms of the function r^p from equation (1).

Having noted this special case, we can proceed to our data requirements.

- (D1) ρ_{ij}^p (or r_{ij}^p) depends only on π_i^p .
- (D1') ρ_{ij}^p (or r_{ij}^p) depends only on π_j^p .
- (D2) ρ_{ij}^p (or r_{ij}^p) depends only on π_i^p and π_j^p .
- (Dn) ρ_{ij}^p (or r_{ij}^p) depends on $\pi_1^p, \dots, \pi_{n^p}^p$, but not on $x_1^p, \dots, x_{n^p}^p$.
- (D+) ρ_{ij}^p (or r_{ij}^p) depends on $\pi_1^p, \dots, \pi_{n^p}^p$ and on $x_1^p, \dots, x_{n^p}^p$.

⁵One can define revision protocols that are neither imitative nor direct: see Lahkar and Sandholm [22] for a discontinuous example.

Protocols in classes (D1) and (D1') require only a single piece of payoff data: either the payoff of the agent's current strategy (under (D1)), or the payoff of the agent's candidate strategy (under (D1')). Protocols in class (D2) are slightly more demanding, as they require agents to know both of these strategies' payoffs.⁶ Protocols in class (Dn) require agents to know the payoffs of additional strategies.

Finally, protocols in class (D+) require information not only about the strategies' payoffs, but also information about the strategies' utilization levels. To preview an example to come (Example 3), note that a protocol that conditions directly on the average payoff $\frac{1}{m^p} \sum_{j \in S^p} x_j^p \pi_j^p$ obtained in the population falls in class (D+). The information about different strategies' utilization levels that is needed to compute average payoffs may be difficult to obtain. Unless information about either these levels or the average payoff itself is provided to the agents by a central planner, we do not expect such information to be readily available in typical large population settings.

2.4. Evolutionary Dynamics

Suppose that a society of agents employ revision protocol ρ during recurrent play of the population game F . If the social state at time t is $x \in X$, the expected change in the number of agents playing strategy $i \in S^p$ from time t to time $t + dt$ can be written as

$$\begin{aligned} & \sum_{j \neq i} x_j^p R dt \cdot \frac{\rho_{ji}^p(F^p(x), x^p)}{R} - x_i^p R dt \cdot \sum_{j \neq i} \frac{\rho_{ij}^p(F^p(x), x^p)}{R} \\ &= \left(\sum_{j \in S^p} x_j^p \rho_{ji}^p(F^p(x), x^p) - x_i^p \sum_{j \in S^p} \rho_{ij}^p(F^p(x), x^p) \right) dt \end{aligned}$$

With this motivation, we define the *mean dynamic* for ρ and F to be the ordinary differential equation

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

The mean dynamic (M), which is defined on the state space X , captures the population's expected motion under protocol ρ in game F . If we fix the revision protocol ρ in advance, the map from games F to differential equations implicitly defined by equation (M) is called the *evolutionary dynamic* induced by ρ .

The form of the mean dynamic (M) is easy to explain. The first term describes the "inflow" into strategy i from other strategies; it is obtained by multiplying the mass of agents playing each strategy j by the rate at which such agents switch to strategy i , and then summing over j . Similarly, the second term describes the "outflow" from strategy i to other strategies. The difference between these terms is the net rate of change in the use of strategy i .

By definition, equation (M) captures the expected motion of the stochastic evolutionary process described at the start of this section. Using appropriate formulations of the law of large numbers (see Kurtz [2]), Benaïm and Weibull [3] and Sandholm [4] prove that solutions to equation (M) closely approximate the sample paths of the underlying stochastic process over finite time spans. Moreover, Benaïm [23] and Benaïm and Weibull [3] prove that over infinite time spans, the stochastic evolutionary process must

⁶Of course, if we take it for granted that an agent is always aware of the payoff of his current strategy, then conditions (D1') and (D2) are equally demanding.

spend the predominant proportion of periods near recurrent points of the dynamic (M).⁷ We therefore leave the stochastic process behind and focus directly on the deterministic dynamic (M).

2.5. Incentive Properties of Evolutionary Dynamics

We now introduce conditions on mean dynamics that link the evolution of aggregate behavior to incentives in the underlying game. The first condition constrains equilibrium behavior, the second disequilibrium adjustment.

- (NS) *Nash stationarity:* $V(x) = \mathbf{0}$ if and only if $x \in NE(F)$.
 (PC) *Positive correlation:* $V^p(x) \neq \mathbf{0}$ implies that $V^p(x)'F^p(x) > 0$.

The condition of central interest here, *Nash stationarity* (NS), requires that the Nash equilibria of the game F and the rest points of the dynamic V coincide. When dynamics satisfy Nash stationarity, one can interpret Nash equilibrium as a requirement of dynamic balance, with individual agents' revisions leaving aggregate behavior in the society fixed. While traditional interpretations of Nash equilibrium play rely on the assumption of equilibrium knowledge, the approach offered here permits an interpretation that can be used when agents have limited information, provided that there are a large number of them.

Nash stationarity can be split into two distinct restrictions. First, (NS) asks that every Nash equilibrium of F be a rest point of V . If state x is a Nash equilibrium, then no agent benefits from switching strategies; in this situation, (NS) demands that aggregate behavior be at rest under V .⁸ Second, Nash stationarity asks that every rest point of V be a Nash equilibrium of F . If the current population state is not a Nash equilibrium, then there are agents who would benefit from switching strategies. (NS) requires that in this situation, the aggregate behavior of the society continues to adjust under the mean dynamic V .

The second condition, *positive correlation* (PC), constrains the directions of evolution from population states that are not rest points: in particular, it requires that strategies' growth rates be positively correlated with their payoffs.⁹ Condition (PC) is useful for studying dynamics derived from hybrid revision protocols, and is essential for proving convergence results—see Sections 5. and 6.

3. Examples

To provide a context for our results, we present some basic dynamics from the evolutionary game theory literature, along with revision protocols that induce them. For simplicity, we focus on the single population case. In what follows, we let $\bar{F} = \sum_{i \in S} x_i F_i(x)$ represent the population's average payoff at state x .

⁷For an application to perturbed best response dynamics, see Hofbauer and Sandholm [24].

⁸In interpreting this restriction, one should bear in mind that rest points of the mean dynamic need not be states at which the underlying stochastic process is itself at rest. Since V describes the *expected* changes in the use of each strategy under the stochastic evolutionary process, it is possible for a rest point of V to represent a balancing of stochastic forces rather than the absence of such forces.

⁹See Sandholm [19] for a derivation of this fact. Versions of condition (PC) were proposed earlier by Friedman [25] and Swinkels [26].

Example 1. The replicator dynamic,

$$\dot{x}_i = x_i(F_i(x) - \bar{F}(x))$$

introduced in the mathematical biology literature by Taylor and Jonker [5], is the most thoroughly studied dynamic in evolutionary game theory. Under this dynamic, the percentage growth rate of each strategy in use is equal to its *excess payoff*—that is, to the difference between its payoff and the average payoff in the population.

While in biological contexts the replicator dynamic describes the process of natural selection, in economic contexts it captures processes of imitation. In fact, all three of the imitative protocols below generate the replicator dynamic, as can be verified by substituting them into equation (M):

$$\rho_{ij}(\pi, x) = x_j(K - \pi_i) \quad (2)$$

$$\rho_{ij}(\pi, x) = x_j(K + \pi_j) \quad (3)$$

$$\rho_{ij}(\pi, x) = x_j[\pi_j - \pi_i]_+ \quad (4)$$

That these protocols capture imitation can be gleaned from the initial x_j term, which represents the idea that an agent who receives a revision opportunity uses the strategy of a randomly chosen opponent as his candidate strategy.

Each of the protocols above has low data requirements. Protocol (2), the *imitation driven by dissatisfaction* protocol of Björnerstedt and Weibull [6], is in class (D1): the agent compares his current payoff to an aspiration level K ; with probability linear in the payoff difference, he switches to the candidate strategy j without checking its payoff.¹⁰ Protocol (3), introduced by Hofbauer [8], is in class (D1'), with switches determined entirely by the payoff of the candidate strategy.

Particularly relevant to our analysis below is protocol (4), *pairwise proportional imitation*, which was introduced in a related context by Schlag [15]. Under this protocol, an agent switches to the candidate strategy only if its payoff is higher than the payoff of his current strategy, switching with probability proportional to the payoff difference. Since it conditions on the payoffs of both the current and candidate strategy, this protocol is of class (D2).

Turning to aggregate incentive properties, it is well known that the replicator dynamic satisfies positive correlation (PC).¹¹ But while the replicator dynamic satisfies Nash stationarity (NS) on the interior of the state space X , it violates this condition on the boundary of X . Since the replicator dynamic is based purely on imitation, unused strategies remain extinct forever; therefore, all monomorphic states, even those that do not correspond to Nash equilibria, are rest points.¹² In fact, evolutionary dynamics derived from a wide range of imitative protocols share these qualitative properties—see Björnerstedt and Weibull [6], Weibull [7], and Hofbauer [8]. §

¹⁰For this protocol to be well-defined, the constant K must be at least as large as the maximal payoff in F . Similarly, for the protocol in (3) to be well-defined, K must be no larger than the minimal payoff in F .

¹¹See, e.g., Proposition 3.6 of Fudenberg and Levine [27].

¹²One might argue that perturbed versions of the replicator dynamic could exclude these non-Nash rest points. However, it is not obvious how one can introduce perturbations that eliminate rest points at non-Nash states without also shifting the remaining rest points away from Nash equilibria. We offer a resolution of this problem in Section 5.

Example 2. The *best response dynamic*, introduced by Gilboa and Matsui [9], is defined by the differential inclusion¹³

$$\dot{x} \in \operatorname{argmax}_{y \in X} y'F(x) - x$$

In words, the best response dynamic always moves from the current state x toward a state y representing a mixed best response to x .

It is easy to verify that the best response dynamic satisfies versions of both of our aggregate incentive requirements, Nash stationarity (NS) and positive correlation (PC). But whether this dynamic provides a credible foundation for the Nash prediction also depends on its informational requirements. The best response dynamic is derived from a revision protocol which has revising agents always switch to a current best response. Since determining a best response requires an agent to know the payoffs to all strategies, this protocol is of class (Dn). Perhaps more importantly, the fact that this protocol uses exact optimization implies that it is discontinuous. This suggests that the best response dynamic, while mathematically appealing, may not provide an ideal foundation for the prediction of equilibrium play in populations of simple agents.¹⁴ §

Example 3. The *Brown-von Neumann-Nash (BNN) dynamic*, defined by

$$\dot{x}_i = [F_i(x) - \bar{F}(x)]_+ - x_i \sum_{j \in S} [F_j(x) - \bar{F}(x)]_+$$

was introduced in the context of symmetric zero-sum games by [10], and subsequently rediscovered by Skyrms [30], Swinkels [26], Weibull [11], and Hofbauer [12]. Like the best response dynamic, the BNN dynamic satisfies both of the aggregate incentive conditions, (NS) and (PC).¹⁵

The BNN dynamic can be derived from the following revision protocol, considered in Sandholm [13]:

$$\rho_{ij}(\pi, x) = [\pi_j - \sum_{k \in S} x_k \pi_k]_+. \quad (5)$$

Under this protocol, an agent who receives a revision opportunity picks a strategy at random, and then compares this strategy's payoff to the average payoff in the population. He only switches to the candidate strategy if its payoff exceeds the average payoff, doing so with probability proportional to the difference.

Unlike the protocol for the best response dynamic, protocol (5) is continuous. But to implement protocol (5), an agent must be aware of the average payoff in the population. This could be the case if agent were told the average payoff by a central planner, or if the agent knew the full payoff vector π and the population state x , and computed their inner product himself. Still, either way, protocol (5) is in class (D+). For this reason, the BNN dynamic does not seem ideal for providing a low-information foundation for the prediction of equilibrium play. §

¹³The formulation of the best response dynamic as a differential inclusion is due to Hofbauer [28].

¹⁴One can avoid the discontinuity of the best response dynamic by turning to dynamics based on perturbed best responses—for instance, the logit(η) dynamic of Fudenberg and Levine [27]. But while introducing payoff perturbations eliminates discontinuities, it does so at the cost of violating the incentive conditions, (NS) and (PC). Indeed, increasing the noise level η increases the degree of smoothing of the best response dynamic, while simultaneously making the failures of conditions (NS) and (PC) more severe. For further discussion of perturbed best response dynamics, see Hofbauer and Sandholm [24,29].

¹⁵As with the replicator dynamic, one can define families of dynamics with similar qualitative properties to the BNN dynamic; see Weibull [11], Hofbauer [12] and Sandholm [13].

4. Pairwise Comparison Dynamics

In this section, we introduce a class of evolutionary dynamics that satisfy incentive conditions (NS) and (PC), but that are based on revision protocols with mild informational requirements. The protocols we define combine key features of protocols (4) and (5) above. Like protocol (4), the new protocols are based on pairwise payoff comparisons, and so have limited data requirements. But like protocol (5), the new protocols rely on direct selection of candidate strategies rather than imitation of opponents, allowing them to satisfy Nash stationarity (NS).

4.1. Definition

We consider revision protocols under which the decision to switch from strategy $i \in S^p$ to strategy $j \in S^p$ depends on the difference between their payoffs:

$$\rho_{ij}^p(\pi^p, x^p) = \phi_{ij}^p(\pi_j^p - \pi_i^p). \quad (6)$$

We assume that the functions $\phi_{ij}^p : \mathbf{R} \rightarrow [0, \infty)$ introduced in equation (6) are Lipschitz continuous and satisfy *sign-preservation*: the conditional switch rate from i to j is positive if and only if the payoff to j exceeds the payoff to i :

$$\text{sgn}(\phi_{ij}^p(d)) = \text{sgn}([d]_+). \quad (7)$$

Evolutionary dynamics generated by such protocols take the form

$$\dot{x}_i^p = V_i^p(x) = \sum_{j \in S^p} x_j^p \phi_{ji}^p(F_i^p(x) - F_j^p(x)) - x_i^p \sum_{j \in S^p} \phi_{ij}^p(F_j^p(x) - F_i^p(x)).$$

We call such dynamics *pairwise comparison dynamics*.

The simplest revision protocol satisfying the restrictions above is semilinear in payoff differences:

$$\rho_{ij}^p(F^p, x^p) = [\pi_j^p - \pi_i^p]_+. \quad (8)$$

Notice that this protocol can be obtained by starting with Schlag's [15] revision protocol (4), and replacing its imitative component with direct selection of candidate strategies. We call the evolutionary dynamic induced by protocol (8),

$$\dot{x}_i^p = \sum_{j \in S^p} x_j^p [F_i^p(x) - F_j^p(x)]_+ - x_i^p \sum_{j \in S^p} [F_j^p(x) - F_i^p(x)]_+,$$

the *Smith dynamic*. As we noted in the introduction, this dynamic is not new: it can be found in the transportation science literature in the work of Smith [14], who uses it to investigate stability of equilibrium in a model of highway congestion.

4.2. Analysis

The protocols (6) that define pairwise comparison dynamics are continuous; being based on pairwise comparisons, they are in data requirement class (D2). Theorem 1 establishes that pairwise comparison dynamics satisfy both of our aggregate incentive conditions.

Theorem 1. Every pairwise comparison dynamic satisfies Nash stationarity (NS) and positive correlation (PC).

The proof of Theorem 1 relies on three equivalences between properties of Nash equilibria and evolutionary dynamics on the one hand, and requirements that sums of terms of the form $\rho_{ij}^p [F_j^p - F_i^p]_+$, or $\rho_{ij}^p [F_j^p - F_i^p]_+$ equal zero on the other. Sign preservation ensures that sums of the three types are identical, allowing us to establish properties (NS) and (PC).

In what follows, $\dot{x} = V(x)$ is the pairwise comparison dynamic generated by the population game F and revision protocol ρ .

Lemma 1. $x \in NE(F) \Leftrightarrow$ For all $i \in S^p$ and $p \in \mathcal{P}$, $x_i^p = 0$ or $\sum_{j \in S^p} [F_j^p(x) - F_i^p(x)]_+ = 0$.

Proof. Both statements say that each strategy in use at x is optimal. ■

Lemma 2. $V^p(x) = 0 \Leftrightarrow$ For all $i \in S^p$, $x_i^p = 0$ or $\sum_{j \in S^p} \rho_{ij}^p(F^p(x)) = 0$.

Proof. ($\Leftarrow$) Immediate.

($\Rightarrow$) Fix a population $p \in \mathcal{P}$, and suppose that $V^p(x) = 0$. If j is an optimal strategy for population p at x , then sign preservation implies that $\rho_{jk}^p(F^p(x)) = 0$ for all $k \in S^p$, and so that there is no "outflow" from strategy j :

$$x_j^p \sum_{i \in S^p} \rho_{ji}^p(F^p(x)) = 0.$$

Since $V_j^p(x) = 0$, there can be no "inflow" into strategy j either:

$$\sum_{i \in S^p} x_i^p \rho_{ij}^p(F^p(x)) = 0.$$

We can express this condition equivalently as

$$\text{For all } i \in S^p, \text{ either } x_i^p = 0 \text{ or } \rho_{ij}^p(F^p(x)) = 0.$$

If all strategies in S^p earn the same payoff at state x , the proof is complete. Otherwise, let i be a "second best" strategy—that is, a strategy whose payoff $F_i^p(x)$ is second highest among the payoffs available from strategies in S^p at x . The last observation in the previous paragraph and sign preservation tell us that there is no outflow from i . But since $V_i^p = 0$, there is also no inflow into i :

$$\text{For all } k \in S^p, \text{ either } x_k^p = 0 \text{ or } \rho_{ki}^p(F^p(x)) = 0.$$

Iterating this argument for strategies with lower payoffs establishes the result. ■

Lemma 3. Fix a population $p \in \mathcal{P}$. Then

- (i) $V^p(x)' F^p(x) \geq 0$.
- (ii) $V^p(x)' F^p(x) = 0 \Leftrightarrow$ For all $i \in S^p$, $x_i^p = 0$ or $\sum_{j \in S^p} \rho_{ij}^p(F^p(x)) [F_j^p(x) - F_i^p(x)]_+ = 0$.

Proof. We compute the inner product as follows:

$$\begin{aligned}
 V^p(x)'F^p(x) &= \sum_{j \in S^p} \left(\sum_{i \in S^p} x_i^p \rho_{ij}^p(F^p(x)) - x_j^p \sum_{i \in S^p} \rho_{ji}^p(F^p(x)) \right) F_j^p(x) \\
 &= \sum_{j \in S^p} \sum_{i \in S^p} (x_i^p \rho_{ij}^p(F^p(x)) F_j^p(x) - x_j^p \rho_{ji}^p(F^p(x)) F_j^p(x)) \\
 &= \sum_{j \in S^p} \sum_{i \in S^p} x_i^p \rho_{ij}^p(F^p(x)) (F_j^p(x) - F_i^p(x)) \\
 &= \sum_{i \in S^p} \left(x_i^p \sum_{j \in S^p} \rho_{ij}^p(F^p(x)) [F_j^p(x) - F_i^p(x)]_+ \right),
 \end{aligned}$$

where the last equality follows from sign-preservation. Both claims directly follow. ■

Now, sign preservation implies that the second conditions in Lemmas 1, 2, and 3(ii) are equivalent. From this observation, Theorem 1 easily follows. In particular, the observation and Lemmas 1 and 2 imply that $x \in NE(F)$ if and only if $V^p(x) = \mathbf{0}$ for all $p \in \mathcal{P}$; this is condition (NS). In addition, the observation, Lemma 2, and Lemma 3(ii) imply that $V^p(x) = \mathbf{0}$ if and only if $V^p(x)'F^p(x) = 0$; this fact and Lemma 3(i) imply that $V^p(x)'F^p(x) > 0$ whenever $V^p(x) \neq \mathbf{0}$, which is condition (PC). This completes the proof of Theorem 1.

5. Hybrid Dynamics

At this point, it might seem that dynamics that have low data requirements and satisfy Nash stationarity are rather special, in that they must be derived from a very specific sort of revision protocol. In actuality, these two desiderata are satisfied rather broadly. To explain why, we consider an agent who uses multiple revision protocols at possibly different intensities. If the agent uses protocol ρ^V at intensity a and protocol ρ^W at intensity b , then his behavior is described by the hybrid protocol $\rho^H = a\rho^V + b\rho^W$. Since mean dynamics are linear in conditional switch rates, the mean dynamic for the hybrid protocol is a linear combination of the two original mean dynamics: $H = aV + bW$.

Theorem 2 derives incentive properties of the hybrid dynamic from those of the original dynamics.

Theorem 2. *Suppose that the dynamic V satisfies (PC), that the dynamic W satisfies (NS) and (PC), and that $a, b > 0$. Then the hybrid dynamic $H = aV + bW$ also satisfies (NS) and (PC).*

Proof. The analysis builds on Section 4 of Sandholm [13]. To show that H satisfies (PC), suppose that $H^p(x) \neq \mathbf{0}$. Then either $V^p(x)$, $W^p(x)$, or both are not $\mathbf{0}$. Since V and W satisfy (PC), it follows that $V^p(x)'F^p(x) \geq 0$, that $W^p(x)'F^p(x) \geq 0$, and that at least one of these inequalities is strict. Consequently, $H^p(x)'F^p(x) > 0$, and so H satisfies (PC).

Our proof that H satisfies (NS) is divided into three cases. First, if x is a Nash equilibrium of F , then it is a rest point of both V and W , and hence a rest point of H as well. Second, if x is a non-Nash rest point of V , then it is not a rest point of W . Since $V(x) = \mathbf{0}$ and $W(x) \neq \mathbf{0}$, it follows that $H(x) = bW(x) \neq \mathbf{0}$, so x is not a rest point of H . Finally, suppose that x is not a rest point of V . Then x is not a Nash equilibrium, and so is not a rest point of W either. Since V and W satisfy condition

(PC), we know that $V(x)'F(x) = \sum_{p \in \mathcal{P}} V^p(x)'F^p(x) > 0$ and that $W(x)'F(x) > 0$. Consequently, $H(x)'F(x) > 0$, implying that x is not a rest point of H . Thus, H satisfies (NS). ■

A key implication of Theorem 2 is that imitation and Nash stationarity are not incompatible. If we combine an imitative dynamic V with any small amount of a pairwise comparison dynamic W , we obtain a hybrid dynamic H that satisfies (NS) and (PC). Thus, if agents usually choose candidate strategies by imitating opponents, occasionally choose these candidate strategies at random, and always decide whether to switch by making pairwise payoff comparisons, then the rest points of the resulting aggregate dynamic coincide with the Nash equilibria of the underlying game.

6. Discussion: Convergence Properties

By linking the notion of Nash equilibrium with the stationary states of evolutionary dynamics, this paper provides an interpretation of Nash equilibrium behavior in large populations that does not make use of equilibrium knowledge assumptions. But to *justify* the use of Nash equilibrium for predicting behavior, this identification is not enough: one must not only consider the dynamics' stationarity properties, but also their convergence properties. Unfortunately, it is known from the work of Hofbauer and Swinkels [31] and Hart and Mas-Colell [32] that no reasonable evolutionary dynamic converges to Nash equilibrium in all games.¹⁶ Therefore, to obtain convergence results one needs to impose additional structure on the games at issue.

Suppose, for instance, that F is a *potential game*: in other words, that there is a scalar-valued function $f : X \rightarrow \mathbb{R}$ whose gradient is F . Because pairwise comparison dynamics satisfy (NS) and (PC), it follows from results of Sandholm [19,34] that in potential games, these dynamics converge to Nash equilibrium from all initial conditions.

Alternatively, consider the class of *stable games*. These games are defined by the property that $(F(y) - F(x))'(y - x) \leq 0$ for all $x, y \in X$, and they include games with an interior ESS, zero-sum games, models of highway congestion, and wars of attrition as special cases. Smith [14] proves that the Smith dynamic converges to Nash equilibrium from all initial conditions in all stable games. Building on this result and on results in the present paper, Hofbauer and Sandholm [35] establish global convergence in stable games for any pairwise comparison dynamic whose revision protocol (6) is not only sign-preserving (7), but also satisfies a symmetry condition called *impartiality*:

$$\rho_{ij}^p(\pi^p, x^p) = \hat{\phi}_j^p(\pi_j^p - \pi_i^p).$$

Impartiality requires that the function of the payoff difference describing the conditional switch rate from i to j does not depend on the current strategy i .

In summary, while pairwise comparison dynamics are subject to the impossibility results of Hofbauer and Swinkels [31] and Hart and Mas-Colell [32], they are known to lead to equilibrium play in two key classes of games.

¹⁶Indeed, Hofbauer and Sandholm [33] exhibit a large class of dynamics under which strictly dominated strategies must survive in some games. This class includes excess payoff dynamics (including the BNN dynamic), pairwise comparison dynamics, and hybrid dynamics of the sort considered in Theorem 2. Since the rest points of these dynamics are Nash equilibria, survival of dominated strategies under these dynamics depends on the existence of solution trajectories that fail to converge to rest points.

Acknowledgements

I thank Carlos Alós-Ferrer, Ronny Freier, Faruk Gul, Larry Samuelson, Jörgen Weibull, two anonymous referees, and many seminar audiences for helpful comments, and Katsuhiko Aiba for able research assistance. Financial support from NSF Grants SES-0092145, SES-0617753, and SES-0851580 is gratefully acknowledged.

References

1. Dekel, E.; Gul, F. Rationality and Knowledge in Game Theory. In *Advances in Economics and Econometrics: Theory and Applications*; Kreps, D.M., Wallis, K.F., Eds.; Cambridge University Press: Cambridge, UK, 1997; Vol. 1, pp. 87–172.
2. Kurtz, T.G. Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes. *J. Appl. Probab.* **1970**, *7*, 49–58.
3. Benaïm, M.; Weibull, J.W. Deterministic Approximation of Stochastic Evolution in Games. *Econometrica* **2003**, *71*, 873–903.
4. Sandholm, W.H. Evolution and Equilibrium under Inexact Information. *Games Econ. Behav.* **2003**, *44*, 343–378.
5. Taylor, P.D.; Jonker, L. Evolutionarily Stable Strategies and Game Dynamics. *Math. Biosci.* **1978**, *40*, 145–156.
6. Björnerstedt, J.; Weibull, J.W. Nash Equilibrium and Evolution by Imitation. In *The Rational Foundations of Economic Behavior*; Arrow, K.J., Colomatto, E., Perlman, M., Schmidt, C., Eds.; St. Martin's Press: New York, NY, USA, 1996; pp. 155–181.
7. Weibull, J.W. *Evolutionary Game Theory*; MIT Press: Cambridge, MA, USA, 1995.
8. Hofbauer, J. Imitation Dynamics for Games. Unpublished manuscript, University of Vienna, Vienna, Austria, 1995.
9. Gilboa, I.; Matsui, A. Social Stability and Equilibrium. *Econometrica* **1991**, *59*, 859–867.
10. Brown, G.W.; von Neumann, J. Solutions of Games by Differential Equations. In *Contributions to the Theory of Games I*; Kuhn, H.W., Tucker, A.W., Eds.; Princeton University Press: Princeton, NJ, USA, 1950; Vol. 24, *Annals of Mathematics Studies*, pp. 73–79.
11. Weibull, J.W. The Mass Action Interpretation. Excerpt from 'The Work of John Nash in Game Theory: Nobel Seminar, December 8, 1994'. *J. Econ. Theory* **1996**, *69*, 165–171.
12. Hofbauer, J. From Nash and Brown to Maynard Smith: Equilibria, Dynamics, and ESS. *Selection* **2000**, *1*, 81–88.
13. Sandholm, W.H. Excess Payoff Dynamics and Other Well-Behaved Evolutionary Dynamics. *J. Econ. Theory* **2005**, *124*, 149–170.
14. Smith, M.J. The Stability of a Dynamic Model of Traffic Assignment—An Application of a Method of Lyapunov. *Transp. Sci.* **1984**, *18*, 245–252.
15. Schlag, K.H. Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-armed Bandits. *J. Econ. Theory* **1998**, *78*, 130–156.
16. Börgers, T.; Morales, A.J.; Sarin, R. Expedient and Monotone Learning Rules. *Econometrica* **2004**, *72*, 383–405.

17. Beckmann, M.; McGuire, C.B.; Winsten, C.B. *Studies in the Economics of Transportation*; Yale University Press: New Haven, CT, USA, 1956.
18. Cooper, R.W. *Coordination Games: Complementarities and Macroeconomics*; Cambridge University Press: Cambridge, UK, 1999.
19. Sandholm, W.H. Potential Games with Continuous Player Sets. *J. Econ. Theory* **2001**, *97*, 81–108.
20. Sandholm, W.H. Negative Externalities and Evolutionary Implementation. *Rev. Econ. Stud.* **2005**, *72*, 885–915.
21. Maynard Smith, J. *Evolution and the Theory of Games*; Cambridge University Press: Cambridge, UK, 1982.
22. Lahkar, R.; Sandholm, W.H. The Projection Dynamic and the Geometry of Population Games. *Games Econ. Behav.* **2008**, *64*, 565–590.
23. Benaïm, M. Recursive Algorithms, Urn Processes, and the Chaining Number of Chain Recurrent Sets. *Ergod. Theory Dyn. Syst.* **1998**, *18*, 53–87.
24. Hofbauer, J.; Sandholm, W.H. Evolution in Games with Randomly Disturbed Payoffs. *J. Econ. Theory* **2007**, *132*, 47–69.
25. Friedman, D. Evolutionary Games in Economics. *Econometrica* **1991**, *59*, 637–666.
26. Swinkels, J.M. Adjustment Dynamics and Rational Play in Games. *Games Econ. Behav.* **1993**, *5*, 455–484.
27. Fudenberg, D.; Levine, D.K. *Theory of Learning in Games*. MIT Press: Cambridge, MA, USA, 1998.
28. Hofbauer, J. Stability for the Best Response Dynamics. Unpublished manuscript, University of Vienna, Vienna, Austria, 1995.
29. Hofbauer, J.; Sandholm, W.H. On the Global Convergence of Stochastic Fictitious Play. *Econometrica* **2002**, *70*, 2265–2294.
30. Skyrms, B. *The Dynamics of Rational Deliberation*. Harvard University Press: Cambridge, MA, USA, 1990.
31. Hofbauer, J.; Swinkels, J.M. A Universal Shapley Example. Unpublished manuscript, University of Vienna, Vienna, Austria and Northwestern University, Evanston, IL, USA, 1996.
32. Hart, S.; Mas-Colell, A. Uncoupled Dynamics Do Not Lead to Nash Equilibrium. *Am. Econ. Rev.* **2003**, *93*, 1830–1836.
33. Hofbauer, J.; Sandholm, W.H. Survival of Dominated Strategies under Evolutionary Dynamics. Unpublished manuscript, University of Vienna, Vienna, Austria and University of Wisconsin-Madison, Madison, WI, USA, 2008.
34. Sandholm, W.H. Large Population Potential Games. *J. Econ. Theory* **2009**, *144*, 1710–1725.
35. Hofbauer, J.; Sandholm, W.H. Stable Games and their Dynamics. *J. Econ. Theory* **2009**, *144*, 1665–1693.