Stochastic Evolution of Rules for Playing Normal Form Games[*]

Fabrizio Germano

Departament d'Economia i Empresa

Universitat Pompeu Fabra

Ramon Trias Fargas 25-27

08005 Barcelona, Spain

`fabrizio.germano@upf.edu`

June 2004

Abstract

The evolution of boundedly rational rules for playing normal form games is studied within stationary environments of stochastically changing games. Rules are viewed as algorithms prescribing strategies for the different normal form games that arise. It is shown that many of the "folk results" of evolutionary game theory, typically obtained with a fixed game and fixed strategies, carry over to the present case. The results are also related to recent experiments on rules and games.

Keywords Rules, evolutionary dynamics, stochastic dynamics, bounded rationality, learning, normal form games. *JEL Classification* C72, C73, D81, D83.

1 Introduction

We consider a framework that is a natural and straightforward extension of the standard framework of evolutionary game theory. Rather than fixing a game that is played repeatedly either in continuous or discrete time, with or without perturbations, and with strategies corresponding to the pure or mixed strategies of the fixed game, we consider an class or set of games G from which a game is randomly and independently drawn each period according to a fixed probability distribution μ. Agents play the drawn games according to rules, which we take to be algorithms prescribing a strategy (possibly mixed) for any game from G that may appear. An interesting feature of the framework is that rules are not simply stategies of a given game, but are algorithms that apply to the entire class G and may have deeper cognitive interpretations, as well applicability beyond the class G. For example, a rule could be the prescription to always play "maxmin," that is, the strategy that maximizes the player's minimum or guaranteed payoffs (with randomizing if ties occur); or it could prescribe the Nash equilibrium strategy maximizing joint payoffs; or simply the strategy that best responds to uniform priors over other players' profiles. The evolutionary process is restricted not by a single game, but by the class of games G and the distribution μ.

Within this setup, we consider evolution of rules from a given set and so model agents learning or updating the probability with which to play the given rules depending on how these perform on the randomly selected games. (In particular, we do not consider here the case where agents may learn or invent rules from some unspecified set; also notice that our rules do not depend on the history of play, but only on the games – the history of play enters indirectly through the evolutionary process.) Rules thus become strategies in a game of rules against rules, where payoffs are simply the μ-expected payoffs of following the given rules against the other players' profiles of rules. This leads to what we call an average game.

Since games are drawn randomly each period, our evolutionary process is a stochastic process in discrete time. For simplicity, we consider aggregate log-monotonic dynamics (see Cabrales and Sobel (1992)) and show that the

1

resulting stochastic dynamics of rules satisfies some basic folk-properties of evolutionary game theory (see Hofbauer and Sigmund (2003)). In particular, we show that rules which are strictly dominated in the average game, are played with probability zero almost surely in the limit; similarly for iteratively strictly dominated rules. Further, we show that, if the stochastic dynamics converges, then the limit rule profile must be a Nash equilibrium of the average game. These results have counterparts, with fixed games, for example in Nachbar (1990), Friedman (1991), Samuelson and Zhang (1992), and Cabrales and Sobel (1992). In our case, they are shown using the strong law of large numbers for sequences of dependent random variables. We also show that pure Nash profiles of the average game that are asymptotically stable under the deterministic aggregate log-monotonic dynamics (on the average game) are stochastically asymptotically stable under the stochastic dynamics. This implies that strict equilibrium rule profiles of the average game are stochastically asymptotically stable under the stochastic dynamics. This also has counterparts for example in Nachbar (1990) and Ritzberger and Weibull (1996). It is not the case, however, that all Nash equilibria are restpoints (or zeros) of the stochastic dynamics; indeed, mixed equilibrium rule profiles need not be; this is thus a folk result that does not carry over to this stochastic setting. Finally, we show that, if empirical frequencies of the stochastic dynamics converge, then the limit point must be in the Hannan set (this is defined in the text; it is a convex polyhedron containing the set of correlated equilibria). It is still open whether empirical frequencies always converge to the Hannan set under this dynamics, but we can exclude that they will in general converge to the set of correlated equilibria.

The present paper relates to several strands of literature. Concerning stochastically changing games (or systems), there is a large literature studying perturbed games (for example, Foster and Young (1990), Fudenberg and Harris (1992), Fudenberg and Kreps (1993), Kandori, Mailath, and Rob (1993), Benaim and Hirsch (1999), Cabrales (2000), Hofbauer and Sandholm (2002)). Most of these papers, rather than modelling players learning algorithms or rules to play in different environments or games, focus on using

noise to obtain either selection or convergence results.

There are also several exceptions to this. Rosenthal (1979) considers agents playing sequences of games with possibly varying opponents and derives conditions that ensure that Markov stationarity properties are satisfied. Also, Rosenthal (1993a,b) studies rules of thumb for playing in random-matching games and characterizes certain steady-state equilibria. Li Calzi (1995) considers agents playing different games and studies what he calls fictitious play by cases, which specifically models how agents may draw on experience from playing similar games in the past via a fictitious play algorithm; he shows almost sure convergence of his process for 2×2 games. Samuelson (2001) considers agents playing different games with rules (or models representing the environment – which in his case are automata with varying states), that are optimal subject to complexity costs; he also studies the evolution of the automata and how they play in equilibrium. Sgroi and Zizzo (2001, 2003) study the process of neural networks having to play randomly drawn games after having been trained to play Nash equilibrium in environments with unique Nash equilibria; they then compare their networks' behavior after the training period is completed with behavior observed in the experimental literature and find potential similarities. Jehiel (2003) defines a notion of analogy-based expectation equilibrium, which involves agents forming analogy classes by bundling nodes at which other players make choices and eventually learning average behavior over analogy classes. Another closely related paper is Heller (2004), which studies the evolution of simple rules vs rules that allow agents to learn their environment; her paper derives conditions on the costs associated with learners' rules guaranteeing that, within changing environments, learners survive in the long run. While our framework allows in principle to address some of the questions arising in these papers, it is much less specific about the concrete learning and cognitive processes involved.

Finally, also closely related are many experimental papers such as Stahl and Wilson (1995), Stahl (1999, 2000), Rankin, Van Huyck, and Battalio (2000), Costa-Gomes, Crawford, and Broseta (2001), Stahl and Van Huyck

(2002), Van Huyck and Battalio (2002), and Selten, Abbink, Buchta, and Sadrieh (2003), who test specifically for the rules agents may use, learn to use, or also learn to develop when playing many different games. Identifying and understanding the rules that may underlie subjects' decisions is an important step towards understanding their strategic behavior in more general environments. Many of the examples throughout the paper are based on some of these experiments.

The paper is organized as follows. Section 2 describes the framework and notation, Section 3 contains all the results and some examples, Section 4 concludes.

2 Framework and Preliminary Notions

Let $I = \{1, .., n\}$ denote the set of players, S_i player i's space of pure strategies, $S = \times_{i \in I} S_i$ the space of pure strategy profiles, and let Σ_i denote the set of probability measures on S_i, $\Sigma = \times_{i \in I} \Sigma_i$ the space of mixed strategy profiles. Let also $S_{-i} = \times_{j \neq i} S_j$, and $\Sigma_{-i} = \times_{j \neq i} \Sigma_j$ and set $K_i = \#S_i$ for the number of i's strategies, $K = \sum_{i \in I} K_i$, and $\kappa = \Pi_{i \in I} K_i$ the number of possible outcomes. In what follows, we consider finite normal form games, that is, where n and each K_i are finite, and fix both the set of players and the set of strategy profiles, so that we can identify a game with a point in Euclidean space $\gamma \in I\!\!R^{\kappa n}$. We denote by $\gamma^i \in I\!\!R^\kappa$ the payoff array of player i and, by slight abuse of notation, also the payoff function of player i at game γ. Finally, $N(\gamma)$ denotes the set of Nash equilibria of γ.

We are interested in rules for playing arbitrary games within given subspaces $G \subset I\!\!R^{\kappa n}$. We view rules as being algorithms that for any game $\gamma \in G$ prescribe a strategy of player i for that game. Formally, we define a **rule** for player i as a map $r^i : G \to \Sigma_i$, $i \in I$. As with strategies for individual games, we let R_i denote a finite list of rules, $R = \times_{i \in I} R_i$ the space of rule profiles, and we denote by $\mathcal{R}_i$ the set of probability measures on R_i; $\mathcal{R} = \times_{i \in I} \mathcal{R}_i$ is the space of mixed rule profiles with generic element ρ.

Given a subset $G \subset I\!\!R^{\kappa n}$ and a probability measure μ on G, we can assess the performance of given rules on games in G by computing the expected

4

payoffs from playing the individual games that are drawn according to the probability measure μ. Throughout the paper we assume the set G to be compact. This leads to the following notion of an average game.

Definition 1 *Let $G \subset \mathbb{R}^{\kappa n}$ be a compact set of games, let μ be a probability measure on G, and let R denote a finite space of rules. The **average game** γ_∞ is defined by:*

$$\gamma_\infty^i(r) = \int_G \gamma^i(r(\gamma)) d\mu(\gamma), \quad for \ r \in R, i \in I.$$

The notion of average game is to be understood not as the average of games in G but as the average of how rule profiles in R perform over games in G. As we will see in the next sections, important properties of the learning behavior of rules, given such an **environment** (G, μ, R), can be derived from the associated average game γ_∞.

Example 1. Take G to be the space of all 2×2 games with payoffs in $[0, 1]$, and take μ to be such that all payoffs are drawn according to independent uniform distribution on $[0, 1]$. Many rules can be defined here. In this and the next example, we consider some of the rules studied in an experimental context by Stahl and Wilson (1995), Stahl (1999, 2000), and Costa-Gomes et al. (2001). In the latter's terminology, three of the simpler (nonstrategic) rules, are: "Naive" (N1), "Pessimistic" (P1), and "Altruistic" (A1), which, for each game $\gamma \in G$, recommend respectively, N1: the strategy that best replies to beliefs assigning equal probability to opponent's actions, P1: the "maxmin" strategy that maximizes the minimum of own payoffs, and A1: the strategy that maximizes joint payoffs.[1] The resulting average game is:[2]

---

[1] In Costa-Gomes et al. (2001) less than 30% of the subjects appear to use such non-strategic rules; between 65 and 90% appear to use "Naive" or "Best Reply to Naive" (N2), and up to 20% "Naive After One Round Dominance" (D1), (see the next example for definitions); their pool of games consists of 2×2, 2×3, and 2×4 games. Stahl and Wilson (1995) and Stahl (1999, 2000) essentially also obtain many subjects playing N1 and N2, and, to a much lesser extent, also Nash equilibrium.

[2] The numbers in this and the following matrices are computed as the average of randomly drawn games, the number of drawn games being such that the standard deviations of the payoffs are less than 10^{-4}.

	N1		P1		A1	
N1	.617		.617		.707	
		.617		.600		.597
P1	.600		.600		.661	
		.617		.600		.596
A1	.597		.596		.712	
		.707		.661		.712

The rule profiles N1 and A1 constitute pure Nash equilibria of the average game. The "maxmin" rule P1 is strictly dominated by N1. Notice how the average game need not have any resemblance with any of the individual games in G.

Given an average game, we can define standard (deterministic) dynamics, which, as we will see, are useful in evaluating limiting properties of the stochastic dynamics (to be introduced in the next section) on the underlying environment (G, μ, R).

Definition 2 *(Discrete) Aggregate Log-Monotonic Dynamics on γ_∞*

$$\rho^i_{k,t+1} = \frac{e^{\alpha^i(\rho_t)(\gamma^i_\infty(r^i_k, \rho^{-i}_t) - \gamma^i_\infty(\rho_t))}}{\sum_{j=1}^{K_i} \rho^i_{j,t} e^{\alpha^i(\rho_t)(\gamma^i_\infty(r^i_j, \rho^{-i}_t) - \gamma^i_\infty(\rho_t))}} \rho^i_{k,t}, \tag{1}$$

where $\alpha^i : \mathcal{R} \to \mathbb{R}_+$ is a continuous function, bounded away from zero.

We refer to the dynamics defined in (1) simply as the **average dynamics**. Though we are not committed to this particular form of dynamics (it is studied, e.g., in Cabrales and Sobel (1992)), we use it in the proofs; (we view Camerer and Ho (1999) and Hopkins (2002) as providing some indirect, empirical and theoretical, support for using such a dynamics).

3   Stochastic Learning of Rules

Next, we consider a process of stochastic learning of rules that occurs over games that are drawn randomly from G according to the probability measure μ. In this context, starting from an initial distribution of rules $\rho_0 \in \mathcal{R}$ within the population of players, we consider a learning process that is an application or extension of the aggregate log-monotonic selection dynamics applied to

this stochastic context. The weights with which rules are played are updated according to the relative performances of the rules on the randomly drawn games; such a rule is also referred to as exponential weighted average rule; it is also closely related to the logistic learning rule (e.g., Camerer and Ho (1999)).

Definition 3 *(Discrete) Stochastic Aggregate Log-Monotonic Dynamics on* (G, μ, R)

$$\rho^i_{k,t+1} = \frac{e^{\alpha^i(\rho_t)(\gamma^i_t(r^i_k(\gamma_t), \rho^{-i}_t) - \gamma^i_t(\rho_t))}}{\sum_{j=1}^{K_i} \rho^i_{j,t} e^{\alpha^i(\rho_t)(\gamma^i_t(r^i_j(\gamma_t), \rho^{-i}_t) - \gamma^i_t(\rho_t))}} \rho^i_{k,t}, \tag{2}$$

where $\alpha^i : \mathcal{R} \to \mathbb{R}_+$, *is a positive continuous function, bounded away from zero.*

We refer to the dynamics defined in (2) as the **stochastic dynamics**. Notice that unlike the (deterministic) average dynamics, where the relative performance of a given rule is evaluated with the fixed average game γ_∞, here the relative performance at t is evaluated with the randomly drawn game γ_t.

3.1 Iterated Strict Dominance

Our first result shows that rules that are strictly dominated in the average game tend to disappear under the stochastic dynamics. This is a stochastic counterpart to Nachbar (1990), Friedman (1991), Samuelson and Zhang (1992), and, in particular, Cabrales and Sobel (1992) and Cabrales (2000). Notice that a rule may be strictly dominated in the average game although it never recommends a dominated strategy in any of the randomly drawn games; such a rule would tend to disappear almost surely in the long run.

Proposition 1 *Let* γ_∞ *be the average game for the environment* (G, μ, R). *If* $r^i_k \in R_i$ *is strictly dominated in* γ_∞ *for some* $i \in I$, *and* $\{\rho_t\}$ *follows some stochastic aggregate log-monotonic dynamics with* $\rho_0 \in int(\mathcal{R})$, *then* $\rho^i_{k,t} \xrightarrow{a.s.} 0$.

Further, if $r^{i'}_{k'} \in R_{i'}$ *is iteratively strictly dominated in* γ_∞ *for some* $i' \in I$, *then* $\rho^{i'}_{k',t} \xrightarrow{a.s.} 0$.

Proof. Suppose r_k^i is strictly dominated in γ_∞ by r_ℓ^i. For simplicity, define $\beta^i : \mathcal{R} \times G \to I\!\!R_+$, $\beta^i(\rho, \gamma) = \sum_{j=1}^{K_i} \rho_j^i e^{\alpha^i(\rho)(\gamma^i(r_j^i(\gamma), \rho^{-i}) - \gamma^i(\rho))}$ and denote $X_t = \rho_{k,t}^i$ and $Y_t = \rho_{\ell,t}^i$. By definition, $X_{t+1} = \frac{e^{\alpha^i(\rho_t)(\gamma_t^i(r_k^i(\gamma_t), \rho_t^{-i}) - \gamma_t^i(\rho_t))}}{\beta^i(\rho_t, \gamma_t)} X_t$. Thus, $X_{t+1} = \prod_{s=0}^t \frac{e^{\alpha^i(\rho_s)(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(\rho_s))}}{\beta^i(\rho_s, \gamma_s)} X_0$. A similar expression holds for Y_t. We will show that the ratio $\frac{X_t}{Y_t}$ converges almost surely to zero. This will complete the proof of the first statement.

We have $\frac{X_{t+1}}{Y_{t+1}} = \prod_{s=0}^t \frac{e^{\alpha^i(\rho_s)\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i})}}{e^{\alpha^i(\rho_s)\gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})}} \frac{X_0}{Y_0}$, which is well defined since $\rho_0 \in \text{int}(\mathcal{R})$. Taking the logarithm of both sides yields, $\log(X_{t+1}) - \log(Y_{t+1}) = \sum_{s=0}^t \alpha^i(\rho_s)\Big(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\Big) + \log(X_0) - \log(Y_0)$. By the strong law of large numbers,

$$\frac{1}{t+1} \sum_{s=0}^t \alpha^i(\rho_s)\Big(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\Big) -$$

$$\frac{1}{t+1} \sum_{s=0}^t E\Big[\alpha^i(\rho_s)\Big(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\Big)\Big]$$

goes to zero almost surely. Since the α^i are bounded away from zero, and since, due to the fact the r_k^i is dominated by r_ℓ^i, $\alpha^i(\rho_s)E\Big[\Big(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\Big)\Big]$ are negative and bounded away from zero. Thus, there is $\varepsilon > 0$ such that

$$\limsup_{t\to\infty} \frac{1}{t+1} \sum_{s=0}^t \alpha^i(\rho_s)\Big(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\Big) < -\varepsilon.$$

Therefore, there is $0 < \delta \leq 1$, such that

$$\limsup_{t\to\infty} \frac{X_{t+1}}{Y_{t+1}} = \limsup_{t\to\infty} \prod_{s=0}^t \frac{e^{\alpha^i(\rho_s)\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i})}}{e^{\alpha^i(\rho_s)\gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})}} \frac{X_0}{Y_0} \leq \lim_{t\to\infty} (1-\delta)^{t+1} \frac{X_0}{Y_0} = 0.$$

To prove the case of iterated dominance, suppose $r_{k'}^{i'}$ is strictly dominated in γ_∞ by $r_{\ell'}^{i'}$ but only if say r_k^i is eliminated first. Denote $X_t' = \rho_{k',t}^{i'}$ and $Y_t' = \rho_{\ell',t}^{i'}$. Because payoffs are continuous on $\mathcal{R}$, for s sufficiently large, $\alpha^{i'}(\rho_s)E\Big[\Big(\gamma_s^{i'}(r_{k'}^{i'}(\gamma_s), \rho_s^{-i'}) - \gamma_s^{i'}(r_{\ell'}^{i'}(\gamma_s), \rho_s^{-i'})\Big)\Big]$ are negative and bounded away from zero almost surely. Thus, there is $\varepsilon' > 0$ such that

$$\limsup_{t\to\infty} \frac{1}{t+1} \sum_{s=0}^t \alpha^{i'}(\rho_s)\Big(\gamma_s^{i'}(r_{k'}^{i'}(\gamma_s), \rho_s^{-i'}) - \gamma_s^{i'}(r_{\ell'}^{i'}(\gamma_s), \rho_s^{-i'})\Big) < -\varepsilon'$$

8

and hence, as above, there is $0 < \delta' \leq 1$, such that $\limsup_{t\to\infty} \frac{X'_{t+1}}{Y'_{t+1}} \leq \lim_{t\to\infty} (1 - \delta')^{t+1} \frac{X'_0}{Y'_0} = 0$. Since the spaces R_i are finite, the argument proves the case for arbitrary rounds of iterated strict dominance. $\square$

Starting from an environment (G, μ, R), one can always add new rules to the set of rules R, thus obtaining an expanded environment (G, μ, R') with $R \subset R'$. An important task is to develop criteria for evaluating rules within an environment, with a view to expanded environments. In this sense, we state the next corollary, which is an immediate consequence of the last proposition. It states that rules that prescribe strategies that are strictly dominated on a subset of games of positive measure μ can be dominated (as rules) in some expanded environment.

Corollary 1 *Let (G, μ, R) be an environment such that $r^i_k \in R_i$ prescribes strategies that are strictly dominated for player i on a subset $G' \subset G$ with $\mu(G') > 0$. Then the space of rule profiles can be expanded to R' with $R \subset R'$ such that, if in the environment (G, μ, R') the process $\{\rho_t\}$ follows some aggregate log-monotonic dynamics with $\rho_0 \in int(\mathcal{R}')$, then $\rho^i_{k,t} \xrightarrow{a.s.} 0$.*

Proof. If player i already has a rule $r^i_\ell \in R_i$ that strictly dominates r^i_k in γ_∞, then this is the case of the previous proposition. If not, then consider a rule that prescribes the same strategies as r^i_k on $G \backslash G'$ and prescribes a strictly dominating strategy on G'. Let R be the space of rule profiles with this rule added to R_i. Then since $\mu(G') > 0$, the described rule strictly dominates r^i_k on the corresponding game γ'_∞ and the previous proposition implies $\rho^i_{k,t} \xrightarrow{a.s.} 0$ if $\rho_0 \in int(\mathcal{R})$. $\square$

Example 2. Adding the rules "Best Reply to Naive" (N2), "Best Reply to Altruistic" (A2), "Naive After One Round Dominance" (D1), and "Risk Dominant Nash" (RDN) to the environment of Example 1,[3] leads to the average game

---

[3] D1 recommends N1 in the game resulting from one round of deletion of strictly dominated strategies; RDN recommends the strategy leading to the (generically unique) risk dominant equilibrium (recall these are 2×2 games).

	N1	N2	A1	A2	D1	RDN
N1	.617 / .617	.617 / .667	.707 / .597	.634 / .626	.617 / .642	.617 / .652
N2	.667 / .617	.625 / .625	.696 / .588	.648 / .618	.667 / .641	.636 / .667
A1	.597 / .707	.588 / .696	.712 / .712	.610 / .751	.595 / .708	.593 / .702
A2	.626 / .634	.618 / .648	.751 / .610	.649 / .649	.625 / .659	.622 / .656
D1	.642 / .617	.641 / .667	.708 / .595	.659 / .625	.641 / .641	.644 / .651
RDN	.652 / .617	.667 / .636	.702 / .593	.656 / .622	.651 / .644	**.651 / .651**

where only the rule RDN survives iterated deletion of strictly dominated strategies.[4] Hence, by Proposition 1 all these rules except for RDN would be played with probability zero in the limit of our stochastic dynamics within this expanded environment.

One may suspect from this that rules that do not recommend Nash strategies on some full measure subset of G can be dominated by appropriately designed rules within some expanded environment. However, as the following simple example demonstrates, this is not true in general.

Example 3. Consider the (degenerate) environment where the game below is played repeatedly every period.

3 3	5 2	0 0
1 1	1 0	1 2
0 0	0 2	2 3

This game has three Nash equilibria: two pure, where players play strategies 1 and 3, and a mixed equilibrium where player 1 mixes between strategies 1 and 2 and player 2 mixes between 1 and 3. In particular, player 2's strategy 2 is not a Nash equilibrium strategy. Next, consider the rules for player 1, N1 and N2 (these coincide with strategies 1 and 3 respectively), and for player

---

[4]N1 and A1 are dominated by A2; A2 and D1 are then dominated by RDN and then N2 is dominated by RDN. The rule P1 was omitted since it is dominated already by N1.

2, N1 and A1 (these coincide with strategies 3 and 2 respectively). These lead to the average game

	N1		A1	
N1	0		5	
		0		2
N2	2		0	
		3		2

which has the property that, no matter how one extends the space of rules, it is impossible to do so in a way that will lead to any one of the above rules of players 1 and 2 being strictly (or even weakly) dominated in the resulting extended average game. In particular, player 2's rule A1 that (always) recommends a non-Nash strategy cannot be dominated in any expanded environment. The reason is that, if a new rule for player 2 ever recommends a different strategy than A1, then the payoff of this strategy against either N1 or N2 (or both) will have to decrease. The same applies to the other rules. Moreover, while the environment here consists of a single game that is drawn every period (with probability one), it is easy to see that one can for example take G to be a (sufficiently small) compact neighborhood of the above 3×3 game with any probability measure μ on G, and the above still goes through.

At the same time it is of course true that a rule profile that recommends playing a Nash equilibrium on some full measure subset of G will itself always be a Nash equilibrium profile of the average game γ_∞ whatever the other rules may be.

3.2   Nash Equilibria

We next consider the issue of convergence of the stochastic dynamics to Nash equilibria. Notice that our log-monotonic dynamics is **uncoupled** in the sense of Hart and Mas-Colell (2003b), i.e., the weight put by player i on rule k at time t depends only on the rule profile ρ_{t-1}^{-i} at $t-1$ and on player i's payoffs in the current game γ_t. In particular, it does not directly depend on the payoffs of any of the other players. Hart and Mas-Colell show that such a dynamics cannot in general (and generically in the space of games)

guarantee convergence to Nash equilibrium. Thus, except for specific games, like two-player potential or zero-sum games, one should not expect the log-monotonic dynamics to converge to Nash equilibrium of the average game (in a deterministic sense). However, analogous to Nachbar (1990), Friedman (1991), and Samuelson and Zhang (1992), we show that if the process does converge, then the limit point must be a Nash equilibrium of the average game.

Proposition 2 *Let γ_∞ be the average game for the environment (G, μ, R) and let $\{\rho_t\}$ follow some stochastic aggregate log-monotonic dynamics with $\rho_0 \in int(\mathcal{R})$. Then, if $\rho_t \overset{a.s.}{\longrightarrow} \rho$, then $\rho \in N(\gamma_\infty)$.*

Proof. Suppose $\rho_t \overset{a.s.}{\longrightarrow} \rho$ but $\rho \notin N(\gamma_\infty)$. Then there exists r_k^i in the support of ρ^i, and there exists r_ℓ^i with strictly higher payoff against ρ^{-i} at γ_∞ than r_k^i. Hence, for s sufficiently large, $\alpha^i(\rho_s) E\left[\left(\gamma_s^i(r_k^i(\gamma_s), \rho_s^{-i}) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\right)\right]$ are negative and bounded away from zero almost surely. Thus, as in the previous proof, denoting $X_t = \rho_{k,t}^i$ and $Y_t = \rho_{\ell,t}^i$, one shows that $\limsup_{t\to\infty} \frac{X_t}{Y_t} \le 0$, which contradicts $\rho_t \overset{a.s.}{\longrightarrow} \rho$ and $\rho_k^i > 0$. $\square$

The next result shows that if the initial rule distribution ρ_0 is sufficiently close to a stable pure strategy Nash equilibrium of the average game, then, in the limit, the stochastic dynamics will converge to that equilibrium profile almost surely. We formalize this using the following definition based on Arnold (1974).

Definition 4 *Let $\rho \in \mathcal{R}$ be a zero of the dynamics $\{\rho_t\}$, then we say ρ is* **stochastically stable** *if for every neighborhood V of ρ and for every $\epsilon > 0$, there exists a neighborhood U of ρ, $U \subset V$ of strictly positive measure, such that $P[\rho_t \in V, t > 0] \ge 1 - \epsilon$ whenever $\rho_0 \in U$; ρ is* **stochastically unstable** *if it is not stochastically stable.*

We say ρ is **stochastically asymptotically stable** *if it is stochastically stable and*

$$\lim_{\rho_0 \to \rho} P[\lim_{t\to\infty} \rho_t(\rho_0) = \rho] = 1.$$

Proposition 3 *Let γ_∞ be the average game for the environment (G, μ, R) and let $\{\rho_t\}$ follow some stochastic aggregate log-monotonic dynamics. If $\rho \in \mathcal{R}$ is a regular pure Nash equilibrium of γ_∞ that is asymptotically stable under the corresponding average dynamics, then ρ is stochastically asymptotically stable under $\{\rho_t\}$.*

Proof. It suffices to show that for every $0 < \epsilon < \bar{\epsilon}$ there exists a neighborhood $U_\delta(\rho)$ such that $P[\lim_{t \to \infty} \rho_t = \rho] \geq 1 - \epsilon, \forall \rho_0 \in U_\delta(\rho)$.

Consider the processes $\{\bar{\rho}_t\}$, $\{\rho_t\}$, and $\{\tilde{\rho}_t\}$ defined respectively by (1), (2), and

$$\tilde{\rho}^i_{k,t+1} = \frac{e^{\alpha^i(\rho_t)(\gamma^i_\infty(r^i_k, \rho^{-i}_t) - \gamma^i_\infty(\rho_t))}}{\sum_{j=1}^{K_i} \rho^i_{j,t} e^{\alpha^i(\rho_t)(\gamma^i_\infty(r^i_j, \rho^{-i}_t) - \gamma^i_\infty(\rho_t))}} \rho^i_{k,t}, \tag{3}$$

all with the same initial distribution ρ_0. In the latter process, $\tilde{\rho}_{t+1}$ is computed based on the average game and assuming other players play according to ρ_t; it will serve as an approximation to the stochastic dynamics.

Since ρ is a pure strategy Nash equilibrium, we may assume $\rho^i = (1, 0, .., 0)$, $i \in I$, and since it is a regular and asymptotically stable zero of $\bar{\rho}_t$ there exists an open neighborhood V of ρ and some $\lambda < 1$ such that

$$\|\bar{\rho}_t - \rho\| \leq \lambda^t \|\rho_0 - \rho\|, \quad \forall \rho_0 \in V, t \geq 0,$$

where $\|\cdot\|$ is the maximum norm. Moreover, as long as $\rho_\tau \in V$, $\tau < t$, we also have

$$\|\tilde{\rho}_t - \rho\| \leq \lambda^t \|\rho_0 - \rho\|, \quad \forall \rho_0 \in V, t \geq 0. \tag{4}$$

Put $\delta_0 = \|\rho_0 - \rho\| \leq \delta$. Then (4) implies $\tilde{\rho}^i_{k,t} \leq \lambda^t \delta_0, \forall k \neq 1, i \in I$.

Since G is compact, the factor by which the weight put on any strategy k can grow from a period to the next is bounded. Next, choose T sufficiently large such that

$$P[\|\rho_{t+1} - \rho_t\| + \|\rho_{t+2} - \rho_t\| \leq (\lambda e^\epsilon)^{\frac{T}{2}} \|\rho_t - \rho\|] = 1. \tag{5}$$

Because of the multiplicative form of the dynamics, for any given $T > 0$ we can choose $\delta > 0$ sufficiently small such that

$$P[\rho_\tau(\rho_0) \in V, \tau \leq T + 1] = 1, \quad \forall \rho_0 \in U_\delta(\rho). \tag{6}$$

13

Finally, the process $\{\log(\rho_t) - \log(\tilde{\rho}_t)\}$ is a martingale difference sequence to which the Hoeffding-Azuma inequality applies, so that assuming $\rho_\tau = \tilde{\rho}_\tau$, $\tau \in I\!N$, we have

$$P[\frac{1}{T_\nu}\|\log(\rho_{\tau+T_\nu}) - \log(\tilde{\rho}_{\tau+T_\nu})\| \leq \epsilon] > 1 - e^{-\frac{T_\nu \epsilon^2}{2}}, \quad \forall \epsilon, T_\nu > 0. \quad (7)$$

In particular, this implies $\rho^i_{k,\tau+T_\nu}/\tilde{\rho}^i_{k,\tau+T_\nu} \leq e^{\epsilon T_\nu}$, $k \neq 1, i \in I$, with probability at least $1 - e^{-\frac{T_\nu \epsilon^2}{2}}$.

We now show that the process $\{\rho_t\}$ starting at $\rho_0 \in U_\delta$ satisfies $P[\lim_{t\to\infty} \rho_t = \rho] \geq 1 - \epsilon$. Notice that by taking $\bar{\epsilon} = \log(\lambda^{-1}) > 0$, we can guarantee that $\lambda e^\epsilon < 1$. We consider time as running in blocks of length $T_\nu = T + \nu$ and show that after each block the process $\{\rho_t\}$ gets, in a precise sense, closer to ρ with a probability that rapidly converges to 1.

Start with $\nu = 1$. Starting the process $\{\rho_t\}$ at ρ_0, we have by (6), $P[\rho_\tau(\rho_0) \in V, \tau \leq T_1] = 1$, and by (4), $\tilde{\rho}^i_{k,T_1} \leq \lambda^{T_1}\delta_0$, $k \neq 1, i \in I$. By Hoeffding-Azuma's inequality (7), we have $\rho^i_{k,T_1}/\tilde{\rho}^i_{k,T_1} \leq e^{\epsilon T_1}$, and therefore

$$\rho^i_{k,T_1} \leq \lambda^{T_1} e^{\epsilon T_1} \delta_0 = (\lambda e^\epsilon)^{T_1} \delta_0, \quad k \neq 1, i \in I,$$

with probability at least $1 - e^{-\frac{T_1 \epsilon^2}{2}}$.

Next, consider $\nu = 2$. Starting the process $\{\rho_t\}$ at the exogenous point $\rho_0(T_1)$ defined by $\rho^i_{k,0}(T_1) = (\lambda e^\epsilon)^{T_1}\delta_0$, $k \neq 1, i \in I$, we have by (5) that after $T_2 = T + 2$ periods, $P[\rho_t(\rho_0(T_1)) \in V, t \leq T + 2] = 1$, (since after two periods $\rho_2(\rho_0(T_1)) \in U_\delta(\rho_0)$, by (5), so that after further T periods $\rho_{T_2}(\rho_0(T_1)) \in V$ a.s. by (6)). Hence by (4) $\tilde{\rho}^i_{k,T_2}(\rho_0(T_1)) \leq \lambda^{S_2}\delta_0$. Again, by (7), we have $\rho^i_{k,T_2}(\rho_0(T_1))/\tilde{\rho}^i_{k,T_2}(\rho_0(T_1)) \leq e^{\epsilon T_2}$, and therefore

$$\rho^i_{k,T_2}(\rho_0(T_1)) \leq (\lambda e^\epsilon)^{S_2}\delta_0, \quad \forall k \neq 1, i \in I,$$

with probability at least $1 - e^{-\frac{T_2 \epsilon^2}{2}}$.

Finally, consider an arbitrary $\nu \in I\!N$. Put $S_\nu = \sum_{\nu'=1}^\nu T_{\nu'}$. Starting the process $\{\rho_t\}$ at the exogenous point $\rho_0(S_{\nu-1})$ defined by $\rho^i_{k,0}(S_\nu) = (\lambda e^\epsilon)^{S_{\nu-1}}\delta_0$, $k \neq 1, i \in I$, we have that after $T_\nu = T + \nu$ periods, by (5), $P[\rho_t(\rho_0(S_{\nu-1})) \in V, t \leq T_\nu] = 1$. With (4) we then have that, $\tilde{\rho}^i_{k,T_\nu}(\rho_0(S_{\nu-1})) \leq$

14

$\lambda^{S_\nu}\delta_0$. Hence, again by (7), we have $\rho^i_{k,T_\nu}(\rho_0(S_{\nu-1}))/\tilde\rho^i_{k,T_\nu}(\rho_0(S_{\nu-1})) \le e^{\epsilon T_\nu}$, and therefore

$$\rho^i_{k,T_\nu}(\rho_0(S_{\nu-1})) \le (\lambda e^\epsilon)^{S_\nu}\delta_0, \quad \forall k \ne 1, i \in I,$$

with probability at least $1 - e^{-\frac{T_\nu \epsilon^2}{2}}$.

But because at each ν the process starts anew at the deterministic point $\rho_0(S_{\nu-1})$, each of the events are independent and hence the Hoeffding-Azuma inequality is applied ν independent times, so that with probability at least

$$\big(1 - e^{-\frac{T_1\epsilon^2}{2}}\big)\big(1 - e^{-\frac{T_2\epsilon^2}{2}}\big)\cdots\big(1 - e^{-\frac{T_\nu\epsilon^2}{2}}\big) = \Pi^\nu_{\nu'=1}\big(1 - e^{-\frac{T_{\nu'}\epsilon^2}{2}}\big)$$

the actual process $\{\rho_t\}$ starting at ρ_0 satisfies

$$\rho^i_{k,S_\nu}(\rho_0) \le (\lambda e^\epsilon)^{S_\nu}\delta_0, \quad \forall k \ne 1, i \in I.$$

It is now easy to see that the process is stochastically asymptotically stable. $\square$

One positive feature of the present framework is that convergence of the process $\{\rho_t\}$ to a pure strategy rule profile can be interpreted as the players learning to play the actual rules (or algorithms) corresponding to the limiting rule profile. In particular, if for example the limiting rule corresponds to say playing the payoff dominant Nash equilibrium strategy, then this means learning (or converging to) the actual algorithm that prescribes playing the payoff dominant Nash equilibrium strategy for every game drawn from G. Moreover, in some cases, rules may be applicable even to games outside of G. The next example is adapted from the experiments of Rankin et al. (2000) and Stahl and Van Huyck (2002), who study the evolution of subjects' behavior within a class of stag hunt games.[5]

Example 4. Take G to be the space of all 2×2 games of the form

$a+e$		e	
	$a+e$		$b+e$
$b+e$		$b+e$	
	e		$b+e$

$b+e$		$b+e$	
	$b+e$		e
e		$a+e$	
	$b+e$		$a+e$

[5]I thank Vince Crawford and John Van Huyck for pointing out these references.

where $a = 1$, $b \in [0, 1]$, $e \in [0, \frac{1}{8}]$, and where b and e are drawn uniformly and independently from their respective ranges and where the left or right payoff matrix is drawn with probability one half. Consider the two rules, "Payoff Dominant Nash" (PDN) and "Risk Dominant Nash" (RDN). The average game is

	PDN		RDN	
PDN	1.063		.562	
		1.063		.937
RDN	.937		.937	
		.562		.937

The rule profiles PDN and RDN are both asymptotically stable under the average dynamics. Hence Proposition 3 says that, if initial propensities to play say PDN are sufficiently high, then the process will converge with high probability to the entire population playing PDN. The probability increases the higher is the initial propensity of playing PDN, and, given an initial propensity, whether the stochastic dynamics converges to PDN or RDN depends on the actual sequence of games drawn.[6]

The following example shows that Proposition 3 does not hold for strictly mixed equilibria that are asymptotically stable in the average game.

Example 5. Consider the following average game obtained from the environment (G, μ, R), where G is the space of all 3×3 games with payoffs in $[0, 1]$ and μ is again uniform.

---

[6]Rankin et al. (2000) obtain that between 80 and 98% of their subjects play PDN after about 75 rounds, starting from initial propensities that seem to be only about 45 to 75%, see Table 1, p. 322; somewhat strikingly, the pattern is repeated for all of their six cohorts; notice that convergence to PDN is not guaranteed with our dynamics from such initial propensities. One possible explanation is that, in their experiment, all six cohorts play the *same* sequence of randomly drawn payoff matrices; as our model suggests, the actual sequence matters for the outcome of the evolutionary process; it could be that their drawn sequence is more likely to lead to PDN than to RDN. Stahl and Van Huyck (2002) obtain *less* convergence to PDN with payoff matrices where essentially $b \in [\frac{1}{2}, 1]$; given initial propenstities, the probability of converging to RDN under our dynamics is also higher for this latter class of games.

	D1	N2
D1	.660	.660
	.660	.724
N3	.650	.750
	.660	.640

This game has a unique mixed equilibrium, which is asymptotically stable under the average dynamics if for example $\alpha^1 = \alpha^2 = 1$. However, it can be checked that, the stochastic dynamics, even if it starts at the mixed equilibrium, leaves any sufficiently small neighborhood with probability one.

Since strict Nash equilibria are always in pure strategies and asymptotically stable under the log-monotonic selection dynamics, we have the following corollary, analogous to Nachbar (1990) and Ritzberger and Weibull (1995).

Corollary 2 *Let γ_∞ be the average game for the environment (G, μ, R) with $\rho \in \mathcal{R}$ a strict Nash equilibrium of γ_∞, and suppose $\{\rho_t\}$ follows some stochastic aggregate log-monotonic dynamics. Then ρ is stochastically asymptotically stable under $\{\rho_t\}$.*

3.3 Empirical Frequencies and the Hannan Set

We next study the evolution of the **empirical frequencies** of the stochastic dynamics. These are defined as $p_{\nu,t} = \frac{1}{t} \sum_{s=0}^{t} \pi_{\nu,t}$, where $\pi_{\nu,t} = \prod_{i \in I} \rho^i(\nu)$ and $\rho^i(\nu) \in [0, 1]$ is the probability player i puts on the rule in R_i that leads to outcome $\nu \in \{1, \ldots, \kappa_R\}$, where $\kappa_R = \Pi_i \# R_i$ is the number of outcomes of the average game γ_∞. A reasonable candidate for the limit set of the empirical frequencies is the **Hannan set** of the game γ_∞, which, following Hart and Mas-Colell (2003a), we define as

$$H(\gamma_\infty) = \{p \in \Delta(R) \mid \gamma_\infty^i(p) \geq \gamma_\infty^i(r_k^i, p^{-i}), \ \forall r_k^i \in R^i, \forall i \in I\},$$

where $p^{-i} \in \Delta(R^{-i})$ is the marginal of p on R^{-i}.[7]

---

[7]The Hannan set is a compact, convex polyhedron defined by (linear) inequalities that are sums of the ones defining the correlated equilibria. It contains the set of correlated equilibria and it can be shown that, unlike the correlated equilibria, outcomes in its support need not be rationalizable and may involve strictly dominated strategies; further, not all rationalizable outcomes are in the support of the Hannan set.

Analogous to Proposition 2, we show that, if the empirical frequencies do converge, then the limit point must be in the Hannan set.

Proposition 4 *Let γ_∞ be the average game for the environment (G, μ, R) and let $\{\rho_t\}$ follow some stochastic aggregate log-monotonic dynamics with $\rho_0 \in int(\mathcal{R})$. Then, if $p_t \xrightarrow{a.s.} p$, then $p \in H(\gamma_\infty)$.*

Proof. Suppose $p_t \xrightarrow{a.s.} p$ but $p \notin H(\gamma_\infty)$. Then there exists r_ℓ^i with strictly higher payoff against p^{-i} at γ_∞ than playing according to p. (Notice that p must place positive probability on at least one outcome where i uses another strategy besides r_ℓ^i.) Hence, again, for s sufficiently large, $\alpha^i(\rho_s)E\left[\left(\gamma_s^i(\rho_s) - \gamma_s^i(r_\ell^i(\gamma_s), \rho_s^{-i})\right)\right]$ are negative and bounded away from zero almost surely. Thus, denoting $X_t = 1 - \rho_{\ell,t}^i$ and $Y_t = \rho_{\ell,t}^i$, one shows that $\limsup_{t\to\infty} \frac{X_t}{Y_t} \le 0$. Hence the probability of outcomes involving strategies other than r_ℓ^i converges to zero, which contradicts $p_t \xrightarrow{a.s.} p$. $\square$

Combining this with Proposition 1 we see that if empirical frequencies converge then the limit will be in the Hannan set and its support will consist of strategies that survive iterated deletion of strictly dominated strategies.

Hart and Mas-Colell (2001) provide a large class of adaptive dynamics (based on average regrets) whose empirical frequencies converge to the Hannan set (i.e., are Hannan consistent, or universally consistent, see Fudenberg and Levine (1998)). It can be checked that our log-monotonic dynamics are based on cumulative and not average regrets, so that Hart and Mas-Colell's proof does not apply directly. It is still an open question whether or not our dynamics are Hannan consistent. What we can say is that empirical frequencies will not in general converge to the set of correlated equilibria.

Example 6. Consider the (degenerate) environment where the game below is played repeatedly every period with the rules corresponding to the pure strategies of the game.

1 1	3 0	0 3
0 3	1 1	3 0
3 0	0 3	1 1

It can be shown that the log-monotonic dynamics, for example with $\alpha^1 = \alpha^2 = 1$ and the initial condition $\rho_0^1 = (\frac{3}{4}, \frac{1}{8}, \frac{1}{8})$ and $\rho_0^2 = (\frac{1}{8}, \frac{3}{4}, \frac{1}{8})$, has the property that the empirical frequencies of the outcomes on the diagonal converge to zero. In particular, the empirical frequencies do not converge to the (unique) correlated equilibrium, which puts probability $\frac{1}{9}$ on every outcome.

While the log-monotonic dynamics is based on (cumulative) regrets it does not keep track of pairwise comparisons of regrets between different strategies, i.e., it does not keep track of conditional regrets (see e.g., Hart and Mas-Colell (2001)), which may be the reason preventing convergence to correlated equilibria.

4   Conclusion

The present framework can be extended in many ways. Some obvious ones consist in dropping some of the stationarity assumptions built into the model. For instance, one could consider rules that are contingent on past behavior (as for example in Stahl (1999, 2000) and Stahl and Van Huyck (2002); it also plays an important role in Jehiel (2003)); one may also allow the distribution μ to change over time. On the other hand, one could generalize the class of dynamics and test whether folk results shown for the present evolutionary dynamics carry over to further classes, for example, to more sophisticated learning or heuristic dynamics like fictitious play type dynamics (see Fudenberg and Levine (1998)) or regret based dynamics (see Hart and Mas-Colell (2001)). In order to extend the results (essentially, a close link between the stochastic and the average dynamics) one needs to check that the dynamics depends in a sufficiently linear fashion on the history of play, so that the law of large numbers can be applied.

However, it seems that some of the main challenges lie in characterizing "good" rules that ideally apply to a wide range of games and environments, and linking them to actual cognitive (or genetic) behavior. We view this paper as a first step towards such a broader and deeper analysis. Another aspect that has not been touched on here is the modelling of the process of learning, conceiving, or developing rules without previous knowledge of the set of rules. The experiments of Selten et al. (2003), where students had to develop algorithms for playing randomly drawn 3×3 games as well as some of the experiments mentioned in the paper, focusing on learning of rules, for example, Rankin et al. (2000) and Stahl and Van Huyck (2002), are useful for this. The neural network approach of Zizzo and Sgroi (2001, 2003), while conceptually quite different from the one of the present paper, may also help understand possible cognitive processes underlying the decision processes of subjects having to play in different environments, especially for intuition-based decision making; Zizzo and Sgroi (2001) also provide evidence that neural networks may play in similar ways to some of the subjects, for example in the Stahl (1999, 2000) and Costa-Gomes et al. (2001) experiments.

References

[1] Arnold, L. (1974) *Stochastic Differential Equations: Theory and Applications,*" Krieger Publishing, Malabar, Florida.

[2] Benaim, M., and M.W. Hirsch (1999) "Mixed Equilibria and Dynamical Systems Arising from Fictitious Play in Perturbed Play," *Games and Economic Behavior*, **29**: 36-72.

[3] Cabrales, A. (2000) "Stochastic Replicator Dynamics," *International Economic Review*, **41**: 451-481.

[4] Cabrales, A., and J. Sobel (1992) "On the Limit Points of Discrete Selection Dynamics," *Journal of Economic Theory*, **57**: 407-419.

[5] Camerer, C., and T.H. Ho (1999) "Experience-Weighted Attraction Learning in Normal-Form Games," *Econometrica*, **67**: 827-874.

[6] Costa-Gomes, M., V. Crawford, and B. Broseta (2001) "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, **69**: 1193-1235.

[7] Foster, D., and H.P. Young. (1990) "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, **38**: 219-232.

[8] Friedman, D. (1991) "Evolutionary Games in Economics," *Econometrica*, **59**: 637-666.

[9] Fudenberg, D., and C. Harris (1992) "Evolutionary Dynamics with Aggregate Shocks," *Journal of Economic Theory*, **57**: 420-441.

[10] Fudenberg, D., and D. Kreps (1993) "Learning Mixed Equilibria," *Games and Economic Behavior*, **5**: 320-367.

[11] Fudenberg, D., and D. Levine (1998) *The Theory of Learning in Games*, MIT Press, Cambridge, MA.

[12] Hart, S., and A. Mas-Colell (2001) "A General Class of Adaptive Strategies," *Journal of Economic Theory*, **98**: 26-54.

[13] Hart, S., and A. Mas-Colell (2003a) "Regret-Based Continuous-Time Dynamics," *Games and Economic Behavior*, **45**: 373-394.

[14] Hart, S., and A. Mas-Colell (2003b) "Uncoupled Dynamics Do Not Lead to Nash Equilibrium," *American Economic Review*, **93**: 1830-1836.

[15] Heller, D. (2004) "An Evolutionary Approach to Learning in a Changing Environment," *Journal of Economic Theory*, **114**: 31-55.

[16] Hofbauer, J., and W.H. Sandholm (2002) "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, **70**: 2265-2294.

[17] Hofbauer, J., and K. Sigmund (2003) "Evolutionary Game Dynamics," *Bulletin of the American Mathematical Society*, **40**: 479-519.

[18] Hopkins, E. (2002) "Two Competing Models of How People Learn in Games," *Econometrica*, **70**: 2141-2166.

[19] Jehiel, P. (2003) "Analogy-Based Expectation Equilibrium," Mimeo, CERAS, Paris, and University College London.

[20] Kandori, M., Mailath, G.J., and R. Rob (1993) "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, **61**: 29-56.

[21] Li Calzi, M. (1995) "Fictitious Play by Cases," *Games and Economic Behavior*, **11**: 64-89.

[22] Nachbar, J.H., (1990) "'Evolutionary' Selection Dynamics in Games: Convergence and Limit Properties," *International Journal of Game Theory*, **19**: 59-90.

[23] Rankin, F.W., J.B. Van Huyck, and R.C. Battalio (2000) "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games," *Games and Economic Behavior*, **32**: 315-337.

[24] Rosenthal, R.W. (1979) "Sequences of games with varying opponents," *Econometrica*, **47**: 1353-1366.

[25] Rosenthal, R.W. (1993a) "Rules of thumb in games," *Journal of Economic Behavior and Organization*, **22**: 1-13.

[26] Rosenthal, R.W. (1993b) "Bargaining rules of thumb," *Journal of Economic Behavior and Organization*, **22**: 15-24.

[27] Samuelson, L. (2001) "Analogies, Adaptation, and Anomalies," *Journal of Economic Theory*, **97**: 320-366.

[28] Samuelson, L., and J. Zhang (1992) "Evolutionary Stability in Asymmetric Games," *Journal of Economic Theory*, **57**: 363-391.

[29] Selten, R., K. Abbink, J. Buchta, and A. Sadrieh (2003) "How to play 3×3 games: A strategy method experiment," *Games and Economic Behavior*, **45**: 19-37.

[30] Sgroi, D., and D.J. Zizzo (2003) "Strategy Learning in 3×3 Games by Neural Networks," Mimeo, University of Cambridge and Oxford University.

[31] Stahl, D.O. (1999) "Evidence based rules and learning in symmetric normal-form games," *International Journal of Game Theory*, **28**: 111-130.

[32] Stahl, D.O. (2000) "Rule Learning in Symmetric Normal-Form Games: Theory and Evidence" *Games and Economic Behavior*, **32**: 105-138.

[33] Stahl, D.O., and J.B. Van Huyck (2002) "Learning Conditional Behavior in Similar Stag Hunt Games," Mimeo, University of Texas, Austin, and Texas A&M University.

[34] Stahl, D.O., and P.W. Wilson (1995) "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, **10**: 218-254.

[35] Van Huyck, J.B., and R.C. Battalio (2002) "Prudence, Justice, Benevolence, and Sex: Evidence from Similar Bargaining Games," *Journal of Economic Theory*, **104**: 227-246.

[36] Zizzo, D.J., and D. Sgroi (2001) "Bounded-Rational Behavior by Neural Networks in Normal Form Games," Mimeo, University of Cambridge and Oxford University.