

Learning in Games by Random Sampling^α

Department of Economics Working Paper 99-04

James W. Friedman

Department of Economics

University of North Carolina

Chapel Hill, NC 27599-3305

email: Jim_Friedman@unc.edu

Claudio Mezzetti

Department of Economics

University of North Carolina

Chapel Hill, NC 27599-3305

email: mezzetti@email.unc.edu

November 16, 1999

Abstract

We study repeated interactions among a fixed set of “low rationality” players. Each player has a status quo action. Occasionally, he randomly samples other actions and changes his status quo if the sampled action yields a higher payoff. This behavior generates a random process, the better-reply dynamics.

^αWe thank seminar participants at Boston and Northwestern Universities for stimulating comments and particularly Russell Cooper, Eddie Dekel and Bob Rosenthal. We thank the associate editor and an anonymous referee for helpful comments.

In the long run the behavior we describe leads to Nash equilibrium in games with the weak finite improvement property. We show that finite, supermodular games and generic, continuous, two-player, quasi-concave games, have this property. If players occasionally change strategies when sampling does not improve payoff (i.e., make mistakes) and if several players can sample at the same time, the resulting better-reply dynamics with simultaneous sampling converges to the Pareto optimal Nash equilibrium in common interest games. In general games, convergence to Nash equilibrium need not occur.

1 Introduction

In many situations people do not strategize, even though they are aware of interacting with others, but instead follow simple decision rules that react to the environment. Consider, for example, an individual who is a “player” in several different “games” at any given time. These games can vary greatly in their importance to the player and in the information he has. People often don’t know others’ payoffs; they may not know the exact nature of their strategic interactions with others; they may not know how smart other players are; and they may not know how many other players there are. Acquiring such information may be possible, but it is rarely costless due to limitations to a player’s intelligence, computational ability, and memory. As a result, a player may not fully understand or find it worthwhile to devote full attention to some of the games he plays; that is, in some games he may choose “not to play games”.

While fully rational behavior is a well understood paradigm, with precise modeling rules, a unified and general approach to modeling bounded rationality has not emerged. Evolutionary models, adaptive behavior, myopic but otherwise rational behavior, have all been studied extensively in recent years and have helped increase our understanding of human behavior, but none of them is immune to criticism.¹ Experimental research in economics and psychology lends support to the belief that people do not always act in accordance with the full rationality paradigm. One of the strongest regularities from experimental studies is that actions that were successful in the past tend to be repeated in the future (e.g., see Roth and Erev [19]). In this paper we study one of the simplest and least sophisticated stimulus-response models of this kind, in which players follow a naive, gradient based decision rule that allows

¹For an overview of the evolutionary game literature the interested reader should consult the surveys by Weibull [24], Samuelson [20], Vega-Redondo [21], and Young [26]. Examples of papers that focus on Bayesian or intelligently adaptive learning are Jordan [8] and [9], Blume and Easley [2], Kalai and Lehrer [10], Milgrom and Roberts [15], Fudenberg and Kreps [6], and Ellison and Fudenberg [4]; two useful overviews of this literature are Fudenberg and Levine [7], and Marimon [12].

for occasional sampling.

In our model, a fixed set of players repeatedly interact; each habitually chooses a particular strategy (the status quo action) and associates with it the payoff resulting from the players' status quo action profile. Occasionally a player samples another strategy, compares the two payoffs, and switches his status quo action if the sampled strategy yields a higher payoff. The behavior of our players generates a stochastic process that moves in the direction of better replies, the better-reply dynamics. Such a process need not be ergodic; however, if a small probability of mistakes is introduced, it is ergodic.

The better-reply dynamics, permitting no more than one player at a time to sample and without the possibility of mistakes, converges to a Nash equilibrium in any finite game having the property that from any action profile there is a finite sequence of single-player improvements leading to a Nash equilibrium (games having the weak finite improvement property). Such games include dominance solvable games, quasi-acyclic games (similar to acyclic games as defined by Young [25]), and games with the finite improvement property (as defined by Monderer and Shapley [16]). We show that finite, supermodular games and generic, continuous, two-player, quasi-concave games also have the weak finite improvement property (in brief, weak FIP).²

In finite games lacking the weak FIP, play need not converge to a Nash equilibrium. We present an example in which a strictly dominated strategy survives, and an example in which play converges to a set of action profiles that does not include the support of the unique, mixed strategy Nash equilibrium of the game.

There are important differences between best-reply and better-reply dynamics. On the one hand, in any game strictly dominated strategies are always eliminated by the best-reply dynamics, but they need not be eliminated by the better-reply dynamics. On the other hand, the best-reply dynamics need not converge to a Nash equilibrium in games with the weak FIP, but the better-reply dynamics always converges to a

²In fact, we prove the stronger result that supermodular games are quasi-acyclic.

Nash equilibrium in such games.

In contrast to most other models of boundedly rational behavior, our players make no attempts at forecasting either aggregate or individual behavior, and thus need little information or memory capacity. In fact, our players need not even know their own payoff functions; they must merely know the payoff they receive from the particular action they currently choose. Thus, the bound on players' rationality is more stringent in our paper than in most of the literature. Evolutionary models typically postulate that a given game is played by large populations of player types, each type always playing in the same role. Our model differs from these evolutionary models in its focus on the behavior over time of individual players, rather than the average behavior of populations of player types.

In the next section we introduce the model and the better-reply dynamics. In Section 3 we show that finite supermodular games and generic, continuous, two-player, quasi-concave games have the weak FIP. In Section 4 we add mistakes to the better reply dynamics by postulating that a player who samples and sees a payoff no better than the status quo payoff may make a mistake and switch. In this section we also introduce a modified better-reply dynamics with mistakes in which more than one player may sample at the same time. We show that in these dynamics, as the sampling and mistake probabilities go to zero, play can only converge to states that belong to a recurrent class of the better-reply dynamics without mistakes, implying that they converge to a Nash equilibrium in finite games with the weak FIP. This section also compares the better and the best-reply dynamics. In Section 5 we prove that in the better-reply dynamics with mistakes and simultaneous sampling play converges to the Pareto dominant Nash equilibrium in common interest games, which include coordination games.³ Section 6 contains concluding remarks.

³Common interest games have a strategy profile whose associated payoff profile strictly dominates that of all other strategy profiles.

2 The Model

We begin by describing the game structure and then introduce the better-reply dynamics. The stage game is $g = (N; A; \mathcal{U}_i)$, where $N = \{1; \dots; n\}$ is the set of players, $\mathcal{U}_i = (\mathcal{U}_{i1}; \dots; \mathcal{U}_{in})$ is the vector of the players' payoff functions, $A = \prod_{i \in N} A_i$ is the set of action profiles with typical element a , and $A_{-i} = \prod_{j \in N \setminus \{i\}} A_j$. In most of the paper we restrict attention to finite strategy spaces, so that $A_i = \{1; \dots; m_i\}$. However, in Section 3.3 A_i is a compact interval on the real line. Throughout the paper certain notational conventions will be followed. Let $X \subseteq \mathbb{R}^k$, $Y \subseteq X$, and $x; y \in X$. Then $X_{-i} = \prod_{j \in N \setminus \{i\}} X_j$ with typical element x_{-i} , $x_{-i} = (x_{-i1}; \dots; x_{-in})$, and $X \times Y = \{x \in X : x \in Y\}$. The inequalities $x \leq y$, $x \geq y$, and $x > y$ mean, respectively, $x_i \leq y_i$ for all i , $x_i \geq y_i$ for all i , and $x \leq y$ but $x \neq y$.

Under the better-reply dynamics, at any time player i has a status quo action a_i , which the player uses almost always. The status quo action profile is $a = (a_1; \dots; a_n)$ and the payoff player i associates with the status quo action profile is $\mathcal{U}_i(a)$. Occasionally, the player samples another action $a_i^E \in A_i \setminus \{a_i\}$. The sampled action becomes the new status quo if it provides a payoff higher than a_i .

Definition 1 Better-Reply Dynamics. At each time period, one player $i \in N$ is randomly selected with all players having the same selection probability; player i randomly samples action $a_i^E \in A_i \setminus \{a_i\}$ with probability $(m_i - 1)^{-1}$ divided equally among the elements of $A_i \setminus \{a_i\}$ and switches his status quo action to a_i^E if and only if $\mathcal{U}_i(a_{-i}; a_i^E) > \mathcal{U}_i(a_{-i}; a_i)$.⁴

The better-reply dynamics generates a Markov process on the finite state space A of action profiles. Let $p(a^0; a^0; \cdot)$ be the transition probability from state a^0 to state

⁴In defining player i 's status quo payoff to be $\mathcal{U}_i(a)$, the better-reply dynamics does not keep track of the occasional observations of a different payoff when some other player is sampling while a is the status quo action. This is a shortcut that captures the intuition that each player i gets accustomed to seeing $\mathcal{U}_i(a)$ when a is the status quo action profile. Taking the long route of precisely formalizing the details of players' behavior is possible, but would require expanding the state space of the stochastic process beyond A . The resulting tedious formulation would not change our results. Indeed, the results in all sections except Section 5 are insensitive to the way that players revise the status quo payoff.

a^0 , and $P(\cdot)$ be the associated Markov transition matrix. A recurrent class $S \subseteq A$ of $P(\cdot)$ is a set of states such that (i) if s and s^0 belong to S , then there is a positive probability of moving from s to s^0 in a finite number of periods; (ii) if $s \in S$ and $s^0 \notin S$, then the probability of moving from s to s^0 in a finite number of periods is zero. A singleton recurrent class is called an absorbing state. A recurrent class with more than one element contains cycles among strategy profiles. Denote by $T \subseteq A$ the set of states, called transitory states, that belong to no recurrent class. If the stochastic process is at any $s \in T$, then it will leave T in a finite number of steps and will never return to any state in T .

If A is the unique recurrent class of $P(\cdot)$, then the stochastic process is irreducible, otherwise it is reducible. For most games, the Markov chain $P(\cdot)$ associated with the better-reply dynamics is reducible and the stochastic process is not ergodic; the recurrent class to which it will converge depends on history.

3 The Weak FIP and Convergence to Nash

This section is divided into three subsections. In Section 3.1 the weak finite improvement property (weak FIP) is introduced and it is shown that the better reply dynamics must converge to a pure strategy Nash equilibrium in any game having the weak FIP. We then note that dominance solvable and potential games have the weak FIP. In Section 3.2 we prove that supermodular games are weakly acyclic which implies they have the weak FIP and in Section 3.3 we prove that generic, continuous, two-player, strictly quasi-concave games have the weak FIP. Thus, the better reply dynamics converges to a pure strategy Nash equilibrium dominance solvable games, potential games, supermodular games, and almost all two-player strictly quasi-concave games.

3.1 The Weak FIP

A state (strategy profile) $a^0 \in A$ is a single player improvement over the state $a \in A$ if it coincides with a in every coordinate except one, say coordinate i , and the payoff to player i is higher under a^0 than under a . Loosely speaking, $a^0 \succ_i a$ is a single player improvement over a if player i would gain by switching from a_i to a_i^0 . This and the weak FIP are formally captured in the following definition.

Definition 2 Let $g = \langle N; A; \{u_i\} \rangle$ be a game. (i) The state $a \in A$ is a single-player improvement over $a' \in A$ if and only if $u_i(a) > u_i(a')$. (ii) The game $g = \langle N; A; \{u_i\} \rangle$ has the weak finite improvement property (weak FIP) if from each action profile $a \in A$ there exists a finite sequence of single-player improvements that ends in a pure strategy Nash equilibrium.

The better-reply dynamics eventually converges to a pure strategy Nash equilibrium in finite games with the weak FIP. Which equilibrium is reached may depend on the initial state of the process; however, any of the pure strategy equilibria may be reached.⁵

Lemma 1 Let $g = \langle N; A; \{u_i\} \rangle$ be a finite game that has the weak FIP. Then the recurrent classes of the Markov process generated by the better reply dynamics are all absorbing states, each of which is a pure strategy Nash equilibrium of g . Every pure strategy Nash equilibrium of g is an absorbing state.

Proof. Let $a \in A$ be a pure strategy Nash equilibrium of g . Then there are no single player improvements from a ; hence, a is an absorbing state. If a is not a pure strategy Nash equilibrium, then a cannot be an absorbing state; however, it remains to be seen whether a can belong to a (non-degenerate) recurrent class. If such a recurrent class

⁵Theorem 5 in Section 4 guarantees that if a game has the weak FIP, then the limiting stationary distribution of the better-reply dynamics with mistakes (see Definition 4) puts positive mass only on states $a \in A$ that are Nash equilibria. Not all Nash equilibria need to have positive mass.

	L	C	R
T	3;3	2;2	2;2
M	2;2	1;1	1;1
B	2;2	1;1	1;1

Table 1: A Game with the Weak FIP, but not the FIP

$S = \{a^1, \dots, a^k\}$ existed, it would correspond to a cycle of the better reply dynamics; there would be a sequence of single player improvements that would go, without loss of generality, from a^1 to a^2 , ..., to a^k to a^1 . Having the weak FIP implies that from at least one of the $a^j \in S$ there is a single player improvement to some $s^0 \notin S$; however, this contradicts S being a recurrent class and completes the proof. ■

A finite game $g = (N; A; u)$ is dominance solvable if there is a unique strategy profile $a^* \in A$ that survives the iterated elimination of strictly dominated strategies.⁶ The surviving strategy profile a^* in a dominance solvable game is the unique Nash equilibrium of g . Clearly, finite, dominance solvable games have the weak FIP.

The game g is an ordinal potential game if there is a function $\phi : A \rightarrow \mathbb{R}$ that satisfies

$$u_i(a) - u_i(a_{-i}^0, a_i^1) > 0 \iff \phi(a) - \phi(a_{-i}^0, a_i^1) > 0$$

for all $i \in N$, $a \in A$, and $a_i, a_i^0 \in A_i$. In an ordinal potential game each player's payoff function is strategically equivalent to some common function ϕ . The game g is a potential game if ϕ satisfies the stronger condition

$$u_i(a) - u_i(a_{-i}^0, a_i^1) = \phi(a) - \phi(a_{-i}^0, a_i^1)$$

for all $i \in N$, $a \in A$ and $a_i, a_i^0 \in A_i$.

⁶A pure strategy $a_i \in A_i$ for player i is strictly dominated by another strategy $x_i \in A_i$ if, for all $a_{-i} \in A_{-i}$, $u_i(a_{-i}, x_i) > u_i(a_{-i}, a_i)$.

Monderer and Shapley [16] defined the finite improvement property (FIP) and proved that any finite ordinal potential game has the FIP.⁷ A game has the FIP if any sequence of single-player improvements ends after a finite number of steps. This is equivalent to saying that there are no single-player improvement cycles. In contrast a game with the weak FIP can have single-player improvement cycles as long as there is a single-player improvement leading out of the cycle. For example, matching pennies does not have the FIP or the weak FIP, since there is an inescapable infinite sequence cycling around the four strategy profiles of the game. The game in Table 1 has the weak FIP, but not the FIP. It is matching pennies augmented by one additional row and column. Although the cycle (MC; MR; BR; BC; MC; ...) implies that the game does not have the FIP, it is possible to make a single player improvement from a state in the cycle to a state outside of it and, from there to the unique Nash equilibrium, (T; L).

3.2 Supermodular Games

We prove below that supermodular games are quasi-acyclic. Quasi-acyclicity is, roughly speaking, the weak FIP with the restriction that one uses strictly payoff-improving best replies rather than better replies. Any quasi-acyclic game then must have the weak FIP. As a result, the better reply dynamics must lead to a pure strategy Nash equilibrium in supermodular games. We begin by defining the appropriate best reply concept and weak acyclicity.

In a game $g = (N; A; u_i)$, the state $ax_i \in A$ is a weak best reply to a if and only if there exists $i \in N$ and $x_i \in A_i$ such that $u_i(ax_i) \geq u_i(ay_i)$ for all $y_i \in A_i$. The state $ax_i \in A$ is a strict best reply to a if and only if it is a weak best reply and

⁷Voorneveld and Norde [23] proved that finite, ordinal potential games are characterized by the absence of weak improvement cycles. A weak improvement cycle is a sequence of strategy profiles $(a^1; \dots; a^{v+1})$ such that (i) $a^1 = a^{v+1}$; (ii) for all $i \in \{1; \dots; v\}$, a^{i+1} and a^i only differ in the strategy of player i ; that is, $a^{i+1} = a^i na_i^{i+1}$ and (iii) $u_i(a^{i+1}) \geq u_i(a^i na_i^{i+1})$ for all i and $u_i(a^{i+1}) > u_i(a^i na_i^{i+1})$ for at least one i .

	L	C	R
T	7; 0	3; 3	0; 7
M	3; 3	5; 5	3; 3
B	0; 7	3; 3	7; 0

Table 2: A Game with the Weak FIP, but not Quasi-Acyclic

$$u_i(ax_i) > u_i(a).$$

Definition 3 The game $g = \langle N; A; u_i \rangle$ is quasi-acyclic if, from all action profiles $a \in A$, there exists a finite sequence of strict best replies that ends in a pure strategy Nash equilibrium.

Acyclic games were introduced by Young [25] who required that any sequence of weak best replies end in a pure Nash equilibrium after a finite number of steps. Quasi-acyclic games are identical to Milchtaich's [13] weakly acyclic games, except that Milchtaich uses weak best replies. Quasi-acyclic games have the weak FIP, but the converse is not true. The game in Table 2 has the weak FIP, but satisfies none of the acyclicity definitions. There is no best-reply sequence starting at $(T; L)$ and ending at the unique Nash equilibrium $(M; C)$; however, $f(T; L); (T; C); (M; C)$ is a sequence of single-player improvements. Thus, the set of games with the FIP is a subset of the set of quasi-acyclic games which, in turn, is a subset of the set of games with the weak FIP.

A finite game $g = \langle N; A; u_i \rangle$ is supermodular if it is possible to order the strategies of each player so that, for any $a; a^0 \in A$ such that $a \geq a^0$, $u_i(a)_i \geq u_i(a; a_i^0) \geq u_i(a^0; a_i)_i \geq u_i(a^0)$. Players' strategies exhibit strategic complementarities in supermodular games; the marginal payoff to an increase in a player's strategy is an increasing function of each of the other players' strategies. Lemma 2 establishes that all finite, two-player, supermodular games have the weak FIP and Theorem 3 extends this result to n -players. In both results the proof is carried out by showing that supermodular games are quasi-acyclic.

Let $A_i(a_i) = \{a_i^0 \in A_i : a_i^0 \leq a_i\}$ be the set of strategies of player i that are less

than or equal to a_i and $A(a) = \prod_{i \in N} A_i(a_i)$. Let $a^0 = (m_1; \dots; m_n)$ be the strategy profile corresponding to the highest strategy for each player and $\mathbf{1}_n = (1; \dots; 1)$ be the n -vector whose elements are all equal to one. For any game $g = (N; A; u)$, any nonempty subset of players $C \subseteq N$, and action profile $a \in A$, define the restricted game $g(C; a) = (C; \{A_i(a_i)\}_{i \in C}; \{u_i\}_{i \in C})$. In game $g(C; a)$, C is the set of active players, $j \in C$ has strategy set $A_j(a_j)$, and, for any given profile of actions $x_C \in \prod_{i \in C} A_i(a_i)$, j 's payoff is $u_j(x_C; \{a_i\}_{i \in N \setminus C})$. Note that $g = (N; A; u) = g(N; a^0)$. To simplify notation when $C = N$ we will write $g(a)$ instead of $g(N; a)$. Finally, let $b_i(a)$ be player i 's highest best reply to the strategy profile a .⁸

There are two elements behind the proof of Lemma 2. The first is that in a two-player, supermodular game g , one can start a process of iterative elimination of strictly dominated strategies beginning from $a^0 = (m_1; m_2)$, the profile with both players using their highest strategy. This process peels off strategies $a_i > a_i^{1st}$ from g , until a Nash equilibrium a^{1st} is reached. Repeating the process in the game where players are restricted to strategies that are strictly less than the strategies in a^{1st} leads to a Nash equilibrium a^{2nd} of the restricted game, and so on. A similar process can be started from $(1; 1)$, the profile with both players using their smallest strategy. The second element consists of showing that, for all $h = 1; 2; \dots$ and $i = 1; 2$, there is a finite sequence of best replies that starts from any profile in which player i uses a_i^{hth} and ends at a Nash equilibrium of the unrestricted game g .

Lemma 2 Any finite, two-player, supermodular game $g = (N; A; u)$ is quasi-acyclic, and thus has the weak FIP.

Proof. Starting at a^0 , the highest strategy profile in A , iteratively eliminate strictly dominated strategies as follows. Define the sequence $a^h = (b_1(a^{h-1}); b_2(a^{h-1}))$ for

⁸More precisely, if $B_i(a)$ is the best reply correspondence of player i , then $b_i(a)$ is the selection of B_i given by the largest element of the set $B_i(a)$.

$h \leq 1$ and note that $a_i^{h_i-1} \leq a_i^h$ for all i . Then for any $a_i > a_i^h$

$$0 > \frac{1}{4}_i(a_i^{h_i-1}na_i) \leq \frac{1}{4}_i(a_i^{h_i-1}na_i^h) \leq \frac{1}{4}_i(a_i^0na_i) \leq \frac{1}{4}_i(a_i^0na_i^h) \text{ for all } a^0 \in A(a_i^{h_i-1}): \quad (1)$$

The first inequality in equation (1) follows from a_i^h being the highest best reply to $a_i^{h_i-1}$; the second follows from $a_i^{h_i-1} \leq a_i^h$ for all i and the supermodularity of g . Equation (1) implies that if $a^h \notin a_i^{h_i-1}$, then any a_i such that $a_i^{h_i-1} \leq a_i > a_i^h$ is strictly dominated by a_i^h in the restricted game $g(a_i^{h_i-1})$, and thus there is a sequence of strict best replies from $a_i^0na_i$ to $a_i^0na_i^h$, where $a^0 \in A(a_i^{h_i-1})$.

Continue this process, going from $g(a_i^{h_i-1})$ to $g(a_i^h)$, until there are no more strictly dominated strategies to remove by this means. Since the sequence $(a^0; a^1; \dots)$ is decreasing and bounded, it must converge and the limit must be a Nash equilibrium a^{1^*} of g .

Define the two faces of $A(a^{1^*})$ by $F_1(a^{1^*}) = \{a^0 \in A : a_i^0 = a_i^{1^*} \text{ and } a_j^0 = a_j^{1^*} g \text{ for } i; j = 1; 2 \text{ and } i \neq j\}$. We have shown that, for any profile $a \in A \setminus A(a^{1^*})$, there is a finite sequence of strict best replies leading to a profile in $F_1(a^{1^*}) \cup F_2(a^{1^*})$. If $(a_i; a_j^{1^*}) \in F_1(a^{1^*})$ then $\frac{1}{4}_i(a^{1^*}) \leq \frac{1}{4}_i(a_i; a_j^{1^*})$; if the inequality is strict, then a^{1^*} is a strict best reply to $(a_i; a_j^{1^*})$; if $\frac{1}{4}_i(a^{1^*}) = \frac{1}{4}_i(a_i; a_j^{1^*})$, then either $(a_i; a_j^{1^*})$ is a Nash equilibrium of g , or there exists $(a_i; a_j) \in A(a^{1^*} \setminus 1_2)$ such that $\frac{1}{4}_j(a_i; a_j) > \frac{1}{4}_j(a_i; a_j^{1^*})$.⁹ This shows that, given any $a \in A(a^{1^*} \setminus 1_2)$, there exists a finite sequence of strict best replies leading either to a Nash equilibrium of g in $A \setminus A(a^{1^*} \setminus 1_2)$ or to a strategy profile in $A(a^{1^*} \setminus 1_2)$.

Using an analogous argument, a process of elimination of dominated strategies can also be started at $(1; 1)$. This process converges to a Nash equilibrium a^{K^*} of g . Let $\mathcal{A}_i(a_i) = \{a_i^0 \in A_i : a_i^0 \leq a_i g\}$ and $\mathcal{A}(a) = \prod_{i \in N} \mathcal{A}_i(a_i)$. Thus, given any $a \in \mathcal{A}(a^{K^*} \setminus 1_2)$, there exists a finite sequence of strict best replies leading either to

⁹Recall that $A(a^{1^*} \setminus 1_2) = \{a^0 \in A : a_i^0 < a_i^{1^*} g\}$. Improvement to somewhere in $A(a^{1^*} \setminus 1_2)$ must be possible, because there is no improvement to somewhere in either face, the point is not a Nash equilibrium, implying improvement is possible in some direction, and the iterated elimination of strictly dominated strategies ensures that improvement in $A \setminus A(a^{1^*})$ is impossible.

a Nash equilibrium of g in $A(a^{k^n})$ or to a strategy profile in $A(a^{k^n} + 1_2)$.

Now consider the restricted game $g(a^{1^n}; 1_2)$ and, as in the original game g , proceed by eliminating strictly dominated strategies to arrive at a Nash equilibrium a^{2^n} of $g(a^{1^n}; 1_2)$. As before, given any $a \in A(a^{1^n}; 1_2) \cap A(a^{2^n}; 1_2)$, there exists a finite sequence of strict best replies leading either to a strategy profile in $A(a^{2^n}; 1_2)$, or to a Nash equilibrium a^n (not necessarily equal to a^{2^n}) of $g(a^{1^n}; 1_2)$. Suppose a^n is not a Nash equilibrium of the unrestricted game g . Then for some i , $u_i(a_i^{1^n}; a_j^n) > u_i(a^n)$.¹⁰ Since a^{1^n} is a Nash equilibrium of g , $u_j(a^{1^n}) \geq u_j(a_i^{1^n}; a_j^n)$ for $j \neq i$ and if $u_j(a^{1^n}) = u_j(a_i^{1^n}; a_j^n)$, then $(a_i^{1^n}; a_j^n)$ is a Nash equilibrium of g , because otherwise there would exist $a_i < a_i^{1^n}$ such that $u_i(a_i; a_j^n) > u_i(a_i^{1^n}; a_j^n) > u_i(a^n)$ which is not compatible with a^n being a Nash equilibrium of $g(a^{1^n}; 1_2)$. Thus, there is a finite sequence of strict best replies from a^n to a Nash equilibrium of g .

We have shown that, for any $a \in A \cap A(a^{2^n}; 1_2)$, there exists a finite sequence of strict best replies leading either to a Nash equilibrium of g in $A \cap A(a^{2^n}; 1_2)$ or to a strategy profile in $A(a^{2^n}; 1_2)$. Repeated application of the same argument shows that from any $a \in A$ there is a finite sequence of strict best replies leading to either a Nash equilibrium of g or to a strategy profile in $A(a^{k^n}; 1_2)$. Since $a \in A(a^{k^n}; 1_2)$ implies $a \in A(a^{k^n} + 1_2)$, we know already that from any profile in $A(a^{k^n}; 1_2)$ there is a finite sequence of strict best replies to either a Nash equilibrium of g or to a strategy profile in $A(a^{k^n} + 1_2) \cap A(a^{k^n})$. Thus, any two-player supermodular game is quasi-acyclic. ■

The proof that an n -player supermodular game is quasi-acyclic also has two elements. The first, iterative elimination of strictly dominated strategies, is essentially the same as in Lemma 2. The second element of the proof uses induction on n .

¹⁰Recall that $u_i(a_i^{1^n}; a_j^n) > u_i(a_i; a_j^n)$ for all $a_i > a_i^n$ since a_i is strictly dominated by $a_i^{1^n}$ in g , and that $u_i(a^n) > u_i(a_i; a_j^n)$ for all $a_i^{1^n} > a_i > a_i^n$ since a_i is strictly dominated by a_i^n in $g(a^{1^n}; 1_2)$.

Theorem 3 Any finite, supermodular game $g = (N; A; u_i)$ is quasi-acyclic, and thus has the weak FIP.

Proof. Suppose that $(n - 1)$ -player, finite, supermodular games are quasi-acyclic. Define the face $F_{N \setminus i}(a) = \{a^0 \in A : a^0 \leq a \text{ and } a_i^0 = a_i\}$. As in the proof of Lemma 2, let $a^h = (b_1(a^{h-1}); \dots; b_n(a^{h-1}))$ for $h \geq 1$ and, starting from a^0 , proceed by iteratively eliminating strictly dominated strategies. The sequence $(a^0; a^1; \dots)$ is decreasing and bounded and must converge to a Nash equilibrium a^{1^*} . Thus, for any profile $a \geq A(a^{1^*})$, there is a sequence of strict best replies leading either to a^{1^*} or to some face $F_{N \setminus i}(a^{1^*})$. We now prove the stronger claim that, given any $a \geq A(a^{1^*} - 1_n)$, there exists a finite sequence of strict best replies leading either to a Nash equilibrium of g in $A(a^{1^*} - 1_n)$ or to $A(a^{1^*} - 1_n)$.

Each face $F_{N \setminus i}(a^{1^*})$ defines an $(n - 1)$ -player supermodular game $g(N \setminus i; a^{1^*})$. Therefore, from any point in $F_{N \setminus i_1}(a^{1^*})$ there is a sequence of strict best replies leading to a Nash equilibrium a^{i_1} of $g(N \setminus i_1; a^{1^*})$. Either a^{i_1} is a Nash equilibrium of g , or there is a strategy $a_{i_1} < a_{i_1}^{i_1} \leq a_{i_1}^{1^*}$ for player i_1 such that $a^{i_1} na_{i_1}$ is a strict best reply to a^{i_1} ; that is, $u_{i_1}(a^{i_1} na_{i_1}) > u_{i_1}(a^{i_1}) \leq u_{i_1}(a^{i_1} na_{i_1}^{1^*})$. In the latter case, either $(a^{i_1} na_{i_1}) \in A(a^{1^*} - 1_n)$, or $(a^{i_1} na_{i_1})$ belongs to a face $F_{N \setminus i_2}(a^{1^*})$ and then there is a finite sequence of strict best replies leading to a Nash equilibrium a^{i_2} of $g(N \setminus i_2; a^{1^*})$. Again, if a^{i_2} is not a Nash equilibrium of the game g , then there is a strict best reply improvement from a^{i_2} either to $A(a^{1^*} - 1_n)$, or to a face $F_{N \setminus i_3}(a^{1^*})$ and from there to a Nash equilibrium a^{i_3} of $g(N \setminus i_3; a^{1^*})$. Repeating this argument generates a sequence of strict best replies that either ends in $A(a^{1^*} - 1_n)$, or that contains a subsequence $(a^{i_1}; a^{i_2}; a^{i_3}; \dots)$. We now show that the subsequence $(a^{i_1}; a^{i_2}; a^{i_3}; \dots)$ is decreasing and thus must converge to a Nash equilibrium of g .

For any $i_h \in \{i_1; i_2; \dots; i_n\}$, let $(a^{i_h} na_{i_h}^0) na_j$ be the first element in the sequence going from $(a^{i_h} na_{i_h})$ to $a^{i_{h+1}}$ involving a player different from i_h . Suppose $a_j > a_j^{i_h}$, then

$$0 \leq u_j(a^{i_h} na_j) - u_j(a^{i_h}) \leq u_j((a^{i_h} na_{i_h}^0) na_j) - u_j(a^{i_h} na_{i_h}^0) > 0;$$

which is a contradiction; hence $a_j < a_j^h$. The first inequality follows from a^h being a Nash equilibrium of $g(N \setminus i_h; a^h)$, the second from $a_{i_h}^0 < a_{i_h}^1 < a_{i_h}^h$ and supermodularity, and the third from $(a^h \setminus a_{i_h}^0) \setminus a_j$ being a strict best reply to $(a^h \setminus a_{i_h}^0)$. Similar reasoning applies to all the other steps in the sequence of strict best replies from $(a^h \setminus a_{i_h}^0)$ to a Nash equilibrium a^{h+1} of $g(N \setminus i_{h+1}; a^{h+1})$. This concludes the proof that, given any $a \in A(a^1 \setminus i_1)$, there exists a finite sequence of strict best replies leading either to a Nash equilibrium of g in $A(a^1 \setminus i_1)$ or to $A(a^1 \setminus i_1)$.

In the restricted game $g(a^1 \setminus i_1)$, iteratively eliminate strictly dominated strategies until a Nash equilibrium a^{2^*} of $g(a^1 \setminus i_1)$ is reached. As before, given any $a \in A(a^1 \setminus i_1) \setminus A(a^{2^*} \setminus i_1)$, there exists a finite sequence of strict best replies leading either to $A(a^{2^*} \setminus i_1)$ or to a Nash equilibrium a^* of $g(a^1 \setminus i_1)$. If a^* is not a Nash equilibrium of the unrestricted game g , then for some i , $\pi_i(a^* \setminus a_i^1) > \pi_i(a^*) > 0$. In this last case, $(a^* \setminus a_i^1) \in F_{N \setminus i}(a^1)$ and there is a sequence of strict best replies from $(a^* \setminus a_i^1)$, and a fortiori from a^* , to a Nash equilibrium a^{**} of $g(N; a^1)$. We first show that $a^* < a^{**}$ and then that a^{**} is a Nash equilibrium of the unrestricted game g .

Let $(a^* \setminus a_i^1) \setminus a_j$ be the first step in the sequence going from $(a^* \setminus a_i^1)$ to a^{**} . Suppose $a_j < a_j^*$, then

$$0 > \pi_j(a^* \setminus a_i^1) > \pi_j((a^* \setminus a_i^1) \setminus a_j) > \pi_j(a^*) > \pi_j(a^* \setminus a_j) > 0;$$

which is a contradiction; hence $a_j > a_j^*$. The first inequality follows from $(a^* \setminus a_i^1) \setminus a_j$ being a strict best reply to $(a^* \setminus a_i^1)$, the second from $a_i^* < a_i^1$ and supermodularity, and the third from $a_j < a_j^*$ and a^* being a Nash equilibrium of $g(a^1 \setminus i_1)$. Similar reasoning applies to all the other steps in the sequence, which implies $a^* < a^{**}$.

To see that a^{**} is a Nash equilibrium of g , suppose the contrary. Then player i would profit by deviating to some strategy $a_i < a_i^1$ and

$$0 > \pi_i(a^{**}) > \pi_i(a^{**} \setminus a_i) > \pi_i(a^{**} \setminus a_i^1) > \pi_i(a^{**} \setminus a_i) > \pi_i(a^*) > \pi_i(a^* \setminus a_i) > 0;$$

which is a contradiction. The first inequality follows from (a^{n+1}_i) being an improvement over a^{n+1} , the second from $a_i^{n+1} = a_i^{n+1}$, $a^{n+1} > a^n$, and supermodularity, the third from $\frac{1}{2} \frac{1}{2} (a^{n+1}_i) > \frac{1}{2} \frac{1}{2} (a^n)$, and the fourth from $a_i < a_i^{n+1}$ and a^n being a Nash equilibrium of $g(a^{n+1}; 1_n)$.

Thus we have proven that, given any $a \in A(a^{n+1}; 1_n)$, there exists a finite sequence of strict best replies leading either to a Nash equilibrium of g belonging to $A(a^{n+1}; 1_n)$ or to $A(a^{n+1}; 1_n)$. Repeated application of this construction shows that given any non-equilibrium $a \in A$, there is a finite sequence of strict best replies from a to a Nash equilibrium a^* of g . This concludes the proof that all finite, supermodular games are quasi-acyclic. ■

Theorem 3 is related to results obtained by Vives [22] and Milgrom and Roberts [14]. Vives studied Cournot tatonnement, in which at any discrete point in time each player switches to a strategy that is a best reply to the past strategy profile. He showed that in supermodular games Cournot tatonnement converges monotonically to a Nash equilibrium if the starting point is below or above all the best reply correspondences of the players. Vives also showed that supermodular games have a greatest and a smallest Nash equilibrium. Milgrom and Roberts [14] showed that the process of iterative elimination of dominated strategies eliminates all strategies that are either greater than those associated with the greatest Nash equilibrium, or smaller than those associated with the smallest Nash equilibrium.

3.3 Quasi-Concave Games

We now examine two-player games $g = (f; 2g; A; \frac{1}{2}i)$ where each player i has a one dimensional, compact, convex strategy set $A_i \subseteq \mathbb{R}$, and a payoff function $\frac{1}{2}i$ that is continuous in $a \in A$ and strictly quasi-concave with respect to $a_i \in A_i$. Let $B : A \rightarrow A$, with $B(a) = (B_1(a_2); B_2(a_1))$ denote the best-reply correspondence of the game g . Strict quasi-concavity of $\frac{1}{2}i$ implies that: (i) B_i is continuous and single-

valued and (ii) player i 's payoff declines as a_i moves away from $B_i(a_{-i})$. We also assume that the best reply functions are transversal; that is, at any Nash equilibrium they must cross and they must not be tangent. The set of Nash equilibria of such a generic quasi-concave game g is a finite, nonempty set $A^{NE} = \{a^1; \dots; a^q\} \subset A$.

We now show that generic, two-player, quasi-concave games have the weak FIP. The intuition for this result is the following. Start from any strategy profile a^0 . Suppose a^0 is not a best reply for some player i . Then the players alternate moves, starting with player i , with each move being a single-player improvement. This continues until one player moves to a point from which the other player can move to a Nash equilibrium. This process can be forced to occur in a finite number of steps.

Theorem 4 Any generic, two-player, quasi-concave game $g = (f_1; \dots; f_q; A; \frac{1}{2}I)$ has the weak FIP.

Proof. The proof proceeds in two parts. In the first part it is shown that a finite number of steps is sufficient to move the two players using only single-player improvements from an arbitrary starting point either to a Nash equilibrium or to a ball of radius $r > 0$ around a Nash equilibrium. In the second part it is shown that it is possible to move the two players using only single-player improvements from an arbitrary starting point in the ball to the Nash equilibrium in the ball.

Part (i). First the strategy space is divided in two separate ways. One way utilizes the Nash equilibria to construct rectangles in A ; the other divides A according to directions of single-player improvement. Each rectangle is closed with a non-empty interior. Two rectangles may share, at most, a side or a vertex, and no rectangle has a Nash equilibrium in its interior. To construct them, draw lines in \mathbb{R}^2 parallel to each axis from all Nash equilibrium points in $A^{NE} = \{a^1; \dots; a^q\}$. That is, draw lines $L_1^h = \{f(a_1^h; a_2) : a_2 \in A_2 \cap Rg\}$ and $L_2^h = \{f(a_1; a_2^h) : a_1 \in A_1 \cap Rg\}$ for all $h = 1; \dots; q$. This divides A into compact rectangles having disjoint interiors.¹¹ A rectangle is

¹¹If all q Nash equilibria are interior to A and the coordinates of all equilibria are distinct, there will be $(q + 1)^2$ rectangles.

external if it intersects the boundary of A and a side of a rectangle that lies in the boundary of A is called an external side; all other sides are called internal sides. From each internal side of a rectangle, it is possible to move to a Nash equilibrium of g via a single-player improvement. Thus, for g not to have the weak FIP there must be at least one point of some rectangle from which it is impossible to reach an internal side through a finite sequence of single-player improvements.

The strategy space A contains four sets, called strict improvement regions, $(+; j)$, $(j; +)$, $(+; +)$, $(j; j)$, according to the direction of improvement for each player. For example, at a point $a = (a_1; a_2)$ in $(+; j)$ player 1's best-reply to a is greater than a_1 (hence, by quasi-concavity, his payoff increases if he unilaterally increases his action), while player 2's best-reply to a is smaller than a_2 (his payoff increases if he unilaterally decreases his action). If a_1 is the best reply of player 1 to a , then a is in the common boundary of $(+; j)$ and $(j; j)$ which is part of the graph of B_1 . Similarly for a_2 and player 2. The closure of a strict improvement region such as $(+; j)$ is called a weak improvement region and is denoted $[+; j]$. If $q > 1$ the sets are each, typically, not connected. These sets will be referred to below in brief as regions.

Now suppose that the point $a^0 = (a_1^0; a_2^0)$ is in a rectangle Z and there is no sequence of single-player improvements leading to a Nash equilibrium of g . Assume without loss of generality that $a_1^0 \notin B_1(a_2^0)$ and consider a move by player 1 to his best reply $a_1^1 = B_1(a_2^0)$. The line segment joining a^0 to $a^0 = (a_1^1; a_2^0)$ must remain in the relative interior of the rectangle, otherwise either it would reach a Nash equilibrium or an internal side from which one step would lead to a Nash equilibrium.¹² Now consider the best reply of player 2 to a_1^1 , $a_2^1 = B_2(a_1^1)$. The line segment joining a^0 to $a^1 = (a_1^1; B_2(a_1^1))$ must remain in the relative interior of the rectangle, reach a Nash equilibrium, or reach an internal side. Only the first of these three possibilities is compatible with never reaching a Nash equilibrium. Continue to find successive best replies by the two players, $a^h = (a_1^h; a_2^h)$, $a^{h+1} = (a_1^{h+1}; a_2^h)$, $a^{h+1} = (a_1^{h+1}; a_2^{h+1})$,

¹²Recall the relative interior of a rectangle includes any external side of that rectangle.

and so forth. All line segments joining these best-reply moves must remain within the relative interior of the rectangle Z . This infinite sequence must either (1) cycle among the four regions, $[i, i]$, $[+, i]$, $[+, +]$, and $[i, +]$, or (2) all the segments must belong to the same region of A .

In case (1), note that only two types of cycles are possible: (i) $[i, i]$, $[+, i]$, $[+, +]$, $[i, +]$, $[i, i]$, and (ii) $[i, i]$, $[i, +]$, $[+, +]$, $[+, i]$, $[i, i]$. Either sequence could only occur if there were a Nash equilibrium in the interior of the rectangle Z . However, by definition of the rectangles, there are no Nash equilibria in the interior of a rectangle.¹³ Thus, (1) is impossible. In case (2), the sequence of movements from best reply to best reply is monotone, since it stays within the same region of A , and hence it must reach the interior of a ball of radius r around a Nash equilibrium in a finite number of steps. This concludes the proof of part (i). The impossibility of (2) is taken up in part (ii) of the proof.

Part (ii). Suppose the ball of radius r is reached at the point a^0 which is, without loss of generality, on the best reply function of player 1 ($a_1^0 = B_1(a_2^0)$). We may change coordinates and suppose, without loss of generality, that $a^0 \in (i, i)$ and the Nash equilibrium is at $(0; 0)$. To have reached such a point, the situation must be as depicted in Figure 1 with a^0 corresponding to the point A . The two best reply functions must be positively sloped with B_1 lying above B_2 as one descends to $(0; 0)$. Dividing \mathbb{R}^2 into quadrants with the Nash equilibrium as the origin, movement from a^0 to the Nash equilibrium will be within the ball and confined to the quadrant containing a^0 . Inside this region the best reply functions are closely approximated by linear functions and the isoprofit curves are closely approximated by identical parabolas, symmetric about each player's best reply function. Write $B_1(a_2) = \beta_1 a_2$ and $B_2(a_1) = \beta_2 a_1$, with $\beta_1, \beta_2 > 0$. Transversality of the best reply functions guarantees that their linear approximations have different slopes, so that $\beta_1 \beta_2 < 1$.

¹³Note that no Nash equilibrium on the boundary of A can be in the relative interior of the rectangle X , because all such equilibria also lie on an interior boundary of some rectangle.

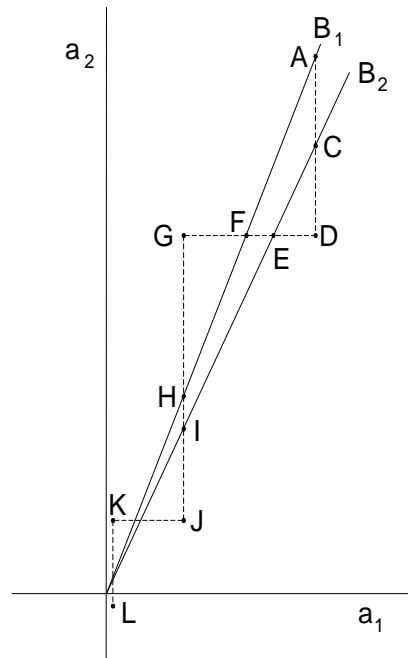


Figure 1: The weak FIP in a Two-Player, Quasi-Concave Game

The path to the Nash equilibrium will be by alternate moves, starting with player 2. Each move will be to a point where the payoff of the moving player is almost unchanged (as opposed to being a best reply). Note that the vertical distance between the best reply functions at some fixed value of a_1 is $a_2 = \frac{1}{1 - \beta_1} \beta_2 a_1$ and the distance from the horizontal axis to the lower best reply function is $\beta_2 a_1$. The ratio of the latter to the former is $\frac{\beta_2}{1 - \beta_1} = \beta_2 (1 - \beta_1)^{-1}$ which is obviously independent of the choice of a_1 (within the ball). The corresponding ratio for the vertical axis is exactly the same.

At the first move, player 2 goes from a^0 to $(a_1^0; a_2^1)$ (from A to D in Figure 1) with a_2^1 chosen so that the payoff of player 2 is only marginally increased. The distance from C to D is almost the same as that from A to C due to the near-symmetry of the isoprofit curves of the players. Player 1 now moves from $(a_1^0; a_2^1)$ to $(a_1^1; a_2^1)$, chosen so that player 1's profit is only marginally increased. Note that DE must strictly exceed EF and that FG is only slightly smaller than FD. That is, with certainty

FG can be chosen to be at least double EF. Now player 2 moves from $(a_1^1; a_2^1)$ to $(a_1^2; a_2^2)$, chosen so that the product of player 2 rises slightly and IJ is more than 3 times the size of HI. At the kth move, the new point is at a distance from the nearer best reply function that is at least k times the distance between the best reply functions. Denote this ratio y_k . After a finite number of moves, $y_{k+1} > \frac{1}{2} = (1 - \frac{1}{2}) > y_k$ which means that the k + 1 move can be chosen to lay either below the horizontal axis or to the left of the vertical axis (illustrated in Figure 1) as the move from K to L. For this move, choose instead the point on the axis. From there it is one step to the Nash equilibrium, which completes the proof. ■

Theorem 4 shows that from all action profiles a 2 A of a generic, quasi-concave game with continuous payoff functions there is a finite sequence of single-player improvements that ends in a pure strategy Nash equilibrium. It is an open question whether Theorem 4 extends to more than two players.

Generic, two-player, quasi-concave games need not have the FIP and need not be quasi-acyclic. Consider, for example, the quasi-concave game $g = (f; A; \frac{1}{4}i)$, where $A_i = [0; 2]$, $\frac{1}{4}_1 = 4 + (4 - 2a_2)a_1 - a_1^2$ and $\frac{1}{4}_2 = 4 + 2a_1a_2 - a_2^2$. The unique Nash equilibrium of g is $(1; 1)$. The game g has the weak FIP, but it does not have the FIP and it is not quasi-acyclic. To see this, note that the best reply functions are: $B_1(a_2) = 2 - a_2$ and $B_2(a_1) = a_1$. Hence there is a cycle of best-reply improvements involving the strategy profiles $(0; 0)$, $(2; 0)$, $(2; 2)$, $(0; 2)$. Table 3 displays an approximation of this game in which each player's strategy set is $f0; 1; 2g$. The weak FIP is satisfied, because it is possible to make a single-player improvement from $(0; 0)$ to $(1; 0)$ and then from $(1; 0)$ to $(1; 1)$.

Not all finite approximations of a generic, two-player, quasi-concave game g need to have the weak FIP. However, the weak FIP will typically hold for approximations of g that have the property that all the strategies corresponding to a Nash equilibrium in the continuous game are present in the finite approximation. For example, the

	0	1	2
0	4; 4	4; 3	4; 0
1	7; 4	5; 5	3; 4
2	8; 4	4; 7	0; 8

Table 3: A Generic Quasi-Concave Game without the FIP and not Weakly Acyclic

weak FIP trivially holds if $A_i = \text{fa}_i^1; \dots; a_i^q$ for all $i \in N$, where $A^{\text{NE}} = (a^1; \dots; a^q)$ is the set of Nash equilibria of g . Table 3 shows that many other approximations will also work. For such approximations of generic, two-player, quasi-concave games the only recurrent classes of the better-reply dynamics are the pure strategy Nash equilibria.

4 Ergodic Variants of the Better-Reply Dynamics

We first introduce two variants of the better-reply dynamics: the better-reply dynamics with mistakes and the better-reply dynamics with mistakes and simultaneous sampling. The advantage of these two variants is that the Markov process they generate is ergodic.¹⁴ Then, in Section 4.1 we show that all versions of the better-reply dynamics can lead to cyclical outcomes in finite games that do not have the weak FIP. Next, in Section 4.2 Theorem 5 makes a connection between the recurrent classes of the better-reply dynamics and the limiting distributions of its two ergodic variants. Following that, in Section 4.3 our results are briefly compared with those of Ritzberger and Weibull [17] and, finally, in Section 4.4 we compare better-reply and best-reply dynamics.

By introducing the possibility that players make mistakes when deciding whether to change status quo action, we obtain the following variant of the better-reply dynamics.

Definition 4 Better-Reply Dynamics with Mistakes. This dynamics is like

¹⁴The better-reply dynamics with mistakes and simultaneous sampling will be used in Section 5.

that of Definition 1 except that when $\frac{1}{L} \sum_{i \in K} \pi_i(a_i^E) > \frac{1}{L} \sum_{i \in K} \pi_i(a)$ player i adopts the sampled action with probability $\frac{1}{L} = \epsilon^L$ for some integer $L > 1$.

In the next variant of the better-reply dynamics we allow sampling by more than one player at the same time. However, for small ϵ , the probability of simultaneous sampling of a new action by exactly k players is k orders of magnitude less likely than the probability that only a single player samples a new action. Let $K \subseteq N$ be the set of players actually sampling in a time period, and let $a_K^E = (a_i^E)_{i \in K}$ be the strategies they sample, with $a_i^E \in A_i$ for all $i \in K$.

Definition 5 Better-Reply Dynamics with Mistakes and Simultaneous Sampling. At any time period each player $j \in N$ randomly samples an action with probability $(m_j + 1)\epsilon$ divided equally among the elements of $A_j \times \dots \times A_j$. Player $i \in K$ switches his status quo action to a_i^E with probability one if $\frac{1}{L} \sum_{i \in K} \pi_i(a_{N \setminus K}; a_K^E) > \frac{1}{L} \sum_{i \in K} \pi_i(a)$; when $\frac{1}{L} \sum_{i \in K} \pi_i(a_{N \setminus K}; a_K^E) < \frac{1}{L} \sum_{i \in K} \pi_i(a)$ he switches with probability $\frac{1}{L} = \epsilon^L$ for some integer $L > 1$.¹⁵

We now need to introduce a few additional notions from the theory of stochastic processes. The resistance between any two states $a^0, a^{00} \in A$, denoted $r(a^0; a^{00}) \in [0; \infty]$, is the extended real number given by $r(a^0; a^{00}) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \log \frac{p(a^0; a^{00}; \epsilon)}{p(a^0; a^0; \epsilon)} = \log \frac{1}{\epsilon}$, with $r(a^0; a^{00}) = \infty$ if $p(a^0; a^{00}; \epsilon) = 0$ for $\epsilon \in [0; \epsilon^*]$. Intuitively, if h players sample and mistakes by ℓ players are required for a transition from a^0 to a^{00} under the better-reply dynamics with mistakes and simultaneous sampling, then $r(a^0; a^{00}) = h + \ell L$.

A tree rooted at a , $T(a)$, is a set of ordered pairs of (or edges between) elements of A , such that starting from each element $a^0 \in A$ there is a unique edge beginning at a^0 and a unique path (or sequence of edges) leading to a . The resistance of a tree rooted at a , $r(T(a))$, is the sum of the resistances of the edges that compose it, $r(T(a)) = \sum_{(a^0; a^{00}) \in T(a)} r(a^0; a^{00})$. The stochastic potential $\tilde{A}(a)$ of a state $a \in A$ is

¹⁵Note that the upper bound on the mistake probability is lower with simultaneous than with isolated sampling. This bound guarantees that the probability of a single mistake ϵ^L is less than the probability that all players sample a new action, $\epsilon^L \prod_{i \in N} (m_i + 1)$.

the minimum resistance over all trees rooted at a . Let $\mathbb{R}(a)$ be the number of trees rooted at a that have minimum resistance.

For any game, the Markov process $P(\epsilon)$ associated with the dynamics in either Definition 4 or Definition 5 is aperiodic and irreducible for all $\epsilon > 0$; hence it has a unique stationary distribution $\pi(\epsilon)$ satisfying $\pi(\epsilon)P(\epsilon) = \pi(\epsilon)$. The stationary distribution $\pi(\epsilon)$ gives the long-run frequency with which each state will be observed. Let $P(0)$ be the stochastic process when $\epsilon = 0$. Results due to Freidlin and Wentzell [5, Chapter 6, Lemma 3.1], Kandori, Mailath and Rob [11, Theorem 1] (KMR henceforth), and Young [25, Theorem 4] imply that (1) the limit stationary distribution $\pi = \lim_{\epsilon \rightarrow 0} \pi(\epsilon)$ exists and is equal to a stationary distribution of $P(0)$; (2) the set of stochastically stable states $A^S = \{a \in A : \pi(a) > 0\}$ is the set of states in A that have minimum stochastic potential; and (3) the mass of any $a \in A^S$ in the limiting distribution is $\pi(a) = \frac{\mathbb{R}(a)}{\sum_{a' \in A^S} \mathbb{R}(a')}$.

4.1 Cycles

The game in Table 4 has a strict Nash equilibrium, $(M; C)$, and a mixed equilibrium $((.25; 0; .75); (.5; 0; .5))$. This game does not have the weak FIP, because there is no single player improvement from $f(T; L); (T; R); (B; R); (B; L)g$, the set of states corresponding to the mixed equilibrium. Recall from Section 2 that the stochastic process under the better-reply dynamics converges to one of its recurrent classes, which in this game are $f(M; C)g$ and $f(T; L); (T; R); (B; R); (B; L)g$. The limit stationary distribution of the better-reply dynamics with mistakes puts positive mass on the states in the set $f(M; C); (T; L); (T; R); (B; R); (B; L)g$, while the stochastically stable states of the better-reply dynamics with mistakes and simultaneous sampling are $f(T; L); (T; R); (B; R); (B; L)g$. Thus all dynamics lead to cycling behavior in this game.

It is clear that the recurrent classes and the stationary distributions of all three better-reply dynamics depend only on ordinal properties of the payoffs. This implies

	L	C	R
T	5; 3	0; 0	3; 6
M	0; 0	2; 2	0; 0
B	3; 4	0; 0	5; 3

Table 4: Strict Nash Equilibrium with Zero Mass

that the stochastic process cannot generally converge to a mixed strategy equilibrium, because the exact probabilities in a mixed strategy equilibrium depend on cardinal properties that are generally altered by order-preserving transformations. For example, in any 2×2 game with a unique, completely mixed Nash equilibrium, the stationary distributions of the dynamics in Definitions 4 and 5 put equal mass on each of the four pure strategy profiles.¹⁶

4.2 Limit Distributions

The following is a simple, general result about convergence of the better-reply dynamics with mistakes and the better-reply dynamics with mistakes and simultaneous sampling.

Theorem 5 Let $g = (N; A; \{u_i\})$ be a finite game. As $\epsilon \rightarrow 0$, the limiting stationary distribution of the dynamics in Definitions 4 and 5 put positive mass on state a only if a belongs to a recurrent class E of the better-reply dynamics. If $a \in E$ has positive mass, then all states in E have positive mass.

Proof. To prove that the limiting stationary distribution puts zero mass on any state a that does not belong to a recurrent class of the better-reply dynamics, it suffices to show that, if $T(a)$ is any minimum resistance tree with root a , then there exists some state $b \in A$ and a tree $T(b)$ with root b whose resistance is smaller than that of $T(a)$.

¹⁶However, it is not true that the stationary distribution always puts equal mass on all states associated with a mixed strategy equilibrium when that equilibrium is unique. See Table 6 for an example.

Consider such a state a and tree $T(a)$. For some b^0 belonging to a recurrent class E of the better-reply dynamics there exists a ...nite sequence $(a_0; a_1; \dots; a_u)$ of single-player improvements, with $a_0 = a$ and $a_u = b^0$. As E is a recurrent class, $T(a)$ must contain at least one transition from a state $b \in E$ to a state $a^0 \notin E$ and the resistance of this transition must be at least 2. Since both b and b^0 belong to E , there is a ...nite sequence $(b_0; \dots; b_z)$ of single-player improvements, with $b_0 = b^0$, $b_z = b$ and $b_h \in E$ for all $h = 1; \dots; z$. Cut the tree $T(a)$ after each transition in the sequence $(a_1; \dots; a_u; b_1; \dots; b_z)$. Reconstitute a tree $T(b)$ from the resulting $u + z$ partial trees by adding transitions from a_h to a_{h+1} for $h = 0; \dots; u - 1$, from a_u to b_1 , and from b_h to b_{h+1} for $h = 1; \dots; z - 1$. Since all the $u + z$ transitions that have been added have resistance 1 and the last of the transitions cut (from b to a^0) has resistance 2, $T(b)$ has a lower resistance than $T(a)$. Since $T(a)$ is a minimum resistance tree with root a and $T(b)$ has a lower resistance than $T(a)$, a must have no mass in the limiting distribution.

It remains to show that if $a \in E$ has positive mass, then all states in E have positive mass. This is done by showing that if $T(a)$ is a minimum resistance tree with root $a \in E$ and $a^0 \in E$ ($a^0 \neq a$), then there exists $T(a^0)$ having the same resistance as $T(a)$. Consider such $a; a^0 \in E$. There is a ...nite sequence $(a_0; a_1; \dots; a_u)$ of single-player improvements, with $a_0 = a$, $a_u = a^0$ and $a_h \in E$ for all $h = 1; \dots; u$. Cut from $T(a)$ all the transitions from each member of the sequence $(a_1; \dots; a_u)$. Add a transition from a_h to a_{h+1} for $h = 0; \dots; u - 1$ to form a new tree $T(a^0)$. Since all the u transitions that have been added have resistance 1, $T(a^0)$ has a resistance no higher than $T(a)$. Thus, all states in E must have the same stochastic potential and if one state in E has positive mass, then all states in E must have positive mass. ■

Theorem 5 implies that the better-reply dynamics with mistakes and the better-reply dynamics with mistakes and simultaneous sampling converge to a Nash equilibrium in ...nite games with the weak FIP. The corollary that follows is also an easy

	L	M	R
T	7; 2	10; 0	3; 1
MT	9; 7	1; 9	50; 5
MB	50; 1	2; 2	1; 3
B	0; 2	0; 0	0; 1

Table 5: No Equilibrium Support in a Recurrent Class

consequence of Theorem 5.

Corollary 6 Let $g = \langle N; A; \mu_i \rangle$ be a finite game. If the limiting stationary distribution of the dynamics in either Definition 4 or Definition 5 puts all mass on a single state $a \in A$, then a is a Nash equilibrium of the game.

Proof. If a alone has positive mass, then the recurrent class to which it belongs must be a singleton. This implies that a is a state from which no player can make a single-player improvement; that is, a is a Nash equilibrium. ■

4.3 The Better-Reply Correspondence

Given a finite game $g = \langle N; A; \mu_i \rangle$, and any $X_i \subseteq A_i$, $i \in N$, with $X = \prod_{i \in N} X_i$, let $\Phi_i(X_i)$ be the set of player i 's mixed strategies with support in X_i and let $\Phi(X) = \prod_{i \in N} \Phi_i(X_i)$. The better-reply correspondence $\sigma : \Phi(A) \rightarrow \Phi(A)$ is the correspondence that maps each profile of mixed strategies into the sets of better-replies by each player.¹⁷ Ritzberger and Weibull [17] defined a set $X = \prod_{i \in N} X_i$ as being closed under the better-reply correspondence if $\sigma(\Phi(X)) \subseteq X$. There are analogies, but also important differences, between sets that are closed under the better-reply correspondence and the recurrent classes of our better-reply dynamics.

Ritzberger and Weibull [17, Theorem 1] showed that for a large class of deterministic, continuous time, population dynamics a set $\Phi(X)$ is asymptotically stable

¹⁷More precisely, $\sigma : \Phi(A) \rightarrow \Phi(A)$ is defined by $\sigma_i(\mu_i) = \{a_i \in A_i : \mu_i(\mu_{-i}; a_i) \geq \mu_i(\mu_{-i}; g) \text{ for all } i \in N \text{ and all } \mu_{-i} = (\mu_{-i}; \mu_i) \in \Phi(X)\}$.

if and only if X is a set closed under the better-reply correspondence. Similarly, the better-reply dynamics eventually converges to one of its recurrent classes.

Ritzberger and Weibull [17, Proposition 4] also proved that if a set X is closed under the better-reply correspondence, then $\Phi(X)$ contains a connected component of Nash equilibria. To the contrary, recurrent classes of the better-reply dynamics need not contain the support of any Nash equilibrium. In the game in Table 5 the better-reply dynamics has a unique recurrent class, $E = \{ (T; L); (MT; L); (MT; M); (MB; M); (MB; R); (T; R) \}$, that does not contain the support of any mixed strategy Nash equilibrium.¹⁸ This game has no pure strategy Nash equilibrium; the set of mixed strategy Nash equilibria is $(\frac{1}{2}; 0; 0; \frac{1}{2}); (p; q; 1-p; q)$, where $q = \frac{3+4p}{16-659p-497}$.

4.4 Best Reply Dynamics

We now discuss the difference between better-reply and best-reply dynamics with and without mistakes.

Definition 6 Best-Reply Dynamics. At the beginning of each time period, a player is randomly selected; with probability α the selected player switches to a best reply to the current state a and with probability $1-\alpha$ he continues playing the status quo action.

Definition 7 Best-Reply Dynamics with Mistakes. This dynamics is like Definition 6, except that the selected player may make a mistake, in which case he changes his status quo to some action at random. The probability of making a mistake, $\frac{1}{L}$, equals α^L for some integer $L > 1$.

The game in Table 6 has a unique equilibrium $(\frac{1}{2}; 0; \frac{1}{2}); (\frac{1}{2}; 0; \frac{1}{2})$. The unique recurrent class of the better-reply dynamics consists not only of the four corners

¹⁸Note that E is not a product set. The definition of closed under the better-reply correspondence only applies to product sets.

	L	C	R
T	0;5	0;0	5;0
M	3;0	1;0	1;0
B	5;0	0;0	0;5

Table 6: Positive Mass on a Strictly Dominated Strategy

(T; L), (T; R), (B; R), and (B; L), but also (M; L). This shows that if a game does not have the weak FIP, then the better reply dynamics may spend a positive fraction of time in a state such as (M; L), that is associated with a strictly dominated strategy (M is strictly dominated by B).¹⁹ In contrast, the four corners constitute the unique recurrent class of the better-reply dynamics.²⁰

More generally, given a finite game $g = (N; A; u)$: (i) states including strictly dominated strategies cannot belong to one of the recurrent classes of the best-reply dynamics; (ii) the limiting stationary distribution of the best-reply dynamics with mistakes puts positive mass only on states that belong to a recurrent class of the best-reply dynamics without mistakes (the proof is essentially the same as the proof of Theorem 5); (iii) if g is quasi-acyclic, then the only recurrent classes of the best-reply dynamics are the pure strategy Nash equilibria. However, in the larger class of finite games with the weak FIP it need not be true that the recurrent classes of the best-reply dynamics coincide with the pure strategy Nash equilibria. For example, in the game in Table 3, the best-reply dynamics has two recurrent classes $f(1; 1)g$ and $f(0; 0); (2; 0); (2; 2); (0; 2)g$ and the limiting distribution of the best-reply dynamics with mistakes puts positive mass on all states belonging to one of these classes. The game, however, has the weak FIP and $f(1; 1)g$ is the only recurrent class of the better-reply dynamics.

¹⁹Hence in games that do not have the weak FIP our players' learning algorithm may not be "adaptive" in the sense of Milgrom and Roberts [15].

²⁰Dekel and Scotchmer ([3]) presented an example of a game in which the discrete replicator dynamics need not eliminate a pure strategy that is strictly dominated by a mixed strategy. Their example can be interpreted as another illustration of the fact that better-reply and best-reply dynamics differ.

5 Common Interest and Coordination Games

In this section we study common interest and coordination games.²¹ The key result is that in such games the limiting distribution of the better-reply dynamics with mistakes and simultaneous sampling puts all mass on a Pareto optimal Nash equilibrium.

Definition 8 The game $g = \langle N; A; \mu \rangle$ is a common interest game if there is a Pareto dominant strategy profile $a^* \in A$ (i.e., a^* such that $\mu(a^*) \geq \mu(a)$ for all $a \in \text{Anf}(g)$).

Theorem 7 Let $g = \langle N; A; \mu \rangle$ be a finite common interest game. The limiting stationary distribution of the better-reply dynamics with mistakes and simultaneous sampling puts all mass on the Pareto dominant state a^* .

Proof. Let $T(a^0)$ be any tree with root $a^0 \in A$. Cut this tree at the transition from a^0 to a^* and add a transition from a^0 to a^* to form a tree $T(a^*)$. The eliminated transition has a resistance exceeding n and the new transition has a resistance no more than n . Thus no tree $T(a^0)$, where $a^0 \in A$, can have minimum stochastic potential. ■

Two features of our model are crucial to obtain convergence to the Pareto dominant equilibrium. First, contrary to models that study the evolution of the average behavior in large populations of players, our players do not play best replies to some empirical distribution of past play. Instead, they follow a naive stimulus-response behavior; they have an habitual action that they play and only change it after experiencing a higher payoff from sampling a new action. Second, it must be possible for several players to sample at the same time. The probability of simultaneous sampling by a group of players must be a decreasing function of the group size, and it must always be larger than the probability of a mistake by a single player. Theorem 7 does

²¹Our definition of common interest games is essentially the same used by Aumann and Sorin [1]. Young [25] uses the term to refer to games in which players' preferences over strategy profiles coincide. Such games are quasi-acyclic and thus also have the weak FIP. However, common interest games as we define them need not have the weak FIP.

not extend to the better-reply dynamics with mistakes of Definition 4 that only allows isolated sampling. For example, in the games in Table 7 the limiting distribution of the better-reply dynamics with mistakes puts positive probability on all the Nash equilibria.

One may object that when simultaneous sampling occurs, a player should contemplate the possibility that the other players may not adopt their sampled actions in the future. A rational player, then, should attach less significance to his payoff when sampling together with other players. Our reply to this objection is that our players are not very rational at all. They need not know that others are sampling at the same time that they are, and they need not know others' payoff functions and available options.

Definition 9 A finite game $g = (N; A; u)$ is a coordination game if: (i) $m_i = m$ for all i ; (ii) the strategies can be ordered so that $a^j = (j; \dots; j)$ is a strict Nash equilibrium for all $j = 1; \dots; m$; (iii) for all $i; j \in N$ and all $h; \ell = 1; \dots; m$, $u_i(a^h) \geq u_i(a^\ell)$ if and only if $u_j(a^h) \geq u_j(a^\ell)$; and (iv) $u_i(a^j) \geq u_i(a)$ for all $a \in A \setminus \{a^1; \dots; a^m\}$.

A game that satisfies only conditions (i) and (ii) of Definition 9 is a partial coordination game. Condition (iii) is called common ranking and condition (iv) is called diagonal dominance. A coordination game is thus a diagonal dominating, partial coordination game with common ranking. Diagonal dominating, partial coordination games have the weak FIP, but partial coordination games need not have it. The battle of the sexes is a diagonal dominating, partial coordination game, but not a coordination game, which illustrates that the sets of partial coordination and common interest games are disjoint.

However, all coordination games are of common interest; therefore, Theorem 7 implies that the limiting distribution of the better-reply dynamics with mistakes and simultaneous sampling puts all mass on the Pareto dominant Nash equilibrium. In the left coordination game in Table 7, all mass goes to the Pareto optimal equilibrium,

	L	R
T	2;2	2;0
B	0;2	3;3

	L	C	R
T	6;6	0;5	0;0
M	5;0	7;7	5;5
B	0;0	5;5	8;8

Table 7: Coordination Games with All Mass on the Pareto Dominant Equilibrium

(B; R), and not to the risk dominant equilibrium (T; L). This game is also a potential game, with potential function $\phi(T; L) = 2$, $\phi(T; R) = \phi(B; L) = 0$, $\phi(B; R) = 1$. Play does not converge to the strategy profile with the highest potential. Both risk dominance and potential functions depend on the cardinal properties of a game. As noted before, in our setup cardinal properties play no role, all choices made by our players depend upon strictly ordinal properties of the payoff functions. The right coordination game in Table 7 is due to Young, who showed that the limiting distribution of his stochastic process puts all mass on (M; C). In the better-reply dynamics with mistakes and simultaneous sampling play converges to (B; R), which is both the risk dominant and the Pareto optimal Nash equilibrium.

In Young [25] and KMR [11], at each point in time players are selected to play a given game from large populations of player types. When selected a player observes a sample of past play and uses the empirical frequencies in this sample as his beliefs about current players' choices. The player then chooses a best reply to these beliefs. They obtained convergence to the risk dominant equilibrium in 2×2 coordination games. Robson and Vega-Redondo [18] studied a similar evolutionary model, but assumed that players tend to imitate the most successful strategy. More precisely, the number of players in a population using the strategy that realizes the highest average payoff among the active strategies has a strictly positive probability of increasing and cannot decrease. They showed convergence to the Pareto optimal equilibrium in two-player common interest games.

Turning to diagonal dominating, partial coordination games, we have:

Theorem 8 Let g be a diagonal dominating, partial coordination game. The limiting

	West			Middle			East		
	L	C	R	L	C	R	L	C	R
T	2; 3; 4	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	1; 1; 1	0; 0; 0	0; 0; 0
M	0; 0; 0	1; 1; 1	0; 0; 0	0; 0; 0	3; 4; 3	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0
B	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	4; 2; 2

Table 8: A Game with no Mass on some of the Pareto Optimal Nash Equilibria

stationary distribution of the better-reply dynamics with mistakes and simultaneous sampling puts positive mass on state a^π only if a^π is a Pareto optimal Nash equilibrium.

Proof. Let a^0 be either an action profile on the main diagonal, or a Pareto dominated action profile on the main diagonal and let $T(a^0)$ be any tree with root a^0 . Since g is a diagonal dominating, partial coordination game, there is a Pareto optimal strategy profile a^π on the diagonal that yields a strictly higher payoff to all players than a^0 . Cut $T(a^0)$ at the transition from a^π and make a transition from a^0 to a^π . The eliminated transition has a resistance exceeding n and the new transition has a resistance no more than n . ■

The converse of Theorem 8 is false; Table 8 is a diagonal dominating, partial coordination game in which a Pareto optimal Nash equilibrium has zero mass in the limiting stationary distribution of the better-reply dynamics with mistakes and simultaneous sampling. The limiting distribution puts all mass on $(M; C; M)$; no mass goes to the two Nash equilibria $(T; L; W)$ and $(B; R; E)$.

Theorem 8 does not generalize to common ranking, partial coordination games. Table 9 displays one such game in which the limiting stationary distribution of the better-reply dynamics with mistakes and simultaneous sampling puts no mass on the Pareto dominant Nash equilibrium $(T; L; West)$; all mass goes to the Pareto dominated equilibrium $(B; R; East)$.

	West			Middle			East		
	L	C	R	L	C	R	L	C	R
T	3; 3; 3	0; 0; 0	0; 0; 0	0; 0; 0	1; 4; 4	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0
M	0; 0; 0	4; 4; 1	0; 0; 0	4; 1; 4	2; 2; 2	0; 0; 0	0; 0; 0	0; 0; 0	0; 3; 3
B	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	0; 0; 0	3; 3; 0	0; 0; 0	3; 0; 3	1; 1; 1

Table 9: A Game with no Mass on the Pareto Dominant Nash Equilibrium

6 Conclusions

We have modeled players having very stringent bounds on their information, memory, and computational ability, and who are both less well informed and less rational than are players in most of the literature on learning, evolutionary games, and bounded rationality. These players follow a very naive gradient behavior. They know what payoff they are typically getting from their “usual” (i.e., status quo) action, they occasionally, randomly sample something else, and they switch when they sample something having a higher payoff than they have been getting. Such players are arguably quite common, due either to inherent limitations on information and the computational complexity of a game, or because they are simultaneously engaged in many games and, having limited intellectual resources, must give short shrift to some games.

In games with sufficient regularity in the payoff functions, such as finite games with the weak FIP, our players’ behavior converges to a Nash equilibrium. We have shown that supermodular and continuous, two-player, quasi-concave games have the weak FIP. In the absence of these regularities in the payoff functions, Nash equilibrium behavior need not result.

It is possible to show that in an environment where the players’ payoff functions randomly change from time to time (i.e., regime shifts), but the probability of a regime change is very small compared to the probability of sampling unused actions, all of our results remain valid. Thus, for example, if all regimes are games with the weak FIP, then the better-reply dynamics will converge to a Nash equilibrium in

each given regime. Given the uncertainties of real life and the knowledge that real life games change character in unexpected ways from time to time, this ...nding provides a rationale for the common sense view that one should occasionally randomly check alternatives that do not appear, on the surface, to be attractive; for a regime shift may have occurred without the player noticing and, as a result, some previously unattractive action may have become very pro...table.

References

- [1] Robert J. Aumann and Sylvain Sorin. Cooperation and bounded recall. *Games and Economic Behavior*, 1:5–39, 1989.
- [2] Larry Blume and David Easley. Rational expectations and rational learning. 1992.
- [3] Eddie Dekel and Suzanne Scotchmer. On the evolution of optimizing behavior. *Journal of Economic Theory*, 57:392–406, 1992.
- [4] Glenn Ellison and Drew Fudenberg. Learning purified mixed equilibria. 1998.
- [5] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 1984.
- [6] Drew Fudenberg and David Kreps. Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367, 1993.
- [7] Drew Fudenberg and David Levine. *Theory of Learning in Games*. MIT Press, Cambridge, 1998.
- [8] James S. Jordan. Bayesian learning in normal form games. *Games and Economic Behavior*, 3:60–81, 1991.
- [9] James S. Jordan. Three problems in learning mixed-strategy equilibria. *Games and Economic Behavior*, 5:368–386, 1993.
- [10] Ehud Kalai and Ehud Lehrer. Subjective equilibria in repeated games. *Econometrica*, 61:1231–1240, 1993.
- [11] Michihiro Kandori, George Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61:29–56, 1993.

- [12] Ramon Marimon. Learning from learning in economics. In David M. Kreps and Kenneth F. Wallis, editors, *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 1*, pages 278–315. Cambridge University Press, Cambridge, 1997.
- [13] Igal Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13:111–124, 1996.
- [14] Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58:1255–1277, 1990.
- [15] Paul Milgrom and John Roberts. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3:82–100, 1991.
- [16] Dov Monderer and Lloyd Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.
- [17] Klaus Ritzberger and Jörgen Weibull. Evolutionary selection in normal-form games. *Econometrica*, 63:1371–1399, 1995.
- [18] Arthur J. Robson and Fernando Vega-Redondo. Efficient equilibrium selection in evolutionary games with random matching. *Journal of Economic Theory*, 70:65–92, 1996.
- [19] Alvin E. Roth and Ido Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate run. *Games and Economic Behavior*, 8:164–212, 1995.
- [20] Larry Samuelson. *Evolutionary Games and Equilibrium Selection*. MIT Press, Cambridge, 1997.
- [21] Fernando Vega Redondo. *Evolution, Games, and Economic Behavior*. Oxford, New York, 1996.

- [22] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19:305–321, 1990.
- [23] Mark Voorneveld and Henk Norde. A characterization of ordinal potential games. *Games and Economic Behavior*, 19:235–242, 1997.
- [24] Jörgen Weibull. *Evolutionary Game Theory*. MIT Press, Cambridge, 1995.
- [25] H. Peyton Young. The evolution of conventions. *Econometrica*, 61:57–84, 1993.
- [26] H. Peyton Young. *Individual Strategy and Social Structure*. Princeton University Press, Princeton, 1998.