WORKING PAPERS

Carlos Alós–Ferrer

Finite Population Dynamics and Mixed Equilibria

October 2000

Working Paper No: 0008

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Finite Population Dynamics and Mixed Equilibria*

Carlos Alós-Ferrer

Department of Economics, University of Vienna
E-mail: Carlos.Alos-Ferrer@Univie.ac.at

This paper examines the stability of mixed-strategy Nash equilibria of symmetric games, viewed as population profiles in dynamical systems with learning within a single, finite population. Alternative models of imitation and myopic best reply are considered and combined with different assumptions about the speed of adjustment. It is found that specific refinements of mixed Nash equilibria serve to identify focal rest points of these dynamics in general games. The relationship between both concepts is studied. In the 2×2 case, both imitation and myopic best reply yield strong stability results for the same type of mixed Nash equilibria.

1. INTRODUCTION

The interpretation of mixed-strategy Nash equilibria is an important problem in Game Theory. The first difficulty lies on the meaning of mixed strategy and whether it is reasonable or not to assume that players randomize between pure strategies with precise probabilities. The second is the issue of indifference. In mixed-strategy equilibria, the players are always indifferent between the mixed strategy they are playing and any of the pure strategies in its support (and actually any other mixed strategy with the same support). While it is true that they have no incentives to deviate, they have no incentives to remain with the same action either.

Many alternative interpretations of such "mixed equilibria" have been proposed. Building on an early suggestion from John Nash [12, pp.32-33], evolutionary game theory has attempted to solve the first difficulty through what could be labeled "the population approach." In the framework of a dynamical system, it is postulated that there is an infinite population of players for each position in the game, and that they are repeatedly randomly matched to play the game. A Nash equilibrium (in mixed strategies)

* I thank Josef Hofbauer for helpful comments and for providing me with examples 4.2 and 4.5.

of the game can be re-interpreted as a giving the proportions of players in each population who play each of the pure strategies available to them.

The second problem, though, becomes specially obvious when dynamic interpretations are postulated, i.e. when Nash equilibria are viewed as rest points of suitable dynamics. Any reasonable dynamic model would have to allow the agents to try out all the alternatives if they are truly indifferent, and hence it is difficult to see in which sense would a mixed-strategy Nash equilibrium be a rest point.

It is often argued that, in a population framework, the population approach is formally equivalent to a situation where all the agents in the population actually play a mixed strategy.[1] However, under both interpretations, the fact remains that every agent is actually indifferent between the strategy he is playing and any pure strategy in the support of the mixed strategy: both would give the same payoff, provided play in the other populations do not change (myopia). Hence, it remains unclear why would any dynamics based on, e.g., myopic best reply select Nash equilibria. Hofbauer [8] rigorously explores this question in the case of the Best Reply continuous-time dynamics for a large population of agents. In a finite population, though, such "indifference" might make the stability of Nash equilibria strongly dependent of the modeling details, and, specifically, of tie-breaking assumptions.

Rather than trying to give a single interpretation for all mixed-strategy Nash equilibria, this work focuses on symmetric games. A natural framework for a dynamic analysis of such games is given not by a multi-population context, but by a single-population one. In such a framework, several models have been proposed in the literature which deal with an explicitly finite population of boundedly rational agents who use behavioral rules based on imitation or best reply. Interaction is not necessarily limited to random matching ([14]), but encompasses also round-robin tournaments ([9, 10]) and general N-player games ([15, 5]).

This paper is specially related to Kandori *et al* [9] and Oechssler [13].

Kandori *et al* [9] study a model where agents imitate highest payoffs when playing a 2×2 game among themselves. In the case where the game has a symmetric mixed-strategy equilibrium but no pure-strategy, symmetric ones, they find that the (potentially high) speed of adjustment of their model makes the mixed profile unstable. Then they make an additional assumption, not on individual behavior but directly on the dynamics (a "contraction" relative to the mixed profile) to stabilize it.

[1] This approach presents technical problems related to, first, the existence of random matching mechanisms for infinite populations (see [7] and [2]), and, second, the aggregation of the results of the random devices underlying the mixed strategies in a large population (see [1]).

Oechssler [13] studies a similar framework (which he calls "the small population case") for best reply, but makes the explicit assumption that, whenever there are several alternative best replies, agents which are already playing one will never change. He then concentrates on convergence issues for games with $n \leq 3$ strategies.

This paper is an attempt to explore why, and if so, when are symmetric mixed-strategy Nash equilibria "stable" in a dynamic context and a finite population, specially when agents are allowed to try all alternatives that they perceive as equally worthwhile.

The first, trivial observation is that, contrary to the multi-population approach, the two "population interpretations" mentioned above are not equivalent. If we consider a population of agents, all of them playing the same mixed strategy, any one of them will remain indifferent between the pure strategies in the support of the mixed one. If the mixed strategy is interpreted as a population profile giving the proportion of agents playing each of the pure strategies in a game, then things change. If an agent changes his strategy, his action would change the population proportions and hence affect his own payoff. Or, to put it in a simpler way, the fact that an agent does not play against himself makes a difference between the population proportions and the profile of strategies that he faces. Thus, it is possible that keeping his current strategy is a strict best response to the situation given by the underlying game and the population framework. Hence, the mixed-strategy Nash equilibrium might be interpreted as a population profile where agents actually play pure strategies, and they remain with them because they actually give larger payoffs than any other alternative.

Once this observation is done, it turns out that several different dynamic approaches are able to sustain population profiles corresponding to mixed-strategy Nash equilibria as stable outcomes or rest points. Specifically, dynamics based on (myopic) best reply and imitation are considered. The importance (or lack of it) of the specific assumptions in the model at hand is illustrated considering both the slow adjustment case (as in [6]) and the quick adjustment one (as in [9]).

First, the simplest case (2×2 games) is analyzed. It is shown that mixed equilibria can be sustained by finite population dynamics when the game has no pure-strategy symmetric Nash equilibria. The exact meaning of "stability" takes different forms in the case of myopic best reply and imitation, but the qualitative features of both models turn out to be essentially identical. Under myopic best reply, the process converges to an absorbing state (i.e. a singleton recurrent communication classes) which either corresponds exactly to the mixed equilibrium or to a state next to it (depending on integer problems). Under imitation (and slow speed of adjustment), the system settles in a narrow (but non-singleton) recurrent communication class centered around a state whose population proportions approach the mixed equilibrium as the population size grows to infinity.

Second, we investigate to which extent these encouraging results extend to the general case. Under myopic best reply, we ask which mixed Nash equilibria will be absorbing states even as the population size grows, and identify a refinement of Nash equilibrium which is unrelated to previous (evolutionary) concepts. We call the associated strategies *BR-focal.*

Under imitation, the analogous attempt results in a different refinement, which we call *Imitation-focal* strategies.

We find that both concepts coincide for 2×2 games. For more general games, Imitation-focal strategies with full support are BR-focal, but the implication fails if the strategy is not completely mixed.

2. THE BASIC MODEL AND ALTERNATIVE DYNAMICS

2.1. How agents play the game

We consider a single finite population of N agents, $i = 1, ..., N$ interacting in discrete time, $t = 1, 2, ...$ to play an underlying symmetric two-player game with finite strategy space $S = \{s_1, ..., s_m\}$ and (symmetric) payoff function $\pi : S^2 \mapsto \mathbb{R}$.

A population state is (summarized by) the number of agents playing each strategy, i.e. the state space is given by $\Omega = \{(n_1, ..., n_m) \in \mathbb{N}^m \ / \ n_1 + ... + n_m = N\}$.

Notation. A typical state is denoted by ω, with $\omega(s)$ denoting its s-th coordinate, i.e. the number of agents playing s in the state ω. We call $supp(\omega) = \{s \in S \ / \ \omega(s) > 0\}$. We also keep the standard notation $supp(\sigma)$ for the strategies in the support of a given mixed-strategy σ.

Each period, each player interacts with all the other agents (*round robin tournament*). Hence, the payoff of an agent playing strategy s when the population state is ω is given by

$$\Pi(s, \omega) = \sum_{s' \in S} \omega(s') \cdot \pi(s, s') - \pi(s, s)$$

The last term ($-\pi(s, s)$) takes care of the fact that an agent will not play against himself. Alternatively, this can be reinterpreted as the expected payoff (times N) when agents are randomly matched in pairs to play the game, with uniform probabilities. It is immaterial which interpretation is taken, as long as agents take decisions according to $\Pi(s, \omega)$.

2.2. How agents learn

After play takes place, some agents will have the opportunity to update their strategies. We call this updating "learning." Whenever an agent is called to learn, he does so according to a *behavioral rule*

Definition 2.1. A Behavioral Rule for agent i is a mapping $B_i : S \times \Omega \longrightarrow \Delta(S)$ where $\Delta(S)$ is the set of probability measures over pure strategies.

$B_i(s, \omega)(s')$ is then the probability with which agent i will play strategy s' after playing strategy s when the population state was ω. This definition (potentially) incorporates the following elements:

- **Myopia.** Agents rely only on experience (B_i depends on current play), and do not perform calculations about what other agents are going to do.
- **Social Learning.** Agents learn from their own actions but also from the actions of others, since B_i also depends on ω.
- **Bounded Memory.** Agents' decisions are influenced only by the last period. See [3] for a discussion of models with longer memory.
- **Probabilistic.** Agents may display partially random behavior. This will be specially important e.g. to incorporate tie-breaking assumptions.
- **Anonymity** Rules cannot depend on the names of other agents. This is formalized through the use of the (summarized) state space Ω.

We consider two focal behavioral rules: *imitation* and *myopic best reply.*

Definition 2.2. The behavioral rule B_i is *imitative* if $B_i(s, \omega)(s') > 0 \iff s' \in supp(\omega)$ and $\Pi(s', \omega) \geq \Pi(s'', \omega)\ \forall\ s'' \in supp(\omega)$. That is, a rule is imitative if it prescribes to imitate the strategy that has given larger payoffs (any of them in case of ties).

Notation. Let ω be an state and let $s \in supp(\omega)$. For every $s' \neq s$, $s' \in S$, we denote $m(\omega, s, s')$ the state such that $m(\omega, s, s')(s) = \omega(s) - 1$, $m(\omega, s, s')(s') = \omega(s') + 1$, and $m(\omega, s, s')(s'') = \omega(s'')\ \forall\ s'' \neq s, s'$. Also, we denote $m(\omega, s, s) = \omega\ \forall\ s$.

$m(\omega, s, s')$ is the state that an agent playing s thinks that he would induce in the population if he would change his strategy from s to s', while everybody else kept his strategy.

Definition 2.3. The behavioral rule B_i is a *myopic best reply* if $B_i(s, \omega)(s') > 0 \iff \Pi(s', m(\omega, s, s')) \geq \Pi(s'', m(\omega, s, s''))\ \forall\ s'' \in S$. That is, the agent computes his best reply to the current strategy profile, (myopically) assuming that no other agents will change their strategy.[2]

[2] If the state space were to be described by S^N, i.e. the strategy profile in the population, and using the standard notation $\omega = (s_i, s_{-i})$, this definition would amount to if $B_i(s, \omega)(s') > 0 \iff \Pi(s', s_{-i})) \geq \Pi(s'', s_{-i}))\ \forall\ s'' \in S$.

It is worth emphasizing the differences between imitation and best reply in the current framework. Imitation requires extremely low computational capabilities, and absolutely no knowledge of the game. Agents merely use the information about the correspondence between actually played strategies and actually observed payoffs. Myopic best reply, on the other hand, requires potentially complex computations and explicit knowledge of the game (payoff function). Agents compare potential, unobserved payoffs that would result from a change in the current situation. In this sense, imitation and myopic best reply represent two extreme, opposite behavioral assumptions. Imitation requires an extremely low degree of rationality. Myopic best reply requires relatively high rationality.

In the present context, however, both rules produce the same results *if the population is large enough*, at least in states where all strategies are present. Since the round-robin tournament induces a continuous function of the stage model payoffs, if the population is (very) large, the payoffs in the state ω and $m(\omega, s, s')$ are arbitrarily close. Hence, there is some confusion in the literature as to whether one model is to be interpreted as imitation or best reply. Technically, a model with imitative behavioral rules is similar to a model with best reply where agents do not take into account the fact that they cannot meet themselves. In the case of a large population, the distinction turns irrelevant.

We will keep a sharp and clear distinction between imitation and best reply, for two reasons. First, we are interested in explicitly finite populations. Second, there is a big conceptual difference between imitation and (myopic) best reply in terms of the degree of rationality they assume.

Further, note that the two rules above incorporate the following

Assumption 1. Whenever a behavioral rule specifies several possible best strategies, the agent chooses all of them with strictly positive probability, without exception.

This assumption is *conceptually* very important. For instance, one might assume that, under myopic best reply, agents already playing a best reply to the current profile do not change strategies even if there are other best replies available, because they have no incentive to deviate. We explicitly depart from such assumptions because they would prevent us from tackling the original problem. The position of this paper is that, if there are several available best replies, an agent has no incentive *not* to deviate, and drift will eventually appear.[3]

[3] Oechssler [13] argues that there are always costs of changing a decision. Even in this case, it can be easily assumed that the probability of remaining with the current, optimal strategy is close to one, but it seems reasonable to allow for drift to other alternative optimal strategies. The results remain unchanged under this approach.

2.3. When agents learn

The concept of *inertia* is standard in learning models. It is assumed that not all agents are able to learn all periods. This is introduced in different ways in different models. Kandori et al [9] argue that inertia is a justification for myopia.[4] In their model, though, the probability of being able to learn is independent across agents. In other models, it is often argued that, if the population is large, it is unrealistic to assume that all agents are able to revise their strategies simultaneously. As an extreme assumption, it is postulated that the probability of two agents learning simultaneously is zero, which gives rise to models where, each period, only one randomly sampled agent is able to revise his strategy.[5]

In all cases, though, it is always assumed that the probability of a given agent being able to revise in a given period is positive. This is merely an anonymity requirement, and does not imply that the dynamics must allow for simultaneous revisions as in [3]. As in [6], it is allowed that only one agent revises each period, provided that any one of them may, potentially, be the chosen one. Symmetrically, it is assumed that no agent has the guarantee of always being able to revise. We explicitly make this assumption:

Assumption 2. For all t, and for each $i = 1, ..., N$, the probability that agent i is able to revise at period t is strictly positive and less than one (although not necessarily independent from that of other agents).

With this *caveat*, we explicitly distinguish between the two approaches mentioned above.

Definition 2.4. We say that a model presents *independent inertia* λ if, every period, each agent is able to revise his strategy with probability $0 < 1 - \lambda < 1$, independent across agents and across periods.

We say that a model presents *non-simultaneous learning* if, every period, a single agent is randomly sampled and this agent is the only one able to revise his strategy.

The interest of these two alternative formulations lies on their relationship to the speed of adjustment of the postulated dynamics. A model with *independent inertia* could be described as one of *quick adjustment*. Each period, all the agents in a fraction of the population (which can be close to one) are able to simultaneously revise their strategies. If we were to think of large populations or short time periods, this implies a large number of revisions per time unit. On the other hand, models with *non-simultaneous*

[4] The introduction of inertia as in [9] has surprising implications in models with memory (see [3]).

[5] This question is related to standard arguments for approximation of discrete-time systems by their continuous-time counterparts. See [6, 4].

learning have *slow adjustment*. Only one agent revises at a time, and, in N periods, always N agents would have revised their strategies.[6]

2.4. Learning processes

We are now able to consider different models. We call (single-population) *Learning Process* any model where:

- **Interaction:** A single finite population of N agents, $i = 1, ..., N$ interacting in discrete time, $t = 1, 2, ...$ to play an underlying finite, symmetric two-player game according to a round-robin tournament (or, alternatively, random matching with evaluation of expected payoffs).
- **Speed of Adjustment:** Each period, there is a specification of when are agents able to learn (e.g. independent inertia or non-simultaneous learning). This specification does not depend on the time index, and *ex ante* the probability of a given agent being able to revise in a given period is strictly positive and less than one.
- **Learning:** When an agent is able to learn, it does so according to a pre-specified behavioral rule (e.g. imitation or myopic best reply). In case of indifference between several options, he chooses all of them with positive probability.

Note that any such Learning Process defines a finite Markov Chain on the state space Ω, and hence can be studied using the standard techniques from the theory of stochastic processes (see e.g. [11]).

The dynamics induces probabilities of transition among states which we denote $P(\omega, \omega')$. The matrix P of transition probabilities is called *transition matrix* of the process. We also denote $P^k(\omega, \omega')$ the probability that the process is at state ω' given that k periods before it was at state ω.

Given a stochastic process with transition matrix P and finite state space Ω, we say that two states ω, ω' *communicate* if there exist $t, t' > 0$ such that $P^t(\omega, \omega') > 0$ and $P^{t'}(\omega', \omega) > 0$. This defines an equivalence relation whose equivalence classes are called *communication classes*. Qualitatively, all states in the same communication class share the same features. For example, the process cannot eventually settle in a strict subset of a communication class, but rather on a full class.

[6] This comparison can be made more rigorously. If we view the dynamical systems we are describing as stochastic processes, it is a standard result that their qualitative features (e.g. recurrent versus transient states, see next section) depend only on the set of positive probability paths. Fix the state space, the set of agents, and the behavioral rules they use, and consider the set of possible dynamics, differing in their speed of adjustment (inertia) assumptions. It is clear that, under Assumption 2, any positive probability path of the dynamics with non-simultaneous learning is a positive probability path of any other dynamics. Also, any positive probability path of any dynamics is a positive probability path of the dynamics with independent inertia. Hence, in this sense, these two dynamics are respectively minimal and maximal.

A communication class C is *transient* if there exist $\omega \in C,\ \omega' \in \Omega \backslash C$ such that $P(\omega, \omega') > 0$. Classes which are not transient are called *recurrent*. The process will eventually leave all transient classes and settle in a recurrent class. If there is just one recurrent class, the process is called it ergodic; if there are more than one, the process exhibits path dependence, i.e., it might settle down in different classes depending on the initial conditions.

DEFINITION 2.5. Consider a learning process with transition matrix P. A population state ω is called *absorbing* if $P(\omega, \omega) = 1$.

An absorbing state forms a singleton recurrent communication class. Once the process gets to an absorbing state, it will never leave it. This is the first condition we will be interested in. Absorbing states are analogous to stationary states of deterministic dynamics.

The *basin of attraction* $B(C)$ of a recurrent class C is the set of states from which there are positive probability paths that reach C in finite time, i.e. $B(C) = \{\omega \in \Omega \setminus C \ / \ \exists t \in \mathbb{N}, \omega' \in C$ such that $P^t(\omega, \omega') > 0\}$. Abusing notation, we also speak of the basin of attraction of an absorbing state, $B(\omega) = B(\{\omega\})$.

3. THE 2 × 2 CASE

Consider a 2×2 symmetric game with payoff matrices given by

	A	B
A	(a,a)	(b,c)
B	(c,b)	(d,d)

where a, b, c, d are real numbers, and there exists a unique symmetric Nash Equilibrium in mixed strategies (this implies that $(a, b) \neq (c, d)$).

Let $(\sigma, 1 - \sigma)$ be the mixed strategy which gives the Nash equilibrium, and let $n^* = \sigma \cdot N$, i.e. the population proportion that corresponds to σ if the mixed strategy is to be interpreted in population terms.

If his opponent is playing the mixed strategy $(\sigma, 1 - \sigma)$, a player is indifferent between playing A and B, i.e.

$$\sigma \cdot a + (1 - \sigma) \cdot b = \sigma \cdot c + (1 - \sigma) \cdot d$$

which implies

$$\sigma = \frac{(d - b)}{(a - b) + (d - c)}$$

For σ to be readily interpreted as a population profile, n^* should be an integer. As we will see, the analysis varies (slightly) depending on whether it is or not. For easy reference, we give a name to this condition (but do not assume it).

Condition (INT). $n^* = \sigma \cdot N$ is an integer, i.e. $n^* \in \{1, ..., N-1\}$.

Notice that, $n^* \cdot a + (N - n^*) \cdot b = n^* \cdot c + (N - n^*) \cdot d$, i.e.

$$n^* \cdot (a - c) + (N - n^*) \cdot (b - d) = 0.$$

Consider a population state ω where exactly n^* agents are playing strategy A and $N - n^*$ agents play B. This population profile corresponds to the mixed strategy Nash equilibrium. Denote $\Pi(s, n)$ the payoff of an agent playing strategy $s = A, B$ when exactly n agents in the population are playing A. It follows that

$$\Pi(A, n) = (n - 1) \cdot a + (N - n) \cdot b$$
$$\Pi(B, n) = n \cdot c + (N - n - 1) \cdot d.$$

Consider any learning process applied to the described game. The state space can be summarized by $\{0, 1, 2, ..., N\}$, where state n is identified with all the situations where exactly n agents play strategy A. Under (INT), the symmetric Nash Equilibrium corresponds then to state n^*.

3.1. Myopic Best Reply

Suppose the learning process is based on myopic best reply. An agent playing A would remain with his current action if

$$\Pi(A, n) > \Pi(B, n - 1)$$

i.e. if the payoff of playing A when there are n agents in the population playing A (including himself) is larger than the payoff he would obtain if he were to switch to B, facing then a state where $n - 1$ agents would play A. Analogously, an agent playing B would remain with his action if

$$\Pi(B, n) > \Pi(A, n + 1).$$

Ties are broken randomly, i.e. we want to explicitly allow all possibilities whenever an agent faces an indifference situation (Assumption 1).

PROPOSITION 3.1. *Consider the* 2×2 *game above, and assume (INT). In any learning process with myopic best reply, the state* n^* *corresponding to the mixed strategy Nash Equilibrium is absorbing if and only if* $c > a$ *and* $b > d$*, i.e. neither (A,A) nor (B,B) are Nash Equilibria.*

Proof. Note that $\Pi(A, n^*) - \Pi(B, n^* - 1) = (n^* - 1) \cdot a + (N - n^*) \cdot b - (n^* - 1) \cdot c - (N - n^*) \cdot d = (n^* - 1) \cdot (a - c) + (N - n^*) \cdot (b - d) = c - a$.

Hence, an agent playing A will keep his action with probability one if and only if $c > a$. Analogously,

$\Pi(B, n^*) - \Pi(A, n^* + 1) = n^* \cdot c + (N - n^* - 1) \cdot d - n^* \cdot a - (N - n^* - 1) \cdot b = n^* \cdot (c - a) + (N - n^* - 1) \cdot (d - b) = b - d$

It follows that an agent playing B will keep his action with probability one if and only if $b > d$. ∎

It is worth noticing that in state n^*, agents do not remain with their current actions because they are indifferent. Quite on the contrary, they do so because their current actions are strictly better than the alternative. Of course, if payoffs are averaged across interactions or re-interpreted as expected payoffs in a random matching framework, the advantage of keeping the current action (e.g. $\frac{b-a}{N-1}$) tends to zero as the population size grows to infinity, but for any fixed N it is still positive. This provides a tempting interpretation for symmetric mixed strategy equilibria at the population level.

Example 3.1. Consider a learning process with independent inertia and myopic best reply. This dynamics is similar to the one in Kandori *et al* [9], with the difference that there agents imitated the action which led to the highest payoffs instead of playing a best reply.

Assume (INT). Consider the state n^*, and suppose $c < a$, $b < d$ (e.g. a Coordination game). On state n^*, all agents will change their current actions if they get the opportunity to revise. Although it could happen that just two agents get revision opportunities and they exchange their strategies, the probability of remaining in state n^* is strictly less than one. Hence n^* is not absorbing.

Proposition 3.2. *Consider the 2×2 game above with $c > a$, $b > d$. In any learning process with myopic best reply, there exists n_A, $n_A - 1 < n^* < n_A$ such that the singleton set $C = \{n \in \Omega \ / \ n_A - 1 < n < n_A\}$ is a recurrent class (unless n_A is exactly an integer). All other states are transient and the process converges to C from any initial condition. Moreover, $\lim_{N\to\infty} \frac{n_A}{N} = \sigma$*

If n_A is exactly an integer, then, under independent inertia the process is irreducible. Under non-simultaneous learning, the set $C = \{n_A - 1, n_A\}$ is a recurrent class and all other states are transient.

Proof. Take any state n. Then,

$$\begin{aligned} \Pi(A, n) - \Pi(B, n-1) &= (n-1) \cdot a + (N-n) \cdot b - (n-1) \cdot c - (N-n) \cdot d \\ &= (n-1) \cdot (a-c) + (N-n) \cdot (b-d) \end{aligned}$$

which is decreasing in n, and equal to zero if and only if

$$n = n_A = \frac{N \cdot (b-d) + (c-a)}{(b-d) + (c-a)}.$$

Hence, for $n > n_A$, A-players which are given the opportunity to revise switch to B with probability one.

Analogously,

$$\begin{aligned} \Pi(B,n) - \Pi(A,n+1) &= n \cdot c + (N-n-1) \cdot d - n \cdot a - (N-n-1) \cdot b \\ &= n \cdot (c-a) + (N-n-1) \cdot (d-b) \end{aligned}$$

which is increasing in n, and equal to zero if and only if

$$n = n_B = \frac{(N-1) \cdot (b-d)}{(b-d) + (c-a)}.$$

It follows that, given opportunity, B-players switch to A whenever $n < n_B$.

Notice $n^* = N \cdot \frac{(b-d)}{(b-d)+(c-a)}$, i.e. $n_B < n^* < n_A$.

On any state n such that $n_B < n^* < n_A$, all agents will keep their current actions regardless of whether they get the opportunity to revise or not. Hence n is absorbing.

Analogously, states out of C are in transient classes. Note that $n_A - n_B = 1$, i.e. C is a singleton unless n_A is exactly an integer. In the first case, the result follows immediately. In the second case, under non-simultaneous learning it suffices to observe that $\{n_A - 1, n_A\}$ is recurrent. Under independent inertia, the fact that $P(n_B, N) > 0, P(n_A, 0) > 0$ makes the process irreducible, i.e. the whole state space is a communication class and the process never settles down.

The proof is completed observing that the quotients n_A/N and n_B/N approach σ as N grows to infinity. ▌

This result tells us that, essentially, under myopic best reply, in a 2×2 game without symmetric pure-strategy equilibria, the mixed strategy equilibrium is the essential prediction. This takes the form of a unique absorbing state which coincides with n^* whenever n^* is an integer, and is next to it when not. The case when n_A is an integer for a given N (and hence n^* is not) is merely an extreme form of an integer problem.

It is interesting to observe that this result is independent of the inertia/speed of adjustment assumptions (except when n_A is an integer). Only the speed of convergence might be affected by such details.

Remark 3. 1. Suppose the learning process of the previous Proposition presents independent inertia λ. Consider any state n. If $n > n_A$, with positive probability $n - n^*$ A-agents are the ones getting opportunity to revise and they will switch to B, therefore the process reaches state n^*. If $n < n_B = n_A - 1$, with positive probability $n^* - n_B$ B-agents switch to A and the process reaches n^*. If $n_B < n < n_A$, neither A-agents nor B-agents will change when given opportunity, hence the state n is absorbing.

Convergence, though, might be slow. The process might keep "overshooting" the set C for a long while, specially if n_A, n_B are close. Technically, the whole set $\Omega \setminus C$ forms a single transient communication class.

Assume instead non-simultaneous learning. For each state $n > n_A$, there is probability $\frac{1}{n}$ that an A-player is given opportunity to revise, hence switching to B and moving the process to state $n - 1$. With the complementary probability, a B-player is given opportunity to revise, which he won't do, and hence the process remains in state n. Hence, the process moves inexorably towards C, without any overshooting possibility. Technically, each state in $\Omega \setminus C$ forms its own transient communication class.

Intuitively, under independent inertia (i.e. with a single, large transient class), convergence is slowed down by the exploding number of possibilities that appear because of simultaneous revisions, and the length of the "long-run" until the process hits C will grow exponentially with the size of the population. Under non-simultaneous learning (i.e. a large number of small transient classes), though, speed of convergence does not essentially depend on the size of the population, or, more precisely, since the number of possibilities from a given state does not depend on the population size, the expected time until C is reached grows linearly with N.

3.2. Imitation

Suppose now that the learning process is based on imitation. Automatically, the states $n = 0$ and $n = N$ are absorbing, since if only one action is observed, no other action can be mimicked. In the "interior" of the state space, an agent playing A would remain with his current action if

$$\Pi(A, n) > \Pi(B, n)$$

i.e. if the payoff of playing A when there are n agents in the population playing A (including himself) is larger than the payoff other agents actually playing B have obtained. Analogously, an agent playing B would remain with his action if

$$\Pi(B, n) > \Pi(A, n).$$

Ties are broken randomly (recall Assumption 1).

PROPOSITION 3.3. *Consider the* 2×2 *game above. In any learning process with imitation, states* 0 *and* N *are absorbing. No state* $n \in \{1, ..., N-1\}$ *can be absorbing.*

Proof. Let $0 < n < N$, i.e. both strategies are played in state n. If $\Pi(A, n) < \Pi(B, n)$, A-players will imitate B, and hence the state is not absorbing. If $\Pi(A, n) > \Pi(B, n)$, B-players will imitate A. If $\Pi(A, n) = \Pi(B, n)$, neither A-players nor B-players have any imitation-based incentive to remain with their current action. Due to the random breaking of ties, it follows that $P(n, n) < 1$ and n can not be absorbing.

States 0 and N are trivially absorbing since an unobserved action can not be imitated. ▮

Under imitation, only *monomorphic* states, i.e. states where all agents are playing the same strategy, can be absorbing. If two different strategies are present and they give different payoffs, then the one giving larger payoffs will be imitated. If they give exactly the same payoffs, then there is always positive probability that an agent drifts from one strategy to the other, hence drawing the process out of the state. It follows that indifference then blurs the intuitive difference between a Hawk and Dove game, where dynamics should lead "in the direction" of n^*, and a Coordination game, where dynamics should lead "away" from n^*.

PROPOSITION 3.4. *Consider the* 2×2 *game above with* $c > a$, $b > d$. *In a learning process with imitation and non-simultaneous learning, there exists* $\hat{n} \in [0, N]$ *such that the set*

$$C = \begin{cases} \{\hat{n}-1, \hat{n}, \hat{n}+1\} & \text{if} \quad \hat{n} \text{ is an integer} \\ \{\lfloor \hat{n} \rfloor, \lceil \hat{n} \rceil\} & \text{if} \quad \text{not} \end{cases}$$

is a recurrent class. The states 0 *and* N *are absorbing, but the process converges to* C *from any initial condition* $1, ..., N-1$. *Moreover,* $\lim_{N \to \infty} \frac{\hat{n}}{N} = \sigma$.

Proof. Take any state n. Then,

$$\begin{aligned} \Pi(A, n) - \Pi(B, n) &= (n-1) \cdot a + (N-n) \cdot b - n \cdot c - (N-n-1) \cdot d \\ &= n \cdot (a-c) + (N-n) \cdot (b-d) + (d-a) \end{aligned}$$

which is decreasing in n, and equal to zero if and only if

$$n = \hat{n} = \frac{N \cdot (b-d) + (d-a)}{(b-d) + (c-a)}.$$

Hence, for $n > \hat{n}$, A-players which are given the opportunity to revise switch to B with probability one, and, for $n < \hat{n}$, B-players will be the ones switching. It follows that, if $n > \hat{n}$, $P(n,n) + P(n,n+1) = 1$, and, if $n < \hat{n}$, $P(n,n) + P(n,n-1) = 1$. Moreover, both probabilities in the sum are always positive, i.e. the process always moves in the direction of $\hat{n}$.

Suppose first that $\hat{n}$ is an integer. In state $\hat{n}$, strategies A and B give the same payoff. Hence, if an A-player is given revision opportunity, he will either keep his strategy or switch to B, and conversely for a B-player. However, $P(\hat{n}-1,\hat{n}) = P(\hat{n}+1,\hat{n}) = 1$, i.e. the process will always return to $\hat{n}$ after one period.

It follows immediately that C is a recurrent class, and the basins of attraction of states 0 and N are empty. Since there are no other recurrent classes, the result follows.

Suppose now that $\hat{n}$ is not an integer. In state $\lfloor \hat{n} \rfloor$, if a B-player is given revision opportunity, he will switch to A and drive the process to state $\lceil \hat{n} \rceil$. Conversely, from state $\lceil \hat{n} \rceil$ an A-player will switch to B and drive the process to $\lfloor \hat{n} \rfloor$. It follows that C is a recurrent class and the process converges to C from any initial condition except 0 and N.

It remains to observe that $\lim_{N\to\infty} \frac{\hat{n}}{N} = \sigma$. ▮

This result shows us that, under imitation and non-simultaneous learning, in a 2×2 game without symmetric pure-strategy equilibria, the mixed strategy equilibrium is again the essential prediction. This takes the form of a very narrow recurrent class around $\hat{n}$, which converges to the appropriate proportion as N grows. This recurrent class fails to be a singleton merely because of either integer problems or exact indifference. Outside the set, there are two other absorbing states created by the very nature of imitation, which have empty basins of attraction.

Remark 3. 2. Note that $\hat{n} \in (n^*, N] \iff d > a$, i.e. $\hat{n}$ is pulled to one or the other side of n^* by the payoffs of the symmetric, pure-strategy profiles.

What happens under independent inertia? Obviously, there are only two absorbing states, 0 and N, and the rest of states are in a single, large transient class. Intuitively, the speed of adjustment is too high and blurs the result. In actual simulations, though, $\hat{n}$ will still play a role, since the actual probabilities of transition still favour a trend towards $\hat{n}$. It is the existence of low but positive probabilities of transition, say, from $\hat{n}$ to 0 or N which (apparently) destroys the result. Kandori *et al* [9] solve the problem by postulating a "contraction" of the dynamics relative to a a mixed profile, i.e. they explicitly assume that the distance to the reference mixed-strategy diminishes. This is an assumption made directly on the dynamics, which is difficult to trace back to individual behavior.

3.3. Extended Example: Hawk and Dove

Consider the well-known “Hawk and Dove” game:

	H	D
H	$(\frac{V-C}{2},\frac{V-C}{2})$	(V,0)
D	(0,V)	$(\frac{V}{2},\frac{V}{2})$

where $C > V > 0$. This game has two asymmetric pure-strategy Nash equilibria, (H, D) and (D, H), and a single symmetric Nash equilibrium where both players play a mixed strategy which gives weight V/C to the strategy H (Hawk). Hence, in our notation, $n^* = (V/C) \cdot N$.

Consider any learning process based on myopic best reply. We can apply Proposition 3.2. Direct computation shows that

$$n_D = (N-1)\frac{V}{C},\ n_H = \frac{(N-1)\cdot V + C}{C} = (N-1)\frac{V}{C} + 1$$

Hence, from any initial condition the process converges to an absorbing state which lies between n_D and $n_D + 1$. If n^* is an integer, then $C = \{n^*\}$. If not, then C is the singleton formed by the closest state to n^*, except in the extreme case when $2 \cdot n^*$ is exactly an integer and n^* is not.

Consider a learning process based on imitation and non-simultaneous learning. We can apply Proposition 3.4. Direct computation shows that

$$\hat{n} = \frac{N \cdot V + C}{C} = N \cdot (V/C) + 1 = n^* + 1$$

Hence, from any initial condition except 0 and N, the process converges to a recurrent class formed by the two closest states to $n^* + 1$ (if n^* is not an integer), or by the states $n^*, n^* + 1, n^* + 2$ (if n^* is an integer).

Both results point to the significance of population profiles where (approximately) a fraction V/C of the population play strategy H, i.e., the profile which corresponds to the mixed-strategy Nash equilibria.

3.4. Coordination Games: a Remark

Suppose that $c < a$, $b < d$, i.e. we have a coordination game. Then, the mixed-strategy equilibria does not correspond to a stable configuration in any sense of the word.

Under myopic best reply, it is still true that $n_B < n^* < n_A$. However, now A-players switch strategies for $n < n_A$, and B-players for $n > n_B$, which means the dynamics points towards the two monomorphic states. The set $\{n \in \Omega \ / \ n_B < n < n_A\}$, which is essentially a singleton next to n^*, is the intersection of both basins of attraction. Hence, n^* marks the boundary between both basins of attraction, even if it fails to be absorbing.

Under imitation, A-players switch strategies for $n < \hat{n}$ and B-players for $n > \hat{n}$, i.e. again the dynamics points towards the monomorphic states.[7] The state $\hat{n}$ marks the exact boundary of the two basins of attraction (technically, because of the tie-breaking assumption, it belongs to both, and hence fails to be absorbing). Since it still approaches n^* as N grows, this still can be taken to point at the significance of the mixed equilibrium.

Regardless of the complications and technical details which spread from a discrete-time, stochastic, behavior-based dynamics, qualitatively the situation is analogous to a well-behaved, continuous-time, deterministic system. For a 2×2 game with two pure-strategy, symmetric equilibria, the mixed equilibrium identifies a repelling point of the dynamics, whereas the two pure-strategy equilibria are stable. In absence of the pure-strategy equilibria, the mixed-strategy one is then globally attracting.

4. THE GENERAL CASE: MYOPIC BEST REPLY

Consider now a general learning process with a single finite population of N agents, $i = 1, ..., N$ interacting in discrete time, $t = 1, 2, ...$, according to a round-robin tournament, to play an underlying symmetric two-player game with finite strategy space $S = \{s_1, ..., s_m\}$ and (symmetric) payoff function $\pi : S^2 \mapsto \mathbb{R}$.

Notation. Let A be the payoff matrix associated to the payoff function π, and denote by e_s a vector of $m = |S|$ coordinates, all of them zero except for a 1 in the position corresponding to strategy s. Notice that $\Pi(s; \omega) = e_s^T \cdot A \cdot (\omega - e_s)$, where e_s^T is the transposed of the vector e_s (vectors are assumed to be column vectors, transposed vectors are row vectors). If we define the associated mixed strategy by $\sigma = \frac{1}{N} \cdot \omega$, then $\Pi(s; \omega) = N \cdot \pi(s, \sigma) - \pi(s, s)$.

4.1. Absorbing states

Lemma 4.1. *Consider any learning process based on myopic best reply. Let $\omega = (n_1, ..., n_m) \in \Omega$, and define the corresponding mixed strategy σ by $\sigma_i = \frac{n_i}{N}$ for all $i = 1, ...m$. Then, ω is an absorbing state if and only if*

$$\pi(s, \sigma) - \pi(s', \sigma) + \frac{1}{N} \left(\pi(s', s) - \pi(s, s)\right) > 0 \; \forall \, s \in supp(\sigma), \; \forall \, s' \neq s$$

Proof. The state ω is absorbing if and only if no agent will change its strategy if given opportunity to revise. Under myopic best reply, this

[7] This dynamics is analyzed in Kandori *et al* [9].

amounts to

$$\Pi(s,\omega) > \Pi(s', m(\omega,s,s'))\ \forall\ s \in supp(\omega),\ \forall\ s' \in S \setminus \{s\}$$

Fix $s \in supp(\omega)$ and $s' \neq s$. This condition is equivalent to

$$\sum_{s''\in S} \omega(s'')\cdot\pi(s,s'') - \pi(s,s) > \sum_{s''\in S} \omega(s'')\cdot\pi(s',s'') - \pi(s',s') - \pi(s',s) + \pi(s',s')$$

or, equivalently,

$$\sum_{s''\in supp(\omega)} \omega(s'')\cdot(\pi(s,s'') - \pi(s',s'')) + \pi(s',s) - \pi(s,s) > 0$$

which yields the required condition. ▌

Remark 4. 1. With the above matrix notation A, e_s, then, a state ω is absorbing if and only if, for all $s \in supp(\omega)$ and for all $s' \neq s$,

$$\begin{aligned} e_s^T\cdot A\cdot\omega - e_s^T\cdot A\cdot e_s > e_{s'}^T\cdot A\cdot\omega - e_{s'}^T\cdot A\cdot e_{s'} &\iff \\ (e_s - e_{s'})^T\cdot A\cdot\omega - (e_s - e_{s'})^T\cdot A\cdot e_s > 0 &\iff \\ (e_s - e_{s'})^T\cdot A\cdot(\omega - e_s) > 0 & \end{aligned}$$

Suppose the game has a mixed-strategy symmetric equilibrium given by the strategy σ. The condition that makes possible to interpret this strategy exactly as a population profile is:

Condition (INT). There exists a state $\omega^* = (n_1, ..., n_m) \in \Omega$ such that $\sigma_i = \frac{n_i}{N}$ for all $i = 1, ...m$.

PROPOSITION 4.1. *Consider any learning process with myopic best reply. If the game has a completely-mixed symmetric Nash equilibrium given by the strategy σ, and assuming (INT), then the corresponding state ω^* is absorbing if and only if*

$$\pi(s',s) > \pi(s,s)\ \forall\ s, s', s' \neq s$$

i.e. each pure strategy is a strict worst reply against itself. In particular, the mixed-strategy Nash equilibrium will never be absorbing if the game has any strict, pure-strategy Nash equilibrium.

Proof. If $supp(\omega) = S$ and σ describes a Nash equilibrium, then $\pi(s,\sigma) = \pi(s',\sigma)$ for all $s, s' \in S$. The result follows then from Lemma 4.1. ▌

Proposition 4.1 was already established in Oechssler [13, Proposition 1], where it is noted that a completely mixed profile as above cannot be invaded by a single mutant playing a pure strategy.

4.2. BR-focal strategies

In the 2×2 case, we saw how an absorbing state exists for a fixed N except under extreme integer problems, and whenever n^* is a well-defined state, it is precisely the absorbing one. This motivates the following definition.

DEFINITION 4.1. A mixed strategy σ is *frequently absorbing under best reply* or simply *BR-focal* if there exists N_0 such that, for all $N > N_0$ such that $N \cdot \sigma(s)$ is an integer for all $s \in S$, the state $\omega = N \cdot \sigma$ is absorbing in any learning process with myopic best reply.

In the Hawk and Dove example, only the Nash equilibrium is BR-focal. We confirm now this intuition for general games.

THEOREM 4.1. *A mixed strategy σ such that $\{\sigma(s)\}_{s\in S}$ are rational numbers is frequently absorbing under best reply if and only if*

(a) (σ, σ) is a Nash Equilibrium of the game.
(b) $\forall\ s' \in S$, if $\pi(s', \sigma) = \pi(\sigma, \sigma)$ then $\pi(s', s) > \pi(s, s)\ \forall\ s \in supp(\sigma)$.

Proof. "Only if." (a) Let σ be frequently absorbing under best reply, and let $s' \in S$ be any pure strategy. It is enough to show that $\pi(s', \sigma) \leq \pi(\sigma, \sigma)$.
Let $s \in supp(\sigma)$. It follows that $\pi(s, \sigma) = \pi(\sigma, \sigma)$. Suppose $\pi(s', \sigma) > \pi(\sigma, \sigma) = \pi(s, \sigma)$. Then, $\pi(s, \sigma) - \pi(s', \sigma) < 0$ and it is possible to find $N > N_0$ such that $N \cdot \sigma$ is a state and

$$\pi(s, \sigma) - \pi(s', \sigma) + \frac{1}{N}\left(\pi(s', s) - \pi(s, s)\right) < 0$$

a contradiction with Lemma 4.1.
(b) Suppose $\pi(s', \sigma) = \pi(\sigma, \sigma)$ and let $s \in supp(\sigma)$. Since $\pi(s, \sigma) = \pi(\sigma, \sigma)$, it follows from Lemma 4.1 that $\pi(s', s) - \pi(s, s) > 0$.
"If." Consider any N such that $N \cdot \sigma$ is a state.
Consider any $s' \in S$. If $\pi(s', \sigma) = \pi(\sigma, \sigma)$, then $\pi(s', \sigma) = \pi(s, \sigma)$ for all $s \in supp(\sigma)$ and hence, by (b), $\pi(s', s) > \pi(s, s)$. It follows that

$$\pi(s, \sigma) - \pi(s', \sigma) + \frac{1}{N}\left(\pi(s', s) - \pi(s, s)\right) = \frac{1}{N}\left(\pi(s', s) - \pi(s, s)\right) > 0$$

If $\pi(s', \sigma) < \pi(\sigma, \sigma)$, then $\exists\ N_0$ such that, $\forall\ s \in supp(\sigma)$, $\forall\ N > N_0$,

$$\pi(s, \sigma) - \pi(s', \sigma) + \frac{1}{N}\left(\pi(s', s) - \pi(s, s)\right) > 0$$

The conclusion follows from Lemma 4.1. ∎

A BR-focal strategy corresponds to a Nash equilibrium and, whenever it is possible to reinterpret it as a finite population profile, it turns out to be absorbing in any dynamics based on best reply. Notice again that the profile is not absorbing because of any tie-breaking assumption, but because agents are actually earning strictly more than they would earn if they were to deviate.[8]

Theorem 4.1 states that a strategy σ is BR-focal if, whenever an s-player faces an alternative s' which is just as good as s against σ, the part of the total payoff against σ that he fails to realize because he cannot play against himself ($\pi(s,s)$) is lower than the part of that same payoff that he will fail to realize after switching to s' due to the fact that, then, there will be one s-player less ($\pi(s',s)$).

Example 4.1. Consider the symmetric game with payoff matrix

	A	B	C
A	0	1	2
B	2	0	1
C	1	2	0

There is a unique symmetric Nash equilibrium, given by the mixed strategy $\sigma = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The payoff of any pure strategy against σ is equal to 1.

Let $N = 3 \cdot k, k \geq 0$, and consider the state $\omega = (k, k, k)$. The average payoff of any pure strategy player is given by

$$\frac{1}{3k-1}\left((k-1)\cdot 0 + k\cdot 1 + k\cdot 2\right) = \frac{3k}{3k-1} > 1$$

The lack of an interaction yielding payoff zero makes the average payoff strictly greater than zero. That is, strategy A is disadvantageous against itself (payoff 0), but this effect is reduced by the fact that an agent does not play against himself.

Consider now, for example, an A-player deciding whether to switch to strategy B. Strategy B is very advantageous against A (payoff 2). If the other agents remain with their strategies, his average payoff would be

$$\frac{1}{3k-1}\left((k-1)\cdot 2 + k\cdot 0 + k\cdot 1\right) = \frac{3k-2}{3k-1} < 1$$

The fact that the agent himself was originally an A-player reduces the potential advantage of a switch to strategy B.

[8] Hence, in the "Small Population Case" considered in [13], it can be argued that the tie-breaking assumption was actually harmless.

Analogously, the average payoff if an A-player would switch to C would be equal to 1. Hence, the strict (myopic) best reply of an A-player is to keep playing A.

The previous theorem shows that σ is frequently absorbing. It turns out, though, that σ is also an ESS.

Oechssler [13] proves that, for 3×3 games with a unique, completely mixed equilibrium, the best reply process (with independent inertia) converges to the corresponding state if and only if said state is absorbing. The next example shows a 4×4 game with a completely-mixed BR-focal equilibrium, which corresponds to an absorbing but unstable state in any learning process with myopic best reply.

EXAMPLE 4.2. Consider the symmetric game with payoff matrix

	A	B	D	D
A	0	1	x	x
B	1	0	x	x
C	x	x	0	1
D	x	x	1	0

Let $0 < x < \frac{1}{2}$. There are three symmetric, mixed-strategy Nash equilibria, corresponding to strategies $\sigma_1 = (\frac{1}{2}, \frac{1}{2}, 0, 0), \sigma_2 = (0, 0, \frac{1}{2}, \frac{1}{2})$, and $\sigma_1 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. By Theorem 4.1, all three of them are BR-focal.

Let $N = 4k$ and $\omega = (k, k, k, k)$. Since σ_3 is BR-focal, no agent will deviate under myopic best reply. However, consider state $\omega' = (k-1, k, k+1, k)$. The payoffs of an A-player are $k + (2k+1)x$, while if he switched to strategy D he would obtain $(2k-1)x + k + 1 > k + (2k+1)x$, that is, the process leads away from ω from states close to it. Qualitatively, the game is similar to a 2×2 Coordination Game, where σ_1, σ_2 correspond to the attracting equilibria and σ_3 to the unstable mixture between them.

4.3. Discussion and comparison with ESS

The similarity of Theorem 4.1 and the original definition of "Evolutionarily Stable Strategy" (ESS) is remarkable. An ESS (a concept which makes most sense in the framework of an explicitly infinite population) is a strategy σ such that (a) (σ, σ) is a Nash Equilibrium, and (b) For any mixed strategy $\sigma' \neq \sigma$ such that $\pi(\sigma', \sigma) = \pi(\sigma, \sigma$, it follows that $\pi(\sigma, \sigma') > \pi(\sigma', \sigma')$. That is, an ESS requires a Nash equilibrium to be resistant to deviations in the sense that, if agents are indifferent between the ESS and another (mixed) strategy, the ESS performs better against the deviating strategy than itself does.

In contrast, a BR-focal strategy σ must build a Nash equilibrium which is resistant to deviations in the sense that, if agents are indifferent between σ and another (pure) strategy, then the deviating strategy performs *better* against strategies in the support of σ than each of them does against itself.

This condition arises from the individual, myopic nature of the considered deviations. If the population profile corresponds to σ, the agent is not really facing a mixed strategy σ, because he can not play against himself. That is, his (non-averaged) payoffs are $N \cdot \pi(s,\sigma) - \pi(s,s)$. Hence, there is a missing term $\pi(s,s)$ in his payoffs. If the strategy s performs badly against itself, this disadvantage will be reduced by the agent's perception of himself being an s-player (recall Example 4.1).

If the agent were to switch to s', because of his own strategy change, the profile in the population does not correspond to σ anymore. His (non-averaged) payoffs after the deviation are $N\cdot\pi(s',\sigma)-\pi(s',s)$. If the strategy s' is advantageous against S, the term $\pi(s',s)$ represents the reduction of this advantage due to the fact that the agent was an s-player himself.

The condition for a strategy σ to be BR-focal is a comparison of these two payoff reductions. If s' is just as good as s against σ, the loss of the possible disadvantage of s against itself ($\pi(s,s)$) must be offset by the reduction in the advantage of s' against s ($\pi(s',s)$).

EXAMPLE 4.3. Frequently absorbing does not imply ESS. Consider the symmetric game whose payoff matrices are given by

	A	B	C
A	0	2	1
B	2	0	1
C	1	1	x

This game has a symmetric mixed-strategy Nash equilibrium given by $\sigma = (\frac{1}{2}, \frac{1}{2}, 0)$, and, if $x \geq 1$ a symmetric pure-strategy Nash equilibrium. Since $\pi(C,\sigma) = 1 = \pi(\sigma,\sigma)$, it is easy to see that σ is an ESS if and only if $\pi(\sigma, C) = 1 > x = \pi(C,C)$, i.e. $x < 1$. However, it follows from the previous Theorem and mere observation of the payoff matrix that σ is frequently absorbing for all possible values of x.

This is a feature of the finite-population framework and the individual-behavior approach. The payoff x is unreachable for an agent which plans to deviate from a profile where strategy C is not present. However, it plays a role for an ESS, where deviations take the form of a small proportion of mutants (which will also play among themselves).

Notice that, in Example 4.2, the completely mixed strategy σ_3 is frequently absorbing but not an ESS.

EXAMPLE 4.4. ESS does not imply frequently absorbing. Consider the symmetric game whose payoff matrices are given by

	A	B	C
A	0	2	1
B	2	0	1
C	-1	3	x

This game has also a symmetric mixed-strategy Nash equilibrium given by $\sigma = (\frac{1}{2}, \frac{1}{2}, 0)$, and, if $x \geq 1$ a symmetric pure-strategy Nash equilibrium. Again, σ is an ESS if and only if $x < 1$. However, since $\pi(C, \sigma) = \pi(A, \sigma)$ and $\pi(C, A) < \pi(A, A)$, it follows from Theorem 4.1 that σ is never frequently absorbing. Indeed, if there are $N = 2 \cdot k$ agents in the population, an agent playing A in state (k, k) will get average payoff $\frac{2k}{2k-1}$, while by switching to C he could get an average payoff of $\frac{1}{2k-1}(-(k-1) + 3k) = \frac{2k+1}{2k-1}$.

Again, the payoff x is unreachable for an agent which plans to deviate from a profile where strategy C is not present. However, in the rationale behind an ESS, a small proportion of mutants will also play among themselves and obtain it. While in the previous example (if $x > 1$) this was to the mutants' advantage, in this case (if $x < 1$) it plays against them.

EXAMPLE 4.5. ESS does not imply frequently absorbing (2). Consider the following case of a "Rock-Scissors-Paper" game.

	R	S	P
R	1	3	0
S	0	1	3
P	3	0	1

The Nash equilibrium given by $\sigma = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is a completely mixed ESS (see [16, pp.101]). However, a mere examination of the payoff matrix reveals that it is not BR-focal (see Proposition 4.1).

5. THE GENERAL CASE: IMITATION

Consider now a general learning process as in the previous section, but endow agents with imitative rules rather than myopic best reply ones.

5.1. No absorbing states

The first negative result (for mixed strategy profiles) is that we have to give up the hope of having absorbing states.

LEMMA 5.1. *Consider any learning process based on imitation. The only absorbing states are the monomorphic states, i.e. states* ω *such that* $\omega(s) = N$ *for some* $s \in S$, $\omega(s') = 0$ *for all* $s' \neq s$. *Moreover, under independent inertia all other states are in transient classes.*

Proof. In any non-monomorphic state ω', there would be different strategies in its support. Let $s \in supp(\omega')$ such that $\Pi(s, \omega') \geq \Pi(s', \omega')\ \forall\ s' \in supp(\omega)$. Then, there is positive probability that an agent not playing s gets the opportunity to revise and imitates s, that is, $P(\omega', m(\omega', s', s) > 0$ for all $s' \in supp(\omega')$, and hence ω' is not absorbing.

Under independent inertia, there is positive probability that all agents not playing s get the opportunity to revise and imitate s. It follows that there is positive probability for a transition from any $\omega' \in \Omega$ to a monomorphic state, and hence any non-monomorphic state is in a transient class. ∎

Two observations follow from the previous lemma. First, independent inertia coupled with imitation is too quick a dynamics for a mixed profile to be sustainable in any sense. Such extreme dynamics can only give rise to monomorphic states.

Second, as we learned in the 2×2 case, for dynamics with slower adjustment this does not preclude mixed profiles from being stable in an appropriately defined sense. However, this will never mean an absorbing state, but, at best, a focal point within a (maybe narrow) non-singleton recurrent communication class.

We will try to pursue the (encouraging) intuition we developed for the 2×2 case and see whether it extends to more general games or not. That is, we try to characterize Nash equilibria which are limits of sequences of mixed strategies which, if reinterpreted as population profiles (barring integer problems), turn out to identify rest points (or, at least, narrow recurrent classes) of imitation dynamics.

5.2. Imitation-absorbing states

The first intuition concerns the existence (for each population size N)of a focal point ($\hat{n}$ in 2×2 games) such that all strategies in its support give the same payoff in the round-robin tournament, and such that it approaches a mixed-strategy Nash equilibrium as the population size grows to infinity. We use a fixed-point argument to show its existence in general games. In the next theorem, we abuse notation writing $\pi(x, y)$ to denote $x^T \cdot A \cdot y$, even if x, y are not mixed strategies but arbitrary vectors.

THEOREM 5.1. *Suppose the game has a symmetric Nash equilibrium given by the mixed strategy σ^*. For any N, there exists a mixed strategy σ_N such that $\pi(s, N \cdot \sigma_N - e_s) = \pi(s', N \cdot \sigma_N - e'_s)\ \forall\ s, s' \in supp(\sigma_N)$ and $supp(\sigma_N) \subseteq supp(\sigma)$. Moreover, if (σ^*, σ^*) is the only symmetric Nash equilibrium on the restricted game with strategy space $supp(\sigma^*)$, then $\lim_{N\to\infty} \sigma_N = \sigma^*$.*

Proof. We can assume without loss of generality that $supp(\sigma^*) = S$. If not, it suffices to define a new game by restricting the previous one to the sub-strategy space $supp(\sigma^*)$. Let Δ be the space of mixed strategies.

Given any mixed strategy σ, define

$$\Pi^R(s, \sigma) = e_s^T \cdot A(N \cdot \sigma - e_s)$$

Extend this function linearly to Δ by $\Pi^R(\sigma', \sigma) = \sum_{s\in S} \sigma'(s) \cdot \Pi^R(s, \sigma)$. Define a correspondence from Δ into itself by

$$B^R(\sigma) = \arg\max \Pi^R(\cdot, \sigma)$$

By continuity of $\Pi^R(\cdot, \sigma)$ on the (compact) set Δ, $B^R(\sigma)$ is nonempty. By linearity of $\Pi^R(\cdot, \sigma)$ on the (convex) set Δ, $B^R(\sigma)$ is convex. By continuity of Π^R, B^R has a closed graph. Kakutani's fixed point Theorem then implies that there exists $\sigma_N \in \Delta$ such that $\sigma_N \in B^R(\sigma_N)$, i.e.

$$\Pi^R(\sigma_N, \sigma_N) \geq \Pi^R(\sigma, \sigma_N)\ \forall\ \sigma \in \Delta$$

In particular, $\Pi^R(\sigma_N, \sigma_N) \geq \Pi^R(s, \sigma_N)\ \forall\ s \in supp(\sigma_N)$, which, by linearity of $\Pi^R(\cdot, \sigma_N)$, implies that $\Pi^R(s, \sigma_N) = \Pi^R(\sigma_N, \sigma_N)\ \forall\ s \in supp(\sigma_N)$

We have shown existence of the sequence $\{\sigma_N\}$. Since this is a sequence of real vectors in a compact set, it has a convergent subsequence $\{\sigma_{N_k}\}$. Let σ_1 be its limit. Then, for all k,

$$\Pi^R(\sigma_{N_k}, \sigma_{N_k}) \geq \Pi^R(\sigma, \sigma_{N_k})\ \forall\ \sigma \in \Delta$$

i.e.

$$\sum_{s\in S} \sigma_{N_k}(s) \left[e_s^T \cdot A \cdot (\sigma_{N_k} - \frac{1}{N_k} e_s)\right] \geq \sum_{s\in S} \sigma(s) \left[e_s^T \cdot A \cdot (\sigma_{N_k} - \frac{1}{N_k} e_s)\right]\ \forall\ \sigma \in \Delta$$

and taking limits when $k \to \infty$,

$$\sum_{s \in S} \sigma_1(s) \left[e_s^T \cdot A \cdot \sigma_1\right] \geq \sum_{s \in S} \sigma(s) \left[e_s^T \cdot A \cdot \sigma_1\right] \ \forall\, \sigma \in \Delta$$

which can be rewritten simply as

$$\pi(\sigma_1, \sigma_1) \geq \pi(\sigma, \sigma_1) \ \forall\, \sigma \in \Delta$$

and means that σ_1 is a Nash equilibrium.

We have proven that every convergent subsequence of $\{\sigma_N\}$ converges to a Nash equilibrium. If the game has only one Nash equilibrium, (σ^*, σ^*), then every convergent subsequence converges to σ^*, which implies that $\{\sigma_N\}$ converges to σ^*. ∎

Remark 5. 1. Actually, σ_N is a symmetric Nash equilibrium of a perturbed game defined as follows. If $A = [a_{s,s'}]$ is the payoff matrix of the original game, for each N we can define a game through the payoff matrix $A^N = [a^N_{s,s'}]$, where $a^N_{s,s'} = a_{s,s'} - \frac{1}{N} a_{s,s}$. Hence, the approximation of a Nash equilibrium (σ, σ) by a sequence (σ_N, σ_N) is actually an approximation through Nash equilibria of perturbed games as the perturbation goes to zero. It is well-known that the Nash correspondence (mapping perturbations to sets of equilibria) is not lower hemicontinuous, and hence, in principle there might be (non-generic) Nash equilibria of the original game that cannot be approximated by any such sequence.

The mixed strategies σ_N are the candidates as focal points of imitation dynamics for any fixed population size. Suppose, analogously to previous (INT) assumptions, that $N \cdot \sigma_N$ is actually a state in the dynamics with N agents. In the 2×2 case, we saw that the set $\{\hat{n} - 1, \hat{n}, \hat{n} + 1\}$ is then a recurrent communication class. The analogous requirement (admittedly quite demanding) would be that, in case an agent drifts away from his strategy under the profile σ_N, in the resulting profile the strategy that the agent has left is the one which now gives the maximal payoff (provided this strategy has not disappeared, i.e. there were more than one agent playing it), prompting a return to the previous profile. This gives rise to the following definition.

Definition 5.1. A non-monomorphic state ω is imitation-absorbing if

(a) $\Pi(s, \omega) = \Pi(s', \omega) \ \forall\, s, s' \in supp(\omega)$

(b) For all $s, s', s'' \in supp(\omega), s \neq s', s''$, $\Pi(s, \omega') > \Pi(s'', \omega')$, where $\omega' = m(\omega, s, s')$

Note that condition (b) implicitly assumes that $\omega(s) > 1 \; \forall \; s \in supp(\omega)$, that is, no strategy can disappear after a single deviation. Since σ_N approaches a mixed strategy equilibrium, and provided $N \cdot \sigma_N$ is a state, this will be true except for small population sizes.

PROPOSITION 5.1. *Consider any learning process with imitation and non-simultaneous learning. If a state ω is imitation-absorbing, then*

$$N_1(\omega) = \{m(\omega, s, s') \; / \; s, s' \in supp(\omega)\}$$

is a recurrent communication class.

Proof. Note that $P(\omega, \omega') > 0 \; \forall \; \omega' \in N_1(\omega)$ by (a) in the definition of imitation-absorbing state. Moreover, with non-simultaneous learning it follows that $P(\omega, \omega'') = 0 \; \forall \; \omega'' \notin N_1(\omega)$.

Consider now a state $\omega' = m(\omega, s, s')$. By (b) in the definition of imitation-absorbing state, it follows that the strategy giving maximum payoff in ω' is s. Suppose $P(\omega', \omega'') > 0$. By non-simultaneous learning, we have that $\omega'' = m(\omega', s'', s)$ for some $s'' \in supp(\omega)$. Note that $\omega'' = m(\omega', s'', s) = m(m(\omega, s, s'), s'', s) = m(\omega, s'', s') \in N_1(\omega)$.

We have proved that $P(\omega', \omega'') = 0 \; \forall \; \omega' \in\in N_1(\omega), \omega' \notin\in N_1(\omega)$. It remains to show that states in $\in N_1(\omega)$ communicate. Note that, for all $\omega' = m(\omega, s, s')$, $P(\omega', \omega) > 0$ and $P(\omega, \omega') > 0$. By transitivity, it follows that $N_1(\omega)$ is a recurrent communication class. ▌

LEMMA 5.2. *Consider any learning process with imitation and non-simultaneous learning. Let $\omega = (n_1, ..., n_m) \in \Omega$, such that $n_i > 1$ for all i. Then, ω is imitation-absorbing if and only if*

(a) $\Pi(s, \omega) = \Pi(s', \omega) \; \forall \; s, s' \in supp(\omega)$
(b) $\forall \; s, s', s'' \in supp(\omega), s \neq s', s''$, $\pi(s, s) - \pi(s, s') < \pi(s'', s) - \pi(s'', s')$

Proof. Fix $s, s' \in supp(\omega)$ and let $\omega' = m(\omega, s, s')$. Let A be the payoff matrix of the underlying game, and let e_s be the vector of $m = |S|$ coordinates, all of them 0 except for a 1 in the position corresponding to strategy s. The condition $\Pi(s, \omega') > \Pi(s'', \omega')$ can be written as

$$e_s^T \cdot A \cdot (\omega' - e_s) > e_{s''}^T \cdot A \cdot (\omega' - e_{s''})$$

and, since $m(\omega, s, s') = \omega - e_s + e_{s'}$, this is equivalent to

$$e_s^T \cdot A \cdot (\omega - 2 \cdot e_s + e_{s'}) > e_{s''}^T \cdot A \cdot (\omega - e_s + e_{s'} - e_{s''})$$

Since $e_s^T \cdot A \cdot (\omega - e_s) = e_{s''}^T \cdot A \cdot (\omega - e_{s''})$, this is if and only if

$$e_s^T \cdot A \cdot (-e_s + e_{s'}) > e_{s''}^T \cdot A \cdot (-e_s + e_{s'})$$

which can be written in compact form as $(e_s - e_{s''})^T \cdot A \cdot (e_s - e_{s'}) < 0$ or, in terms of the payoff function π:

$$\pi(s, s) - \pi(s, s') < \pi(s'', s) - \pi(s'', s')$$

as required. ▌

Condition (b) above is clearly an *spite* requirement. Whenever an s-player switches to s', an arbitrary agent, playing strategy s'', has a loss of $\pi(s'', s) - \pi(s'', s')$, since he faces one s-player less and one s'-player more. The condition above requires the loss experienced by the remaining s-players to be minimal.

5.3. Imitation-focal strategies

Condition (b) above is independent of population size, while condition (a) is not. This motivates the following definitions.

Definition 5.2. Let σ be a mixed strategy. An *imitation sequence approaching* σ is a pair $(\{N_k\}, \{\sigma_k\})$ such that $\{N_k\}$ is an strictly increasing sequence of population sizes and $\{\sigma_k\}$ is a sequence of mixed strategies $\{\sigma_k\}$ such that

(i) For all $s, s' \in \mathit{supp}\,\sigma$,

$$\pi(s, \sigma_k) - \frac{1}{N_k}\pi(s, s) = \pi(s', \sigma_k) - \frac{1}{N_k}\pi(s', s')$$

(ii) $\lim_{k\to\infty} \frac{\omega_k}{N_k} = \sigma$.

An imitation sequence approaching σ is *n*on-degenerate if there exists at least one σ_{N_k} such that $\mathit{supp}(\sigma_{N_k}) = \mathit{supp}(\sigma)$ and $N_k \cdot \sigma_{N_k}$ is a well-defined state of the dynamics with N_k agents.

Degeneracy may appear because of integer problems, e.g. games with some rational and some irrational payoffs.

Definition 5.3. A mixed strategy σ is *frequently imitation-absorbing*, or simply *imitation-focal* if

(i) There exists an imitation sequence approaching σ.

(ii) For all imitation sequences $(\{N_k\}, \{\sigma_k\})$ approaching σ, there exists $\hat{N}$ such that, whenever $N_k > \hat{N}$ and $N_k \cdot \sigma_k$ is a well-defined state of the dynamics with N_k agents, then it is an imitation-absorbing state.

The strategy is *n*on-degenerate if the sequence in (i) is non-degenerate.

THEOREM 5.2. *Consider any learning process with imitation and non-simultaneous learning. Let σ be a mixed-strategy.*

(a) If σ is frequently imitation-absorbing, then (σ, σ) is a Nash Equilibrium of the game obtained by restricting the strategy space to supp σ.

(b) If σ is frequently imitation-absorbing and non-degenerate, then, for any $s, s', s'' \in supp(\sigma), s \neq s', s''$,

$$\pi(s,s) - \pi(s,s') < \pi(s'',s) - \pi(s'',s') \tag{1}$$

(c) Suppose that, either (σ, σ) is the only Nash equilibrium of the game with restricted strategy space supp σ, or there exists an imitation-sequence approaching σ. If condition (1) holds, then σ is frequently imitation-absorbing.

Proof. Without loss of generality, restrict the strategy space to $supp(\sigma)$.

(a) Let σ be frequently imitation-absorbing, and let $(\{N_k\}, \{\sigma_k\})$ be an imitation sequence approaching σ. By (i),

$$\pi(s, \sigma_k) - \frac{1}{N_k}\pi(s,s) = \pi(s', \sigma_k) - \frac{1}{N_k}\pi(s', s') \; \forall \; s, s' \in supp\, \sigma$$

Taking limits when $k \to \infty$, we obtain

$$\pi(s, \sigma) = \pi(s', \sigma) \; \forall \; s, s' \in supp\, \sigma$$

which proves that (σ, σ) is a Nash equilibrium.

(b) Follows directly from Lemma 5.2.

(c) If (σ, σ) is the only Nash equilibrium in the game with strategy space $supp(\sigma)$, by Theorem 5.1, there exists an imitation sequence $(\{N\}, \{\sigma_N\})$ approaching σ. Hence, for $N > N_0$, $supp(\sigma_N) = supp\, \sigma$. Whenever $N \cdot \sigma_N$ is an state, Lemma 5.2 proves that it is almost absorbing.

If (σ, σ) is not the only Nash equilibrium of the restricted game but one appropriate sequence exists, the proofs follows analogously. ∎

EXAMPLE 5.1. Consider the symmetric game with payoff matrix

	A	B	C
A	1	1	1
B	2	0	1
C	0	2	0

This game has a symmetric Nash equilibrium given by the mixed strategy $\sigma = (\frac{1}{2}, \frac{1}{2}, 0)$. It is easy to check, through Theorem 5.2, that this strategy is imitation-focal.

We first compute the sequence σ_N approaching σ. Given a population size N, we want to find a profile $\omega_N = N \cdot \sigma_N = (k, N-k, 0)$, where maybe k is not an integer, such that, abusing notation, $\Pi(A, \omega_N) = \Pi(B, \omega_N)$. This equation amounts to $(k-1) \cdot 1 + (N-k) \cdot 1 = k \cdot 2 + (N-k-1) \cdot 0$, which implies $N = 2k+1$ or

$$\sigma_N = \left(\frac{N-1}{2}, \frac{N+1}{2}, 0 \right)$$

This sequence does indeed approach σ and, moreover, $N \cdot \sigma_N$ is a well-defined state of the dynamics with N agents whenever N is odd.

Let, thus, $N = 2k+1$ with $k > 2$ and $\omega = (k, k+1, 0)$. This state is imitation-absorbing. We check now the dynamic consequences of this fact.

In state ω, both A and B yield a payoff of $2k$. If an A-player imitates B, the process moves to $m(\omega, A, B) = (k-1, k+2, 0)$. In the new state, strategies A and B earn respectively $2k$ and $2k-2$, and hence the only possible transition will happen when a B-player imitates A and the process goes back to ω. Analogously, if a B-player imitates A, the process moves to state $m(\omega, B, A) = (k+1, k, 0)$. Strategies A and B earn then $2k$ and $2k+2$ respectively, and hence the only positive-probability transition occurs when an A-player imitates B and the process goes back to ω.

5.4. Comparison of BR-focal and Imitation-focal strategies

First, we prove the intuition we obtained in the 2×2 case and show that for this class of games, both concepts are equivalent.

Corollary 5.1. *Consider a 2×2 symmetric game with rational payoffs and a unique Nash equilibrium (σ, σ), with σ a mixed-strategy. Then, σ is frequently absorbing under best reply if and only if it is frequently imitation-absorbing.*

Proof. It suffices to compare Theorems 4.1 and 5.2.[9] Let the game be given by

	A	B
A	(a,a)	(b,c)
B	(c,b)	(d,d)

[9]It is easy to check in this case that the imitation sequence approaching the mixed equilibrium is non-degenerate if the payoffs are rational.

Condition (b) in Theorem 5.2 reduces to $a + d < b + c$ which is obviously implied by condition (b) in Theorem 4.1, i.e. $a < c$ and $d < b$.

Conversely, if $a + d < b + c$ but $a \geq c$ (analogously if $d \geq b$), then $d < b$ and it follows that strategy A weakly dominates strategy B, which contradicts the existence of a (completely) mixed symmetric Nash equilibrium. ▮

Now we show that for completely mixed strategies, Imitation-focal strategies are always BR-focal strategies, i.e. the former are a sharper refinement of Nash equilibria that collects stability properties both from imitation and best-reply dynamics.

PROPOSITION 5.2. *Let σ be a completely mixed strategy with rational coordinates. If σ is a (non-degenerate) frequently imitation-absorbing strategy, then it is frequently absorbing under best reply.*

Proof. By the previous result, it suffices to consider games with at least three pure strategies. Let σ be frequently imitation-absorbing, and suppose $supp(\sigma) = S$. Suppose further that σ is not frequently absorbing under best reply. By Theorem 4.1, there exist $s, s' \in S$ such that $\pi(s, s) \geq \pi(s', s)$.

Consider any $s'' \in S \setminus \{s, s'\}$. By Theorem 5.2 (interchanging s', s'' in condition (b)),

$$\pi(s, s) + \pi(s', s'') < \pi(s, s'') + \pi(s', s)$$

and, since $\pi(s, s) \geq \pi(s', s)$, it follows that $\pi(s, s'') > \pi(s', s'') \ \forall \ s'' \neq s, s'$.

Applying condition (b) in Theorem 5.2 to s', s'', s for any $s'' \neq s, s'$, we obtain

$$\pi(s', s') + \pi(s, s'') < \pi(s', s'') + \pi(s, s')$$

Since $\pi(s, s'') > \pi(s', s'')$, it follows that $\pi(s, s') > \pi(s', s')$.

In summary, strategy s (weakly) dominates strategy s'. This is a contradiction with the fact that both are in the support of the completely mixed strategy σ, and (σ, σ) is a Nash equilibrium. ▮

Remark 5. 2. It follows from this result and Proposition 4.1 that if σ is an Imitation-focal strategy, the game restricted to the strategy subspace *supp* σ cannot have any symmetric Nash equilibria in pure strategies.

The requirement of σ to be completely mixed is necessary. Since, under imitation dynamics, anything not in the support of the current state is irrelevant, it is easy to see that an imitation-focal strategy might easily fail to be BR-focal if the strategies outside its support are attractive enough.

Example 5.2. Imitation-focal and less than full support does not imply BR-focal. Consider the symmetric game with payoff matrix

	A	B	C
A	1	3	3
B	3	1	3
C	3	3	3

The mixed strategy $\sigma = (\frac{1}{2}, \frac{1}{2}, 0)$ is imitation-focal by Theorem 5.2. While (σ, σ) is a Nash equilibrium of the game restricted to $\{A, B\}$, it is not even a Nash equilibrium of the complete game and hence fails to be a BR-focal strategy.

Now we show that the reverse implication fails, even in the full-support case.

Example 5.3. BR-focal does not imply Imitation-focal. Consider the symmetric game with payoff matrix

	A	B	C
A	0	1	1
B	1	0	1
C	1	3	0

This game has a unique symmetric Nash equilibrium, given by the mixed strategy $\sigma = (\frac{1}{5}, \frac{1}{5}, \frac{3}{5})$.

This strategy is BR-focal by Theorem 4.1. Indeed, consider a state ω reproducing exactly the proportions in σ, i.e. $\omega = (k, k, 3k)$ with $N = 5 \cdot k$. An A-player is obtaining payoff $\Pi(A, \omega) = (k-1) \cdot 0 + k \cdot 1 + (3k) \cdot 1 = 4k$. If he switched to B, he would obtain payoff $(k-1) \cdot 1 + k \cdot 0 + (3k) \cdot 1 = 4k - 1 < 4k$. If he switched to C, he would obtain $(k-1) \cdot 1 + k \cdot 3 + (3k) \cdot 0 = 4k - 1 < 4k$. Hence, he has an incentive to keep his current strategy. Analogously, B-players and C-players will also keep their strategies under myopic best reply.

Note, however, that $\pi(A, A) + \pi(C, B) = 0 + 3 > 1 + 1 = \pi(A, B) + \pi(C, A)$, and hence, by Theorem 5.2, σ fails to be imitation-focal.

To see why, we compute the sequence σ_N. Let $\omega_N = (n_A, n_B, N - n_A - n_B)$ such that

$$\Pi(A, \omega_N) = \Pi(B, \omega_N) = \Pi(C, \omega_N)$$

From the first equation, it follows that $n_A = n_B = k$. From the second equation, $N = 5k$. It follows that $\sigma_N = (\frac{1}{5}, \frac{1}{5}, \frac{3}{5})$, i.e. in this case the sequence is constant and equal to σ.[10]

Consider, then, $\omega = (k, k, 3k)$ with $N = 5 \cdot k$ and $k > 2$. All agents are obtaining payoff $4k$, i.e. all strategies are perceived as equally worthwile. Suppose an A-player decides to imitate strategy B. The resulting state is $m(\omega, A, B) = (k-1, k+1, 3k)$, with payoffs $4k+1, 4k-1, 4k+2$ respectively for strategies A, B, and C. Hence, the maximum payoff corresponds now to strategy C and the process may drift further away from σ. Since an state ω is imitation-absorbing only if $N_1(\omega)$ is recurrent, we see that $(k, k, 3k)$ is never imitation-absorbing.

6. CONCLUSION

This paper identifies two concepts, called *BR-focal* and *Imitation-focal* strategies. Summarizing the analysis, these concepts could be roughly defined as follows.

A BR-focal strategy is a mixed strategy which corresponds to a symmetric Nash equilibrium and such that, whenever (barring integer problems) it is re-interpreted as a (finite) population profile of agents playing pure strategies, it turns out to be an absorbing state of any discrete-time, stochastic dynamics where agents play the underlying game in a round-robin tournament and update their strategies according to myopic best reply.

An imitation-focal strategy is a mixed strategy which corresponds to a symmetric Nash equilibrium (of the game where the strategy space is restricted to be its support) such that it is the limit of a sequence of strategies indexed by population size, which, if (barring integer problems) re-interpreted as population profiles, turn out to be in a narrow recurrent communication class of the finite-population, discrete-time dynamics where agents play the underlying game in a round-robin tournament, update their strategies according to imitation rules, and adjustment is non-simultaneous.

In spite of the technical complexities which make the analysis proceed in this way, both BR-focal and imitation-focal strategies can be characterized through conditions which depend only on the payoff matrix and not on any population size. It turns out that both concepts coincide in the case of 2×2 games, and that every imitation-focal, completely mixed strategy is also BR-focal, but the reverse implication fails for general games.

Technical difficulties associated both to finite population and discrete time make necessary a careful consideration of the appropriate "rest point" notion, and integer problems force to keep track of when mixed strategies

[10]The trick is of course the zero entries in the diagonal.

can be reinterpreted as population profiles. The analysis, however, shows that the above mentioned refinements of symmetric, mixed-strategy Nash equilibria can be considered rest points of finite-population, discrete-time dynamics in a meaningful, well-defined sense.

It is specially important to notice that the mixed strategy equilibria studied in this paper exhibit what we could call stability properties without resorting either to continuous-time approximations or to explicitly infinite populations, at the price of, e.g., tracing approximations and facing integer problems. In the words of John Nash [12, p.33], "The populations need not be large if the assumptions still hold. [...] Actually, of course, we can only expect some sort of approximate equilibrium, since the information, its utilization, and the stability of the average frequencies will be imperfect."

REFERENCES

1. C. Alós-Ferrer. Individual Randomness in Economic Models with a Continuum of Agents. Working Paper 98-07, Department of Economics, University of Vienna, 1998.
2. C. Alós-Ferrer. Dynamical systems with a continuum of randomly matched agents. *Journal of Economic Theory*, 86:245–267, 1999.
3. C. Alós-Ferrer. Learning, Memory, and Inertia. Working Paper 00-03, Department of Economics, University of Vienna, 2000.
4. C. Alós-Ferrer. Population Size and the Modeling of Time in Learning Processes. Mimeo, Department of Economics, University of Vienna, 2000.
5. C. Alós-Ferrer, A. Ania, and K. R. Schenk-Hoppé. An Evolutionary Model of Bertrand Oligopoly. *Games and Economic Behavior*, forthcoming, 2000.
6. M. Benaim and J. Weibull. Deterministic Approximation of Stochastic Evolution in Games. WP 534, The Research Institute of Industrial Economics, Stockholm, 2000.
7. R.T. Boylan. Laws of large numbers for dynamical systems with randomly matched individuals. *Journal of Economic Theory*, 57:473–504, 1992.
8. J. Hofbauer. Stability for the Best Response Dynamics. Mimeo, Department of Mathematics, University of Vienna, 1995.
9. M. Kandori, G. Mailath, and R. Rob. Learning, Mutation, and Long-Run Equilibria in Games. *Econometrica*, 61:29–56, 1993.
10. M. Kandori and R. Rob. Evolution of equilibria in the long run: a general theory and applications. *Journal of Economic Theory*, 65:383–414, 1995.
11. S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes, 2nd Ed.* Academic Press, San Diego, 1975.
12. J.F. Nash. *Essays on Game Theory.* Edward Elgar, Cheltenham, UK, 1996.
13. J. Oechssler. An Evolutionary Interpretation of Mixed-Strategy Equilibria. *Games and Economic Behavior*, 21:203–237, 1997.
14. A. J. Robson and F. Vega-Redondo. Efficient Equilibrium Selection in Evolutionary Games with Random Matching. *Journal of Economic Theory*, 70:65–92, 1996.
15. F. Vega-Redondo. The Evolution of Walrasian Behavior. *Econometrica*, 65:375–384, 1997.
16. J.W. Weibull. *Evolutionary Game Theory.* The MIT Press, Cambridge, Massachusetts, 1995.