WORKING PAPERS
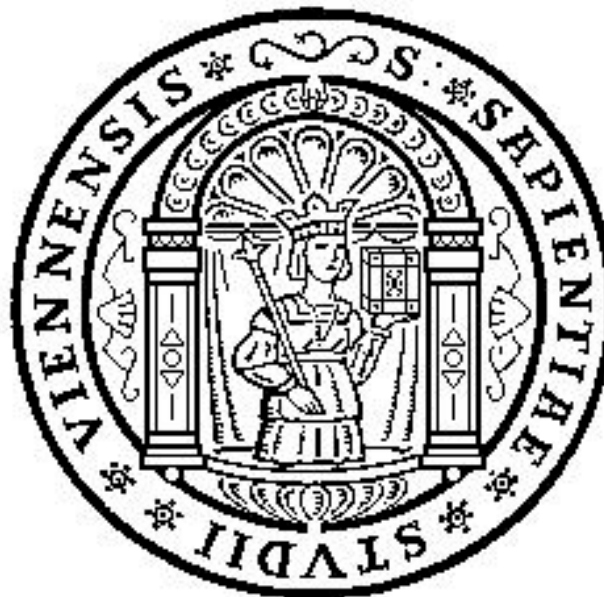
Learning, Memory, and Inertia

Carlos Alós-Ferrer

August 2000

Working Paper No: 0003

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

All our working papers are available at: http://mailbox.univie.ac.at/papers.econ

Learning, Memory, and Inertia

Carlos Alós-Ferrer[¤]
Department of Economics, University of Vienna

This version: August 2000

Abstract

This paper explores the impact of memory in standard models of learn-ing and evolution in games where a ⁻nite number of agents display imita-tion behavior, focusing on coordination games (as in Kandori et al (1993)) and N-player games where spiteful behavior allows to discard Nash equi-libria. It is shown that the way inertia is modeled in such examples actu-ally entails a strong \no-memory" assumption. Once inertia is removed (or modeled otherwise), the addition of bounded memory changes the pre-dictions dramatically. The analysis highlights the stability properties of Nash outcomes in purely evolutionary contexts with a ⁻nite population of agents.

1   Introduction

Evolutionary models rely on the idea that, in a population framework, agents obtaining higher payo®s would thrive at the expense of those who fare worse.

It is by now well understood that, in a ⁻nite population, this idea can have surprising implications. For instance, a Nash equilibrium can be quickly dis-carded in favor of other outcomes because of the e®ect of spite. An agent deviating from a Nash equilibrium will have worse payo®s than before, but it is still possible that the payo®s of other agents after this deviation are even worse, leaving the deviant with the highest payo®s.

While this approach is certainly sensible in a biological framework, where agents live one period and are replaced by their o®spring, certain doubts arise when the same arguments are applied to learning models in economics. If the agents do not die but try to learn, the e®ects of spite rely on the assumption that previous outcomes (in the above example, the payo®s of the Nash equilibrium) are immediately forgotten and only the current comparison of payo®s matter.

If players adjust their actions through a myopic best reply process, it might be reasonable to ignore the e®ects of memory, since the possibility to compute a best reply implies a degree of knowledge about the game that outweighs any information gained by a single observation. Many evolutionary models, however, are interpreted as lead by imitation and trial-and-error considerations on the side of agents who do not know the structure of the game they are playing. In such a framework, neglecting to consider the information gained at even

---

1

the most recent periods of play might impose too restrictive constraints on the considered learning processes.

This paper explores the impact of adding ﬁnite memory to what by now are standard models of learning and evolution in games. The ﬁrst result is that an apparently innocent component of such models, the presence of inertia, actually precludes memory from having any eﬀect in most situations, hinting that inertia, at least in the way it is usually modeled, carries a strong \no-memory" assumption.

The second result is that, once inertia is excluded, the addition of memory does in general change the results of the models considered, proving that the exclusion of memory is not a harmless assumption.

The third result is that, for its eﬀects to have full strength, memory must be long enough. This required length, however, is actually quite short in standard examples.

The eﬀects of memory and inertia are analyzed in three frameworks. In the ﬁrst, two boundedly rational agents play an inﬁnitely repeated game where standard models without memory predict a non-Nash, strictly dominated outcome. The addition of memory reverses this result, illustrating the evolutionary value of the Nash property. However, inertia would cancel out this eﬀect. The second framework is a generalization of the model of Kandori, Mailath, and Rob [3] for coordination games. It is shown that, if memory is long enough (and inertia is excluded), the long run predictions are linked to Pareto E±ciency, and not to Risk Dominance as in the original model without memory. The third framework is a stylized N-player game with a unique, strict Nash equilibrium and a \spiteful outcome" which would be selected by imitation if no memory is allowed. This is essentially the situation in a Cournot Oligopoly as in Vega-Redondo [7]. It is shown that, in the presence of memory, both the Nash equilibrium and the spiteful outcome are stochastically stable (i.e. long-run predictions). This illustrates the tension between \evolutionary stability" and Nash strategies in an evolutionary framework and clariﬁes their respective stability properties. This last example is developed for the case of a Cournot Oligopoly in a companion paper (see Section 7).

2 The Model(s)

Consider a ﬁnite population of N agents who repeatedly interact according to a pre-speciﬁed stage model. This stage model might be an N-player game, a round-robin tournament to play a two-person symmetric game, or the agents might be randomly matched to play such a bilateral game (formally, all these settings could be formalized as games, but the expression \stage model" is preferred for clarity).

Agents face a symmetric situation. Each of them has a ﬁnite strategy set S (e.g. the strategies of the games mentioned above). At the beginning of each period each agent chooses a pure strategy. The strategies chosen and the stage model will determine the payoﬀs agents receive at the end of the period.

Given N and S, a stage model can be summarized by a mapping

$$\pi : S^N \longrightarrow R^N$$

such that $\pi_i(\sigma(s)) = \pi_{\sigma(i)}(s) \ \forall s \in S^N, i = 1, \dots, N, \sigma \in \Sigma_N$, i.e. the situation

2

is symmetric. S_N is the symmetric group of permutations of N elements, and $\pi(s)$ denotes the vector $(s_{\pi(1)}, ..., s_{\pi(N)})$.

For example, the stage model can simply be an N-player game (playing the field situation, as in [7]), or it might specify that agents play a bilateral finite game sequentially against each other agent in the population (round robin tournament, as in [3]). A stage model could have agents randomly matched in pairs to play such a game, as in [6], although this introduces an additional modeling complication that we will discuss below.

Agents have finite memory, recalling exactly $K \geq 0$ periods in addition to the current one. The particular case $K = 0$ represents the \no-memory" case.

A vector $p = (s_1, ..., s_N) \in S^N$ is called a population profile.

The profile of strategy choices in the population in the last $K + 1$ periods is represented by a state $\omega \in \Omega = S^{N(K+1)}$, whose interpretation is as follows:

$$\omega = (s_1^{-K}, ..., s_N^{-K}, ..., s_1^0, ..., s_N^0)$$

where s_i^{-k} is the strategy that agent i chose k periods ago.

The specification above does not include the payoffs realized in each state. If the stage model ι is deterministic (e.g. a round-robin tournament or an N-player game), this is justified, since the payoffs are uniquely determined by the strategies. In other cases, e.g. random matching, the payoffs might depend on some stochastic element and hence there will be a need to specify them. Without memory, as e.g. in [6], this difficulty can be bypassed since only one realization of the payoffs matter, hence it can be included in the dynamical model afterwards instead of in the stage model itself. With memory, $K \geq 1$, the payoffs actually realized must be recorded.

We consider deterministic models where $\iota(s)$ is a fixed vector of payoffs. Stage models including random matching or games against nature could be included in the formalization making $\iota(s)$ a distribution, at the expense of formal complexity. This task is left for future research.

With this convention, given a state ω, where s_i^{-k} is the strategy that agent i chose k periods ago, we denote by μ_i^{-k} the payoff that he obtained.

After payoffs are realized, agents update their strategies according to behavioral rules which prescribe an action based on the information available in the current state and some stochastic element. Formally, a behavioral rule for agent i is a mapping $B_i : \Omega \to \Phi(S)$, where $\Phi(S)$ represents the probability distributions on S. For example, a behavioral rule may prescribe a best reply to the current profile, or to an average of the last periods' play, randomly selecting a choice in case of ties. It may also prescribe to imitate the strategy which yielded the highest payoffs, again randomly breaking ties.

In the sequel, we will consider behavioral rules such that

$$B_i((s^{-K}, ..., s^0), (\mu^{-K}, ..., \mu^0))(s_j^{-k}) > 0 \ (\Rightarrow) \ \mu_j^{-k} \geq \mu_{j'}^{-k'} \forall k', j'.$$

i.e. rules which prescribe to mimic strategies that have yielded the highest remembered payoffs. This rules particularize to the ones used in e.g. [3], [6], and [7], when $K = 0$. One interpretation is that they capture (in an extreme way) the standard evolutionary argument that strategies which fare better should thrive at the expense of other strategies. Another, more learning-oriented interpretation is that they are imitation rules, used by boundedly rational agents to try and single out the best strategies.

Additionally, the process might exhibit inertia. With a fixed probability $0 \le \gamma < 1$, independent across time and across agents, each agent might "not get the learning draw," i.e. he might be forced to stay with his current strategy. If $\gamma = 0$, we have the no-inertia case.

The system described is characterized by $N; S; \phi; \{B_i\}_{i=1}^N; K$, and γ. Abusing notation, and since we will assume imitating behavioral rules as explained above, we simply call such a system a $(\phi; \gamma; K)$-dynamics.

Any such dynamics gives rise to a finite Markov chain, to which standard treatment applies. Given two states $\omega; \omega'$ it is possible to compute the probability of transition from ω to ω' in one period. Call it $P(\omega; \omega')$. The transition matrix of the process is given by $P = [P(\omega; \omega')]_{\omega; \omega' \in \Omega}$.

The techniques used to analyze these processes are by now standard. A summary is presented in Appendix A.

The "solution concept" used is an invariant distribution, i.e. a distribution of strategies which, if adopted by the population, would be reproduced after updating (more precisely, a vector of probabilities π such that $\pi \cdot P = \pi$). If such an invariant distribution is unique { we call then the process ergodic {, and under standard technical conditions, this distribution summarizes the limit time averages with which each strategy is played along any sample path, independently of initial conditions.

In general, however, the learning processes postulated above give rise to non-ergodic Markov chains, with several (sometimes many) invariant distributions (whose support corresponds to sets of states called absorbing sets, or, more properly, "recurrent communication classes" - see Appendix A). The situation is analogous to one of multiplicity of equilibria.

3   The Equilibrium Concept

In order to solve the multiplicity problem, and following the literature (see e.g. [3], [8], or [2]), we perform a stability analysis. The basic process P is perturbed by a mutation process in the following way. With fixed probability $\varepsilon > 0$, independent across agents and across time, each agent "mutates" and chooses a strategy by a different process than that prescribed by his behavioral rule. For instance, it might simply "tremble" and pick up any strategy at random, which would make the perturbed process irreducible.

The mutation process carries enough randomness to make the perturbed Markov chain, denoted $P(\varepsilon)$, ergodic (see Appendix A), although not necessarily irreducible. The corresponding (unique) invariant distribution is denoted $\pi(\varepsilon)$.

We want to consider small perturbations. It is by now a well-established result that $\pi^* = \lim_{\varepsilon \to 0} \pi(\varepsilon)$ exists and is an invariant distribution of the unperturbed process P. It singles out an stable prediction of this original process, in the sense that, whenever an arbitrarily small perturbation is applied to it, the play approximates that described by π^*.

We call π^* limit invariant distribution. The states in its support, $\{\omega = \pi^*(\omega) > 0\}$ are called stochastically stable states (sometimes also long-run equilibria). It is easy to see that the set of stochastically stable states is a union of recurrent communication classes ("absorbing states") of the unperturbed process P.

There exists a quite convenient characterization of the stochastically stable states which we summarize and particularize to our setting in Appendix B.

4 Unperturbed Dynamics

The first step of the analysis of a $(\iota; \frac{1}{2}; K)$-dynamics is to determine the recurrent communication classes of the unperturbed process.

Definition 4.1. Let $s \in S$. We denote by $\text{mon}(s; K) \in S^{N(K+1)}$ the state $[(s; ....; s); ....; (s; ....; s)]$, where all players choose the strategy s for $K + 1$ periods. We call these states K-monomorphic or, when there is no possible confusion, simply monomorphic

Lemma 4.2. Consider any $(\iota; \frac{1}{2}; K)$-dynamics. The recurrent communication classes of the unperturbed process are the singletons $\{\text{mon}(s; K)\}$.

Proof. These singletons are recurrent aperiodic communication classes, since there is probability one of remaining in them once they are reached (the only remembered strategy is s, so that is the only possibility to imitate). In any other state, consider one strategy that has yielded the maximum payoffs in the last $K + 1$ periods. There is positive probability that all agents revise their strategies and choose exactly that strategy.

If $\frac{1}{2} > 0$, there is positive probability that they will not revise for the next K periods. This proves that all other states are in transient classes.

If $\frac{1}{2} = 0$, then all agents will get the opportunity to revise during $K + 1$ periods (there is no inertia), and there is positive probability that, each period, all of them imitate the same action. Hence, at the end of those periods, the state is one where all agents have been doing the same that other agents for $K + 1$ periods (although not necessarily the same in different periods). For the next $K + 1$ all of them will imitate the (remembered) strategy that gives larger payoffs when all agents choose it, which completes the proof. ¥

5 Non-Nash outcomes and memory

Let us consider a first, simplified example. Fix $N = 2$ and consider the stage game given by:

	E	D
E	(4,4)	(1,2)
D	(2,1)	(0,0)

With only two agents, many dynamic models of evolutionary type, including those in [3] and [6], reduce to an infinitely repeated game played by boundedly rational players. This is precisely the situation analyzed here. Hence, this example can be considered a particular case of those models.

The example is designed to highlight the differences introduced by memory in as sharp and simple a manner as possible. The game above has a dominant strategy, E, which leads to a Nash equilibrium which presents the larger payoffs in the table. In contrast, the other symmetric outcome, (D,D), is dominated and leads to the worst payoffs.

Lemma 4.2 implies that, for any dynamics with memory $K \geq 0$ and inertia $\frac{1}{2} \geq 0$, the recurrent communication classes of the unperturbed dynamics are the monomorphic states corresponding to the Nash equilibrium (E,E) and to the dominated outcome (D,D). We would expect then any reasonable dynamics to select the Nash outcome.

In any $(\gamma; \lambda; K)$-dynamics for the above game, it is easy to see (see Appendix B) that the only relevant quantities for the analysis are the number of mutations required for the transition from mon(D; K) to mon(E; K) to occur. The state requiring more mutations to destabilize in favor of the other will be the stochastically stable one.

First consider the case without memory, $K = 0$. If the players are playing (E; E) and one mutation occurs, the new state is (E; D) (or (D; E)). In this state, the mutant earns payoff 2, while the non-mutant earns only 1. Hence, the mutant will be imitated and we will reach state (D; D) with only one mutation.

It is true that the player deviating from (E; E) (which is a Nash equilibrium) has his payoff reduced, but the other player has his payoff reduced even more. In relative terms, the mutant is better off. In absence of memory, the previous payoffs are irrelevant.

How many mutations does it take to destabilize (D; D)? Two mutations obviously suffice. If only one mutation occurs, again the new state is (E; D), and the mutant (playing E) finds himself with worse payoffs than the non-mutant. So state (D; D) can not be left with only one mutation.

It is true that the player deviating from (D; D) plays a best response and has his payoff raised, but, in doing so, he manages to increase the payoff of the non-mutant even more. Notice that both payoffs (for the mutant and the non-mutant) are higher than the original ones (at (D; D)), so, in this case, whether those payoffs are remembered or not is irrelevant.

The last effect, which gives stability to non-Nash outcomes, is usually called spite. To differentiate it from the first effect, i.e. the one making unstable a Nash equilibrium, we will tentatively label the former negative spite.

The occurrence of both effects in this game lead to the result that the transition from (E; E) requires only one mutation, whereas the reverse one needs two. Hence, the only stochastically stable state for the $(\gamma; \lambda; 0)$-dynamics is the non-Nash, strictly dominated, Pareto-inefficient outcome (N; N).[1]

Now we introduce memory. Specifically, consider any $(\gamma; \lambda; K)$-dynamics for the game above, with $K > 0$. Furthermore, to see the interplay between memory and inertia, consider $\lambda > 0$.

The transition from mon(D; K) to mon(E; K) can still be done in two mutations. If both players mutate to E, they both will get payoff 4, higher than the previous one of 0, and they will stick to E. Because of spite, one mutation is not enough.

At first glance, it would seem that now the reverse transition can not be made with only one mutation. If players are playing (E; E) and one mutates to D, they both will remember that the previous period they obtained a payoff of 4 with strategy E, larger than both after-mutation payoffs, and hence they should revert to it.

This is not the case, because inertia is strictly positive. The following transition path requires only one mutation. First, one player mutates from E to D, and the profile (E; D) is reached. For the next K periods, this player acts by inertia, i.e. does not get any revision opportunity. The state then is [(E; D); (E; D); ....; (E; D)], i.e. the profile (E; E) and its associated payoffs have been forgotten. If the second player gets his revision opportunity for the next $K + 1$ periods, he will copy strategy D (which brings a higher payoff in

[1] Rhode and Stegeman [5] present a similar example for the N-agent case analyzed in [3]

6

$(E; D))$ and the state $mon(D; K)$ will be reached.

Hence, even after introducing memory, as long as inertia is positive, the selection result does not change. The trick is transparent. Inertia allows us to \freeze" agents until the relevant information has been forgotten. Rather than seeing any literal signi¯cance in this result, we take it as the ¯rst hint that there is a certain degree of contradiction between inertia and memory. The standard way to introduce inertia in a model implicitly includes a \no-memory" assumption.

With this interpretation, we proceed to study the prediction of dynamics with memory but without inertia for the game above, i.e. the $(\frac{1}{2}; 0; K)$-dynamics, with $K \geq 1$.

Reasoning exactly as above, the transition from $mon(D; K)$ to $mon(E; K)$ requires two mutations, independently of K.

The reverse transition can not occur with only one mutation. If the players are playing $(E; E)$ and one of them deviates to N, then next period the higher payo® remembered will be 4, associated to E in the pro¯le $(E; E)$. Hence, players will revert to E. Even two simultaneous mutations would not su±ce, by the same argument. In the absence of further mutation, players will just copy E. The only way around this is a chain of mutations, at least one per period, until the pro¯le $(E; E)$ is forgotten. If for $K + 1$ periods an agent mutates to D, ¯nally the state will be $[(E; D); (E; D); :::; (E; D)]$ and then D will be imitated. This requires $K + 1$ mutations.

Suppose $K = 1$, i.e. players have a \short" memory of one period. Then the transition from $mon(E; K)$ to $mon(D; K)$ requires two mutations, exactly as the reverse one, implying that both states are stochastically stable. If $K > 1$, then the ¯rst transition requires strictly more than two mutations, which means that $mon(E; K)$ is the only stochastically stable state.

This simple example gives us the following intuitions. First, the introduction of memory in a model is of little relevance as long as inertia is maintained. If the e®ect of memory is to be appreciated, then inertia should be removed from the modeling, or modeled otherwise (see Section 8). Second, once memory is introduced in this way, the predictions of the model can be dramatically altered, clarifying for instance the evolutionary value of a Nash equilibrium as above. Third, for its full e®ects to be observed, memory has to be long enough (in the above example meaning at least two periods).

6 Coordination Games and the KMR model

We turn to the analysis of one simple and well-known model in order to try and pin down the e®ects of the introduction of memory, pursuing the intuitions acquired in the previous example in a more general framework.

Kandori et al [3] pose a model without memory, where agents in a ¯nite population play a coordination game. The most interesting case is when this game has two pure-strategy, strict Nash equilibria, such that one is Risk Dominant and the other is Pareto E±cient. They predict that the limit invariant distribution will be concentrated at the state where all agents play the Risk Dominant equilibrium (and not the Pareto E±cient one).

Let the stage game G have payo® matrices given by

such that:

1. $a > c$ and $d > b$. Hence, (P,P) and (R,R) are strict Nash equilibria, and we have a problem of equilibrium selection.

2. $a > d$ and $a + b < c + d$. Hence, (P,P) is Pareto Efficient and (R,R) is risk dominant, which is the most interesting case.

The KMR model is best interpreted as one where agents play with each other, each period, in a round-robin tournament. The unperturbed dynamic can be understood as follows. Each period, each agent independently gets the opportunity to revise his strategy with probability $1 - \lambda$. Then he mimics the strategy that has yielded larger payoffs[2] (randomly choosing one in case of ties, with positive probability for all of them).[3] In their original model, inertia is strictly positive, $\lambda > 0$, and there is no memory, $K = 0$.

Let $\pi^{RR}(G)$ denote the stage model where agents play the game G sequentially against each other. The model described is then a $(\pi^{RR}(G); \lambda; 0)$-dynamics with $\lambda > 0$.

Since there are only two strategies in the underlying stage game, the state space can be represented in a reduced form as $\Theta = \{0, \ldots, N\}$, where a typical state n is identified with the number of agents playing strategy P. Kandori et al [3] assume N even for simplicity.[4]

Let $\pi(P; n)$ and $\pi(R; n)$ be the payoffs of an agent playing P or R when exactly n agents are playing P ($\pi(P; 0)$ and $\pi(R; N)$ are not defined).

$$\pi(P; n) = (n - 1)a + (N - n)b$$
$$\pi(R; n) = n \cdot c + (N - n - 1)d$$

The next lemma analyzes the difference of payoffs when both strategies are present, which will turn out to be the key quantity for the analysis.

Lemma 6.1. Consider the game G. For $1 \le n \le N - 1$, $\pi(P; n) - \pi(R; n)$ is increasing in n. Moreover, $\pi(P; n) = \pi(R; n)$ if and only if $n = n^\#$, where

$$n^\# = \frac{N(d - b) + (a - d)}{(a - c) + (d - b)} > \frac{N}{2}$$

Proof.

$$\pi(P; n) - \pi(R; n) = (n - 1)a + (N - n)b - n \cdot c - (N - n - 1)d =$$
$$= n \cdot (a - c) + (N - n)(b - d) + (d - a)$$

---

[2] This approximates a best reply dynamics if th population size is very large, but is qualitatively different for small N. The conditions on the dynamics given in [3] can be readily interpreted as coming from imitative behavior, but, strictly speaking, not as coming from best-reply behavior

[3] In their original paper, Kandori et al [3] did not rule out the possibility that even if every agent was playing the same strategy, some agent would switch to the other, now unobserved strategy, if this happened to yield larger payoffs when played by a nonexistent, \ghost mutant." However, Rhode and Stegeman [5] observed that this possibility introduces serious anomalies.

[4] Rhode and Stegemann [5] drop this requirement, showing that in some extreme cases, with N odd, the predictions of Kandori et al [3] may change. We keep N even for simplicity.

which is increasing in n since $(a - c) > 0$ and $(b - d) < 0$. Moreover,

$$\Pi(P;n) = \Pi(R;n) \iff n \cdot [(a - c) + (d - b)] = N(d - b) + (a - d)$$
$$\iff n = \frac{N(d - b) + (a - d)}{(a - c) + (d - b)}$$

All that remains to show is that $n^\ast > \frac{N}{2}$. This follows from the Risk Dominance of R. Since $(a - c) < (d - b)$,

$$n^\ast = \frac{N(d - b) + (a - d)}{(a - c) + (d - b)} > \frac{N(d - b) + (a - d)}{2(d - b)} > \frac{N}{2} + \frac{a - d}{2(d - b)} > \frac{N}{2}$$

$$\yen$$

6.1 Dynamic analysis

Lemma 4.2 implies that, for all $0 \le \frac{1}{2} < 1$ and all $K \ge 0$, the only recurrent communication classes of the unperturbed $(\Pi^{RR}(G); \frac{1}{2}; K)$-dynamics are the singletons $\{(0; K+1; 0)\}$ and $\{(N; K+1; N)\}$. Moreover, these classes are aperiodic. Hence, every $(\Pi^{RR}(G); \frac{1}{2}; K)$-dynamics will eventually settle in one of the two Nash equilibria (see Appendix A).

Kandori et al [3] show that their dynamics (strictly positive inertia and no memory) selects the Risk Dominant one. Their result translates to our framework as follows. For any $\frac{1}{2} > 0$, the only stochastically stable state of the $(\Pi^{RR}(G); \frac{1}{2}; 0)$-dynamics is the singleton $\{0\}$, i.e. the Risk Dominant equilibrium.

The question arises immediately. How can the Pareto E±cient outcome, which has strictly larger payo®s, be more unstable than the other equilibrium? Without memory, the answer is clear. Once a mutation occurs, the payo®s of the Pareto E±cient equilibrium are no longer observed. Neither are they remembered. Hence, their magnitude is irrelevant. Even though after a single mutation both the mutant and the non-mutants earn lower payo®s than before, all that matters is the e®ect of spite. When a mutant meets a non-mutant, the mutant earns higher payo®s than the other agent. And still it does not seem plausible that agents forget strictly superior payo®s that easily.

Hence, it is reasonable to allow them to remember previous payo®s, i.e., allow then to have (bounded) memory. However, the next Proposition shows that the results will not change as long as inertia is present in the model.

Proposition 6.2. Let G be as above and N even. For any $\frac{1}{2} > 0$ and $K \ge 0$, the only stochastically stable state of the $(\Pi^{RR}(G); \frac{1}{2}; K)$-dynamics is $(0; K+1; 0)$, i.e. the Risk Dominant equilibrium.

Proof. By Lemma 4.2, states $(0; ....; 0)$ and $(N; ....; N)$ are the only ones which the process will not leave without a mutation. Thus, it is clear that the minimal-cost tree is one of the following two. In the first, state $(0; ....; 0)$ is connected to state $(N; ....; N)$ with the appropriate chain of mutations, and all other (transient) states are connected at zero cost to the tree. In the second, state $(N; ....; N)$ is connected to state $(0; ....; 0)$, other states connected at zero cost. The cost of the first tree is the number of mutations required for a transition from $(0; ....; 0)$ to $(N; ....; N)$, and vice versa. If the first is higher, then $(0; ....; 0)$ is the stochastically stable state, and reciprocally. Let us compute those costs.

9

How many mutations are required for a transition from $(0, \ldots, 0)$ to $(N, \ldots, N)$ to occur? Imagine n mutations happen. The difference in payoffs between mutants and non-mutants is $\pi(P, n) - \pi(R, n)$.

By Lemma 6.1, the mutant strategy P will yield lower payoffs than the non-mutants (playing R) if and only if $n < n^*$. Since $n^* > \frac{N}{2}$, $\pi(P, n) - \pi(R, n) < 0 \; 8 \; n < \frac{N}{2}$.

In terms of the dynamics, this means that in any state where in the last $K + 1$ periods no more than $\frac{N}{2}$ agents have simultaneously played strategy P, the other agents will never imitate them. Quite on the contrary, when given opportunity, the P-players will switch to R.

An immediate implication is that the cost of the transition from $(0, \ldots, 0)$ to $(N, \ldots, N)$ is strictly greater than $\frac{N}{2}$.

Another implication is that the cost of the reverse transition is smaller than or equal to $\frac{N}{2}$. Consider the state $(N, \ldots, N)$. If exactly $\frac{N}{2}$ to R occur, then there is positive probability that, for K consecutive periods, no agent receives the opportunity to revise. After those K periods, we will be in the state $(\frac{N}{2}, \ldots, \frac{N}{2})$, and there will be only two payoffs in the agents' memories. Since, by Lemma 6.1, $\pi(P, \frac{N}{2}) < \pi(R, \frac{N}{2})$, if now all the agents receive revision opportunities for $K + 1$ periods, they will imitate R (the only other payoff appearing in intermediate states from this point on, $\pi(R, 0)$, would also lead to imitation of R if it were higher), completing the transition to $(0, \ldots, 0)$.

Hence, the only stochastically stable state is $(0, \ldots, 0)$, i.e. the Risk Dominant equilibrium is selected. ¥

In the previous proof, transitions are constructed bypassing memory. Agents are just "frozen" through inertia until previous knowledge disappears.

There are two ways to interpret the previous result. The first, naive one, would be to say that the result of Kandori et al [3] is generalized to models with finite memory. The second is to take it as the hint that there is a fundamental "no-memory" assumption built into the inertia assumption. We want to argue that the only way to study the effect of memory is to either suppress inertia from the model or to change the way it is modeled. We proceed now to illustrate that, without inertia, the introduction of memory can actually reverse the results in a model, showing that memory has indeed "a bite."

Theorem 6.3. For K large enough, the only stochastically stable state of the $(\pi^{RR}(G), 0, K)$-dynamics is $(N, \ldots, N)$, i.e. the Pareto Efficient equilibrium.

Proof. Analogously to the previous Proposition, the only relevant figures are the costs of the transition between the two states, $(N, \ldots, N)$ and $(0, \ldots, 0)$.

Consider the state $(N, \ldots, N)$. If $N - n$ mutations (to R) occur, three payoffs will be observed and remembered: $\pi(R, n)$, $\pi(P, n)$, and $\pi(P, N)$. Unless $\pi(R, n)$ is the highest of all these three payoffs, P will be imitated. The question is, hence, how many mutations are required for $\pi(R, n)$ to be the highest payoff.

In particular, $\pi(P, N) < \pi(R, n) \iff (N - 1)a < n \mathcal{C} c + (N - n - 1)d$, which is impossible since $a > c$ and $a > d$. Hence, after any number of mutations, all agents will return to strategy P. The transition from $(N, \ldots, N)$ to $(0, \ldots, 0)$ cannot occur as long as the profile where everybody was playing P is remembered. Moreover, agents will immediately revert to this profile in absence of further mutation. Hence, for this profile to be "forgotten," mutations

must occur during $K + 1$ consecutive periods. This means that the cost of the transition cannot be lower than $K + 1$. (In fact, by Lemma 6.1 at least $\lceil N - n^{\pi} \rceil$ mutations must occur in at least one period).

Consider now the state $(0; ....; 0)$. If n mutations (to P) occur, three payoffs will be observed and remembered: $\pi(R; n)$, $\pi(P; n)$, and $\pi(R; 0)$.

By Lemma 6.1, $\pi(P; n) - \pi(R; n) \geq 0 \ (\Leftrightarrow) \ n \geq n^{\pi} = \frac{N \cdot (d - b) + (a - d)}{(a - c) + (d - b)}$

Analogously,

$$\pi(P; n) - \pi(R; 0) = (n - 1)a + (N - n)b - (N - 1)d \geq 0 \ (\Leftrightarrow)$$

$$n \cdot (a - b) - N \cdot (d - b) + (d - a) \geq 0 \ (\Leftrightarrow) \ n \geq \frac{N \cdot (d - b) + (a - d)}{(a - b)} =: \dot{n}$$

Hence, the transition from $(0; ....; 0)$ to $(N; ....; N)$ will occur if $\max(n^{\pi}; \dot{n})$ simultaneous mutations occur,[5] meaning that the cost of the transition is lower than or equal to $\lceil \max(n^{\pi}; \dot{n}) \rceil$. The maximum is equal to n^{π} if $c > d$ and to $\dot{n}$ if $c < d$.

In summary, the cost of the transition from $(N; ....; N)$ to $(0; ....; 0)$ is greater than $K + 1$, while the cost of the reverse one is lower than or equal to $\lceil \max(n^{\pi}; \dot{n}) \rceil$. Obviously, for K long enough, the first is larger, hence $(N; ....; N)$ becomes the only stochastically stable state, i.e. the Pareto Efficient equilibrium is selected. ¥

6.2 The length of memory

How long must memory be? A justified concern would be the possibility that, as N grows, so does the required memory. This is not the case.

In the previous proof, the condition $K + 1 > \max(n^{\pi}; \dot{n})$ is sufficient, but not necessary. To refine this bound, we can observe that the transition from $(N; ....; N)$ to $(0; ....; 0)$ can only occur if mutations occur during $K + 1$ consecutive periods, but one mutation per period will in general not suffice. Call n_i the number of mutations occurred in period $t - i$, being t the last period, for $i = 0; ....; K$.

Hence, the state prior to successful imitation of the mutant strategy R is $(N - n_K; ....; N - n_1; N - n_0)$: At least in one of the periods, the payoffs of the mutants (playing R) must be larger than the payoffs of the non-mutants (playing P), which, by Lemma 6.1, implies that at least $\lceil N - n^{\pi} \rceil$ mutations must occur in at least one period). Without loss of generality, suppose this event happened last period, i.e. $n_0 \geq \lceil N - n^{\pi} \rceil$.

Moreover, for the mutation to be ultimately successful, the payoff attained by these n_0 mutants, $\pi(R; N - n_0)$, must be the largest remembered payoff.

Suppose, for the sake of simplicity, that n^{π} is an integer number and $n_0 = N - n^{\pi}$. Then, if $n_i < n_0$, and since $\pi(P; n)$ is increasing in n,

$$\pi(P; N - n_i) > \pi(P; N - n_0) = \pi(P; n^{\pi}) = \pi(R; n^{\pi}) = \pi(R; N - n_0)$$

a contradiction. Hence, in this case at least $N - n^{\pi}$ mutations must occur every period. Moreover, if these mutations do occur, the transition will happen with positive probability, and hence we have identified the cost of the transition.

---

[5] In case this quantity is larger than N, N mutations would suffice, yielding the same result

11

On the other hand, the cost of the reverse transition is $\max(n^*; \hat{n})$, which means that a (sharper) sufficient condition for the result in the previous theorem is $(K + 1) \cdot (N - n^*) > \max(n^*; \hat{n})$.

Consider the case $c > d$ (which corresponds to games of the "Stag Hunt" type). This condition reduces to $(K + 1) > \frac{n^*}{(N - n^*)}$, which is actually necessary and sufficient (because at least $\lceil n^* \rceil$ mutations must occur in a period for the payoff of the P-players to exceed that of the R-players). After some algebra, this condition amounts to

$$(K + 1) > \frac{N(d - b) + (a - d)}{N(a - c) - (a - d)}$$

This is a much sharper bound. When N grows, it is easy to see that the lower bound for K actually decreases, approaching the constant $\frac{(d-b)}{(a-c)} - 1$, which, interestingly enough, is derived exclusively from the relative "risk jump" between both equilibria (the Risk Dominance criterion amounts to $a - c < d - b$). If the Pareto Efficient equilibrium is "much riskier" than the Risk Dominant one ($d - b$ much larger than $a - c$), then more memory is needed to reverse the result. The result of Kandori et al [3] fits then nicely. "No memory" will never suffice to reverse the result.

This analysis shows that the lower bound for the memory decreases when the population size increases along a sequence of values such that n^* is an integer. It can be shown that, if N grows to infinity without this provision, we obtain the same qualitative behavior of the bound (with mere "jumps" due to integer problems).

If $c < d$, the condition is $(K + 1) > \frac{\hat{n}}{(N - n^*)}$, which reduces to

$$(K + 1) > \frac{N(d - b) + (a - d)}{N(a - c) - (a - d)} \cdot \frac{(a - b) + (d - c)}{a - b}$$

When N grows, the lower bound for K is also decreasing, and approaches the constant $\frac{(d-b)}{(a-c)} \cdot [1 + \frac{d-c}{a-b}] - 1$, which is a larger bound than in the other case but still is related to the "risk jump," amplified by a factor which comes from the fact that, in this case, the number of mutations required for the payoffs of P-players to overcome those of R-players is still not enough to overcome the (remembered) payoffs of the profile where everybody was coordinating on R. In a sense, the Risk Dominant equilibrium is "stickier" in this kind of games. The factor $[1 + \frac{d-c}{a-b}]$ could be tentatively interpreted as a positive externality: the gain of staying with R when other players switch from P to R, relative to the analogous quantity for P. In the previous case, this was not a gain but a loss.

Example 6.4. Take $a = 8; b = 0; c = 6; d = 3$. Notice that $a - c = 2$, $d - b = 3$, i.e. the Risk Dominance is not strong. The Pareto Efficient equilibrium is a 50% riskier than the other one ($\frac{3}{2} - 1 = 0.5$). In this case, $n^* = \frac{3N+5}{5}$, and both the requirement that N is even and that n^* is integer are fulfilled for $N = 10; 20; \ldots$. If n^* mutations occur, then, by Lemma 6.1, the mutants obtain larger payoffs than the non-mutants, and, since $c > d$, the mutation is successful. The opposite transition requires at least $N - n^* = \frac{2N-5}{5}$ mutations per period. The conclusion of the theorem holds if $\frac{2N-5}{5}(K + 1) > \frac{3N+5}{5}$, i.e. if $K > \frac{N+10}{2N-5}$. If $N = 10$, this amounts to $K > 1$. For $N = 20$ or larger, the condition is merely $K > 0$, i.e. any memory.

Example 6.5. Take $a = 5; b = 0; c = 4; d = 2$. Notice that $a − c = 1$, $d − b = 2$, i.e. the Risk Dominance is stronger in relative terms. The Pareto Efficient equilibrium is a 100% riskier than the other one ($\frac{2}{1} − 1 = 1$). In this case, $n^\# = \frac{2N+3}{3}$. If $n^\#$ mutations occur, then, by Lemma 6.1, the mutants obtain larger payoffs than the non-mutants, and, since $c > d$, the mutation is successful. The opposite transition requires at least $N − n^\# = \frac{N−3}{3}$ mutations per period. The conclusion of the theorem holds if $\frac{N−3}{3}(K + 1) > \frac{2N+3}{3}$, i.e. if $K > \frac{N+6}{N−3}$. If $N = 6$, this amounts to $K > 4$. If $N = 12$, the condition is $K > 2$. For $N = 18$ or larger, the condition is merely $K > 1$.

7 Memory and Spite

Consider any symmetric N-player game with finite strategy space S and payoff functions π_i. We adopt the standard notation $\pi_i(s_i; s_{−i}) = \pi_i(s_1; ....; s_n)$ for convenience.

The game itself can be considered a stage model (playing the field). Suppose this game has a unique symmetric Nash equilibrium which is strict, i.e. there exists s^N such that

$$\pi_i(s^N; (s^N; ....; s^N)) > \pi_i(s; (s^N; ....; s^N)) \forall s \in S; s \neq s^N; \forall i = 1; ....; N$$

This Nash equilibrium describes a (monomorphic) situation such that, if a player deviates, the payoffs he gets after the deviation are lower than his own payoffs before the deviation. That is, a mutant will always lose in absolute terms.

Suppose there exists another strategy s^P with the following properties:

1. Spite: $\pi_i(s^P; (s; s^P ....; s^P)) > \pi_i(s; (s^P; s^P ....; s^P)) \forall s \in S; s \neq s^P; \forall i$

2. Negative Spite: $\pi_i(s^P; (s; s; ....; s)) > \pi_i(s; (s^P; s....; s)) \forall s \in S; s \neq s^P; \forall i$

We call such a strategy spiteful. The interpretation of Spite is the following: whenever a player deviates from the symmetric (monomorphic) profile where everybody plays s^P, his payoffs after deviation are lower than the payoffs of the other players (also after his deviation). That is, a mutant will always lose in relative terms.

The interpretation of Negative Spite is related to the opposite idea. Given any monomorphic situation, if a single player deviates to s^P, then after the deviation he earns more payoffs than the other players (also after the deviation).

Under mild technical assumptions, Cournot Oligopolies feature both a unique Cournot-Nash equilibrium and an spiteful outcome. Vega-Redondo [7] shows that the spiteful outcome corresponds to the Walrasian equilibrium, and proves that, in a model with imitation and no memory, this is precisely the unique long-run prediction.

This can be easily generalized to any game with a unique spiteful outcome. Denote by $(\pi(G); \frac{1}{2}; 0)$-dynamics the dynamics given by the game G as stage model, inertia $\frac{1}{2}$, and no memory.

Proposition 7.1. Consider a symmetric game G with a unique spiteful outcome s^P. Then the only stochastically stable state of the $(\pi(G); \frac{1}{2}; 0)$-dynamics is $mon(s^P)$.

Proof. By Lemma 4.2, the only recurrent communication classes of the unperturbed dynamics are the singletons $mon(s)$, for all possible $s \in S$. Let Z denote the number of monomorphic states, i.e., the cardinality of the strategy set S. Take any $s \neq s^P$. By Spite, the arrow leaving $mon(s^P)$ in any $mon(s)$-tree has cost at least 2, and all other $Z - 2$ arrows have at least cost 1 by definition. Hence, the stochastic potential of $mon(s)$ is larger than or equal to $(Z - 2) + 2 = Z$.

Consider now $mon(s^P)$. By Negative Spite, the transition from $mon(s)$ to $mon(s^P)$ can be done with a single, direct mutation. This allows us to construct a $mon(s^P)$-tree of cost exactly $Z - 1$, which proves that the stochastic potential of $mon(s^P)$ is $Z - 1$, implying the claim. ¥

Remark 7.2. By [2, Theorem 1], it follows that the expected wait until $mon(s^P)$ is reached is of order ε^{-1}, which means that the speed of convergence is the maximum achievable one.

What drives the previous result is that, by Spite, the spiteful outcome cannot be destabilized with a single mutation, whereas, by Negative Spite, all other outcomes (including any Nash equilibrium) are destabilized with a single mutation. This is precisely the stylized situation captured by the example in Section 5.

Suppose the game G does have a strict, symmetric Nash equilibrium. By virtue of Negative Spite, this outcome is deprived of any significance. It is easy to argue that, since Nash Equilibria are fixed points of best-reply correspondences, and we are not assuming rationality-demanding best reply behavior but rather simple-minded imitation behavior, there is no reason at all for Nash equilibria to play any role. However, we shall see that still Nash might play a role, once we allow for memory (which, as in previous section, implies suppressing inertia).

Definition 7.3. Consider any $(\beta; 0; K)$-dynamics. A state ω is pre-stable if it cannot be destabilized by a single mutation.

Whenever an spiteful outcome exists, it is automatically pre-stable by Spite. If $K = 0$, there is no other pre-stable outcome by Negative Spite.

Proposition 7.4. Consider a symmetric game G with a strict, symmetric Nash equilibrium where all players play strategy s^N. Then, $mon(s^N)$ is pre-stable in any $(\beta(G); 0; K)$-dynamics with $K \geq 1$.

Proof. Consider the state $mon(s^N)$, and consider a single mutation to another strategy $s \neq s^N$. Since $mon(s^N)$ describes a Nash equilibrium, it follows that the mutant is earning less payoffs than the payoff that all individuals in the population experienced the previous period. Hence, the maximum remembered payoff is either the one earned the previous period or the one earned by the non-mutants in the present one. Both possibilities lead to imitation of s^N and hence the mutation is unsuccessful. ¥

The interesting observation is the fact that the concept of Nash might play any role at all in an imitation model. Nash equilibria are defined in a way that requires players to have a full knowledge of the game, in order to be able to compute a best reply. In an imitation model, players do not need to know the game, since they simply compare observed payoffs. The role of memory is allowing for intertemporal comparisons, and the key is the equivalence of the following two payoff comparisons.

1. The actual comparison of the payoff that a boundedly rational deviant experiences after deviating and his former payoff.

2. The potential comparison of the payoff that a rational but myopic player experiences, and the one that he would earn if he deviated.

This observation allows us to pin down precisely in which sense is a Nash equilibrium stable in a model of imitative behavior, and why this stability is ignored in the absence of memory. One can not take this implication any further, as a set of examples easily shows.

We saw in the example of Section 5 that the addition of sufficiently long memory switches the prediction away from the spiteful outcome and allows selection of the Nash equilibrium. The following example shows that this is not a general result.

Example 7.5. Consider a 4-player symmetric game where each player has to choose among two strategies, A, and B. As in Section 6, let $\pi(A;n)$ and $\pi(B;n)$ be the payoffs of an agent playing A or B when exactly n agents are playing A. We specify all the payoffs in the following matrix:

	4	3	2	1	0
A	2	0	3	1	{
B	{	1	3	2	0

It is straightforward to check that $(A;A;A;A)$ is a strict Nash equilibrium and $(B;B;B;B)$ is a spiteful outcome. Hence the latter will be selected by any dynamics with imitation and no memory.

If we allow for K periods of memory, then we know that both monomorphic outcomes are pre-stable, i.e. one mutation is never enough. However, from any of them two mutations suffice to obtain a payoff of 3 for the mutants, which is the largest possible payoff. Imitation follows with positive probability, hence the cost of both transitions is always 2, irrespective of the length of memory. Both outcomes are stochastically stable for any $K \geq 1$.

This example features two interesting properties. On the one hand, the length of memory is irrelevant, as long as it is non-trivial. The symmetric Nash equilibrium is always stable. On the other hand, the spiteful outcome is also always stable. While the Nash equilibrium bases its stability on the maximization of absolute payoffs (reinterpreted as intertemporal comparisons), the spiteful outcome bases its stability on the maximization of relative payoffs (i.e. payoffs experienced simultaneously). In absence of more structure, there is no reason for one stability criterion to prevail over the other.

Pre-stability, while being an intuitive signal of stability, does not imply stochastic stability. Still, in the presence of memory, Nash equilibria are never irrelevant in imitation models. In a companion paper [1], it is shown that the analysis of a Cournot Oligopoly with imitating firms and non-trivial memory is quite complicated. The main result, however, is that in a wide range of oligopolies, both the Cournot and the Walras outcomes are stochastically stable. In other oligopolies, the set of stochastically stable states is formed by the monomorphic states corresponding to quantities in a proper subinterval of the interval defined by the Cournot and Walras quantities, including the Walras one but not the Cournot one. Even in this case, though, it is still true that it

is precisely the pre-stability of the Cournot outcome which creates other stable outcomes di®erent from the Walrasian (i.e. spiteful) one.

8   Discussion

This paper shows that the introduction of memory radically changes the long-run predictions in standard models of evolutionary type. The importance of this result comes from the observation that a \no-memory" assumption is actually too strict for even boundedly rational economic agents.

8.1   Memory and Inertia

An interesting observation is the observed struggle between assumptions of inertia and memory. It is argued that, to be able to study the e®ects of memory, it is necessary to remove inertia from the models. This can be discussed at an intuitive level.

The assumption of inertia builds into the agents' behavior a hardwired preference to keep on doing what they did in the past. In the studied framework, agents are assumed to imitate those actions which yield larger payo®s, among those that they observe and remember. Once the available set of actions is enlarged to include past ones, the assumption of memory simply introduces in the agents' behavior a preference to keep on doing what worked in the past. The presence of inertia as modeled in [3] and others would simply eliminate this preference by undiscriminatedly favoring all past actions, irrespectively of whether they worked or not.

This observation (\to keep on doing what worked in the past") shows that memory incorporates a certain idea of inertia. The di®erence is that this kind of inertia is then an endogenous result, rather than an exogenous behavioral assumption.

It is worth going back to the original reasons for the introduction of inertia in the considered examples. Kandori et al [3] present an alternative interpretation of their dynamics as based on myopic best reply considerations. Then, they argue that this myopia can be justi¯ed in a framework with inertia, where agents sometimes do not adjust. Agents who do adjust might expect others not to do so.[6]

Under the interpretation of the considered dynamics as models for imitation behavior, agents are centered on learning from their experiences, and are boundedly rational in the sense that strategic considerations escape to them. Hence, it can be argued that inertia is not needed to justify their behavior.

The complete removal of exogenous inertia is, however, not strictly necessary. The technical problem arising between inertia and memory reduces to the following simple observation. When studying limit behavior as the mutation (experimentation) probability " goes to zero, any behavior postulated with a ¯xed probability, such as inertia with probability ½, becomes arbitrarily more

---

[6]There is a conceptual problem with this argument. Given inertia, and once these rational considerations are brought into the picture, it is not clear why individuals should compute a best-reply, that is, maximize the current payo®, once they get to choose. If agents foresee that they will be unable to change their strategy for a while, this adjustment is not just their decision for a single period. They should then rationally forecast the trajectory of the system and optimize against it.

probable than experimentation. In a sense, the current modeling of inertia is exaggerating its role. An alternative would be to study also the limit behavior as inertia becomes small, or simply to link its probability to the probability of experimentation, e.g. agents might uphold their strategies with probability $\frac{1}{2} = c\epsilon$", with c an arbitrary constant. It is interesting to note that this possibility is already incorporated in the current models. Since a mutation to the previously used strategy is always possible, all what is left is to reinterpret such a mutation as \inertia." Making the constant c arbitrarily high (but ¯xed) allows to emphasize this behavior, without a®ecting the limiting results.

8.2 Memory and Learning

The presence of memory allows agents to behave as if they were able to \experiment conditionally." When a mutant tries a new strategy, he will observe its success relative to the payo®s obtained the previous period. If the mutation brings payo®s down, the mutant will be able to \correct" his mistake and go back to the previous action. This observation, which is of intertemporal nature in an explicitly dynamic framework, naturally reintroduces better-response considerations into models of bounded rationality without explicitly assuming that the agents compute any best reply function. However, the di®erent situations analyzed in this paper show that this is not the only consideration. It is the interplay between better response (to do better than yesterday) and relative success (to do better than the others) which creates a rich dynamic in which two properties arise as determining the long-run outcomes. The ¯rst property is the one associated to Nash equilibria, reinterpreted as an intertemporal comparison of own payo®s. The second, the already familiar e®ects of spite and negative spite.

In the ¯rst, simpli¯ed example, we have shown that the introduction of memory reverses a quite counter-intuitive prediction, thanks to the better-response considerations introduced by memory. In the Kandori et al model, we have seen the interplay between the two properties and how they change the equilibrium selection as the memory length increases. In the examples mentioned in the last section, we see the two forces clash, giving us examples where two focal points are selected: the ¯rst thanks to its Nash-equilibrium condition, the second thanks to its spiteful properties. As already mentioned, in absence of more structure, there is no reason for one force to prevail over the other.

A   Techniques: Unperturbed Dynamics

This section presents a brief summary of the techniques associated with the analysis of unperturbed Markov Chains. See e.g. [4] for details.

Let – be the state space of a ¯nite Markov chain with transition matrix P, and denote $P^T(\omega; \omega^0) = [P^T]_{\omega;\omega^0}$, i.e. the probability of transition from state ω to state ω^0 in exactly T periods. Two states communicate if there exists $T_1 > 0; T_2 > 0$ ¯nite such that $P^{T_1}(\omega; \omega^0) > 0$ and $P^{T_1}(\omega; \omega^0) > 0$. This de¯nes an equivalence relation and hence a partition of –. The equivalence classes are called communication classes. The process is called irreducible if there is only one communication class, i.e., all states communicate.

It is a standard result in stochastic processes theory that states in the same communication class have essentially the same properties. For instance, heuristically, a system may settle in a communication class in the long run, but it cannot settle in a proper subset of that class.

A communication class C is transient if there exists $\omega \in C$; $\omega' \notin C$ such that $P(\omega, \omega') > 0$. A communication class is recurrent if it is not transient. In ¯nite time, the system will always leave the transient classes. Once the system gets into a recurrent class, it will never leave it. Hence, the recurrent communication classes (sometimes called \limit sets" or \absorbing sets") contain the states where the system will be in the long run.

An invariant distribution is a vector $\nu \in \mathbb{C}(S)$ such that $\nu = \nu \cdot P$. There exists an invariant distribution for each recurrent communication class, which describes the time-average behavior of the system once (and if) it gets into that class. That is, $\nu(\omega)$ is the limit of the average time that the system spends in state ω, along any sample path that eventually gets into the corresponding recurrent class.

The set of all possible invariant distributions is then the convex hull of the invariant distributions associated to the recurrent communication classes.

A state ω is aperiodic if the greatest common divisor of the numbers $\{T > 0 = P^T(\omega, \omega) > 0\}$ is one. If one state in a communication class is aperiodic, then all states in the class are aperiodic, and the class is then itself called aperiodic. If a recurrent class is aperiodic, then the associated invariant distribution describes also the long-run probabilities of the states in the class, $\lim_{T \to \infty} qP^T = \nu$ for all probability distributions q whose support is contained in the recurrent class. This is not true for non-aperiodic classes (intuitively, because they contain cycles and hence the limit probabilities do not exist).

We call the process ergodic if it has a unique recurrent communication class (every irreducible process is trivially ergodic). The (unique) invariant distribution constitutes the long-run prediction for any ergodic process, since it represents the limit behavior of the process independently of initial conditions.

If the process is not ergodic, then several invariant distributions exist, describing the long-run behavior along di®erent sample paths, i.e. the prediction depends on the initial conditions.

B   Techniques: Perturbed Dynamics

The following section presents a brief overview of by-now-standard techniques used to perform stability analysis by perturbing a process with vanishing probability ε. A more complete exposition can be found in [3] or [8]. See also [2] for a recent analysis of this type of models.

Given two states ω, ω', we call distance from ω to ω' the minimal number of strategy changes needed for the transition to occur in one period. Since we have memory, this distance is formally in¯nite if the memory does not \agree" in both states, i.e. if $s_i^k \neq s_i'^{k-1}$ for some $k \in \{-K + 1, \ldots, 0\}$; $i \in \{1, \ldots, N\}$. Otherwise, the distance is $d(\omega, \omega') = |\{i \in \{1, \ldots, N\} = s_i^0 \neq s_i'^0\}|$.

The cost of the transition is then $c(\omega, \omega') = \min\{d(\omega'', \omega) = P(\omega, \omega'') > 0\}$, i.e. the minimal number of mutations for the transition to occur in one period under $P(\varepsilon)$. This is then extended by addition to directed paths.

18

Given a state $\omega \in \Omega$, an ω-tree is a collection of ordered pairs of elements of Ω, such that every state in $\Omega \setminus \{\omega\}$ is the first element of exactly one ordered pair, the state ω is not the first element of any pair, and from any state in $\Omega \setminus \{\omega\}$ there is a sequence of pairs leading to ω.

Alternatively, in a more graph-theoretic language, an ω-tree is a directed graph in Ω, such that every state in $\Omega \setminus \{\omega\}$ is the initial point of exactly one arrow, no arrow departs from ω is not the first element of any pair, and from any state in $\Omega \setminus \{\omega\}$ there is a directed path leading to ω.

Given an ω-tree h, we can define its cost as $C(h) = \sum_{(\omega',\omega'') \in h} c(\omega';\omega'')$. Let H_ω be the set of ω-trees, and let $\grave{A}(\omega) = \min_{h \in H_\omega} c(h)$. Then, the stochastically stable states (i.e. those in the support of the limit invariant distribution μ^*) are those in $\mathrm{argmin}_{\omega \in \Omega} \grave{A}(\omega)$. The quantity $\grave{A}(\omega)$ is sometimes called stochastic potential of ω.

References

[1] C. Alós-Ferrer. Cournot vs. Walras in Oligopoly Models with Memory. Mimeo, Department of Economics, University of Vienna, 2000.

[2] G. Ellison. Basins of Attraction, Long Run Stochastic Stability, and the Speed of Step-by-Step Evolution. Review of Economic Studies, 67:17{45, 2000.

[3] M. Kandori, G. Mailath, and R. Rob. Learning, Mutation, and Long-Run Equilibria in Games. Econometrica, 61:29{56, 1993.

[4] S. Karlin and H. M. Taylor. A First Course in Stochastic Processes, 2nd Ed. Academic Press, San Diego, 1975.

[5] P. Rhode and M. Stegeman. A comment on \Learning, Mutation, and Long-Run Equilibria in Games". Econometrica, 64:443{449, 1996.

[6] A. J. Robson and F. Vega-Redondo. E±cient Equilibrium Selection in Evolutionary Games with Random Matching. Journal of Economic Theory, 70:65{92, 1996.

[7] F. Vega-Redondo. The Evolution of Walrasian Behavior. Econometrica, 65:375{384, 1997.

[8] P. Young. The Evolution of Conventions. Econometrica, 61:57{84, 1993.