CARESS Working Paper 98-05
LEARNING CORRELATED EQUILIBRIA IN POTENTIAL GAMES

By Antonella Ianni[*][†]

Department of Economics

University of Southampton

Southampton S017 1BJ

U.K.

February, 1998

Abstract

The paper develops a framework for the analysis of finite n-player games, recurrently played by randomly drawn n-tuples of players, from a finite population. We first relate the set of equilibria of this game to the set of correlated equilibria of the underlying game, and then focus on learning processes modelled as Markovian adaptive dynamics. For the class of potential games, we show that any *myopic-best reply* dynamics converges (in probability) to a correlated equilibrium. We also analyze *noisy best reply* dynamics, where players behaviour is perturbed by payoff dependent mistakes, and explicitly characterize the limit distribution of the perturbed game in terms of the correlated equilibrium payoff of the underlying game.

1

1 Introduction

A branch of recent literature in economic theory has studied models of learning and evolution in an attempt to provide a rationale for commonly used equilibrium notions in game theory. Clearly, any equilibrium notion characterizes epistemologically equilibrium behaviour on the part of players. However, if agents behaviour is taken as a primitive of the analysis, then not all types of behaviour are necessarily consistent with an equilibrium notion. Hence, the focus of this literature on convergence and stability properties of dynamic processes of learning and evolution that may lead to equilibrium.

One line of research builds upon the characterization of equilibrium play in terms of *aggregate* behaviour, as in the mass-action interpretation, that J. Nash himself suggested in his Ph.D. Dissertation (Princeton University, 1950). The idea is fairly simple, in that it relates any of the players of an underlying game with a set of agents who could potentially ful ll the role of that player in the interaction. Optimality of equilibrium choices must then hold for each of these agents.

Focusing on a dynamic process of learning and/or evolution in this setting requires the study of the dynamics of strategy choices in the aggregate. On the grounds of the complexity of a potentially very rich and articulate form of aggregate interaction, this formalization relaxes the assumption of rationality of players, to various forms of modelled reasonable behaviour. An equilibrium notion is then characterized in terms of the steady state of such dynamic process.

As many games admit multiple equilibria, though each equilibrium may be justi ed on the grounds of a dynamic process that leads to it, it is unclear how we are to predict what is the likely outcome of the interaction. One way of addressing multiplicity issues is to describe agents behaviour in probabilistic terms, formalizing the idea that there is a clear and systematic rational element in their decisions, but choices as such are not always consistent with it, as sometimes players make mistakes, and blindly imitate other players or choose actions in an entirely random fashion. By guaranteeing ergodicity properties of the dynamics, mistakes can recover unique predictions, as some equilibria might be more robust to perturbations than

2

others.

This paper contributes to this literature in three respects. First it provides a simple rationale for the notion of correlated equilibrium. The focus on the wider set of correlated equilibria (rather than on Nash equilibria) stems naturally from considering a very general pattern of interaction among players. Such generality of the interaction structure constitutes a second contribution, as it allows to relate different plausible models in a unified framework. Third, the formulation of the model allows for a straight-forward characterization of the long-run properties of an underlying stochastic dynamics and provides a useful tool for analyzing equilibrium selection issues for the class of potential games.

In particular, the paper develops a framework for the analysis of finite n-player games, recurrently played by randomly drawn n-tuples of players, from a finite population. The framework of the interaction is exogenously given, in the sense that it is formalized in terms of a probability distribution that is common knowledge among players. It accounts for uniform random matching (where each player is equally likely to be matched with any other $(n-1)$ players in the population) and local matching (where each player s opponents are randomly chosen among her neighbours) as special cases. Though the analogy is not pursued here, as noted in Morris (1997), there exists a close relation between the interpretation we provide in terms of local interaction and that of a game of incomplete information.

Section 3 analyzes the relation between the set of equilibria of the aggregate game (defined as a population game in Section 2) and the set of equilibria of the underlying game, i.e. the game that once matched, players play. It is shown that the identification of the former in terms of replicated version of the latter, does not necessarily hold, as choices may be correlated due to the pattern of interactions. The Section builds upon the results of Mailath, Samuelson and Shaked (1997). By constructing a *replica* game, it is shown that if matching is uniform, then any of the resulting correlated equilibria corresponds to a Nash equilibrium. In the model, the random matching may provide a correlation (or convexification) device, that acts in an analog way as past play does in Hart and Mas-Colell (1997), though the analogy between the two papers stops here.

In Section 4, we formalize a class of learning models as *myopic best-reply* behaviour. The dynamics rely upon the assumption that players hold static expectations on the environment they are called to interact in, and, whenever they are able to

3

adjust their strategies, they act so as to maximize their expected payoff. In a framework that focuses on the evolution of play over time, these dynamics are particularly appealing in that, whenever the environment is stationary, players behaviour is optimal. We show that if the underlying game is a potential game (roughly speaking, a game where players behave as if they had identical payoff functions), the population game, played under any arbitrary matching, also converges to an absorbing state. As the potential function we use is the average expected payoff in the population, the result (stated in Theorem 1) is essentially a generalization of the fundamental theorem of natural selection, to a nite population - nite game setting. Hence, we are able to characterize correlated equilibria of the underlying game as being local maximizers of the potential. The result we obtain here is related to Neyman (1997), though we do not need any speci c assumption of convexity.

In Section 5, we take the view that ergodicity properties of the process make it a good predictor of the long run behaviour of the population game. To this aim we analyze perturbed best-reply dynamics, where players adopt a myopic best-reply to the previous period state of the system with high probability, but with some small, though positive, probability they do something else instead. Speci cally, we assume that mistakes are payoff dependent, in that the probability with which any of the available actions is adopted depends on the expected payoff achievable with that speci c action. For the same class of games for which we are able to prove convergence in Section 4, we show that the probability with which each state is observed in the limit is a function of the average expected payoff in the population, in that state. This allows us to address equilibrium selection issues over the set of correlated equilibria of the underlying game. Though technically less demanding, the result generalizes the model of Blume (1993) to a nite population - n-player game, where interaction is not necessarily restricted to be on a lattice.

2 The Model

As it will become apparent in what follows, the line we take in the formalization is to keep the model as simple as possible, though generality is achieved at the cost of at times heavy notation. The model consists of three key ingredients: a speci cation of the interactive setting, a de nition of a notion of equilibrium in this setting and a characterization of a dynamic learning process through which agents can potentially

learn to play the equilibrium, by interacting repeatedly over time. We address these issues in exactly this order.

2.1 The game:

Let Ω be a finite set; μ be a probability measure defined over subsets of elements of Ω, and G be a normal form game. We take Ω to represent the set of players, μ to formalize the probability with which players are matched and G to be the game that, once matched, players play. A *population game* is simply defined by these three elements:

Definition 1 (Population Game) *Given a population of players Ω, a random matching technology μ, and an underlying game G, the triple $\Gamma = (\Omega, G, \mu)$ defines a population game if Ω, μ and G satisfy[1]:*

1. *$2 \leq \sharp \Omega = N < \infty$.*

2. *For $2 \leq n \leq N$, μ is defined over $P = \{(\omega_1, \omega_2, ..., \omega_n) : \omega_i \in \Omega\}$ and is such that for all ω_i in Ω, $\mu(\omega_i) = \sum_{-i} \mu(\omega_i, \omega_{-i}) > 0$;*

3. *$G = (\{A_i\}_{i=1}^n, \{\Pi_i\}_{i=1}^n)$ is an n player-normal form game, with action spaces $\{A_i\}_{i=1}^n$ and payoff functions $\{\Pi_i\}_{i=1}^n : \prod_{i=1}^n A_i \to \Re$ respectively.*

We refer to μ as the *random matching* among players, and we assume this is exogenously given to players, in the sense that μ is common knowledge among players.

The general specification of the matching that we adopt encompasses standard formalizations of the interaction pattern used in the recent literature on random matching models, though it is by no means limited to those. In order to see this, let us denote the support of μ by S, and state the following:

Definition 2 (Uniform Matching) *Matching is uniform if $\forall$ $(\omega_1, \omega_2, ..., \omega_n)$ in S, $\mu(\omega_1, \omega_2, ..., \omega_n) = (\sharp S)^{-1}$.*

If $S = P$, we shall denote μ as μ_U and refer to it as uniform population matching. If $S \subset P$, we shall denote μ as μ_L and refer to it as uniform local matching.

In particular μ_U characterizes a model, commonly referred to in the literature as a *random matching model*, where each player is equally likely to be matched with any subset of $(n-1)$ players in the population. μ_L instead formalizes the idea that the set

of potential opponents for player ω depends on ω itself, and player ω is equally likely to meet any player *within* that set. Since one possible way to distinguish player ω_1 from player ω_2 in the population is to provide each player with an identity in terms of a speci c *location* on an appropriately characterized space , this formalization has often been used in the recent literature (for example in Ellison (1993) and (1995), Blume (1993), Anderlini and Ianni (1996a and 1996b), Morris (1996), Ely (1996) and others). Unlike in most of these contributions, the results we obtain here do not typically depend on any particular topological characterization of the set .

In general, an underlying game requires the assignment of a role to each of the players who take part in the interaction. In our setting, a random matching that has this property satis es the following De nition:

De nition 3 (Role identi cation) *Matching assigns roles to players if μ partitions   in $\{$ $_i\}_{i=1}^n$, i.e. $\mu(\omega_1, ..., \omega_n) > 0$ only if, for all $i = 1, 2, ..., n$, each $\omega_i \in$ $_i$.*

We will conventionally assume that if the above de nition holds, then all ω s in $_i$ who take part in the interaction will play role i in G (i.e. they will have set of actions A_i and payoff function Π_i)[2].

In the examples, we will nd it convenient to represent   in terms of a graph, $\mathcal{G}($ $, E)$, where E is the set of edges connecting any n-tuple of elements of   to which μ assigns positive probability.

2.2 Equilibria

As G is a normal form game, action spaces are nite. Let A_ω be the set of actions available to player ω in G and let any $\mathbf{a}_\omega \equiv \{a_\omega^1, a_\omega^2, \cdots, a_\omega^{\sharp A_\omega}\}$ such that $a_\omega^i \geq 0$ and $\sum_{i=1,\cdots,\sharp A_\omega} a_\omega^i = 1$ denote any mixed strategy adopted by player ω. We think of a pure strategy as a degenerate probability distribution and we denote it by a_ω (or a_{ω_i} if we want to specify that $a_\omega^i = 1$). Player ω s opponents are denoted by $-\omega$ and are chosen at random, among those players for which $\mu(\omega, -\omega) > 0$. If Π_ω denotes ω s payoff function in G, ω s expected payoff from choosing $\mathbf{a}_\omega$ is denoted by $E_\mu[\Pi_\omega(\mathbf{a}_\omega, \mathbf{a}_{-\omega}) \mid \mu]$. Strategy $\mathbf{a}_\omega^*$ maximizes ω s expected payoff, if and only if $\mathbf{a}_\omega^* \in Arg\max_{\mathbf{a}_\omega} E_\mu[\Pi_\omega(\mathbf{a}_\omega, \mathbf{a}_{-\omega}) \mid \mu]$.

In the population, a mixed (vs. pure) strategy pro le $\mathbf{a} \in \prod_\omega \mathbf{a}_\omega$ (vs. $a \in \prod_\omega a_\omega$) associates a mixed (vs. pure) strategy $\mathbf{a}_\omega$ (vs. a_ω) to each player $\omega \in$. A subset of

the set of all profiles are the *equilibrium profiles* of the population game Γ, defined as follows:

Definition 4 (Equilibria of Γ) *Given a population game Γ, $\Theta(\Gamma) \subseteq \prod_\omega \mathbf{a}_\omega$ is the set of its equilibrium profiles. Each element $\mathbf{a}^* \in \Theta(\Gamma)$ is such that $\mathbf{a}^* \equiv \prod_\omega \mathbf{a}_\omega^*$, where $\mathbf{a}_\omega^* \in Arg \max_{\mathbf{a}_\omega} E_\mu[\Pi_\omega(\mathbf{a}_\omega, \mathbf{a}_{-\omega}) \mid \mu]$ for all ω.*

$\theta(\Gamma) \subseteq \Theta(\Gamma)$ is the set of strict[3] *equilibrium profiles of Γ. Each element $a^* \in \theta(\Gamma)$ is such that $a^* = \prod_\omega a_\omega^*$, where $a_\omega^* = Arg \max_{a_\omega} E_\mu[\Pi_\omega(a_\omega, a_{-\omega}) \mid \mu]$ for all $\omega \in$.*

We note at this point that the above definition of equilibrium reminds of that of a Bayesian Nash equilibrium in a game of incomplete information. Though it will not be pursued here, the analogy is not a mere coincidence. In Anderlini and Ianni (1996a) such interpretation was used to analyze a model where interaction takes place on a lattice of Von Neumann-Morgestern neighbourhoods and players repeatedly play a 2-by-2 coordination game. Morris (1997) provides a careful general characterization of this analogy, that hinges upon the formalization of the pattern of interaction among players.

2.3 Learning Processes

The behavioural assumption underlying the learning model we analyze is that each player aims at maximizing her expected payoff. Expectations as to potential opponents behaviour are adaptive. This is motivated by the idea that future play will not be different from what was observed in the recent past. In the jargon of the recent literature on learning, the dynamic processes we shall study are those of *myopic best-reply* and a noisy version of it.

Time is discrete. We shall assume that, at each time t only one n-tuple of players actually play the game (though, by assumption, all players have a strictly positive probability to take part in the interaction[4]). The latter will involve players $(\omega_1, ..., \omega_n)$ with probability $\mu(\omega_1, ..., \omega_n)$. Before the random draw takes place, each player ω has chosen an action from the set A_ω, by appealing to a *rule*. Within the same model, all players appeal to the same rule, which is simply a mapping from what she knows to what she would do, were she chosen to play and takes the form of a probability distribution over the action space. We shall describe different rules below.

At the beginning of time, nature assigns an action to each player. The rst interaction takes place at time zero. Between time zero and time one, only one

player, chosen at random in the population, receives an *updating opportunity*; the latter consists of the possibility to change the original action according to the rule. Then time one s interaction takes place. As time rolls by, the above story is repeated with exactly the same timing. As two players that consecutively update their actions have information sets that differ by the action of at most one player, we believe that this specification is more likely than others to capture the inherent stochastic nature of information gathering[5].

The specification of the behavioural rules and of the allocation of updating opportunities, completely defines a stochastic process over $\prod_\omega a_\omega$. We will refer to $\wp$ as the process that governs the dynamics under myopic best reply, to $\wp^\sigma$ as the process under noisy best reply, and we shall denote by $P(a' \mid a)$ and $P^\sigma(a' \mid a)$ the probability with which the system transits from a to a' under the two processes respectively.

2.3.1 Behavioral Rules

As expectations are adaptive, we need to explicitly account for a dynamic element in the specification of the behavioural rules we postulate. We assume that at the beginning of period t, all players are informed about the configuration of actions in the population at period $t-1$ (namely a_{t-1}).

MBR: The rst behavioral rule we analyze is known as the *myopic best-reply:* players hold static expectations about their opponents behavior and, whenever they have the opportunity to do so, they choose the action that maximizes their expected payoff. As a result, if it is player ω s turn to update, she will choose an action such that:

$$\mathbf{a}_\omega^t \in Arg \max_{\mathbf{a}_\omega} E_\mu[\Pi_\omega(\mathbf{a}_\omega, a_{-\omega}^{t-1}) \mid \mu]$$

NBR: The second behavioral rule we analyze is based on the idea that the probability with which a player adopts each action available to her depends on the expected payoff to that action. To this aim, we assume that, if it is player ω s updating turn, she will choose an action such that:

$$\mathbf{a}_\omega^t \equiv \{a_\omega^{it} = \frac{\exp[\ \sigma E_\mu[\Pi_\omega(a_{\omega_i}, a_{-\omega}^{t-1}) \mid \mu]]}{\sum_{i \in A_\omega} \exp[\sigma\ E_\mu[\Pi_\omega(a_{\omega_i}, a_{-\omega}^{t-1}) \mid \mu]]}, \ i = 1, ..., \sharp A_\omega\} \tag{1}$$

where $\sigma \geq 0$ is a parameter. The rule is not meant to model any specific behavioural assumption, but instead formalizes the idea that the ratio between

the (logarithm of the) probabilities with which a player chooses any two actions is proportional to the difference in the expected payoff[6]. If we are to motivate probabilistic behaviour in terms of mistakes, the above rule postulates that the probability with which such mistakes occur is *payoff-dependent*, in that very costly mistakes are less likely to occur than relatively less costly mistakes.

3 Characterization of Equilibria

The first objective we have is to characterize the set of equilibria of the population game Γ in terms of the set of equilibria of the underlying game G. Being finite, G admits at least one Nash-equilibrium in pure or in mixed strategies. In general, we shall denote by $N(G)$ the set of Nash-equilibria of G and by $\Psi(G)$, the set of correlated equilibria of G, as introduced in Aumann (1974).

In order to look at a random matching model where only a subset of individuals are drawn to play the underlying game, we first build on the mass action interpretation verbally described by John Nash himself and borrow the terminology used elsewhere in the literature. This is done by constructing a *replica game* for the underlying game G. A replica game is constructed from G, by replicating say $2 \leq m = N/n$ agents with exactly identical preferences, for each role i in G. We then assume that the n players who are going to play G, are drawn at random, each of the m replicas of i having equal probability. Within our model each m-replica of G is obtained as follows:

Definition 5 (Replica of G) $\Gamma_m = (\ , \mu_R, G)$ *is an* m*-replica of* G *if* $\mu \equiv \mu_R$ *is such that a) it satisfies Definition 3, with* $\sharp\ _i = m$ *for all* $i = 1, 2, ..., n$*, and b) it satisfies Definition 2 and it has full support.*

Condition a) requires the matching to define an n-partition of the set or, equivalently, to generate an n-partite graph. Condition b) requires the probability distribution that defines the matching to be uniform, i.e. $\mu_R(\omega_1, ..., \omega_n) = (m)^{-n}$ for all $(\omega_1, ..., \omega_n)$ in the support of μ_R. Hence roles are known to players, while opponents are uncertain and randomly chosen in the population.

The idea behind the replica game is that each of the roles i available in G is played by an individual chosen at random among all players in $\ _i$. It is clear that for $m = 1$ the replica game Γ coincides with G. Definition 4 would characterize the set

9

of Nash-equilibria and the subset of strict Nash-equilibria respectively. It is natural at this point to ask whether this is also true for any m-replica ($m > 1$), and whether the relation would hold if we relaxed conditions a) and/or b).

In order to address these questions we need some further notation to relate the set of equilibria of the population game, previously defined in terms of action profiles in the population, to probability distributions over sets of actions, or Cartesian products of them. We proceed as follows. First note that a generic profile, $\mathbf{a} \in \prod_\omega \mathbf{a}_\omega$, induces a partition of the *set of n-tuples of players* according to the actions that are chosen in $\mathbf{a}$. Let us denote the restriction of μ over this partition $\mu_{\mathbf{a}}(\prod_i A_i)$. By construction, this is a probability distribution over the Cartesian product of the action space: any $\mu(i_1, i_2, ...i_n)$ is the frequency with which an outside observer would observe (or rather foresee, if mixed strategies are chosen) the combination of actions ($i_1 \in A_1, i_2 \in A_2, ...i_n \in A_n$), if the action profile in the population was $\mathbf{a}$ and matching was defined by μ. Let us denote by $\mu_{\mathbf{a}}(A_i)$ for $i = 1, 2, ..., n$ its corresponding marginals (for example $\mu(i_1) = \sum_{i_2 \in A_2} .... \sum_{i_n \in A_n} \mu(i_1, i_2, ...i_n)$).

Then the following holds:

Proposition 1 *Let $\Gamma_m = (\quad, \mu_R, G)$ be an m-replica of G. Then for all $\mathbf{a}^* \in \Theta(\Gamma_m)$,* $\mu_{\mathbf{a}}(A_i) \in N(G)$.

Proof. See Appendix.

Relaxing condition a) in the Definition 5 leads to a remark about the consistency of the interpretation of the model.

Suppose μ is uniform and has full support over P, but does not assign roles to players (this is the case if Definition 2 holds, but Definition 3 does not). Then roles in the underlying game, as well as opponents in the interaction are uncertain to each player. If matching is uniform, one plausible interpretation is that any player is equally likely to play any of the roles available in G. If actions are chosen *ex-ante*, i.e. before playing the game, a consistent specification requires all action spaces to consist of exactly the same labeled elements. In this case, however, the relation between the set of equilibria of the population game, Γ, and the underlying game, G, is not obvious[7]. Now suppose instead that, before the interaction takes place, each player can choose a strategy that consists of the choice of an action for each of the possible roles that could be assigned to her in the interaction (in other words, suppose that such strategies can be made conditional on future roles). Then the equilibria of

the population game as in Definition 4 can be interpreted in an *ad interim* sense, i.e. after the uncertainty about roles resolves and only the uncertainty about opponents has to be accounted for. The class of games for which this distinction does not matter is the class of *symmetric games*, i.e. games for which the payoff functions are independent of the roles[8].

We now focus on the implications of condition b) used in Definition 5, and show that the characterization of the set of equilibria of Γ in terms of replicated version of the Nash equilibria of G does not necessarily hold if matching is not uniform. This particular feature of local random matching models was first noted in Mailath, Shaked and Samuelson (1997), where the authors characterize the set of equilibria of a model formalized in terms of a number of finite populations of players, one for each of the roles of an underlying normal form game, randomly grouped in casts to play one shot of the underlying game. It is not difficult to see that such a model is analog to what we previously defined as an m-replica of G in all respects, apart from condition b), since matching is not necessarily uniform. The authors show i) that each equilibrium of the model defines a correlated equilibrium of the underlying game, and ii) that for any correlated equilibrium of an underlying game one can design an interaction mechanism that reproduces that equilibrium in terms of observed frequencies over the Cartesian product of the action spaces. The Proposition that follows builds upon their result.

Proposition 2 (Mailath, Samuleson, Shaked (1997)) *Let $\Gamma = (\ ,\mu,G)$.*

If μ satisfies Definition 3, then i) for any $\mathbf{a}^* \in \Theta(\Gamma)$, $\mu_{\mathbf{a}}(\prod_i A_i) \in \Psi(G)$ and ii) for any $\psi \in \Psi(G)$ there exists an a and a μ such that $\mu_{\mathbf{a}}(\prod_i A_i) = \psi$.

Proof. See Appendix.

As pointed out in Mailath, Samuelson and Shaked (1997) and in Ianni (1997), if matching does not assign roles to players (i.e. if Definition 3 is not satisfied), the ω_is are exchangeable in the definition of μ. As a result $\mu_{\mathbf{a}}(\prod_i A_i)$ will itself be symmetric (since, roughly speaking, any permutation of the n players would be observationally equivalent), and so will the correlated equilibrium[9]. As not all correlated equilibria are symmetric, in the Sections that follow we shall assume that the random matching allows for role identification among players, or in other words, that Definition 3 holds throughout.

The main conclusion that we can draw from the analysis so far is that, in general, the set of equilibria of the population game might look very different from the

(replicated version of the) set of equilibria of the underlying game. This is in general the case if each player is not equally likely to interact with any other player in the population. Some illustrative examples follow.

Example 1 *Let C_A and C_C be 2-by-2 underlying games with payoff matrices given by:*

$$C_A \equiv \begin{bmatrix} 3/2, 3/2 & 0,0 \\ 0,0 & 1,1 \end{bmatrix} \qquad C_C \equiv \begin{bmatrix} 3,3 & 1,1 \\ 0,0 & 4,4 \end{bmatrix}$$

and consider $\Gamma_A = (\ , C_A, \mu)$ and $\Gamma_C = (\ , C_C, \mu)$ where $\sharp\ = 4$ and μ is speci ed as in the following graphs (vertices are players, edges are elements in the support μ and the number above each edge is the probability with which each matching occurs).

A: Uniform nearest neighbours   B: Local nearest neighbours   C: Local nearest neighbours

The setting on the left (Figure A) is a simpler version of the model analyzed in Ellison (1993), where players are located on a circle and interact only with their nearest neighbors. It is not difficult to see that, for the game on the left, the only strict equilibria of Γ_A are those where every player adopts exactly the same action, i.e. replicated versions of the strict Nash equilibria of the underlying game. Any con guration of play where different actions co-exist cannot be supported in equilibrium. The picture in the middle (Figure B) shows why this analogy does not hold in general, as for example a con guration of play where α and β adopt say the rst of the two actions available, and δ and γ adopt the other is also a strict equilibrium of Γ, though it does not correspond to any of the Nash equilibria of G. Analog reasoning applies to underlying asymmetric games, such as C_C, played under the matching pattern of which in Figure C. It is not difficult to check that, in this latter case, one possible strict equilibrium of the population game has α and β playing the Top-Left Nash equilibrium strategies, while γ and δ play the Bottom-Right. Proposition 2 shows that any such equilibrium induces a correlated equilibrium of the underlying game. For

the C_A-Figure B and the C_C-Figure C games, these are respectively:

$$\begin{bmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{bmatrix} \qquad \begin{bmatrix} 1/3 & 1/6 \\ 1/8 & 9/24 \end{bmatrix}$$

Example 2 *Let Z be a 2-player - 3-action game of the Rock, Scissors, Paper kind:*

$$Z \equiv \begin{bmatrix} 0,0 & 2,1 & 1,2 \\ 1,2 & 0,0 & 2,1 \\ 2,1 & 1,2 & 0,0 \end{bmatrix}$$

and consider $\Gamma = (\ , Z, \mu_L)$ *where* μ_L *is specified as follows:*

Uniform Local Matching

Though the underlying game Z admits only one Nash equilibrium in mixed strategies (that assigns equal probability to each of the actions), one strict equilibrium of the corresponding Γ has players α and δ playing R, players β and ε playing S and players γ and ϕ playing P, inducing the following correlated equilibrium:

$$\begin{bmatrix} 0 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 \\ 1/6 & 1/6 & 0 \end{bmatrix}$$

Though all the configurations in the above Examples yield a strict equilibrium of the corresponding population game, we will argue in the Sections that follow, that the equilibria of Example 1 are, in a way, more robust than that of Example 2. In order to address issues of dynamic stability of equilibria of the population game, we now move on to analyze the learning processes we study.

4   Will we ever see an equilibrium?

The question we address next is the interpretation of the equilibria in terms of the learning process we examine. We shall do so by studying the equivalence classes of

the process[10] $\wp$ that, we recall, ranges over $\prod_\omega a_\omega$. It should be clear that, given an arbitrary matching μ, by lumping the process in terms of the fractions of players in the population adopting each of the available actions, we would inevitably miss out relevant information. Given a population game, Γ, let $E^\wp(\Gamma)$ be the set of ergodic sets of the process under $\wp$ and $A^\wp(\Gamma) \subseteq E^\wp(\Gamma)$ be the set of absorbing states. Given that the state space is finite, we know that $E^\wp(\Gamma) \neq \emptyset$. We also know that state a is absorbing if and only if $P(a \mid a) = 1$ and therefore, only those states that are strict equilibria of Γ can be absorbing (or in other words $A^\wp(\Gamma) = \theta(\Gamma)$). In order to fully characterize the process under MBR dynamics, we need to show that such absorbing states can be reached, possibly in finite time, by the process.

It is well known that MBR dynamics can cycle (i.e. admit ergodic sets different from absorbing states) over simple normal form games, the Rock-Scissors-Paper being a leading example. It is also known that for some classes of normal form games cyclic behaviour can be ruled out if we are able to show that the game admits a potential, i.e. a real valued function that ranges over the Cartesian product of the action spaces and behaves monotonically along any path of the dynamics. Normal form games that exhibit this property are defined in Monderer and Shapley (1996a), as potential games. In that paper the authors explicitly characterize this class of games by defining a set of conditions on the relation between the underlying payoff functions, that are sufficient to guarantee convergence of any dynamics that moves along an improvement path. In a population game of the kind analyzed here, as expected payoffs depend on the payoffs of the underlying game, as well as on an arbitrary definition of the matching, checking for those conditions is at best not operative, as it would involve studying functions defined over $\prod_\omega a_\omega$ (rather than $\prod_i A_i$).

The logic we follow is then to construct a tractable potential function for the population game and show that convergence to an absorbing state obtains whenever such potential exists.

For each profile $\mathbf{a}$, we define the *average expected payoff* per interaction by aggregating expected payoff from a round, over all players:

$$\Pi_\mu(\mathbf{a}) = \frac{1}{n} \sum_{(\omega, -\omega)} \Pi_\omega(\mathbf{a}_\omega, \mathbf{a}_{-\omega}) \mu(\omega, -\omega)$$

This quantity clearly changes, as the MBR dynamics unfolds, and it is clear from the construction that $\Pi_\mu(\mathbf{a})$ is real valued and bounded over its domain. In what follows, we shall show that if the underlying game G is a potential game, than this

14

function behaves monotonically along any path of MBR dynamics.

One problem we need to address is the occurrence of ties that might make a player indifferent between different actions. To this aim, recall that, as the population is nite, each player can only see a nite number of distributions of play among her potential opponents. Hence, for most underlying games, we can deal with ties by perturbing the payoff functions and/or the matching in such a way that observed frequencies do not reproduce any of the mixed equilibria of G^{11}. As a working tool, we de ne a population game for which this holds as having *no ties*:

De nition 6 Γ *has no ties if* Π_ω *and* μ *are such that for all* $\omega \in$ *and for all* $a_\omega, a'_\omega \in A_\omega$: $E_\mu[\Pi_\omega(a_\omega, a_{-\omega}) \mid \mu] \neq E_\mu[\Pi_\omega(a'_\omega, a_{-\omega}) \mid \mu]$.

The result that follows identi es a sufficient condition for global convergence for MBR dynamics played in a population game Γ. The condition requires the underlying game G to have identical interests (as in Monderer and Shapley (1996b)).

De nition 7 (Game with Identical Interests) $G = (\{A_i\}_{i=1}^n, \{\Pi_i\}_{i=1}^n)$ *is a n player-normal form game with identical interests, if* $\{\Pi_i\}_{i=1}^n \equiv \Pi : \prod_{i=1}^n A_i \to \Re$ *for all* i *s.*

If this condition is satis ed, then the following Theorem shows that (1) the average expected payoff in the population, in a given state, $(\Pi_\mu(a))$ corresponds exactly to the average payoff of the underlying game, taken with respect to the restriction of μ induced by the state a over the Cartesian product of G s action space (previously denoted as $\mu_\mathbf{a}(\prod_i A_i)$ and here lightened to $\mu_\mathbf{a}$); (2) that the function $\Pi_\mu(a)$ is non-decreasing along any path of the myopic best-reply dynamics, and that there exists a relation between the set of its maximizers and the set of absorbing states of the underlying process.

Theorem 1 *Let* $\Gamma = (\ , \mu, G)$, *and assume* G *has identical interests. Then the following holds:*

(1) for any given a, $\Pi_\mu(a) = E_{\mu_\mathbf{a}}[\Pi(i_1, i_2, ..., i_n)]$ *for* $i_i \in A_i$;

(2) along any path of $\wp$:

    a) $\Pi_\mu(a)$ *is non-decreasing;*

b) $\Pi_\mu(a)$ *is (locally) maximized at any* $a \in A^\wp(\Gamma) \equiv \theta(\Gamma)$, *and, whenever* Γ
has no ties, the only maximizers are to be found in $\theta(\Gamma)$.

Proof. See Appendix.

The above theorem asserts that for the class of games with identical interests (that includes for example pure coordination games, as in Kandori and Rob (1993) or Robson and Vega-Redondo (1996)), asymptotics of the process can be analyzed by looking at the set of states that maximize the average payoff. Since we know from Proposition 2 that every equilibrium a induces through $\mu_\mathbf{a}(\prod_i A_i)$ a correlated equilibrium of the underlying game, we think of $\Pi_\mu(a) = E_{\mu_\mathbf{a}}[\Pi(i_1, i_2, ..., i_n)]$ as the average payoff of the underlying game, taken with respect to the probability coefficients that de ne the corresponding correlated equilibrium. Being the image of a convex polyhedron under a linear map, the set of correlated equilibrium payoffs is also a convex set in $\Re^n$. As the function $\Pi_\mu(a)$ behaves monotonically along the dynamics, the above result also implies convergence, over time, of μ_a to the set of correlated equilibria of the underlying game G. Hence each of the local maximizers of the potential $\Pi_\mu(a)$ corresponds to an equilibrium of Γ and to a correlated equilibrium of G. This is stated next.

Corollary 1 *Consider* $\Gamma = (\quad, \mu, G)$, *where* G *has identical interests and* Γ *has no ties. If* Γ *is played under MBR, for any given initial condition* a_0,

$$\lim_{t\to\infty} \Pr[a_t \in \theta(\Gamma)] = 1 \quad and \quad \lim_{t\to\infty} \Pr[\mu_{a_t} \in \Psi(G)] = 1$$

Proof. See Appendix.

In general, equilibria of the population game depend on the interaction pattern μ, as well as on the underlying game, and unless G admits a Nash equilibrium in strictly dominant actions, the set $\theta(\Gamma)$ will not be a singleton. In this case, as it is well known, the probability with which each state is observed in the limit will depend on the initial condition. In the Section that follows, we exploit the ergodicity properties of a perturbed process to analyze the long run behaviour of the system.

5   Which equilibria are we likely to see?

In this Section we analyze the process generated under noisy best reply. We assume mistakes are independent across players and over time, but do explicitly depend on

payoff considerations, as formalized in equation (1). It is clear, that this process $\wp^\sigma$ also ranges over $\prod_\omega a_\omega$ and, for any $0 < \sigma < \infty$, it admits a unique ergodic set that includes all states in the state-space (i.e. $E^{\wp^\sigma}(\Gamma) = \prod_\omega a_\omega$).

The result that follows shows that, for the same class of population games for which we proved global convergence in the previous Section, we are also able to fully characterize the limit distribution of the process $\wp^\sigma$.

Theorem 2 *Consider $\Gamma = (\ ,\mu, G)$. For any $0 < \sigma < \infty$:*

(1) *$\wp^\sigma$ is ergodic. Each entry is such that $\lim_{\sigma \to \infty} P^\sigma(a^1 \mid a) = P(a^1 \mid a)$ for all $a, a^1 \in \prod_\omega a_\omega \equiv \mathcal{A}$.*

    If G has identical interests, then:

(2) *$\wp^\sigma$ admits the following probability vector as a unique ergodic limit distribution:*

$$P_\sigma = [P_\sigma(a) = \frac{\exp[\sigma \Pi_\mu(a)]}{\sum_{a \in \mathcal{A}} \exp[\sigma \Pi_\mu(a)]}, \quad a \in \mathcal{A}]$$

Proof. See appendix.

The above result is to be interpreted in the following way. Recall that in Section 4 we speci ed conditions under which the average expected payoff in the population was strictly increasing, and as such could be taken to represent a potential, or an energy function for the system under myopic best-reply dynamics. Sufficient for this is that the underlying game has identical interests, or that it is a potential game. In this case at each step of the noiseless process the average expected payoff increases. As previously pointed out, the (noiseless) myopic best reply dynamics can be obtained for the model we analyze in this Section with $\sigma = \infty$. Mistakes occur for any nite value of σ. In this case the system can loose energy at random in amounts that are proportional to the value of σ. Hence some of the local maxima of the energy function, i.e. the average expected payoff, can be destabilized. If actions are chosen in an entirely random fashion, that is for $\sigma = 0$, the system wanders randomly among all its possible states. As a result, for any nite value of the parameter σ, we cannot identify states as attractors for the dynamics. However, we can identify attractor *probability distributions* over states, that take the form speci ed in the above Theorem.

As part (1) and (2) of the Theorem are true for each nite σ, and transition probabilities are continuous in σ, we can then consider the limit as σ becomes in nite.

17

It is a direct implication of the Theorem that, as this happens, the system spends most of its time in the state for which the average expected payoff achieves a global maximum. As (from part (1)) $\wp^\sigma$ converges to $\wp$ entry-by-entry as $\sigma \to \infty$, the limit distribution concentrates all of its mass on one of the absorbing states of $\wp$ that are absorbing states of the population game.

It is interesting to interpret the ndings of Theorem 2 in terms of equilibrium selection over the set of correlated equilibria of the underlying game. In particular, a direct implication is that, if the underlying game admits a Pareto-dominant equilibrium, then it will be selected by the learning process.

Corollary 2 *Consider* $\Gamma = (\ , \mu, G)$, *where* G *has identical interests, and suppose* G *admits* $i = (i_1, i_2, ..., i_n)$ *as a Pareto-dominant outcome. If* Γ *is played under NBR, then there exists an* $a \in \mathcal{A}$ *and a* μ *for which:*

$$\Pi_\mu(a) = E_{\mu_a}(\Pi) = \Pi(i_1, i_2, ..., i_n) \quad and \quad \lim_{\sigma \to \infty} P_\sigma(a) = 1.$$

Proof. See Appendix.

Games that fall into this class are, for example, *common interest* games, as in Aumann and Sorin (1989). The above Corollary then postulates that a population game, played under noisy best-reply dynamics, is able to select the one that yields the highest correlated equilibrium payoff within that set.

It is not difficult to see that the results of Theorems 1 and 2 can be generalized to the class of better reply dynamics (where each player myopically adopts an action that only *increases* her expected payoff) and to the class of games that are best-reply equivalent to games with identical interests. As the rst generalization would not substantially add to the analysis, and in the second the explicit relation with the average payoff (and correlated payoff in equilibrium) would be lost, this line is not pursued.

As it turns out in Theorem 2, the full characterization of the process can be done very simply in terms of Markov chains. For the class of games that we analyze, an equivalent speci cation of the stochastic process can be given in terms of a Markovian random eld[12]. This characterization requires, as a further assumption in our model, that the set of players, , be equipped with a distance (for example in a lattice of Von Neumann or Moore neighbourhoods). Such characterization turns out to be particularly useful if the model is extended by allowing for a countably[13] in nite population of players, all other things being equal. This would make the version of

the model we analyze here under uniform local matching (i.e. for $\mu \equiv \mu_L$), close to that studied in Blume (1993), where the author provides a characterization for an underlying two-player m-by-m symmetric game in terms of an Ising model (for $m = 2$), or a Pott s model for $m > 2$). As pointed out in that paper, complications might arise as the partition function, i.e. the normalizing factor of which in Theorem 2, part (2), may not be summable. Our results for $n = 2$ and for an underlying symmetric game would constitute an analog of Theorems 6.1 and 6.3 in that paper.

It should be clear that, though the speci cation of the model is different, the logic followed in the analysis is very close to that of Kandori, Mailath and Rob (1993). Namely, the equilibrium selection results are obtained in two steps: rst, the asymptotics (over time) of a noisy dynamics is characterized in terms of an ergodic limit distribution expressed in terms of a parameter (σ, in our model), and second, a limit is taken over this parameter to formalize a process of learning by players. Clearly, if the order of the limits is reversed, the logic fails to hold, as so do the ergodicity properties. Besides, the caveat of Bergin and Lipman (1996) applies to this model as well: if the model allows for state-dependent mistakes, for example if σ depends on a (the state) or on ω (the player), and the speed at which each σ reaches in nity differs, then the equilibrium selection result of the Corollary would not necessarily hold.

Perhaps surprisingly, the closest analog of our convergence result is that obtained for an in nite population / random matching model where the dynamics are modelled as replicator dynamics. As Hofbauer and Sigmund (1988) show, for a two-player symmetric potential game (a partnership game in their terminology), the replicator equation is a gradient vector eld, with the mean payoff as potential function in a suitable Shahshahani metric. A full understanding of this analogy warrants future research.

6   Conclusions

The paper provides a starting point for the analysis of the dynamic properties of equilibria for nite n player games, repeatedly played by randomly drawn n-tuples of players from a nite population. The general framework adopted throughout accounts for standard assumptions (as random matching, or local matching) about the random way players are matched, as particular cases. On one hand such generality helps to

19

capture aspects of the articulate way in which information is gathered and transmitted in real life, and is probably more conducive to applications of this kind of theoretical modeling to market structures. On the other hand, this generality is needed also at a theoretical level, if equilibrium convergence and equilibrium selection results that have been the focus of a branch of recent literature, are to be generalized to settings more complex than underlying coordination games.

It emerges from the analysis that the way matching among individuals takes place determines a variety of equilibria that do not correspond to replicated versions of the equilibria of the underlying game. Clearly, the specific characterization of each of them relies on topological properties that might not be obvious. However, in terms of the measure they induce, the focus is to be directed on the set of correlated equilibria of the underlying game, rather than on that of its Nash equilibria. Pursuing this line of research could also be of practical help in understanding the structure of the set of correlated equilibria, beyond the results of Cripps (1993) and Evangelista and Raghavan (1996). Furthermore, the learning process studied in the paper provides a much needed rationale for the notion of correlated equilibrium, that hinges upon the local nature of information gathering. Along these lines it seems appealing to combine a spatial characterization of the model with some consideration of the history of play in terms of past play. A more plausible model would combine these two dimensions (space and time) in the formalization of the learning process. As for the class of games with identical interests, we know (from Monderer and Shapley (1996b)) that learning dynamics like ctitious play converge, and we know from Hart and Mas-Colell (1997) that history *per se* can work as a convexi cation device to support a correlated equilibrium, results could be obtained along these lines.

The convergence and equilibrium selection results of this papers are shown to hold for potential games. Clearly this is only a sufficient condition, that may be replaced by others in particular settings. In the class of 2-player - 3-actions coordination games for example assumptions like the band-wagon property or supermodularity have been shown to work in Kandori and Rob (1993) and (1995) for a game played under uniform population matching. However, if matching is local, these conditions are neither necessary, nor sufficient to rule out cyclic behaviour.

In our setting potential games are of interest because they allow us to derive convergence and equilibrium selection results that do not depend on the particular speci cation of the interactive structure. As payoff functions are identical, these

equilibria are robust to permutations of the roles in the underlying game, and the distinction between *ex-ante* and *ad-interim* interpretation of the equilibrium conditions does not bite. This matters when analyzing convergence and stability properties of the dynamics we study, as well as when studying a noisy version of the dynamics: first because equilibria of the population game are robust to any one player deviation (this plays the same role that strictness of Nash equilibrium plays in the underlying game); second because it guarantees reversibility properties of the noisy dynamics we study. This latter point clearly depends on the specific formalization of the mistakes that affect players choices in the model we use.

It is difficult to provide a convincing motivation (other than technical convenience) for adding mistakes to an otherwise myopically optimal behaviour. However, preliminary results show that the formalization used here can be motivated in terms of limited information on the part of scarcely informed rational players, who make inference on the aggregate play by sampling some observations. Motivation aside, one thing noted in an example in Ianni (1997) is that, for a population game played under local matching and for the underlying game in Example 1 of this paper, this specification of mistakes yields different equilibrium selection results than the one used in many papers that originated from Kandori, Mailath and Rob (1993) and Young (1993). In these papers mistakes do not explicitly depend on expected pay-offs, but only on the ranking between the expected payoff achievable with different actions. It would be interesting to investigate further the dependence of equilibrium selection results on the specification of the process of mistakes. Blume (1995) constitutes a first step in that direction, for an underlying 2 player game, played under uniform population matching. One may conjecture that the two specifications might be related to properties of properness and perfection of the equilibria.

University of Pennsylvania, 2/28/98.

Appendix:

As the result of Proposition 1 can easily be derived from Proposition 2, we invert the order of the proofs.

Proof of Proposition 2

i) The proof extends the result of Proposition 1, part (1.1) in Mailath, Samuelson and Shaked (1997) by allowing players to use mixed strategies. The extension is straight-forward, as it suffices to notice that player ω s expected payoff from choosing $\mathbf{a}_\omega$ is a linear function of ω s mixing probabilities a_ω^i, for $i = 1, 2, ..., \#A_\omega$ and of ω s opponents mixed strategies, $a_{-\omega}^{-i}$:

$$E_\mu[\Pi_\omega(\mathbf{a}_\omega, \mathbf{a}_{-\omega}) \mid \mu] = E_\mu[\sum_{i \in A_\omega} \sum_{-i \in A_{-\omega}} a_\omega^i a_{-\omega}^{-i} \Pi_\omega(a_i, a_{-i}) \mid \mu]$$

Suppose $\mathbf{a}^* \in \Theta(\Gamma)$. Then for all ω for which $a_\omega^i > 0$ and for all $j \neq i$ the following holds:

$$
\begin{aligned}
0 \quad &\leq \quad \sum_{a_{-i}} \sum_{-\omega \,:\, a_{-\omega}^{-i} > 0} \mu(\omega, -\omega) a_\omega^i [\Pi_\omega(a_i, a_{-i}) - \Pi_\omega(a_j, a_{-i})] a_{-\omega}^{-i} \\
&\leq \quad \sum_{\omega \,:\, a_\omega^i > 0} \sum_{a_{-i}} \sum_{-\omega \,:\, a_{-\omega}^{-i} > 0} \mu(\omega, -\omega) a_\omega^i a_{-\omega}^{-i} [\Pi_\omega(a_i, a_{-i}) - \Pi_\omega(a_j, a_{-i})] \\
&\equiv \quad \sum_{a_{-i}} [\Pi_i(a_i, a_{-i}) - \Pi_i(a_j, a_{-i})] \sum_{\omega \,:\, a_\omega^i > 0} \sum_{-\omega \,:\, a_{-\omega}^{-i} > 0} \mu(\omega, -\omega) a_\omega^i a_{-\omega}^{-i} \\
&\equiv \quad \sum_{a_{-i}} [\Pi_i(a_i, a_{-i}) - \Pi_i(a_j, a_{-i})] \; \mu_{\mathbf{a}}(a_i, a_{-i})
\end{aligned}
$$

As this holds for any i and j, and for any ω the above inequalities reproduce exactly those that a correlated equilibrium of G must satisfy. Hence the assert follows.

Explicitly, the probability distribution induced by $\mathbf{a}^*$ over the Cartesian product of the action space, $\mu_{\mathbf{a}^*}(\prod_i A_i)$ has components:

$$\mu_{\mathbf{a}^*}(i_1, i_2, ...i_n) \equiv \sum_{(\omega_1, \omega_2, ..., \omega_n)} a_{\omega_1}^{i_1} a_{\omega_2}^{i_2} .. a_{\omega_n}^{i_n} \mu(\omega_1, \omega_2, ..., \omega_n)$$

for any combination of actions $(i_1 \in A_1, i_2 \in A_2, ...i_n \in A_n)$.

ii) As for the converse of the Proposition, if matching satis es De nition 3, the proof is as straight-forward as in part (1.2) of the quoted paper, hence omitted. ∎

Proof of Proposition 1

From Proposition 2 we know that (for each given m and) for each equilibrium of the m-replica of G, $\mathbf{a}^* \in \Theta(\Gamma_m)$, $\mu_{\mathbf{a}^*}(\prod_i A_i)$ de nes a correlated equilibrium of G. Hence, to prove the assert we only need to show that the correlated equilibrium is the product measure of its marginals:

$$\mu_{\mathbf{a}^*}(i_1, i_2, ...i_n) \equiv \sum_{(\omega_1, \omega_2, ..., \omega_n)} a_{\omega_1}^{i_1} a_{\omega_2}^{i_2} .. a_{\omega_n}^{i_n} \mu(\omega_1, \omega_2, ..., \omega_n) =$$

$$= \sum_{\omega_1 \in \; 1} a_{\omega_1}^{i_1} \sum_{\omega_2 \in \; 2} a_{\omega_2}^{i_2} ... \sum_{\omega_n \in \; n} a_{\omega_n}^{i_n} (m)^{-n}$$

This concludes the proof, once we notice that its corresponding marginals are exactly $\mu(i_l) = (m)^{-1} \sum_{\omega_l \in \; l} a_{\omega_l}^{i_l}$ for all $l = 1, 2, ..., n$. $\blacksquare$

Proof of Theorem 1

(1): It suffices to notice that $\Pi(a_\omega, a_{-\omega}) = \sum_{i_1 \in A_1} ... \sum_{i_n \in A_n} a_{\omega_1}^{i_1} ... a_{\omega_n}^{i_n} \Pi_\omega(i_1, i_2, ...i_n)$. Hence:

$$\Pi_\mu(a) = \sum_{i_1 \in A_1} ... \sum_{i_n \in A_n} \left[\sum_{(\omega_1, \omega_2, ..., \omega_n)} a_{\omega_1}^{i_1} ... a_{\omega_n}^{i_n} \mu(\omega_1, \omega_2, ..., \omega_n) \right] \Pi(i_1, i_2, ...i_n)$$

$$= \sum_{i_1 \in A_1} ... \sum_{i_n \in A_n} \mu_{\mathbf{a}}(i_1, i_2, ...i_n) \Pi(i_1, i_2, ...i_n)$$

$$= E_{\mu_{\mathbf{a}}}[\Pi(i_1, i_2, ...i_n)].$$

(2). a): Let a^0 and a^1 be such that $P(a_1 \mid a_0) > 0$. In order to prove the statement, it suffices to show that, given μ, $\Pi(a^1) - \Pi(a^0) \equiv \Delta\Pi(a^0) \geq 0$. As previously de ned:

$$\Pi_\mu(a) = \frac{1}{n} \sum_{(\omega, -\omega)} \Pi_\omega(a_\omega, a_{-\omega}) \mu(\omega, -\omega)$$

$$= \frac{1}{n} \left[\sum_{(\omega, -\omega)} \mu(\omega, -\omega) \Pi_\omega(a_\omega, a_{-\omega}) + \sum_{(\omega', -\omega') \neq (\omega, -\omega)} \mu(\omega', -\omega') \Pi_{\omega'}(a_{\omega'}, a_{-\omega'}) \right]$$

Hence, if ω is the player who changes action between a^0 and a^1:

$$\Delta\Pi_\mu(a^0) = \frac{1}{n} \left[\sum_{(\omega, -\omega)} \mu(\omega, -\omega) [\Pi_\omega(a_\omega^1, a_{-\omega}) - \Pi_\omega(a_\omega^0, a_{-\omega})] \right] =$$

$$= \frac{1}{n} \left[n \sum_{-\omega} \mu(\omega, -\omega) [\Pi(a_\omega^1, a_{-\omega}) - \Pi(a_\omega^0, a_{-\omega})] \geq 0 \right.$$

as the only transitions that occur with strictly positive probability in $\wp$ are those where ω switches to a best-reply.

(2): b) The first statement is true since absorbing states of the process are all (and only) those states that are strict equilibria of Γ. In any of these, no player can achieve a higher expected payoff by choosing a different action, given the actions of all other players. Hence the average over all n-tuples of players cannot be increased in a single step of the dynamics.

To prove the second statement, notice that for an underlying potential game:

$$
\begin{aligned}
\Pi_\mu(a) &= \frac{1}{n} \sum_{(\omega, -\omega)} \Pi_\omega(a_\omega, a_{-\omega})\mu(\omega, -\omega) \\
&= \frac{1}{n} \sum_\omega \mu(\omega) \sum_{-\omega} \Pi_\omega(a_\omega, a_{-\omega})\mu(-\omega \mid \omega) \\
&= \frac{1}{n} \sum_\omega \mu(\omega) \sum_{a_{-i}} \Pi_i(a_i, a_{-i}) \sum_{-\omega: a_{-\omega}=a_{-i}} \mu(-\omega \mid \omega) \\
&= \frac{1}{n} \sum_\omega \mu(\omega) \sum_{a_{-i}} \Pi(a_i, a_{-i})\mu(a_{-i} \mid a_i)
\end{aligned}
$$

i.e. it is a weighted average of smooth functions of the form $\sum_{a_{-i}} \Pi(a_i, a_{-i})\mu(a_{-i} \mid a_i)$. Hence $\Pi_\mu(a)$ is maximized if and only if each single function is maximized. By Definition 4, for all ω s and for a given μ, any a_i adopted in equilibrium is a maximizer for a given a_{-i}. As Γ has no ties, for each player the best-reply is unique. Hence a is absorbing. ∎

Proof of Corollary 1

It suffices to notice that the stochastic process under MBR is an absorbing Markov chain over the state space $\prod_\omega a_\omega$. Hence, no matter where the process starts, the probability after t steps that the process is in an ergodic state tends to 1 as t tends to infinity. The function $\Pi_\mu(a_t)$ behaves monotonically over t and is locally maximized at any $a \in \theta(\Gamma) \subseteq \Theta(\Gamma)$, that we know from Proposition 2 defines a correlated equilibrium of G. ∎

Proof of Theorem 2

(1): Clearly the process $\wp^\sigma$ ranges over the same state space as $\wp$ does, and from Definition 1, it is clear that the perturbation does not alter the Markovian properties of the process. Also, from the same Definition, each player adopts a mixed strategy that has full support. As a result any transition between any two states

that differ only by player ω s action takes place with strictly positive probability. Hence any state a communicates with any other state a^1 in a nite number of steps. In the terminology introduced at the beginning of Section 4, $\wp^\sigma$ contains a unique ergodic set that includes all states in the state-space. As a standard result (see, for example, Kemeny and Snell (1976), Theorem 4.1.6), $\wp^\sigma$ admits a unique limit distribution, i.e. a unique probability vector $[P_\sigma(a), \quad a \in \mathcal{A}]$ such that $P_\sigma(a) = \sum_{a^1 \in \mathcal{A}} P_\sigma(a^1) P^\sigma(a \mid a^1)$. Besides, the latter is ergodic, in that for any arbitrary probability vector $[q(a), \quad a \in \mathcal{A}]$, $\lim_{t \to \infty} q(a)(\wp^\sigma)^t = P_\sigma(a)$.

As for the second statement, suppose ω is adopting a suboptimal action a_{ω_j} in state a, and his best-reply, given state a, is action a_{ω_i}, that he adopts in state a^1. Under $\wp^\sigma$ the transition between a and a^1 will occur with probability:

$$P^\sigma(a^1 \mid a) = \frac{1}{\sharp} \frac{\exp[\,\sigma E_\mu[\Pi_\omega(a_{\omega_i}, a_{-\omega}^{t-1}) \mid \mu]]}{\sum_{i \in A_\omega} \exp[\sigma\, E_\mu[\Pi_\omega(a_{\omega_i}, a_{-\omega}^{t-1}) \mid \mu]]}$$

Hence $\lim_{\sigma \to \infty} P^\sigma(a^1 \mid a) = 1 = P(a^1 \mid a)$. Accordingly, $\lim_{\sigma \to \infty} P^\sigma(a \mid a^1) = 0 = P(a \mid a^1)$. As ω, i and j are chosen arbitrarily, the assertion is proved.

(2): As the underlying game is a potential game, we are able to explicitly derive its limit distribution by simply noticing that the chain is reversible, i.e. the probability vector $[P_\sigma(a), \quad a \in \mathcal{A}]$ is such that $P_\sigma(a) P(a^1 \mid a) = P_\sigma(a^1) P(a \mid a^1)$ for all $a, a^1 \in \mathcal{A}$ (see, for example, Liggett (1985), Proposition 5.7). Again, suppose ω is adopting action a_{ω_j} in state a, and action a_{ω_i} in a^1. Then the following holds:

$$
\begin{aligned}
\frac{P_\sigma(a^1)}{P_\sigma(a)} &= \exp[\sigma \Delta \Pi_\mu(a)] \\
&= \exp[\sigma \sum_{-\omega} \mu(\omega, -\omega)[\Pi(a_{\omega_i}, a_{-\omega}) - \Pi(a_{\omega_j}, a_{-\omega})]] \\
&= \left[\frac{1}{\sharp} \frac{\exp[\,\sigma \sum_{-\omega} \mu(\omega, -\omega)\Pi(a_{\omega_i}, a_{-\omega})}{\sum_{i \in A_\omega} \exp[\sigma \sum_{-\omega} \mu(\omega, -\omega)\Pi(a_{\omega_i}, a_{-\omega})]} \right] \cdot \\
&\quad \cdot \left[\frac{1}{\sharp} \frac{\exp[\,\sigma \sum_{-\omega} \mu(\omega, -\omega)\Pi(a_{\omega_j}, a_{-\omega})]}{\sum_{i \in A_\omega} \exp[\sigma \sum_{-\omega} \mu(\omega, -\omega)\Pi(a_{\omega_i}, a_{-\omega})]\,]} \right]^{-1} \\
&= \frac{P_\sigma(a^1 \mid a)}{P_\sigma(a \mid a^1)}
\end{aligned}
$$

[14] as the relative advantage, for player ω, of action i with respect to action j is exactly equal to the ratio between the one step transition probabilities. ∎

Proof of Corollary 2

As G has identical interests, a degenerate probability distribution that has mass on $i = (i_1, i_2, ..., i_n)$ is a (strict) Nash and correlated equilibrium of G. Hence (from Proposition 2, part ii)) we can construct a pro le of actions a in a population game Γ, for which $\mu_a(i_1, i_2, ..., i_n) = 1$ and $\Pi_\mu(a) = E_{\mu_a}[\Pi(i_1, i_2, ..., i_n)] = \Pi(i_1, i_2, ..., i_n)$. As $\Pi_\mu(a) > \Pi_\mu(a')$ for all $a' \neq a$, the remaining part of the statement follows from part (2) of Theorem 2, once we notice that $P_\sigma(a)$ is continuous in σ. ∎

Notes

[1]Notational conventions we keep throughout are: the use of greek letters (other than μ) to index players and latin letters to index actions; the symbol $\sharp W$ to denote the cardinality of the nite set W and $\phi(W)$ to denote any permutation of its elements; $E_* \bullet$ to denote the expected value of the random variable $\bullet$ taken with respect to the measure $*$. Furthermore, we abide to the convention of denoting a nite set $(w_1, w_2, ..., w_n)$ as (w_i, w_{-i}) where the subscript $-i$ refers to all indices other than i in the set.

[2]A companion paper of this, Ianni (1997), studies an analog model where the random matching does not satisfy the above de nition.

[3]As in Harsanyi (1973), Strict simply means that, given μ, each player has a unique best response to her rival s strategies.

[4]This assumption is only introduced for convenience, as, in general, there is no guarantee that a *complete* matching (involving $(n)^{-1}N$ groups) exists. As it might sound a little bit odd that a player learns, even though she does not play the game in every period, we notice that, for $n = 2$, an equivalent speci cation of the model, where *all* players play the game at *each* time t, could be obtained by assuming that a) μ satis es De nition 3 and b) the induced graph $\mathcal{G}$ satis es a condition known as Hall s condition. See Biggs (1990) for a graph-characterization of existence conditions. Alternatively, one could plausibly postulate that, within each time period, very many matchings take place, so that the expected payoff to a player can be taken to correspond to the average actual payoff from that round of interactions.

[5]A motivation for postulating this sort of asynchronous updating is provided in Anderlini and Ianni (1996a) in terms of *noise at the margin*. The idea is that strategy changes are costly; as a result whenever a learning rule prescribes an action that is different from the action previously adopted, a player does not necessarily follow the prescription and with some positive probability, does not change her action. Hence the dynamics of the system reproduce any of the possible paths of asynchronous updating.

[6]This is obvious if we re-write the rule as follows:

$$\frac{\ln a_\omega^i}{\ln a_\omega^j} = \sigma [E_\mu[\Pi_\omega(a_{\omega_i} a_{-\omega}) \mid \mu] - E_\mu[\Pi_\omega(a_{\omega_j}, a_{-\omega}) \mid \mu]]$$

A behavioural speci cation of this rule could be given in terms of conditional logit speci cation, as in McFadden (1974). McKelvey and Palfrey (1995) t an analog model to a variety of experimental data sets. For our purposes, the idea we want to convey is that, for a player wanting to go to the 20th oor by elevator, the probability with which she presses wrong buttons is decreasing in the number of levels below oor 20.

[7]To see this, consider the simplest case where $n = N = 2$ and $G = (A, A, \Pi_1, \Pi_2)$, where we read $\Pi_i : A \times A \to \Re$, $i = 1, 2$ as the payoff matrices. Unless the game is symmetric, which is the case if $\Pi_2' = \Pi_1$, the set of equilibria characterized by De nition 4, would correspond to the set of

Nash-equilibria of a different game, namely $\bar{G} = (A, A, \frac{1}{2}(\Pi_1 + \Pi_2'), \frac{1}{2}(\Pi_1 + \Pi_2'))$.

[8]$G = (\{A_i\}_{i=1}^n, \{\Pi_i\}_{i=1}^n)$ is a symmetric n player-normal form game, if $A_i = A$ for all i s and payoff functions $\{\Pi_i\}_{i=1}^n : \prod\limits_{i=1}^n A_i \to \Re$ are such that $\Pi_i(a_i, a_{-i}) = \Pi_i(a_i, a_{\phi(-i)}) = \Pi(a_i, a_{-i})$ for all i s.

[9]A correlated equilibrium ψ is symmetric if $\psi_{i_1 i_2, ... i_n} = \psi_{\phi(i_1 i_2, ... i_n)}$ for all permutations $\phi(i_1 i_2, ... i_n)$ of the indices $(i_1 i_2, ... i_n)$.

[10]Recall that two states belong to the same equivalence class if they communicate , i.e. if the process can go from one state to the other. The resulting partial ordering shows the possible directions in which the process can proceed. The minimal elements of the partial ordering of equivalence classes are called ergodic sets, i.e. sets that, once entered, cannot be left by the dynamics. Ergodic sets that contain only one element are called absorbing states.

[11]Note that this property is generic either in the payoff space, or in the speci cation of the matching.

[12]In Ianni (1996), Section 6.3.2 this is done for the case of a two-player underlying game.

[13]If the population is nite, the characterization in terms of Markov elds (instead of Markov chains) would only provide a different interpretation of the model, at the cost of technical complexity. Though used to argue in favour of technicality, such remark also appears as Remark 6 in Allen (1982), where the author studies a model of stochastic technological diffusion.

[14]The condition $\frac{P(a^1)}{P(a)} = \frac{P(a^1|a)}{P(a|a^1)}$ is known as Detailed Balance Condition.