Evolutionary Game Theory: Why Equilibrium and Which Equilibrium

Hamid Sabourian*and Wei-Torng Juang†

November 5, 2007

1 Introduction

Two questions central to the foundations of game theory are (Samuelson 2002):

(i) *Why equilibrium*? Should we expect Nash equilibrium play (players choosing best response strategies to the choices of others)?

(ii) If so, *which equilibrium*? Which of the many Nash equilibria that arise in most games should we expect?

To address the first question the classical game theory approach employs assumptions that agents are rational and they all have common knowledge of such rationality and beliefs. For a variety of reasons there has been wide dissatisfaction with such an approach. One major concern has to do with the plausibility and necessity of rationality and the common knowledge of rationality and beliefs. Are players fully rational, having unbounded computing ability and never making mistakes? Clearly, they are not. Moreover, the assumption of common (knowledge of) beliefs begs another question as to how players come to have common (knowledge of) beliefs. As we shall see in later discussion, under a very broad range of situations, a system may reach or converge to an equilibrium even when players are not fully rational (boundedly rational) and are not well-informed. As a result, it may be, as

*School of Economics, Mathematics and Statistics, Birkbeck College Univerity of London and Faculty of Economics, University of Cambridge

†Institute of Economics, Academia Sinica, Taipei 115, Taiwan

Samuelson (1997) points out, that an equilibrium appears not because agents are rational, "but rather agents appear rational because an equilibrium has been reached."

The classical approach has also not been able to provide a satisfactory solution to the second question on selection among multiple equilibria that commonly arise. The equilibrium selection problem is particularly acute when the game has many strict Nash equilibria. This is because such equilibria are robust with respect to small perturbations and as a result they survive most refinements of Nash equilibrium proposed in the classical approach (one possible exception is Harsanyi and Selten 1988). Such dissatisfaction with the classical approach has spurred the emergence of the evolutionary approach.

Since Maynard Smith and Price (1973) and Maynard Smith (1982) proposed a new way of thinking about the evolution of animal behavior, sex and population genetics, evolutionary game theory has enjoyed a vibrant development. In the last 20 years, game theorists (longer for mathematical biologists) have addressed these questions in the context of evolutionary/learning frameworks. The hope has been that such an evolutionary approach can provide: first, a justification for Nash equilibrium (by coordinating beliefs so that they are consistent with strategies); second, new insights into which equilibrium will be played.

The biological idea behind the evolutionary approach is that more successful players/strategies will have a better chance to survive or prosper in the population. One typical mechanism consistent with the concept is replicator dynamics in which the population frequency for each strategy changes as a function of the difference between the performance of the strategy and that of the population average. Replicator dynamics are too specific to capture complex human behaviour, however. To better fit human environments, other dynamics such as monotonic dynamics that include replicator dynamics as a special case have been proposed (see below for the definition of the different dynamics). Such selection dynamics in human setting represent the process of learning and imitation (the adjustment rules) of the individuals. We shall give a detailed discussion on these dynamics in the next section.

In this paper, we would like to survey some of the literature on the two questions (why equilibrium and which equilibrium) within the *dynamic* evolutionary framework, paying special attention to the behavioural rules (the type space) in each population and their possible evolution. What follows will not be an exhaustive survey. There are some excellent textbooks and surveys on evolutionary game theory (for example Fudenberg and Levine 1998, Hofbauer and Sigmund 1998 and 2003, Samuelson 1997, Vega-Redondo 1996 and 2003, Weibull 1995, Young 1998).

In the first part, we shall discuss the conditions under which a deterministic evolutionary dynamics process converges to a Nash equilibrium and potential failure for such convergence. The dynamics in these models are usually determined by some fixed behavioural rule that all agents follow (for example reinforcement learning, imitation, best response, fictitious play).

We focus on Nash equilibrium because firstly it is the central equilibrium concept in non-cooperative games and secondly, in evolutionary models, irrespective of the particular dynamics (adjustment rules agents have), there is strong justification for Nash equilibria in terms of the steady states of these systems and their local dynamics. On the other hand, as Hofbauer and Sigmund (2003) point out, "a central result on general 'adjustment dynamics' shows that every reasonable adaptation process will fail, for some games, to lead to a Nash equilibrium" globally. We shall briefly outline Hart and Mas-Collel's (2003) explanation for the lack of global convergence to Nash equilibria in such set-ups. However, almost all deterministic evolutionary models that have looked at the issue of convergence to Nash have considered dynamics in which all agents adopt one fixed behavioural rule. We shall also consider the question of convergence to Nash when multiple rules coexist, agents can adopt different rules and the selection is at the level of rules. Here, we will argue that convergence to Nash may be easier to establish if the set of rules allowed are sufficiently rich.

In the second part of this paper, we shall address the "which equilibrium" question by introducing small perpetual random shocks to the evolutionary dynamics in finite populations. This approach has been very successful in selecting between different strict Nash equilibria in specific applications. However, with this approach, it turns out that the specific equilibria selected depend very much on the adjustment rules allowed. For example, best response dynamics seems to select the risk dominant equilibrium in a 2×2 coordination game (Kandori, Malaith and Rob 1993 and Young 1993) while the imitation one seems to favour the efficient equilibrium (Robson and Vega-Redondo 1996). Therefore, to answer the "which equilibrium" question, we need to explore the dynamics in which multiple rules coexist and compete with each other. This is important not only for addressing the question of equilibrium selection but also because in a world with heterogeneous individuals it is not plausible that a single adjustment rule can capture all important properties of human behaviour. As Young (1998, p 29) points out, "we would guess that people adapt their adaptive behavior according to how they classify a situation (competitive or cooperative, for example) and reason from personal experience as well as from knowledge of others' experiences in analogous situations. In other words, learning can be very complicated indeed." For the most of this section we compare the issue of equilibrium selection in the context of specific games for the different dynamics/rules. We then conclude the paper by considering some specific models with small perpetual random shocks in which multiple rules coexist, and in particular we discuss some of the justification provided for some specific rules such as the imitation one.

2 Why Equilibrium? Deterministic Evolutionary Models

The standard evolutionary game with discrete or continuous time consists of n large (often infinite) populations of myopic and unsophisticated agents playing some underlying one-shot (normal form) N-player game $G = (A_i, \pi_i)_{i=1}^n$ infinitely often, where, for each player $i = 1, .., n$, A_i is a finite set of (pure) strategies (henceforth also called actions) and $\pi_i : \times_i A_i \rightarrow \mathbf{R}$ is the payoff function at each date.

Each population i consists of a set of (countable) types. In the most of the literature a type in population i is usually identified with an action $a_i \in A_i$ in the one-shot game G. That is, he is programmed to *always* execute a_i.

Before describing the dynamics, we introduce some standard definitions. Let $A \equiv \times_i A_i$ be the set of action profiles for the one-shot game. We adopt the convention that for any profile $y = (y_1, .., y_n)$, y_{-i} refers to y without its i-th component. Next, for any strategy profile $a \in A$, let $B_i(a) = \{a'_i \in A_i \mid \pi_i(a'_i, a_{-i}) \geq \pi_i(a''_i, a_{-i}) \; \forall a''_i \in A_i\}$ be the set of best responses for i given a. Then a profile of strategies $a^* \in A$ is a Nash equilibrium if $a^*_i \in B_i(a^*)$ for all i. A Nash equilibrium $a^* \in A$ is strict if $B_i(a^*)$ is unique for all i. Similarly, we can define a Nash equilibrium for the game G in the space of mixed strategies.

Dynamics. At each date t each member of population i is randomly matched with one member of every other population to play the game G and receives a payoff depending on his action and the actions taken at t by those with whom he is matched.

Time could be discrete and with $t = 0, 1, 2, ...$ or continuous where $t \in [0, \infty)$.[1] At each date t the state of the system is the proportion (probability) of each type/strategy/action in each population; thus it can be described by n probability distributions $P_i^t = \{P_i^t(a_i)\}_{a_i \in A_i}$ where $P_i^t(a_i)$ denotes the proportion (probability) of strategy a_i in population i at date t. The state of the system evolves according to some dynamics describing selection (and mutation in models with randomness). In the discrete case, the dynamics (with no mutation) is given by

$$P^{t+1} = \Gamma(P^t) \tag{1}$$

where $P^t = (P_1^t, .., P_n^t)$. With continuous time the law of motion is described by

$$\dot{P^t} = \Gamma(P^t) \tag{2}$$

The selection dynamics for the case of discrete time are such that at each $t + 1$ the proportion of each strategy changes as a function of how well the type has done on average at t, in terms of payoff, relative to the payoff of the other strategies. Note that the proportions at each date $t + 1$ has a stringent Markovian property of depending only on the environment (payoffs) in the previous period t and not on outcomes prior to t. The most canonical model of evolutionary dynamics that embodies the idea of Darwinian selection is the replicator dynamics (RD). Here, the growth of each strategy in any population i is assumed to be an increasing function of its average (expected) payoff minus population i's average payoff. Thus, in the case of continuous time the replicator dynamics is described by

$$\dot{P_i^t} = P_i^t(a_i)(E\pi_i^t(a_i) - E\overline{\pi}_i^t) \tag{3}$$

where, with some abuse of notations, $E\pi_i^t(a_i)$ denotes the average (expected) payoff to strategy a_i at time t and $E\overline{\pi}_i^t = \Sigma_{a_i' \in A_i} E\pi_i^t(a_i')P_i^t(a_i')$ is the average payoff in population i at date t. Note that since each population is 'large' number and there is no aggregate uncertainty, $E\pi_i^t(a_i)$ is simply $\Sigma_{a_{-i} \in A_{-i}} \pi_i(a_i, a_{-i})P_{-i}^t(a_{-i})$. We can also describe a discrete analogue for RD.

[1] Since any dynamic evolutionary system involves playing a one-shot game infinitely often, such models, by definition, have an infinite aspect about them. In addition to this infinity of the time horizon, in this survey paper we also consider cases in which the size of the population, the set of types and/or strategies are infinite.

There are several properties of RD to note. First there is no birth of new types or death of existing types in finite time: $P_i^0(a_i) > 0 \Leftrightarrow P_i^t(a_i) > 0, \forall t$. Second, the dynamics is determined by relative payoffs in the sense that types that do better in terms of average payoffs grow faster than those that do less well.

In the biological literature, the selection dynamics represents reproduction based on fitness. In fact, if the evolutionary model is such that agents live for only one period, reproduction is asexual, $E\pi_i^t(a_i)$ represents the number of offspring of type a_i, each offspring of a_i type does exactly the same as a_i and there is no mutation, then RD, defined by (3), precisely describes the dynamical system for the continuous time. In a non-biological (e.g. economics, political science, social) setting, the selection represents learning and imitation. Ideally, one would like to build up the selection dynamics from a precise and acceptable theory of how individual players switch between different actions/types. The evolutionary approach often avoids such a difficult task and instead it places assumptions directly on the selection dynamics and hopes that these properties are general enough to include processes produced by a variety of learning and imitation theories.(See Fudenberg and Levine 1998 and Young 2004).

Nachbar (1990), Bjornerstedt and Weibull (1996), Schlag (1998), Borgers and Sarin (1997) and others show that one may also be able to provide a justification for RD based on imitation and learning. In general, however, these learning/imitation models are too specific and there is often no compelling reason for adopting such a precise imitation/learning specifications. Thus, RD framework simply cannot capture the complexity of learning and imitation in social and economic contexts and one needs a broader type of dynamics than RD.

Fortunately, it turns out in the deterministic evolutionary set-ups a large number of dynamics share many essential features. One very large class of dynamics that includes RD as a special case is defined by the property that strategies that do better in terms of payoffs grow faster relative to those that do less well. More precisely, a selection evolutionary dynamics is called monotonic if at every date t the growth rates of the different strategies are ranked by their average payoffs:

$$G_i^t(a_i') > G_i^t(a_i'') \ \Leftrightarrow E\pi_i^t(a_i') > E\pi_i^t(a_i'') \ \forall a_i', a_i'' \in A_i \tag{4}$$

where $G_i^t(a_i)$ is the growth rate of strategy a_i at t, and thus satisfies $G_i^t(a_i)P_i^t(a_i) =$

$P_i^{t+1}(a_i) - P_i^t(a_i)$ in the discrete time model and $G_i^t(a_i)P_i^t(a_i) = \dot{P}_i^t(a_i)$ when time is continuous.

Monotonic dynamics is consistent with fairly general learning and imitation models. Consider for example an imitation dynamics in which at any date a player has the chance to sample another player at random and adopt his behaviour; and the rate at which a player of a certain type switches behaviour to another type depends on the current average payoffs of the two types. Monotonic dynamics is consistent with such imitation models as long as the rate at which players switch depends positively on the success of the sampled type's payoff relative one's own.

Often in describing monotonic dynamics it is assumed that the dynamics satisfy two further properties. First it is assumed that the function describing the adjustment, Γ is continuous (Lipschitz continuous for the continuous time case) and second there is no birth of new types or death of existing types: $P_i^0(a_i) > 0 \Leftrightarrow P_i^t(a_i) > 0, \forall t$. The first condition is for technical reasons; however, note that it excludes certain type of discontinuous imitation dynamics in which players switch their behaviour if the sampled behaviour has a strictly better payoff. The no birth assumption is however very much consistent with the imitation story because imitative behaviour has the property that players only sample amongst existing types; as a result only types that existed in the past survive. The assumption that types are not created or destroyed can be relaxed in some cases; however, in order to simplify the discussion we shall also assume, unless stated otherwise, that monotonic dynamics has this property.

The properties of monotonic dynamics, discussed in the next section, broadly holds, for even larger classes of dynamics. For example, one generalisation of monotonic dynamics is *weakly positive dynamics*: for any population, if there exists a type that has a higher payoff than the average of the population, then some type must have a positive growth.

Another type of dynamics is the (myopic) best response dynamics (Matsui, 1992). Here, at any date a fixed proportion of each population chooses a best response strategy (one that maximises the player's one-period payoff) given the average strategy in every other population in the previous period. Another important variation of best response dynamics is the fictitious play dynamics where each player chooses a best response to the historical frequency of past plays. Such dynamics are often described in the context of finite populations. They are clearly inconsistent with monotonic dynamics because they require players to switch to best responses (when the opportunity arises) whereas monotonic dynamics require the ranking of growth rates of *all* strategies according to payoffs. Moreover, these dynamics may require discontinuous shift in behaviour. Nevertheless, it turns out that many of the qualitative features of these dynamics in terms of Nash are similar to those of the monotonic dynamics in the deterministic settings.

Before turning to properties of the dynamics there are three further points to note. First, the literature on evolutionary game theory often considers single population models in which all agents play a *symmetric* one-shot game repeatedly. Here, we described a more general set-up with multiple populations to allow for the possibility that the underlying game G may be asymmetric. (See Weibull 1995 for some discussion of the difficulties with multi-population models that are not present in single population models.) For some specific analysis, to simplify the discussions, we shall at times limit ourselves to single population models. Second, much of the literature on deterministic evolutionary dynamics considers continuous time. The results for the discrete and continuous dynamics sometimes differ (the former dynamics is often less "nice"). For ease of exposition, we shall at times limit the discussion to one kind of time dynamics. However, often there are analogous results for the other kind of time dynamics. Finally, the models in this section are mainly deterministic while those in the "which equilibrium" section are stochastic. The reason is that deterministic environments suffice for our discussion on the "why equilibrium" part while introducing mutations here only complicates unnecessarily without adding much insight. In contrast, in the "which equilibrium" part in the next section, to select between equilibria, we introduce mutations or noises to perturb a dynamic system in such a way that the underlying system may drift among each equilibrium; this then allows us to examine at which equilibrium the system will spend most of the time.

2.1 Convergence and Stability in Standard Deterministic Models

Folk Theorem Before describing the properties of deterministic evolutionary models, we need to introduce two *local* stability concepts that are common in the literature: Lyapunov stability and asymptotic stability. A stationary state is *Lyapunov stable* if small perturbations do not result in dynamics that takes the system far from it. A stationary state is *asymptotically stable* if it is Lyapunov stable and there exists a neighbourhood of it such that for any initial state belonging to this neighbourhood the path converges to it. When the underlying game G is finite (A_i is finite) and a type refers to a pure action a_i in G we have the following Folk Theorem of the evolutionary game theory.

Theorem 1 *(Hofbauer and Sigmund 1998) For any 'reasonable dynamics':*

1. *Any (pure or mixed) Nash equilibrium of G is a stationary state of the evolutionary dynamics.*[2]

2. *Any Lyopunov stable stationary state is a Nash equilibrium (possibly mixed).*

3. *A strict Nash equilibrium is an asymptotically stable stationary state.*

4. *The limit of any interior convergent path is a Nash equilibrium (possibly mixed).*

There are a few remarks concerning this theorem. First, the term 'reasonable dynamics' refers to most of the dynamics found in the literature. However, for the purpose of illustrating the ideas here, define it to be dynamics that are continuous and monotonic (with discrete or continuous time).

Second, at a stationary state of such a dynamic all strategies played by various members must give the same payoffs; since at any Nash equilibrium any strategy that is played with a positive probability must be optimal (best response) it follows that any Nash equilibrium is a stationary state (condition 1 in the Theorem). However a stationary state need not be a Nash equilibrium in models with no birth (as in the case with imitation) because there may be a better strategy that is not played and hence its population may not grow. However, a Lyapunov stable stationary state must be a Nash equilibrium because any superior strategy that is not played in the stationary state will be played once a perturbation occurs; once this occurs the superior strategy will grow leading the system away from the stationary state (condition 2 in the Theorem).

[2] Note that although here agents do not randomise and choose pure strategies, the proportion of each population that take different strategies (actions) corresponds to a mixed strategy.

Third, if a state corresponds to a strict Nash equilibrium then it must be asymptotically stable (condition 3 in the Theorem) because in the case of a strict Nash equilibrium there exists a neighbourhood of the equilibrium such that the equilibrium strategy for each player is the unique best strategy .

Fourth, even when the dynamics are continuous and monotonic there is no guarantee that such dynamic will converge. If the path does converge, however, then the limiting state of this convergent path is a Nash equilibrium (condition 4 in the Theorem); otherwise in the limit of this path there is some strategy that does better than others and hence must grow, but this contradicts the path being convergent.

Fifth, note that Nash equilibrium in the Folk Theorem can equally refer to mixed ones even though no agent randomises. This is simply a reflection of the fact that in an evolutionary setting the state of the system refers to a probability distribution over the pure strategies.

Global Convergence The above arguments that an outcome is a Nash equilibrium if it is stable, or a Nash equilibrium is the limit of any convergent path, or the path converges to a strict Nash equilibrium if the process starts from states that are sufficiently close to that equilibrium, provide a somewhat qualified answer to "*why equilibrium*". We may want a stronger result of the form that an evolutionary dynamics must produce *global* convergence to a Nash equilibrium. However, starting with Shapely's (1964) early example on fictitious play, this stronger result that the evolutionary dynamics converges to Nash globally from any (interior) initial state is missing. Indeed, the dynamic trajectories of evolutionary processes in general do not converge; periodic cycles as well as all the other complexities of arbitrary dynamical systems (limit cycles and chaos) are all possible.

One simple example for which the unique Nash equilibrium may not be even asymptotically stable is the 2-player zero-sum generalised "Rock-Paper-Scissors" game in which each player has 3 strategies R, S, P and the payoff for the row player is described by the following matrices:

$$\begin{bmatrix} 1 & 2+\epsilon & 0 \\ 0 & 1 & 2+\epsilon \\ 2+\epsilon & 0 & 1 \end{bmatrix}$$

where ϵ is some real number. This game has a unique mixed Nash equilibrium μ^* in which each strategy is chosen with a probability 1/3. The simplest evolutionary dynamics to consider is a single population RD with continuous time. Here it is not difficult to verify that at any t the derivative of the product of the probabilities of the three strategies, $P^t(R)P^t(S)P^t(P)$, is positive (negative,zero) if ϵ is positive (negative, zero). Thus, depending on the value of ϵ this product term is always increasing, constant or decreasing. But this implies the following.

- For $\epsilon > 0$ all trajectories from all interior initial state spiral inwards towards the unique Nash equilibrium μ^*; thus the dynamics is globally stable.

- For $\epsilon = 0$ all paths are cycles; thus the Nash equilibrium is Lyapunov stable but not asymptotically stable.

- for $\epsilon < 0$ all trajectories from all interior initial state except μ^* spiral outwards; thus the Nash equilibrium is unstable.[3]

The above example is of course very specific. However, the possibility of cycles, global instability and complex dynamics is even more prevalent when there are multiple populations, time is discrete, each player has a large number of strategies or other standard dynamics are considered.

Hart and Mas-Collel (2003), hence called HM, try to provide an answer for why it is difficult to formulate sensible dynamics that always guarantee (global) convergence to a Nash equilibrium. They consider an adaptive stationary dynamics (not strictly a standard evolutionary model) with a finite population in which the adjustment in a player's strategy does not depend on the payoff function of other players (it may depend on the other player's strategies and his own payoff function). They call this dynamics uncoupled and claim that any sensible heuristic dynamics must have this property (most standard dynamics based on agents with limited rationality found in the literature, such as best-reply, fictitious play, better reply, payoff-improving, monotonic are all uncoupled). They show that stationary uncoupled dynamics cannot be guaranteed to converge to a Nash equilibrium in a deterministic setting with continuous time.[4] They therefore conclude that the lack of con-

[3] See Weissing (1991) for detailed discussion on the dynamics of more generlaised "Rock-Paper-Scissors" games.

[4] There exist uncoupled dynamics that are not convergent to Nash equilibria but are close to them most of the time; see Foster and Young (2002).

vergence is due to the informational requirement of uncoupledness which precludes too much coordination of behaviour amongst agents.[5]

To illustrate the nature of HM's result consider the general dynamics for the continuous time described in (2). The precise dynamics clearly depend on the payoff functions and can thus be rewritten, with some abuse of notation, as

$$\dot{P}_i^t = \Gamma_i(P^t, \pi) \text{ for all } i = 1, .., n$$

HM call the above dynamic system uncoupled if $\dot{P}_i^t$ does not depend on π_j for all $j \neq i$. Thus $\Gamma_i(P^t, \pi) = \Gamma_i(P^t, \pi')$ for any two profile of payoffs π and π' such that $\pi_i = \pi'_i$.

To facilitate the analysis they restrict the dynamics $\Gamma = (\Gamma_{1,..}, \Gamma_n)$ to be C^1. Suppose further that the dynamics is hyperbolic (the eigenvalues corresponding to the Jacobian of Γ have non-zero real parts), so that it behaves locally like a linear system $\dot{P}^t = D\Gamma(P^t, \pi)P^t$, where for any P, $D\Gamma(P, \pi)$ is the Jacobian matrix of $\Gamma(., \pi)$ computed at P. Then (asymptotic) stability imposes conditions on the Jacobian matrix $D\Gamma$. Uncoupledness also imposes conditions on the Γ. Unfortunately, these two sets of conditions may not be consistent for many games.

One example provided by HM to illustrate such inconsistency is a family of three player games in which each player has two strategies, and the payoffs are given by the following two matrices:

$$\begin{bmatrix} 0,\, 0,\, 0 & \epsilon^1, 1, 0 \\ 1,\, 0,\, \epsilon^3 & 0{,}1,\, \epsilon^3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0,\, \epsilon^2,\, 1 & \epsilon^1, 1, 0 \\ 1,\, \epsilon^2,\, 0 & 0,\, 0,\, 0 \end{bmatrix},$$

where all ϵ^i are some numbers in a small neigbourhood of 1, and player 1 chooses the row, player 2 the column and player 3 the matrix. (Thus, the family of games considered is defined by the above two matrices where for each i, $\epsilon^i \in (1-\eta, 1+\eta)$ for some small $\eta > 0$.)

Denote the game (the profile of payoff functions) when $\epsilon^i = 1$ for all i by π^*. Next note that for any ϵ^i, each such game in this family has a unique Nash equilibrium; in particular the Nash equilibrium of the game π^* is given by each player randomising with probability half.

[5] Hart and Mas-Collel (2000, 2001) and others have studied specific learning processes that allow stochastic moves (regret-matching) that always converge to a correlated equilibrium. They argue that learning from past history induces correlation in behaviour and therefore one cannot expect dynamic processes of this kind to induce to Nash equilibrium. See also Young (2004) for further discussion of these issues.

Next, consider any dynamics Γ that is uncoupled. The state of the dynamics of this game is described by $P_1, P_2, P_3 \in [0,1]$ denoting respectively the probability of top row, left column and left matrix. Now, by way of contradiction suppose that the dynamics converges to Nash in this family of games. Then since $(1/2, 1/2, 1/2)$ is the Nash equilibrium when $\epsilon = 1$ it must hold that

$$\Gamma_1(1/2, 1/2, 1/2, \pi^*) = 0. \tag{5}$$

Now for any a close to $1/2$ consider another game π^a where $\epsilon^1 = \epsilon^2 = 1$ and $\epsilon^3 = \frac{a}{1-a}$; thus the payoffs of players 1 and 2 are the same as in π^* and the payoff of 3 is different. By simple calculation one can show that the state $(a, 1/2, 1/2)$ is the unique Nash equilibrium of π^a. Again since the dynamics is convergent it must be that $\Gamma(a, 1/2, 1/2, \pi^a) = 0$. But since the dynamics is uncoupled and player 1 has the same payoff in π^a and π^*, it follows that

$$\Gamma_1(a, 1/2, 1/2, \pi^*) = 0. \tag{6}$$

Since (6) holds for any a close to $1/2$ it follows from (5) that $\partial\Gamma_1/\partial P_1 = 0$ at $(1/2, 1/2, 1/2, \pi^*)$. By the same argument one can show that entire diagonal of the Jacobian of function $D\Gamma(., \pi^*)$ vanishes at $(1/2, 1/2, 1/2)$. Now if the dynamics is hyperbolic this implies that the Jacobian has an eigenvalue with positive real part; but then we have a contradiction to the claim that in the game π^* the dynamics converges ($(1/2, 1/2, 1/2)$ is asymptotically stable).

The negative results on convergence are for an arbitrary set of games. There are, however, some noteworthy families of games that do have stable equilibria (at least with continuous time) for some of the standard (uncoupled) dynamics such as RD, monotonic, best response, fictitious play (see Hofbauer, 2000 and Sandholm, 2007).

One such family of games are those that are strictly dominance solvable (games in which there is unique solution to iterative deletion of strictly dominated strategies). Here, with monotonic dynamics the strictly dominance solution is globally stable for any initial interior state. This is simply because with monotonic dynamics it can be shown that the share of any pure strategy that does not survive iterative deletion of strictly dominated strategies goes to zero asymptotically. The basic idea for this result is that if a pure strategy is strictly dominated by another pure strategy then its growth rate is always less than the other and hence its share must go to zero.[6] By

[6]This result does not hold for mixed strategies. Even for the case of continuous time

the same argument, the result extends to iterative deletion of strictly dominated strategies: Once the shares of the strictly dominated strategies are close to zero then strategies that are removed in the second round of iterative deletion of strictly dominated strategies must have a lower payoff than some other strategies and so their share starts to shrink to zero and so on.

The above results on strict domination do not extend to weak dominance and iterative deletion of weakly dominated strategies. Samuelson and Zhang 1992 give an example of weakly dominated pure strategy that is not eliminated with even continuous time replicator dynamics.

Other families of games with globally stability for some standard dynamics include zero-sum, two-player games, nondegenerate 2×2 games and (weighted) potential games. For the fictitious play dynamics, the beliefs of players (empirical distribution of the actions of the individual players) converge to a Nash equilibrium in finite, zero-sum, two-player games (Robinson, 1951) and in nondegenerate 2×2 games (Miyasawa, 1961, and Monderer and Shapley, 1996a). Also, for best response and fictitious play dynamics the solution path in (weighted) potential games (Monderer and Shapley, 1996b) is globally stable and converges to Nash. The class of potential games include many noteworthy games such as common interest games and classical Cournot Oligopoly with constant marginal cost. In such game games players effectively strive to maximize a common function - the potential function. Thus, in potential games, as well as in two-person zero-sum games, the payoff of one player can be used to determine the payoff of other players. This may explain why HM's negative result on convergence, based on the informational content of uncoupled dynamics, does not apply to such games.

To obtain global convergence for a more general class of games we may have to consider other plausible evolutionary processes that are not adaptive and/or uncoupled as in HM. Next, we shall do precisely this by allowing richer set of types (in particular types that can choose different actions at different dates) and allowing selection at the level of rules. Before turning to this issue, would like to make a few brief remarks on the evolutionary dynamics above when the strategy space A_i of the underlying game is infinite.

dynamics stronger conditions than monotonicity are needed to ensure that strategies that are dominated by mixed strategies are eliminated (e.g. Aggregate monotonicity in Samuelson and Zhang 1992, convex monotonicity in Hofbauer and Weibull 1996). In the discrete case even these stronger conditions do not suffice. Dekel and Scotchmer (1992) give an example to show that discrete time replicator dynamics does not eliminate all strategies strictly dominated by a mixed strategies.

Stage-games with infinite strategies Infinite (one-shot) games are of interest for many reasons. First, when mixed strategies are allowed the strategy space is trivially continuous. Second, there are many games in economic contexts that are modeled with an infinite or continuous strategy set (e.g. bargaining games, the War of Attrition and Cournot duopoly, to name a few). Third, one may be interested in the question of whether the infinite case is the limit of successively finer finite approximations (the infinite case may then be regarded as a reference point).

It turns out that whether the strategy set for the one-shot game is infinite or finite does matter in terms of developing the evolutionary dynamics on the set of probability distributions over the set of strategies (e.g. the replicator equation), ensuring that the solutions are well-defined, stability of the stationary states and its relationship with static game-theoretic equilibrium concepts such as Nash. There are some recent papers on this topic (for example, Seymour 2000, Oechssler and Riedel 2001,2002 and Cressman 2004). Most of this recent literature assumes continuous time. The immediate issue here is, not surprisingly, what constitutes an appropriate notion of closeness and/or convergence for probability distributions on the set of strategies. Here there are several different possible topological extensions of the finite strategy space set-up, and conclusions on stability depend critically on what definitions are adopted. Here, for reasons of space we shall not consider this issue any further except to mention that even one of the results in Theorem 1 does not extend to the infinite case. Oechssler and Riedel (2002) provide an example showing that strict equilibria need not be asymptotically stable or even Lyapunov stable when the stage game is infinite.

Global Convergence and Richer Type Space In standard evolutionary model types in a population refer to strategy in the underlying one shot-game. Anderlini and Sabourian (1996), henceforth called AS, have looked at the question of global convergence with richer type space. In particular, they allow a type to be a rule of behaviour taking different actions at different dates (or histories) and selection is applied to the set of rules. More formally, a type $x_i : H \to A_i$ for population i is now a rule mapping from the set of past information H into the set of pure actions A_i. It assumed that there are a countable number of types/rules. If H is sufficiently informative then all standard adjustment rules such as best-response or imitation can be included in the set of allowable types. The critical assumption in AS is that the informational content of H is such that when a type makes a decision the dates are known; thus types may condition their actions at least on time and thus can take different actions at different times.

The evolutionary dynamics is the same as the discrete dynamics described above except that it operates by selecting amongst types (rules) according to how well the type has done at that date relative to the average payoffs of others. Formally, let $Q_i^t(x_i)$ be the proportion type x_i in population i at date t. The dynamics prescribes how Q_i^t evolves according to how well x_i does at t. Since payoffs depend on actions, without any loss of generality one can write the growth of each type at each date t as function of the action he takes and the distributions of actions in all the populations at t; thus the growth rate of type x_i at t, $g_i^t(x_i) = \frac{Q_i^{t+1}(x_i) - Q_i^t(x_i)}{Q_i^t(x_i)}$, can be written as

$$g_i^t(x_i) = \gamma_i(P^t, a_i^t(x_i))$$

where, as before, P^t is the distribution over the set of action profiles at t, $a_i^t(x_i)$ denotes the strategy taken by x_i at t and γ_i is the growth function for population i.

Here, as in the standard model, every limit point of the distributions of actions of the players corresponds to a Nash equilibrium (possibly mixed) of the one-shot game if the dynamics, defined by the function γ_i, is continuous and monotonic. This result is similar to the existing Folk Theorem of evolutionary game theory mentioned above (condition 4 in Theorem 1): if a path converges it must be to Nash.

The main result of AS concerns the global convergence of the distributions over types. First, define a type to be *feasible* if it has a positive initial probability. AS show that for every i and x_i, $Q_i^t(x_i)$ converges if for each population i there exists a feasible type $\overline{x}_i$ ($Q_i^0(\overline{x}_i) > 0$), referred to as the 'smart' type, which can grow as fast as any other type in population i at *each* date.[7] The existence of such a smart type with a positive initial probability ensures that the system converges. The intuition for this is as follows. Since $\overline{x}_i$ has the maximum growth rate it follows that $Q_i^t(\overline{x}_i)$ is increasing. Since $Q_i^t(\overline{x}_i)$ is also bounded, it then follows that the growth rate of $\overline{x}_i$ must go to zero. But the growth rates of other types in population i do not exceed that of the smart type and the sum of the growth rates of all types in population

[7] The assumption of positive initial probability for the smart types is more than is needed to establish the result. In fact, AS assume only that for each population a smart type is born at some finite date with a positive probability.

i at each date is zero. Therefore, it follows that the growth rate of others must also go to zero. (Effectively, since $Q_i^t(\overline{x}_i)$ is monotonically increasing, it is like a Lyapunov function.)

There are four immediate points to note concerning this result. First, since the system is deterministic, it is always possible to construct a smart type as long as types are allowed to condition their actions at least on time. This is because at any date t, a smart type needs to take an action $\overline{a}_i^t$ that is best in terms of growth:

$$\gamma_i(P^t, \overline{a}_i^t) \geq \gamma_i(P^t, a_i), \forall a_i. \tag{7}$$

Now, given the parameters of the system Q_i^0 and the selection function γ_i, for any t it is possible, by computing recursively forward (simulating) the system, to compute first P^t and then find an action $\overline{a}_i^t$ that satisfies (7).

Second, note that each smart type may need to take different actions at different times to ensure that it grows faster than others in the same population.

Third, the result does not depend on any particular assumption about the shape of the selection dynamics.

Fourth, note that although probabilities over types converge, probabilities over actions do not necessarily converge. This is because each type can condition on the past and can take different actions at different times. Therefore, even when probabilities over types converge, if the underlying game G has more than one Nash equilibrium, it is possible that in the limit different types could be switching between different Nash equilibria of G. An immediate corollary is that if G has a unique Nash equilibrium then probabilities over actions converge as well.

The main assumption in AS needed to establish the convergence result is the existence of a *feasible* smart type for each population. This assumption, however, is not primitive.[8] The identity of any smart type depends on the initial profile of distributions over types $Q^0 = (Q_1^0, .., Q_n^0)$; but then there is no guarantee that any smart type $\overline{x}_i$ for population i corresponding to Q^0 is such that $Q_i^0(\overline{x}_i) > 0$. An attractive set-up would be a set of assumptions on the of feasible types for the initial distributions Q_i^0 for each i that would automatically ensure that each population has a smart type with a positive

[8] Note that this assumption has some similarity with the grain of truth assumption in the Bayesian learning literature (Kalai and Lehrer 1993).

initial probability. This would require a closure property on the set of feasible types.

Clearly, we cannot ensure that the smart types are feasible if we restrict the set of types in population i to be a finite set (e.g. a finite automata with a bound) because the smart types (corresponding to any arbitrary initial distribution Q^0 with this finite support) for population i may not belong to this finite set. Another possible approach is to make the set of feasible types sufficiently rich so that it includes the smart types. However, this cannot go too far: if there is a continuum of feasible types, then it may not be possible to have an initial distribution for each population which attaches positive weights to a smart type. Thus, some restriction on the cardinality of the set of rules seems necessary. An obvious restriction is to make the space of allowable types to be *countable infinity*.

AS explore the case in which the set of types is restricted to be Turing-computable machines. The set of (Turing-) computable functions have nice closure properties (in particular the existence of a computable Universal program that can 'simulate' other computable functions); moreover, the cardinality of this set is countable infinite. However, restricting types to be Turing machines introduces new difficulties all related to the halting problem in the space of Turing machines. In particular, now one needs to construct a computable smart type that halts; but the existence of such a halting machine may be problematic because simply simulating the system and taking the action that maximizes the growth rate at each date may not be possible: in simulating the system the smart type may end up simulating other types that do not halt.

This problem cannot be trivially avoided by simply assuming that all types always halt. This is because to simulate the system the initial distribution Q_i^0 needs to be computable but this may not be consistent with the support of Q_i^0 being the same as the set of all halting machines.

3 Which Equilibrium: Stochastic Model with a Finite Population

The critique that standard game theory imposes too strong rationality on players is one reason for the interest in evolutionary game theory. Another reason, as we mentioned before, has been that the standard game theory does not have good predictive power; a game may have too many Nash equilibria.

Does the evolutionary approach help to select amongst the multiplicity of Nash equilibria? By only considering asymptotic stability of a Nash equilibrium in a deterministic model the answer may have to be "no" in many interesting classes of games. In particular, according to the Folk Theorem of evolutionary game theory (condition 3 in Theorem 1) every strict Nash equilibrium is asymptotically stable in the standard deterministic evolutionary model. Therefore, such models cannot be used to select between Nash equilibria (and thus cannot predict which equilibrium will occur) when the game has multiple strict Nash equilibria.

However, asymptotic stability is about one-shot mutation, it stays silent for the cases in which (small) mutants invade perpetually. Utilizing techniques from Freidlin and Wentzell (1984), Foster and Young (1990) consider a selection dynamics that is subject to stochastic shocks at every date. These perpetual stochastic shocks provides momentum for the system to escape from one strict equilibrium to another. However, as the probability that a random shock occurs becomes smaller, some equilibria may be visited more often. When such probability approaches zero, the system will be attracted into a much smaller set of equilibria, the stochastically stable set.

Kandori, Malaith and Rob (1993) and Young (1993a), Robson and Vega-Redondo (1996) and many others have successfully applied the same approach (with small perpetual mutations) to select amongst the set of strict Nash equilibria in evolutionary models with a finite population. In these papers the rules the agents follow are fixed (e.g. best reply, imitation). However, it turns out that the particular equilibrium selected depends precisely on the rules allowed. Specifically, in the 2×2 coordination game with one equilibrium being risk dominant and the other efficient, the best reply rule tends to select the former and the imitation rule tends to select the latter. As in the deterministic case above, we ask how the predictions of these models are affected if multiple rules are present in the system and rules that evolve are endogenously determined through some evolutionary selection. Juang (2002) provides a result that favors imitation when both the best response and imitation co-exist.

In this section we will discuss mainly the question of equilibrium selection in stochastic evolutionary models in the specific contexts of the 2×2 coordination game mentioned above and the classical Cournot oligopoly. We shall explain how the selection results depend crucially on the rules allowed and we will describe some attempts at resolving the issue of which rule by allowing multiple rules and applying evolutionary selection to the set of rules. We shall also briefly discuss other models with multiple rules and in particular some of the justifications provided for some specific rules such as imitation.

3.1 Mutation, Long-Run Equilibrium and Stochastic Stability

3.1.1 Mathematical background

Consider an evolutionary model in which a population of M (M is even) players are randomly paired to play a stage game repeatedly. After observing the outcome of the plays at the previous period, a fixed share of the players is randomly chosen to revise their actions. Under quite general set-ups this dynamics can be described by a stationary Markov process P on a finite state space S.

In general if there exist multiple strict Nash equilibria in the stage game then there exist multiple recurrent classes in the underlying Markov process P. (Henceforth, if a recurrent class consists of exactly one state, we shall refer to it as a recurrent state.) In this case we say such dynamics are non-ergodic and the limit points of the dynamics are described by the set of recurrent classes of P, denoted by $RC = \{X_1, .., X_r\}$. Specifically, depending on the initial starting state and the random process that follows, the dynamics may be absorbed into one of the recurrent classes and the system will stay there forever. Thus, we cannot say anything about the "which equilibrium" question.

To deal with the non-ergodicity problem, we introduce mutation by letting players take, with small probability, actions that differ from what they intend to take originally. This may represent players making mistakes, doing experimentation or population renewal. More specifically, suppose the dynamic is perturbed by allowing each agent i to mutate independently with a small probability $\epsilon > 0$ such that the perturbed process P^ϵ is also a stationary Markov chain with the state space S. Here mutation is parameterised by ϵ and P^ϵ converges to P as $\epsilon \to 0$.

The introduction of mutation allows the system to switch between any two states (and between any two recurrent classes of P) with a positive probability. As a result, for any $\epsilon > 0$ the perturbed process P^ϵ is ergodic; and therefore it has a *unique* invariant distribution μ^ϵ that summarises the "long-run" behaviour of the perturbed process from any initial condition.

We are interested in cases where the amount of mutation is arbitrarily small. It turns out that as the mutation rate ϵ approaches zero, the corresponding invariant distribution μ^{ϵ} converges to an invariant distribution μ of the unperturbed dynamics P. This often results in selection among multiple recurrent classes (equilibria) of the unperturbed dynamics P. That is, with mutation we allow the system to drift among different recurrent classes/states of P, but with the mutation rate going to zero, the system will concentrate on a much smaller set of states.

The states that μ attaches a positive probability are called the stochastically stable states and the set of such states is denoted by

$$\Omega \equiv \{s \mid \mu(s) > 0\}.$$

Clearly, the set of stochastically stable Ω is a subset of the set of recurrent states $RC = \{X_1, .., X_r\}$ of P and are the states that would be observed in the long-run with arbitrarily small amounts of mutation. As we shall see in later discussion, in many cases, the stochastically stable set Ω is singleton; in such cases this technique of perturbing the dynamics is very powerful in selecting a unique state.

In general, the characterization of Ω involves locating the recurrent class(es) with the least *stochastic potential*, as proposed by Freidlin and Wentzell (1984) and Foster and Young (1990). When there are only two recurrent classes in the system, this can be done by just counting the minimum number of mutations necessary for the system to transit from one recurrent class to the other. For illustration, suppose there are two recurrent classes X_1 and X_2. Denote the least number of mutation for the system to switch from X_2 to X_1 (from X_1 to X_2) as γ_1 (γ_2). Suppose $\gamma_1 < \gamma_2$ ($\gamma_1 > \gamma_2$). This means it is easier for the system to transit from X_2 to X_1 (from X_1 to X_2) than the other way round. We can then conclude that the recurrent class X_1 (X_2) is stochastically stable and we expect to observe X_1 (X_2) almost all the time in the long run.

For cases with more than two recurrent states, we need to "count" the number of mutations in a way that is a little more complicated. Suppose, for example, that there are four recurrent states (S_1, S_2, S_3, S_4) in a system. To compute the stochastic potential of any recurrent state, say, S_4, we have to construct a "tree" with root S_4. To be specific, for any other recurrent class, find a unique directed path from that state to S_4. Here we illustrate in Figure 1 below four different trees with root S_4. Sum up the total number of mutations for each feasible X_4-tree. The least one among all the X_4-trees is the stochastic potential of S_4. Similarly we can compute the stochastic potential for all other recurrent classes. The recurrent states/classes that have the least stochastic potential are stochastically stable.

[INSERT FIGURE 1 HERE]

3.1.2 Risk Dominance versus Efficiency

The classical example of equilibrium selection is a 2×2 coordination game with two strict Nash equilibria; one is risk dominant and the other efficient. The payoff matrix of such a game is given by

	A	B
A	a, a	b, c
B	c, b	d, d

Figure 2

where

$$a > c, d > b, a > d \text{ and } a + b < c + d. \tag{8}$$

Note that given (8) the game has two pure strategy strict Nash equilibria (A, A) and (B, B), the former is efficient and the latter is risk dominant.

Best response rule Kandori, Malaith and Rob (1993, hereafter KMR) construct a model with a single finite population of players M.[9] The players are repeatedly matched to play the 2×2 coordination game as in Figure 2. Before they play the game in each period, players can revise their action over time. In particular, at each date the players are assumed to adopt the best response rules (actions that are best responses to the distribution of actions in the previous period.

The state of the system s is defined to be the number of players that choose action A, thus $s \in \{0, 1, ..., M\}$. For any state s realized in the last period, we can define the best response rule followed by all the players by

$$r^{BR}(s) = \begin{cases} A & \text{if } \frac{s}{M} > p^*; \\ B & \text{if } \frac{s}{M} < p^*; \\ \text{randomize between } A \text{ and } B & \text{if } \frac{s}{M} = p^*, \end{cases}$$

where $p^* = (d - b)/(a + d - b - c)$.

Without mutation, it is straightforward to see that the distribution of actions converges to one of the two strict Nash equilibria (to states 0 or M), depending on the initial state. To see this, suppose the system starts at some state s_0 such that $s_0 > Mp^*$. Then playing A is a best response and whenever players are revising their actions they will play A and the system will converge to $s = M$. On the other hand, if the system starts at $s_0 < Mp^*$, then the system will converge to the state in which all play B ($s = 0$).

Next KMR introduce mutation by allowing players to make small mistakes *independently* in choosing their actions. Thus, for any given mutation rate ε, they obtain a perturbed Markov process that is ergodic and thus has a unique stationary distribution μ^ε. KMR find that in the long-run (as $\varepsilon \to 0$) the risk-dominant equilibrium B will be selected. More formally, they show that with best response rules the unique stochastically stable state is $s = 0$.

To see this note that there are exactly two recurrent classes/states 0 and M. Since the players follow the best response rule, it is easy to check that one needs $[(1 - p^*)M]^+$ action mutations to transit from state M to state 0 and $[p^*M]^+$ action mutations for the other way around, where the notation $[x]^+$ denotes the smallest integer that is not less than x. Since $[(1-p^*)M]^+ < [p^*M]^+$ it follows that it is easier for the system to transit from state M to state 0 than the other way round (0 has a lower stochastic potential). Thus, $s = 0$ is stochastically stable.

[9] Note that here since the game is symmetric there is a single population of agents.

Young (1993a) has an analogous set of results for a similar setup except that players randomly choose a best response to a sample from finite histories.

Before considering rules of behaviour other than the best response rule, note that in the above (and in the other models or perturbed dynamics we consider below) the probability that a mutation/error occurs is independent of the date, the state and the identity of the players. In the absence of a good explanation for mutations such a uniformity assumption on the mutation rates seems like a reasonable starting point.

However, it should be noted that the above sharp predictions of the KMR (and the models below) depend crucially on this uniformity assumption. Bergin and Lipman (1996) introduce mutations rates that are state-dependent in the KMR set-up and show that any strict Nash equilibrium can be selected in the long-run as the mutation rates go to zero, depending on the relative probability of the mutation rates in the different states. The intuition behind this result is easy to see. The stochastic stability of a recurrent state depends on its stochastic potential, which involves "counting" the number of mutations for relevant transitions. Suppose now that the mutation rates are state-dependent; in particular suppose that the mutation rate in each state $s \in \{[p^*M]^+, ..., M\}$ is ε and the mutation rate in each state $s \in \{0, 1, ..., [p^*M]^+ - 1\}$ is ε^α, for some constant α. The system with no mutation has, as before, two recurrent states 0 and M. But now with mutation, the probability of the transition from state 0 to state M is of the order $\varepsilon^{\alpha[p^*M]^+}$ while the probability of the transition from state M to state 0 is of the order $\varepsilon^{[(1-p^*)M]^+}$. Also, denote the unique invariant distribution by μ^ε. Then it is not difficult to show that there exist α^* such that as $\varepsilon \to 0$ the ratio $\mu^\varepsilon(s = 0)/\mu^\varepsilon(s = M)$ goes to infinity if $\alpha > \alpha^*$ and to zero if $\alpha < \alpha^*$. In other words, the assumption of state-dependent mutations allows us to support any of the two recurrent states by adjusting the appropriate mutation rate for each state.

Imitation rule Robson and Vega-Redondo (1996, hereafter RV) construct a model similar to KMR except that players adopt the imitation rule to play the 2×2 coordination game described in Figure 2. Specifically, each player, whenever revising actions, chooses the one that performed the best (in terms of average payoff) in the previous period. The state of the system $s \in \{0, 1, ..., M\}$ is again the number of players that choose action A. The imitation rule is given by

$$r^I(s) = \begin{cases} A & \text{if } \pi^A(s) > \pi^B(s); \\ B & \text{if } \pi^A(s) < \pi^B(s); \\ \text{randomize between } A \text{ and } B & \text{if } \pi^A(s) = \pi^B(s), \end{cases}$$

where $\pi^A(s)$ and $\pi^B(s)$ are the average payoffs for using action A and action B respectively in state s.

Again without mutation, the distribution of actions converges to one of the two strict Nash equilibria (to states 0 or M). With small mutations RV show that the efficient equilibrium, rather than the risk dominant one, is selected in the long run. More formally, they show that if players adopt the imitation rule then there exist some $\overline{M} > 0$ such that for all $M > \overline{M}$ the unique stochastically stable state is $s = M$.

To see this note that with the imitation rule there are also exactly two recurrent classes 0 and M. Moreover, the system can switch from state 0 to state M if two agents mutate to playing A and are paired together to play the game in the same period. Since $c > b$ it also follows that to transit from state M to state 0 requires strictly more than two mutations if $M > 2(a-b)/(a-c)$. Thus state M is stochastically stable.

The striking difference between the results in KMR and RV lies in the role of the rule adopted by players. In KMR's model, players best respond to the action frequency in the last period. For an action to be a best response, there must be at least a fixed share of the population playing the same action. In a 2×2 coordination game illustrated in Figure 2, if all players are playing A, then they will continue to do so unless at least $[p^* M]^+$ players mutate to B, in which case playing B is the best response rather than playing A. Since B is risk dominant ensures it follows that $p^* < 1/2$. Thus it is easier to switch from all playing A to all playing B than the other way round. When players imitate each other in RV's model, players only care about the performance of an action, rather than its frequency. This gives the efficient equilibrium an advantage. Consider the transition from all playing B to all playing A. We actually need only two mutations plus some "luck" in random matching: Two players mutate to A and these two are paired to play the stage game. In this case, the average payoff from A is a while the average payoff from B remains d. Although such "luck" in pairing occurs with a small probability, it is far more likely than players mutate to the other action as the mutation rate ε approaches zero. As for the transition the other direction, it is not difficult to see that a number of mutations proportional to the population size ($\frac{(a-c)}{(a-b)}M$) is indispensable. Thus if the population is reasonably large, the former transition is more likely to occur.

Rule Evolution and Equilibrium Selection The KMR and RV results illustrate the effect that different rules have on the selection amongst multiple equilibria. What if players are free to choose from different rules? As a preliminary step towards rule evolution and equilibrium selection, Juang (2002) looks at a model very similar to those in KMR and RV except that players may be able to choose either best response or imitation rules to play the stage games. Specifically, at each period one randomly chosen player can revise his rule by choosing one of the existing rules according to how each rule has performed in the previous period. Thus, he will make no change if only one rule is remaining.

Juang (2002) adopts a fairly general set-up of rule selection. The only assumption he makes on the rule selection criterion is what he calls *experimental*: If at any date t all players play some Nash equilibrium then at period $t+1$ the selection assigns a positive probability to each existing rule.

There are two points to note concerning the this assumption. First, it does not impose any restrictions when agents are not playing a Nash equilibrium. Second, at any Nash equilibrium all agents (and thus rules) are doing as well as each other; hence selection has no force and therefore it is reasonable to assume that every existing rule receives a positive probability as in the experimental assumption.

Juang introduces "mutation" at two levels. The first one refers to the action level. As before, a player makes mistakes with probability $\varepsilon > 0$ when choosing actions (implementing his rules). The second is at the rule level. It is assumed that, when updating rules, each player is also prone to making mistakes with probability ε^{η} where $0 < \eta < \infty$. Thus, in term of probability, one rule mutation is equivalent to η action mutations.

In this set-up, a state is a 3-tuple denoting the number of players playing A; the number of players playing A and adopting the imitation rule; and the total number of players adopting the imitation rule.

Note that with no mutation, because of the assumption that rule selection criteria are experimental, any state in which both rules coexist cannot be recurrent. This is because in any recurrent class players must be playing a Nash equilibrium and therefore if both rules coexisted, by the experimental property, both will be chosen with a positive probability. This allows the total number of players adopting the imitation rule to drift between 1 and $M-1$ until it reaches either 0 or M.

Since any state in which more than one rule is present is not recurrent and in any recurrent class agents must be playing the same Nash equilibrium, it follows that the system with no mutation has four recurrent states S_1, S_2, S_3 and S_4. In state S_1 all adopt the imitation rule and play A; in S_2 all adopt the imitation rule and play B; in S_3 all adopt the best response rule and play A; and in S_4 all adopt the best response rule and play B.

[INSERT FIGURE 3 HERE]

Figure 3 demonstrates the relationship between the setup of Juang with and those in KMR and RV. The transitions between S_3 and S_4 describe the ones in KMR and the transitions between S_1 and S_2 describe the ones in RV.

Juang's demonstrate the following result on the long-run (as $\epsilon \to 0$) behaviour of the system. The states S_1 and S_3 in which all agents play A (the action corresponding to the efficient equilibrium) are the only stochastically stable states in the above model provided the population size M is sufficiently large.

A sketch of the proof of Juang's result is as follows. First, consider the transition from any state where all agents adopt one rule to any state where all agents adopt the other, while using the same Nash equilibrium action profile. Given the assumption that all rule selection criteria are *experimental*, one rule mutation is sufficient for any such transition. This is because once one player mutates to adopt a rule that is different from the one adopted by all players, then both rules are present in the population. Since all players are playing the same Nash equilibrium, the rule selection criterion will prescribe both rules with a positive probability next period. The above arguments illustrate that one rule mutation suffices for a system to transit from one state in which all players adopt one rule to another state in which all players adopts the other rule, while they are playing the same Nash equilibrium.

Second, with four recurrent states, we need to compute the stochastic potential for each recurrent state in the way stated in subsection 3.1.1. Figure 3 describes the transitions between recurrent states and corresponding number of mutations. It is easy to see that states S_1 and S_3 are stochastically stable provided the population size M is sufficiently large. The intuition behind this is as follows. For transition from a state where all agents play A to another state where all agents use the same rule and play B, more than two action mutations are needed if M is above some parameter. But, as in RV, when all players are adopting the imitation rule, two action mutations are sufficient for the transition from all B's to all A's. Therefore any tree with the least stochastic potential must contain the transition from S_2 to S_1. This property gives the efficient equilibrium an advantage over the risk-dominant one. (It is easy to see that both states S_1 and S_3 have the stochastic potential $2\eta + 2$ while both states S_2 and S_4 have the stochastic potential $2\eta + \underline{m}$ where $\underline{m}$ refers to the least number of action mutation needed for the system to switch from all A's to all B's, when all players are adopting the imitation rule).

3.1.3 Cournot Oligopoly model: Nash or Walrasian equilibrium

Classical Cournot Oligopoly set-up has been another area in which the technique of stochastic evolutionary dynamics described above has been applied. Here, again the theory has sharp predictions which depend on the nature of the rules of behaviour allowed.

A symmetric discretised version of the Cournot model is as follows. There is a single of population consisting of n identical firms. In the one-shot game they set their output quantities simultaneously. They face an inverse demand function given by $P(.)$; thus if the total output brought to the market is Q the price will be given by $P(Q)$. Each firm has an constant marginal cost $c > 0$. The strategy of firm i is simply the quantity of output q_i it produces. To keep the finiteness of the model we assume that the set of output choices is finite.

In this set up the payoff of firm i is simply

$$\pi_i(q_i, q_{-i}) = q_i \left\{ P(\sum_{j=1}^{n} q_j) - c \right\}$$

Two reference benchmark outcomes for this one-shot game are the symmetric Cournot-Nash equilibrium and the Walrasian (also called Competitive) equilibrium. The former, denoted by a symmetric n-tuple of outputs $\mathbf{q}^c = (q^c, .., q^c)$ refers to the Nash equilibrium of the one-shot game. The latter refers to the outcome in a competitive market in which firms cannot influence prices and is defined by a symmetric n-tuple of outputs $\mathbf{q}^w = (q^w, .., q^w)$ where q^w satisfies

$$q^w \{P(nq^w) - c\} \geq q_i \{P(nq^w) - c\} \text{ for all output levels } q_i.$$

Assume that these two equilibria exist and are unique. Assume also that $P(.)$ is decreasing so that $q^c < q^w$.

Now suppose that the above Cournot game is played repeatedly and at any date each firm has an opportunity to adjust its behaviour with a positive probability.

best response dynamics First, let us consider the case in which the firms follow a myopic best response rule. The selection problem in this case is trivial. Since the Cournot game is a potential game, the best response dynamics for this game with no mutation/perturbation is globally stable, and therefore the output decisions of the firms converge to the unique Cournot-Nash equilibrium $\mathbf{q}^c$. When the dynamics is perturbed the set of stochastically stable states belong to the limit points (recurrent classes) of the unperturbed dynamics. Since the latter is unique, the stochastically stable state corresponds to the Cournot-Nash equilibrium.

Imitation dynamics If all firms follow the imitation rule then whenever a firm receives the opportunity to revise its actions it follows the action of one of the firms which obtained a higher payoff in the previous period. Imitation implies that the dynamics with no mutation converges a.s. to a monomorphic state in which all firms produce the same output. The recurrent classes of the dynamics are simply the set of all symmetric n-tuples $\mathbf{q}= (q, q, .., q)$.

Next, consider the same model with small mutation. Vega-Redondo (1997) shows that the stochastically stable set of this imitation dynamics is unique and is given by all firms producing the Walrasian output q^w. The basic idea here is as follows. First at any recurrent state $\mathbf{q}$ a single mutation by one firm to q^w is sufficient to induce the unperturbed dynamics towards the Walrasian equilibrium $\mathbf{q}^w$ with a positive probability. The intuition for this is that after the single mutation to q^w the mutant earns a higher profit than the other firms. But then by imitation all firms can end up producing q^w. Second, it requires more than one mutation to switch from $\mathbf{q}^w$ to any other recurrent state $\mathbf{q}$. This is because if in state $\mathbf{q}^w$only one firm mutates to another output level q, the mutant ends up earning a lower profit that the other firms; therefore the unperturbed imitation dynamics must lead the process back to state $\mathbf{q}^w$.

The Imitation rule in Vega-Redondo (1997) has one period memory. Alos-Ferrer (2004) considers the case of imitators with long but finite memory in Vega-Redond0 (1997) set-up. He shows that the process converges to a set of monomorphic states in which the firms produce an output between the Cournot and the Walrasian outcome.

Imitation and best response rules What happens if different firms adopt different rules. Schipper (2002) considers a model in which a population of imitators and best responders play repeatedly the above symmetric Cournot game. He shows that the long run distribution converges to a recurrent set of states in which imitators are better off than are best responders. His finding is robust even when best responders are more sophisticated. In his model the players can not change their rules. Thijssen (2005) relaxes this restriction and studies a similar environment as Schipper but allows players to change their rules. At the rule evolution level, he assumes that firm imitate the rule of the firm with the highest profit; thus the rule selection criterion is imitating in this set-up. Thijssen (2005) shows that Walrasian behavior is the unique stochastically stable state in this set-up with competing rules. He establishes his result by using the theory of nearly-complete decomposability to disentangle the action evolution and the rule evolution. Intuitively, this means one uses the limit distribution of the action evolution to obtain the limit distribution of rule evolution.

3.1.4 Other applications of stochastic stability

There are also many other applications of the stochastic stability approach in economic models such as bargaining (Young 1993b), social contracts (Young 1998) and social networks (Goyal and Vega-Redondo 2005), to name a few. In most of these applications players choose a single rule - best response rule. (Some exceptions are Saez-Marti and Weibull (1999), and Matros (2003) who extend Young's (1993b) result on selection in 2-player Nash bargaining problem to two rules set-up in which some individuals use the best response rule and others use best response to best response).

Hurken (1995) and Young (1998) extend Young (1993a) to general finite games and characterise the stochastic stable states of the perturbed dynamics in which at any date all individuals have finite memory and choose a (myopic) best response to a sample of past plays of their opponents. The difficulty with dealing with more general games is that the recurrent classes of the unperturbed games with minimal stochastic potential may involve cycles and stochastic stable states need not be Nash equilibria of the underlying games. For generic games Hurkens (1995) and Young (1998) characterise the stochastic stable states by showing that they belong to pure-strategy profiles in minimal sets closed under best response.

Josephson and Matros (2004), on the other hand, characterise the equilibrium of the same model as those in Hurkens (1995) and Young (1998) but with a different behavioural rule from best response. They assume that all individuals follow an imitation rule of choosing the most attractive strategy in their sample. First, not surprisingly, they show, in contrast to the results with best-reply dynamics, that with imitation rules a state is a limit point of the unperturbed dynamics if and only if it is a repetition of a single pure-strategy profile, a monomorphic state. Second, they show that all stochastically stable states of the perturbed dynamics are monomorphic and belong to a minimal set closed under single better replies (a minimal set of strategy profiles such that no single player can obtain a weakly better payoff by deviating unilaterally and playing a strategy outside the set).

Young (1993a, 1998), Hurkens (1995) and Josephson and Matros (2004) consider general games with single rules. Matros (2004) allows introduces of multi-rules in the same evolutionary setting. Here, as in AS discussed in the previous section, a rule refers to an action choice for each sample of past (finite) observations. He defines *weakly rational rules* as the rules that are "consistent with an equilibrium." This set includes all belief-based rules (those that are best responses to some belief) as well all imitation rules. He then shows that if players use only weakly rational rules including the best response rule then any limit point of the unperturbed dynamics belong to a minimal set closed under best response. Thus, very informally, without perturbation the dynamics with the set of weakly rational rules select the same outcome as does the dynamics with only the best response rules as in Hurkens (1995) and Young (1998). Matros then considers the perturbed dynamics and restricts his analysis to a subset of the weakly rational set, which he calls boundedly rational rules. The definition of this subset is quite complicated and rather ad hoc; nevertheless it still includes all belief-based and imitation rules. Very informally, he shows that when players use boundedly rational rules, including the best response and the imitation rule, the long-run outcomes of the perturbed dynamics as the mutation becomes small is the same those that will be obtained when agents choose only best response and imitation dynamics.

3.2 Why Imitate?

The literature on evolutionary framework with small but perpetual random shocks has been successful in selecting amongst multiplicity of equilibria in some well-known applications but as we have explained above the precise selection in many instances has been dependent on which rules of behaviour are allowed. In particular, as we have discussed, the best response rule and the imitation rule often have very different predictions. The important question is to ask why one rule is chosen rather than another. A less ambitious question may be why players imitate, rather than best respond, as the latter seems to make more sense in terms of rationality and expected utility maximization. Juang (2002) and Thijssen (2005), discussed above, address the question of selection between best response and the imitation rules in the context of specific games by introducing a selection dynamics between the two rules and letting them compete. However, the games they consider are very specific. In the remainder of this section we will briefly discuss some justifications of why individuals imitate in contexts other than the standard evolutionary game.

3.2.1 Imitate to maximize expected payoff

Schlag (1998) shows that through imitation, most individuals will learn to choose an expected payoff maximizing action with probability arbitrarily close to one, provided the population is sufficiently large. In his model finitely many agents repeatedly choose among actions that yield uncertain payoffs (multi-armed bandits). Periodically, new agents replaces some existing ones. Before entering the population each individual must commit to a behavioral rule that determines his next choice. When choosing an action in each period, an individual, guided by his rule, will sample another individual and is informed of the action chosen and the payoff received by that sampled agent in the last period. By comparing the performance of that sample with his own, the player decides to either switch to the action chosen by that sampled player or stick to his own action. Such a rule is "*imitating*" in that the player only selects either the action he chose in the last period or the action adopted by the individual he samples.

Schlag shows that an imitating rule is improving (payoff increasing in each bandit) if and only if, when two agents using different actions happen to sample each other, the difference in the probabilities of switching is proportional to the difference in their realized payoffs[10]. By allowing agents to select among improving rules, the author characterises the unique optimal *Proportional Imitation Rule*. Schlag then shows that, for any initial state with all actions present, the above optimal rule will guide most individuals to select the expected payoff maximizing action after finite periods.

3.2.2 Imitate to survive

The above discussion illustrates that through imitation individuals may "learn" to choose the optimal action. Imitation performs well even when competing with others. Blume and Easley (1992) studies a model in which investors distribute their wealth among a bundle of assets. Asset s pays off a positive amount of revenue if and only if state s occurs. It is easy to see that any investor putting zero share of his wealth on any asset s with state s occurring with positive probability will almost surely be bankrupt in finite time. Moreover, if we assume all traders have identical saving rates, then to survive in the market each investor must learn to distribute his wealth on each asset s with a share proportional to the probability that state s occurs. Blume and Easley try to find out what kinds of rules could survive in the market. Obviously the simple rule that invest in asset s with a fixed wealth share equal to the probability that state s occurs will survive. That is, an investor holding the correct belief on the states of the world and invest according to his belief will survive in the market.

They also show that some other rules may survive in the market. For example, a Bayesian learning rule whose prior has finite support containing the true model may survive. This is not surprising as a Bayesian will learn the true model if his prior contains the true model and he could observe all histories and update his belief according to his observation. Another

[10] Since *Imitate if Better* cannot distinguish between lucky and certain (or highly probable) payoffs, this rule might generate negative expected improvement and is not improving.

rule that may survive is a search rule. An investor who adopts the search rule selects at date t the portfolio rule that would have maximized his date t wealth share given the actual prices and states up to date t. Both the Bayesian learning rule and the search rule involve relatively sophisticated behavior. For example, a Bayesian's prior must contain the true model and, he must update his posterior probability on each model in each date. For the search rule, the investor must remember the history of the state-price process, compute the objective function and maximize it. Consider an imitator as follows. Use at date t the rule adopted at date $t-1$ by the trader with the largest wealth share. It turns out that such a imitation rule will survive: Imitating the trader with the largest wealth share is equivalent to imitating the most successful trader in the long run. The trader with the largest wealth share may not be adopting the most successful rule initially, but then his wealth share will shrink and will eventually be dominated by another trader who performs better. In the long run the investor with the largest wealth share must adopt the most successful rule. Thus an imitator will eventually settle down to the most successful rule and will survive in the market.

3.2.3 Imitation to efficiency

Ellison and Fudenberg (1993) construct a social learning model in which a continuum of agents repeatedly select between two competing alternatives with one being superior to the other but, subject to random shocks. Therefore, the inferior choice may do better than the superior one some of the time. In this model, if only a share of the agents may revise their choice each time and when revising, they just choose the alternative that performed better in the last period, then the society will settle down to a state in which the share of the agents adopting the superior choice is equivalent to the probability that the better alternative outperforms the other, the so-called *probability-matching* behavior.

Ellison and Fudenberg then assume that an agent, when revising his choice, also take into account how many agents adopting each alternative. Specifically, he considers not only the performance of both alternatives but also their popularity. Thus the weight an agent assigns to the popularity of an alternative plays an important role in the social learning environment. Intuitively, if the weight is too small, then we shall observe a similar outcome as the probability-matching case, and thus there will always be a non-zero share of the society who adopt the inferior choice. On the other hand, if the weight is too large, then the society will settle down to either the efficient choice or the inferior one, depending on the initial state and the realized random shocks. Only with the weight within an optimal range will a society learn to choose the efficient choice.

Juang (2001) consider a heterogeneous society in which agents may have different popularity weights and introduce a replicator-like dynamic to differentiate agents according to their performance. He characterises two conditions, called *precision* and *diversity,* each of which guarantees all agents in the society herd on the superior alternative. Precision requires that at least non-negligible share of agents use popularity weight that is within the optimal range in Ellison and Fudenberg (1993). This is straightforward as agents with such a weight will learn the superior choice precisely in the long run and those who fail to do so will be outperformed and will vanish in the long run. The diversity condition requires that at least one "low" popularity weight and one "high" popularity weight co-exist in the society (low and high are relative to the optimal range stated above). Intuitively, agents with lower (higher) weight care more about performance (popularity). If the inferior choice is more popular, then agents with low weight will outperform those with high weight, and help the society escape from inefficient herding and reach the state in which the superior choice is more popular. Once this occurs, agents with high weight will adopt the superior choice with greater proportion and outperform those with low weight and lead the society into the efficient herding in the limit. An interpretation for diversity is that "*rational*" agents locate the optimal choice for the society while it is "imitating" agents who make the profits.

References

Alos-Ferrer A. (2004): "Cournot versus Walras in Dynamic Oligopolies with Memory," *International Journal of Industrial Organisation*, 22, 193-217.

Anderlini, L. and Sabourian, H. (1996): *The Evolution of Algorithmic Learning Rules: A Global Stability Result*, European University Institute working paper.

Bergin, J. and Lipman, B. L. (1996): "Evolution with State-Dependent Mutations," *Econometrica,* 64, 943-56.

Blume, L. and Easley, D. (1992): "Evolution and Market Behavior," *Journal of Economic Theory*, 58, 9-40.

Borgers, T. and Sarin, R. (1997): “Learning through Reinforcement and Replicator Dynamics,” *Journal of Economic Theory*, 77, 1-14.

Bjornerstedt, J. and Weibull, J. (1996): “Nash Equilibrium and evolution by imitation,” in K. Arrow and E. Colombatto(Eds) *The Rational Foundations of Economic Behaviour*. New York Macmillan, 155-171.

Cressman, R. (2004): *Stability of the Replicator Equation with Continuous Strategy Space*, Mimeo.

Dekel, E. and Scotchmer, S. (1992): “On the Evolution of Optimizing Behavior,” *Journal of Economic Theory*, 57, 392-406.

Ellison, G. and Fudenberg, D. (1993): “Rules of Thumb for Social Learning,” *Journal of Political Economy*, 101, 612-643.

Foster, D. and Young, P. (1990): “Stochastic Evolutionary Game Dynamics,” *Theoretical Population Biology,* 38, 219-232.

Freidlin, M. and Wentzell, A. (1984): *Random Perturbations of Dynamical Systems*, Springer Verlag, New York.

Fudenberg, D. and Levine, D. (1998): *The Theory of Learning in Games*, MIT Press.

Goyal, S. and Vega-Redondo, F. (2005): “Learning, Network Formation, and Coordination,” *Games and Economic Behavior*, 50, 178-207.

Harsanyi, J. and Selten, R. (1988): *A General Theory on Equilibrium Selection in Games*. MIT Press.

Hart, S. and Mas-Collel, A. (2000): “A Simple Adaptive procedure Leading to a Correlated Equilibria” *Econometrica,* 68, 1127-50.

Hart, S. and Mas-Collel, A. (2001): “A General Class of Adaptive Strategies,” *Journal of Economic Theory*, 98, 26-54.

Hart, S. and Mas-Collel, A. (2003): “Uncoupled Dynamics do not Lead to a Nash Equilibrium” *American Economic Review,* 93, 1830-6.

Hofbauer, J. (2000): “From Nash and Brown to Maynard Smith: Equilibria, Dynamics, and ESS, *Selection* 1, 81-88.

Hofbauer, J. and Sigmund, K. (1998): *Evolutionary Games and Population Dynamics*, Cambridge University Press.

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

Hofbauer, J. and Weibull, J. (1996): “Evolutionary Selection against Dominated Strategies,” *Journal of Economic Theory*, 71, 558-573.

Hurkens S. (1995): "Learning by Forgetful Players," *Games and Economic Behavior*, 11, 304-329.

Josephson J. and Matros, A. (2004): "Stochastic Imitation in Finite Games," *Games and Economic Behavior*, 49, 244-259.

Juang, W. (2001): "Learning from Popularity," *Econometrica*, 69,735-747.

Juang, W. (2002): "Rule Evolution and Equilibrium Selection," *Games and Economic Behavior*, 39, 71-90.

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

Matros, A. (2003): "Clever Agents in Adaptive Learning," *Journal of Economic Theory*, 111, 110-124.

Matros, A. (2004): *Simple Rules and Evolutionary Selection*, Mimeo, Pittsburgh.

Matsui, A. (992): "Best Response Dynamics and Socially Stable Strategies," *Journal of Economic Theory*, 57, 343-362.

Maynard Smith, J. (1982): *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.

Maynard Smith, J. and Price, G. R. (1973): "The Logic of Animal Conflicts," *Nature,* 246**,** 15-18.

Miyasawa, K. (1961): *On the Convergence of the Learning Process in a* 2×2 *Non-Zero-Sum Two Person Game*, Research Memorandum No. 33, Economic Research Program, Princeton University.

Monderer, D. and Shapley, L. (1996a): "Fictitious Play Property for Games with Identical Interests," *Journal of Economic Theory*, 68, 258-265.

Monderer, D. and Shapley, L. (1996b): "Potential Games," *Games and Economic Behavior*, 14, 124-143.

Nachbar, J. (1990): "Evolutionary Selection Dynamics in Games: Convergence and Limit Properties," *International Journal of Game Theory*, 19, 59-89.

Oechssler, J. and Riedel, F. (2001): "Evolutionary Dynamics on Infinite Strategy Spaces," *Economic Theory*, 17, 141-162.

Oechssler, J. and Riedel, F. (2002): "On the Dynamic Foundation of Evolutionary Stability in Continuous Models," *Journal of Economic Theory*, 107, 223-252.

Robinson, J. (1951): "An Iterative Method of Solving a Game," *Annals of Mathematics*, 54, 296-301.

Robson, A. J. and Vega-Redondo, F. (1996): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory,* 70**,** 65-92.

Saez-Marti M. and Weibull J. (1999): "Clever Agents in Young's Evolutionary Bargaining Model," *Journal of Economic Theory*, 86, 268-279.

Samuelson, L. (1997): *Evolutionary Games and Equilibrium Selection*, MIT Press.

Samuelson, L. (2002): "Evolution and Game Theory," *Journal of Economic Perspectives*, 16, 2, 47-66.

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

Sandholm, W. (2007): *Population Games and Evolutionary Dynamics*, to be published by MIT Press.

Schipper, B. (2002): *Imitators and Optimizers in Cournot Oligopoly*, mimeo.

Schlag, K. (1998): "Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-armed Bandits," *Journal of Economic Theory*, 78, 130-56.

Seymour, R. (2000): *Dynamics for Infinite Dimensional Games*, mimeo, University College London.

Shapley, L. (1964): "Some Topics in Two-Person Games," in *Advances in Game Theory*, edited by Melvin Dresher, L. S. Shapley and A. W. Tucker. Annals of Mathematics Studies No. 52. Princeton: Princeton University Press.

Thijssen, J. (2005): *Nearly-complete Decomposability and Stochastic Stability with an Application to Cournot Oligopoly*, Mimeo.

Vega-Redondo, F. (1996): *Evolution, Games, and Economic Behavior*, Oxford University Press.

Vega-Redondo, F. (2003): *Economics and the Theory of Games*, Cambridge University Press.

Weibull, J. (1995): *Evolutionary Game Theory*, MIT Press.

Weissing, F. (1991): "Evolutionary Stability and Dynamic Stability in a Class of Evolutionary Normal Form Games," in *Game Equilibrium Models I: Evolution and Game Dynamics*, edited by R. Selten, Springer-Verlag.

Young, P. (1993a): "The Evolution of Conventions," *Econometrica,* 61**,** 57-84.

Young, P. (1993b): "An Evolutionary Model of Bargaining," *Journal of Economic Theory*, 59**,** 145-168.

Young, P. (1998): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton University Press.

Young, P. (2004): *Strategic Learning and its Limits*, Oxford University Press.