EVOLUTIONARY EQUILIBRIUM WITH FORWARD-LOOKING PLAYERS

Lawrence E. Blume

March 1995
Current Version: 27 June 1995

This research is supported by the National Science Foundation. The author is also grateful for support from the Institute for Advanced Studies at Hebrew University and the Department of Economics at Tel Aviv University.

Abstract

The population game literature builds upon an implicit model of player interaction, pairwise random matching, and three behavioral postulates: noisy decisionmaking, myopic decisionmaking, and the random arrival of choice opportunities (inertia or lock-in). By now the role of noise is understood, This paper investigates the role of the player interaction model and the other two behavioral postulates by building two distinct fully intertemporal population models where players have rational expectations about the future and discount future expected payoff streams. Not surprisingly, myopic play emerges inboth models as the discount rate becomes large. In one model it also arises as lock-in increases. The two models exhibit distinct myopic behavior. Specialized to coordination games, only in one model is there selection of the risk-dominant equilibrium. The most surprising result is that in neither model does patient play lead to payoff-dominant equilibrium selection. Quite the contrary. If players are patient enough, the basin of attraction for the risk-dominant equilibrium state enlarges to include the entire state space.

JEL Classification: C78

Correspondent:

Professor Lawrence Blume
Department of Economics
Uris Hall
Cornell University
Ithaca, NY 14853
LB19@CORNELL.EDU

Eugene Delacroix's (1828) image of Mephistopheles galloping off with Faust illustrates the consequences of impatience. This paper determines if patience has its own rewards.

1. Modelling the Evolution of Strategic Choice

The canonical stochastic population model assumes a population of players who from time to time are randomly paired with each other in a strategic relationship described by a two-person game. They enter the match with a strategy choice already locked in, and are rewarded according to their choice and the choice of their opponent. Again from time to time players have opportunities to revise their choices. When a revision opportunity occurs players will best-respond to their beliefs with high probability, but not surely. These models admit several possible interpretations of the event that a player doesn't best respond, but the important fact is that this random component of play is unmodelled noise. That is, its existence is simply assumed.

The four characteristics which describe all models in this class are *pairwise random matching, myopia, inertia* and *noisy choice. Pairwise random matching* describes how players in the population interact with one another. Although it is never formally modelled, the idea is that frequently players are randomly matched with other players, and these matches are the source of all payoffs. From each match a player gets a payoff which is determined by her strategy and the strategy of her opponent.

The remaining characteristics describe individuals' choice behavior. *Myopia* has to do with the modelling of the term "best-response" in the preceding description. It means that players respond to the expected payoff that would result from matches given the distribution of play in the current state. *Inertia* is the supposition that players cannot revise their strategies before each match,but instead only occasionally. And of course *noisy choice* describes the unmodelled random noise. Random experimentation, random utility and random replacement of players have all been used to justify the stochastic perturbation of best-response behavior.

Most of the existing literature fixes a myopic decision rule and a specification of inertia, and studies the impact upon the model's long-run behavior of changes in the characteristics of the unmodelled noise, especially as that noise is made small. In this paper I propose to study the effects of alternatives to myopic best response, variations in the rate of strategy revision, and assumptions about the nature of player interaction alternative to pairwise random matching.

The matching technology brings two players together for an instant of time, just as two billiard balls might collide while both are in motion. The payoffs from the match are determined by the choices the two players are locked in to at the instant of the match. This story has generated some interest as part of a stylized model of the evolution of choice in a population of players, but describes almost no interesting economic phenomena. [1] I will

[1] The Menger-Kiotaki-Wright model of monetary exchange (Menger, 1892 and Kiyotaki and Wright, 1989) and some versions of Diamond's search equilibrium model (Diamond, 1982) are the only exceptions I have found so far, and even these models go beyond the simple formalisms of the contemporary evolutionary game theory literature.

compare this interaction model with a model in which players meet and are bound together in an ongoing relationship for a random period of time during which the players receive a continuous payoff flow. During the match players may have opportunities to revise their choice of strategy. The billiard-ball model I will refer to as the *discrete match model,* and the alternative model I call the *continuous flow model.*

In both interaction models I will study how inertia and impatience interact to determine the short- and long-run behavior of the stochastic process describing players strategies. For both interaction models I prove the existence of equilibrium play when players are forward-looking and have "rational expectations" about the course of future play. This equilibrium concept is the most restrictive method for introducing forward-looking behavior, and is certainly not in harmony with the evolutionary paradigm, which eschews complex rationality requirements. Nonetheless it is of interest for four reasons. First, upon any boundedly rational notion of forward-looking behavior that did not include the possibility of fully rational outcomes sits the burden of justification for its systematic biases. Second, it must be admitted that a rationality requirement that players understand some aggregate process is different from the inductive approach to rationality through common knowledge and common belief that models of bounded rationality in games are most interested in avoiding. Third, this is the way to consider forward-looking behavior without entering the interesting but tangential argument of exactly what kinds of forward-looking choice models would be most interesting to study. Forth, full rationality is particularly interesting in the models presented here because in both interaction models the consequences of patient rational play are striking.

The myopia hypothesis is motivated by bounded rationality considerations. A central concern of this paper is the ways in which myopic play can emerge from the dynamic choice models suggested by the foregoing description of equilibrium choice, and what different form myopia takes in the discrete match and continuous flow models.

The significance of the inertia hypothesis is two-fold. First, random matching opportunities remove certain equilibrium possibilities, such as "all play up in even periods and down in odd periods" in a coordination game. Second, inertia affects the tradeoff between present and future rewards in strategic choice. When choice is myopic this is not important; but when players are forward-looking, they will trade off the short term benefits of playing optimally in the current choice environment against the long term cost of possibly being locked into a bad choice for the environment of the future.

The characterizations of the emergence of myopic behavior in both interaction models are valid for all symmetric games and do not depend on the magnitude of the noise. The results on patient play are limited in three ways. First, I investigate only two-by-two coordination games, the benchmark strategic problem of the stochastic evolution literature. Second, the results assume that the stochastic noise is small. In other words, all results are "equilibrium selection" results. Third, I study only Markov, symmetric and "monotonic equilibria", equilibria in which all players use the same Markovian decision rule, and in which that rule is of the form "play A if enough players are choosing A, otherwise choose B". There is reason to believe that for generic coordination games, monotonic equilibria are the only equilibria when players are patient and tremble rarely. But for the moment this claim must remain a conjecture.

Both interaction models are described by three parameters: r, the intertemporal discount rate, δ the arrival rate of new matches, and σ, the arrival rate of strategy revision opportunities. As expected, myopic play arises as players become impatient, as r becomes large. More surprising is that in the discrete match model, myopic behavior emerges for any discount rate when the arrival rate of revision opportunities σ, is sufficiently small. This includes situations wherein all players are extremely patient. The consequences of myopic play are quite different in the two models. While myopic play in the discrete match model affords the usual equilibrium selection results in coordination games, there is no equilibrium selection result in the myopic continuous flow model. Both equilibria of a coordination game are stochastically stable states (and these are the only such), and the invariant distribution depends upon the distribution of choice in the event of a tremble.

Although the myopic play of the two models differs, their behavior when players are patient is identical. The surprising result here is the strong emergence of risk-dominant equilibrium selection. Just as when preferences are myopic, the risk-dominant equilibrium has the larger basin of attraction, but in the limit case of extremely patient players the basin of attraction expands to the entire state space. Thee conclusions are independent of the level of noise, and so in fact no noise is needed to get the now-conventional risk-dominant equilibrium selection result. When players are patient, decreasing the noise decreases rather than increases the expected waiting time for a transition from coordination on the risk-dominated equilibrium to coordination on the risk-dominant equilibrium.

Models in the population game literature have been implemented in three ways. Canning (1990), Kandori, Mailath and Robb (1993), Young (1993) and Samuelson (1994) all implement the evolutionary story as a discrete time Markov chain. Foster and Young (1990) and Fudenberg and Harris (1992) implement this model as a Brownian motion. Blume (1993a, 1993b and 1994) implements this model as a continuous-time jump process. In particular, when all players are ex-ante identical, this formulation is a continuous-time birth-death process. Binmore, Samuelson and Vaughan (1993) use a birth-death chain approximation to analyze a model with more complicated Markovian evolution. Birth-death models have an advantage over those models in which at any date many players may revise their strategy; they are significantly easier to analyze. In particular, they make possible the study of the dynamic programming problems necessary to understand forward-looking behavior. Consequently I will use the formulation of evolutionary models with random matching as developed in Blume (1994).

2. The Models

The two models share a common structure concerning strategy revision opportunities, instantaneous payoffs and the like. They differ only in the matching technology — how it is that matches actually generate payoffs. This section first describes the common elements. Then the matching technologies and the resulting dynamic programming problems are taken up in turn for the discrete match and continuous flow models.

2.1. Common Elements

A *stochastic strategy revision process* is a population process which describes the distribution among a set of strategies of a population of players. The population has N players, named player 1 through player N. Given too is a payoff matrix for a $K \times K$ symmetric game G. Without loss of generality we will assume that all entries in G are strictly positive. Each player is labelled with a strategy. The *state* of the population process at time t is a vector which describes the distribution of strategies in the player population.

From time to time each player has an opportunity to revise her strategy. A *policy* for a player is a map that assigns to each state a probability distribution over from strategies from which she would draw the strategy she would choose if give the opportunity to revise her strategy at an instant when the process is in that state. Such policies are called *Markovian* in the dynamic programming literature. In general, of course, one would want to allow dependence upon the entire observable history of the process. But it will become apparent that if all other players are using Markovian policies, then any one player has a Markovian optimal response.

The phrase "from time to time" has a very specific meaning. An event that happens "from time to time" is an event which happens at random intervals whose evolution is described by a Poisson alarm clock. Consider the arrival of strategy revision opportunities for player n. Associated with player n is a collection $\{x_{nl}\}_{l=1}^{\infty}$ of independent random variables, each distributed exponentially with rate parameter σ (mean $1/\sigma$). The variable x_{n1} is the waiting time until the first strategy revision opportunity, x_{n2} is the interarrival time between the first and second strategy revision opportunity, and so forth. Thus the waiting time until the mth strategy revision opportunity for player n is $\sum_{l=1}^{m} x_{nl}$.

Finally, when a player choose a strategy k, that strategy may not in fact be implemented. Some other choice l may be realized instead. Let $q = (q_1, \ldots, q_K)$ be a completely mixed probability distribution on the set of strategies. Also fix an ϵ strictly between 0 and 1. With probability $1 - \epsilon$ the player's choice is realized, and with probability ϵ the players new choice is chosen by a draw from the distribution q.

Each player discounts the future at rate r. Given the policies of other players, the expected present discounted value of the payoff stream from a policy for any player can be computed. Each player chooses a policy to maximize that expected present discounted value. A stochastic strategy revision process is an *equilibrium strategy revision process* if each player's policy maximizes that expected present discounted value given the policies of all other players.

2.2. The Discrete Match Model

From time to time a pair of players are brought together in a match. They receive an instantaneous payoff from the match which is determined by the strategy each player is employing at the time of the match and the payoff matrix G.

"From time to time" means the following: For each pair of players (m, n) with $m < n$ there is a collection of independent rate $\delta/(N-1)$ exponentially distributed random variables $\{x_{mnl}\}_{l=1}^{\infty}$, where x_{mnl} is the interarrival time between the $l-1$th and lth match of players m and n. It follows from the properties of independent, distributed random variables that the interarrival time between the $l-1$th and lth matches of player n with anybody is exponentially distributed with rate parameter δ. Notice that matching cannot be independent across players (although for large N it is approximately so). This does not matter. The only independence requirements are the independence of revision opportunities and matches, and the independence of revision opportunities across players.

Choose a player, say, player 1. Let Δ_{N-1}^{K} denote the set of vectors of integers in $\mathbf{R}^{\mathbf{K}}$ of the form $a = \{a_1, \ldots, a_K\}$ where the a_k are nonnegative and $\sum_k a_k = N-1$. Let Δ^K denote the set of probability distributions on the K pure strategies available to each player. A policy is a map $\pi : \Delta_{N-1}^{K} \to \Delta^K$. The argument of a policy is a list of the number of player 1's $N-1$ fellow players who are currently playing each strategy. The stochastic process which describe the evolution of the opponents' behaviors is called the *opponent process.* Typically the current strategy of player 1 will affect the evolution of opponents' play, and so the evolution of the opponent process will be contingent upon the current choice of player 1.

If all players employ a common Markovian policy, the decision problem facing each player will set up as that of controlling a birth-death process. Suppose that all players other than player 1 have settled upon a common policy π. The opponent process can change state only when one of the opponents has a strategy revision opportunity. Thus the only allowable changes of state are those in which one a_k decreases by 1 and another a_l increases by 1. This corresponds to an opponent currently playing k switching to l. The rate at which this happens depends upon the current choice of player 1. Suppose player 1 is currently playing strategy m. Let e_k, e_l and e_m denote the kth, lth and mth unit vector in $\mathbf{R}^{\mathbf{K}}$, respectively. Let a denote the current state and $b = a - e_k + e_l$ denote the new state. Then the transition rate is

$$\lambda_{ab}^{m} = a_k\sigma\big((1-\epsilon)\pi(a + e_m - e_k)_l + \epsilon q_l\big)$$

There are a_k opponents currently choosing k, so strategy revision opportunities arrive collectively to the collection of k players at rate $a_k\sigma$. If the state of player 1s opponent process is a, the state of the opponent process of an opponent playing k is $a + e_m - a_k$: $a + e_m$ describes what all players are doing; then subtract off what the k-player is doing. Consequently, the term inside the parentheses describes the probability that a k-player will choose l. All transitions requiring two or more players to change strategies run at rate 0.

Notice that, given the choice of player 1, the opponent process is a multi-type birth-death process. Player 1 can affect the transition rates by changing strategies, although this effect becomes negligible as the number of players grows.

When player 1 plays strategy k, her opponents (aggregate) behavior is described by a multitype birth-death process on the set Δ^K_{N-1} with transition rates $\lambda = \{\lambda^k_{ab}\}_{a,b\in\Delta^K_{N-1}}$. Let $\lambda^k_a = \sum_{b\neq a}\lambda^k_{ab}$. This is the rate at which the opponent process changes state when player 1 plays strategy k.

Given the policy choice of her opponents, player 1's optimization problem is a conventional discounted stochastic dynamic programming problem. She has finitely many controls with which to control a Markov jump process. To describe the Bellman equation still more notation is needed. When player 1 is matched, the probability that she meets an opponent playing strategy l is proportional to the number of players choosing strategy l in the current state a. Thus her expected return from a match when the opponent process is in state k and she plays strategy k is

$$v(k,a) = \sum_{l=1}^{k} \frac{a_l}{N-1} G(k,l). \tag{2.1}$$

Suppose now that player 1 is currently playing k and the opponent process for player 1 is in state a. At some random moment in the future an "event" will take place. That event can be one of several things: (1) The opponent process can change state. (2) Player 1 can be matched with an opponent. (3) Player 1 can have a strategy revision opportunity. The waiting time from now until the occurrence of the next event is a random variable $\tilde{\tau}$ which is exponentially distributed with parameter $\lambda^k_a + \sigma + \delta$. When the event occurs, with probability $\lambda^k_{ab}/(\lambda^k_a + \sigma + \delta)$ it will be a transition of the opponent process to state b, with probability $\sigma/(\lambda^k_a + \sigma + \delta)$ player 1 will have a strategy revision opportunity, and with probability $\delta/(\lambda^k_a + \sigma + \delta)$ she is matched.

A state for player 1's decision problem is a pair (k,a) where k is the strategy player 1 is using and a is the state of the opponent process. Let $V(k,a)$ denote the value of the state (k,a). The value function is given by the Bellman equation

$$\begin{aligned} V(k,a) &= E\Big\{e^{-r\tilde{\tau}} \sum_{b\neq a} \frac{\lambda^k_{ab}}{\lambda^k_a + \sigma + \delta} V(k,b) + \frac{\delta}{\lambda^k_a + \sigma + \delta}\big(v(k,a) + V(k,a)\big) + \\ &\qquad \frac{\sigma}{\lambda^k_a + \sigma + \delta} \max_{\pi} \sum_l \big((1-\epsilon)\pi_l(a) + \epsilon q_l\big) V(l,a)\Big\} \\ &= \sum_{b\neq a} \frac{\lambda^k_{ab}}{r + \lambda_a + \sigma} V(k,b) + \frac{\delta}{r + \lambda_a + \sigma} v(k,a) + \qquad (2.2) \\ &\qquad \frac{\sigma}{r + \lambda_a + \sigma} \max_{\pi} \sum_l \big((1-\epsilon)\pi_l(a) + \epsilon q_l\big) V(l,a) \end{aligned}$$

Standard arguments show that the Bellman equation has a unique solution which is the value function for player 1's dynamic programming problem.

2.3. The Continuous Flow Model

In the continuous flow model players are matched, and the result of the match is a continuous flow which lasts until the next match. There are several ways of introducing a matching technology. For instance, matches might break up and re-form one by one. In this case some continuous flow value would be assigned to the event of being unmatched. Another approach is to assume that everyone is always matched, but at random moments the entire population is re-matched. The second approach will be taken here. The main difference in the setup of the two models is that in the discrete match model only frequencies of players choices matter. But in the payoff flow model it is important to keep track of the names of the players playing each strategy.

At the beginning of time each player is matched with another player. Let M denote the set of all possible matchess. Suppose that from time to time a new match $\mu \in M$ is chosen from M. Let $\{x_n\}_{n=1}^{\infty}$ denote a collection of independent random variables, each distributed exponentially with parameter δ. At time 0 the players are matched. The first rematch occurs at time x_1, the second at time $x_1 + x_2$, and so forth. Let μ_n denote player ns partner in the match μ. There are $C = (N-1)(N-3)\cdots 1$ possible matches, and each match is equally likely. Thus the probability of realizing any given match is $1/C$, and the probability that a given player will be matched with any particular opponent, is $1/(N-1)$.

A configuration of the population is a specification of each player's choice of action. Let $\mathcal{S}$ denote the set of all possible configurations, with typical element s. When population play is described by configuration s, $s(n)$ is the play of player n. For a given configuration s, s_n^k is the configuration that results from changing player n's action to k and leaving all other players unchanged. Such transitions are the only transitions that can occur with positive probability. any other transition would require that two or more players have a strategy revision opportunity at the same instant, which is a 0-probability event.

Now a policy is a function $\pi : \mathcal{S} \times M \to \Delta^K$, which assigns to each configuration and match a probability distribution on the pure strategies.

Again only symmetric equilibria will be investigated. Let $\lambda_n^k(s, \mu)$ denote the transition rate from s to s_n^k.

$$\lambda_n^k = \sigma\big((1-\epsilon)\pi_k(s, \mu) + \epsilon q_k\big)$$

An "event" occurs when one of three things happen: (1) an opponent has a strategy revision opportunity;[2] (2) a new match ocurs; and (3) when the player has a strategy

[2] In the discrete match model this first event had the opponent process changing state. In this model it is more convenient to record all strategy revision opportunities and not just those which result in a change of state. Nonetheless the principle is the same.

revision opportunity. The arrival rate of an event is $\rho = N\sigma + \delta$. Let τ denote the waiting time to the next event. τ is distributed exponentially with parameter ρ. When an event occurs, the probability of the event being of any particular kind is proportional to its rate parameter. Thus the Bellman equation takes the following form:

$$\begin{aligned} V(s,\mu) = E_\tau\Big\{ & \int_0^\tau e^{-rt} G\big(s(n), s(\mu_n)\big)\, dt + \\ & e^{-r\tau} \frac{\delta}{N\sigma + \delta} \frac{1}{C} \sum_{\nu \in M} V(s,\nu) + \\ & e^{-r\tau} \frac{\sigma}{N\sigma + \delta} \sum_{m \neq n} \sum_k \lambda_n^k(s,\mu) V(s_m^k, \mu) + \\ & e^{-r\tau} \frac{\sigma}{N\sigma + \delta} \max_\pi \sum_k \big((1-\epsilon)\pi_k(s,\mu) + \epsilon q_k\big) V(s_n^k, \mu)\Big\} \end{aligned}$$

Computing the expected value gives

$$\begin{aligned} rV(s,\mu) = & \frac{r}{N\sigma + \delta + r} G\big(s(n), s(\mu_n)\big) + \\ & \frac{1}{N\sigma + \delta + r} \Big\{ \frac{\delta}{C} \sum_{\nu \in M} rV(s,\nu) + \\ & \sum_{m \neq n} \sum_k \lambda_n^k(s,\mu) rV(s_m^k, \mu) + \\ & \sigma \max_\pi \sum_k \big((1-\epsilon)\pi_k(s,\mu) + \epsilon q_k\big) rV(s_n^k, \mu)\Big\} \end{aligned} \tag{2.3}$$

The main difference between this Bellman equation and the Bellman equation for the discrete match model has to do with the treatment of matches. The first term in the continuous flow Bellman equation is the flow of utility from the current match. There is no such term in the discrete match model. The second term is different, too. The second term in the discrete match model is "making a match". The payoff from this is the immediate reward to the match plus the value of the current state. In the continuous flow model this term is just the expected value of the value function over all possible next matches.

3. Equilibrium

This section briefly covers the necessities of the equilibrium concept. The existence arguments are all straightforward exercises, and so little detail is provided.

Definition 3.1: A *rational population equilibrium* for the discrete match model is a policy function π such that for all states a of the opponent process and policy functions π', $E_{\pi(a)}V(\tilde{k},a) \geq E_{\pi'(a)}V(\tilde{k},a)$. A *rational population equilibrium* for the continuous flow model is a policy function π such that for all configurations and matches (s,μ), and policy functions π', $E_{\pi(s,\mu)}V(s_{-n},\tilde{s}_n,\mu) \geq E_{\pi'(s,\mu)}V(s_{-n},\tilde{s}_n,\mu)$. An *equilibrium strategy revision process* is the stochastic strategy revision process which results when all players use a rational population equilibrium policy.

The equilibrium strategy revision process for the discrete match model is constructed from the equilibrium policy function as follows. Let a denote a state and $b = a - e_k + e_l$ denote a state reachable from a in one transition. To reach b, a player choosing strategy k switches to strategy l. The transition rate from a to b is

$$\lambda_{ab} = \sigma a_k\big((1-\epsilon)\pi_l(a) + \epsilon q_l\big).$$

The term σa_k is the arrival rate of strategy revision opportunities to the group of players choosing k, and the second term is the probability that the strategy revision opportunity will conclude with the selected player choosing l.

The equilibrium strategy revision process for the continuous flow model is slightly more complicated than that for the discrete matching model because the process $\{s_t\}_{t=0}^{\infty}$ of configurations is not Markov. However the process $\{s_t,\mu_t\}$ of configurations and matches is Markov, and the construction is basically the same. The transition rate between (s,μ) and (s',μ) is 0 unless $s(m) = s'(m)$ for all but one player n. If this condition is met, then the transition rate is

$$\sigma\big((1-\epsilon)\pi_{s'(n)}(s,\mu) + \epsilon q_{s'(n)}\big)$$

The rate of change from (s,μ) to (s,μ') is δ/C.

Theorem 3.1: *For both models a rational population equilibrium exists.*

Proof: For both dynamic programs the solution correspondence will be non-empty-, convex- and compact-valued, and upper hemi-continuous with respect to the parameters of the opponent processes. These parameters are in turn continuous with respect to the opponents' policy, so a standard fixed-point argument proves the existence of equilibrium. □

One interesting extension of this analysis is to consider "mistake" or "experimentation" schemes where the probability of taking an action depends upon its rank order. This is a simple extension of the preceding analysis. Set $\epsilon = 1$. Let q_1 denote the probability of choosing the top-ranked element, q_2 the second best, and so forth. Now make the objects of choice rank orderings. The existence theorem for "proper" trembles is proven in exactly the same way as as Theorem 3.1.

In Blume (1993a) I suggested that one source of random noise is random utility. That is, in evaluating outcomes each player draws from a payoff distribution. Distributions are known, but the value of a draw is observed only by the player making the draw, and only at the moment of a match. The numbers in the payoff matrix are means of the payoff distributions. For the specific cases studied in Blume (1993a, 1994) the random utility model sets up nicely, but I do not have an existence proof at any level of generality. The problem is with the single-person decision problem. The random utility model would require that the probability distribution q depend upon the values $V(k,a)$. One would hope that something like the Bellman equation would determine these values. Unfortunately, with dependence on the $V(k,a)$ the Bellman equation need no longer have a unique fixed point. Fundamentally, the problem lies in developing a random utility choice theory which is consistent with some form of backwards induction — an interesting and perhaps important exercise, but beyond the scope of this paper.

4. Myopic Behavior

Myopic behavior emerges when the value function becomes proportional to the payoff function in every state. This section shows why myopic behavior emerges in both models as the discount factor r gets large. In addition, myopic behavior emerges in the discrete match model, but not in the payoff flow model, as the arrival rate of strategy revision opportunities becomes small. We will also see that the consequences of myopic behavior in the two models are quite different.

4.1. The Discrete Match Model

In the Bellman equation, the transition rate from state a to b is equal to the probability that a randomly chosen opponent will cause the transition from a to b times the arrival rate of strategy revision opportunities. That is, $\lambda^k_{ab} = \sigma\rho^k_{ab}$ where ρ^k_{ab} is greater than 0 and less than $N-1$, and $\lambda^k_a = \sigma\rho^k_a$ where $\rho^k_a = \sum_b \rho^k_{ab}$. Then the Bellman equation becomes

$$\frac{rV(k,a)}{v(k,a)} = \frac{r\delta}{r+\sigma(1+\rho^k_a)} + \sum_{b\neq a}\frac{\sigma\rho^k_{ab}}{r+\sigma(1+\rho^k_a)}\frac{rV(k,b)}{v(k,a)} + \frac{\sigma}{r+\sigma(1+\rho^k_a)}\max_\pi\big((1-\epsilon)\pi(a)_l\frac{rV(l,a)}{v(k,a)} + \epsilon\sum_l q_l\frac{rV(l,a)}{v(k,a)}\big)$$

The myopic limit is reached when the left-hand side is independent of k and a.

Theorem 4.1: *For any discount rate r and matching rate δ, the myopic limit $V(k,a)\propto v(k,a)$ is reached as σ becomes sufficiently small. For any σ and δ, the myopic limit is reached as r becomes large.*

The myopic limit is reached by making σ, the arrival rate of strategy revision opportunities, sufficiently small or r, the discount rate, sufficiently large. As σ shrinks the first term on the right-hand side converges to δ and all other terms converge to 0. One might argue that this is a back door into impatience. Decreasing σ increases the number of payouts between revision opportunities, which effectively changes players' rate of time preference. This argument is not correct. Formally, it states that if (1) the revision rate is changed, (2) time is rescaled to return the revision rate to its old value, and (3) the interest rate r is changed so that the present value of an arbitrary payoff stream is the same before and after the rescaling of time, then the value of the $V(k,a)/v(k,a)$ should change proportionately, and this is simply not the case.

Notice too that slowing down the rate of strategy revision opportunities is not the same thing as speeding up the matching rate. As δ grows, $V(k,a)/v(k,a)\Big/V(l,a)/v(l,a)$ converges to $(r+\sigma(1+\rho_a^l)/(r+\sigma(1+\rho_a^k)$. This ratio will typically not be constant.

Proof: Fix r and δ. The $V(k,a)$ are bounded from above as σ goes to 0 by the present value of the payoff stream that would result by making each match pay off at the highest payoff of the game, regardless of the state, and from below by the same construction applied to the lowest possible payoff. Clearly $rV(k,a)/v(k,a)$ converges to δ.

Now fix σ and δ, and consider the expression $rV(k,a)/v(k,a)$ as r grows large. Again the $rV(k,a)$ terms are bounded above and below uniformly in r. By recursive substitution we can see that for large r,

$$\frac{rV(k,a)}{v(k,a)} = \delta + O(r^{-1})$$

and so, in the limit, all left-hand sides converge to δ. □

Theorem 4.1 does not say that long-run considerations cease to matter in the myopic limit. In any state where there is not a unique myopic best-reply, the choice within the myopic best-reply set will be governed by long-run considerations, no matter how large r or how small σ. However it will generally be the case that such states are rare, and so information about the dynamics of equilibrium strategy revision processes can still be inferred from studying best-response to the instantaneous reward. In the case of two-by-two coordination games myopic play leads exactly to the play described by Kandori, Mailath and Rob (1993). Blume (1994) works out that model in the context of continuous time birth-death processes, and demonstrates a number of features of the strategy revision process, including its large-ϵ behavior and the robustness of the equilibrium selection result to a variety of perturbations of the model. All of those results carry over into the large r and small σ versions of the present model. In 2×2 coordination games there can be at most one state, the boundary state between the two basins of attraction, in which the immediate returns to the two strategies are equal. Moving the boundary backward or forward by 1 unit is not enough to change any part of the analysis of the model.

4.2. The Continuous Flow Model

In this model, the short run return for player n to the current state is $G\big(s(n), s(\mu_n)\big)$. Myopic behavior emerges when the ratio $V(s,\mu)/G\big(s(n), s(\mu_n)\big)$ is independent of s and

μ. Since

$$\min_{k,l} G(k,l) \leq rV(s,\mu) \leq \max_{k,l} G(k,l)$$

so the argument used for the discrete match model works here as well, and gives the following result:

Theorem 4.2: *Myopic behavior emerges for large discount rates:*

$$\lim_{r\to\infty} \frac{rV(s,\mu)}{G\big(s(n), s(\mu_n)\big)} = 1$$

uniformly in s and μ. If in the payoff matrix G, against each pure strategy there is a unique best response, then for sufficiently large r, each player's optimal policy is to maximize the return against the action of the current opponent.

If two $K \times K$ payoff matrices G have the same unique best responses to pure strategies, then myopia induces the same equilibrium strategy revision processes for the two matrices. All two-by-two coordination games have the same best responses to pure strategies, so they all have the same myopic behavior.

Consider now the special case of two-by-two coordination games, and denote the two strategies α and β. When players always best-respond, $\epsilon = 0$, the strategy revision process is not ergodic. There are only two limit states, $s(n) \equiv \alpha$ and $s(n) \equiv \beta$, and the process settles into one of these states in finite time and remains there forever. The probability of reaching a given limit state depends upon the initial configuration of play only through the initial distribution of play in the configuration. To see this, consider one copy of the process. Construct a new stochastic process from the first process simply by relabelling the names of players according to a fixed permutation at every step. Since the probabilities of any given player having a strategy revision opportunity or being matched with another particular player are all independent and equiprobable, it is easily seen that this second process is another version of the first. Since any two initial configurations of play with the same frequency of play are related through a permutation of names, we see thatonly frequencies of play matter. Let p denote the fraction of players choosing α in the initial configuration, and let $\phi(p)$ denote the probability that the limit configuration $s(n) \equiv \alpha$ is reached.

Theorem 4.3: *The function $\phi(a)$ has the following properties:*

1. *$\phi(p)$ is strictly increasing in p;*
2. *$\phi(p) = 1 - \phi(1-p)$;*
3. *$\phi(1/2) = 1/2$, $\phi(0) = 0$ and $\phi(1) = 1$.*

Proof: Fact 2 follows because the states α and β are symmetric: A process constructed from the strategy revision process by replacing every α with a β and every β with an α gives another version of the strategy revision process. Fact 3 is a consequence of fact 2.

Fact 1 is shown by a standard coupling argument, which goes as follows. We are going to build two copies of the strategy revision process that move together. Whenever a strategy revision opportunity arrives in the first process, it arrives for the same player in the second process and vice versa. Whenever a new match is drawn for the first process, the same match is drawn for the second process and vice versa. If the two processes started from the same initial configuration, they will move identically. Suppose now we assign initial configurations so that the set of players initially playing α in process 1 is a strict subset of the set of players playing a in process 2. The initial fractions of α players are $p_1 < p_2$, respectively. Whenever a player n in process 1 chooses α at a strategy revision opportunity, it will be because she is matched against an a player in process 1. By construction, this opponent is also an a player in process 2, so player n will choose α in process 2 as well. Consequently, whenever process 1 converges to "all play a", so does process 2: $\phi(p_1) \leq \phi(p_2)$. To see that the inequality is strict, consider a player who plays β in process 1 and a in process 2, and consider the following possibility: Of the first $2(N-1)$ events, the odd ones give her a partner she has never met before. The even events are strategy revision opportunities for her partners. At the end of these $2(N-1)$ events, everyone will have adopted the action she is playing, β in process 1 and α in process 2. This possibility occurs with positive probability, and so the inequality is strict. □

The $\epsilon = 0$ behavior for the myopic discrete match model is well-known. There is a p^* such that $v(p^*N, \alpha) = v(p^*N, \beta)$. Then

$$\phi(p) = \begin{cases} 1 & \text{if } p > p^*, \\ 0 & \text{if } p < p^*. \end{cases}$$

For small positive ϵ the joint distribution of configurations and matches is ergodic, and therefore so is the projection down onto configurations. But the limit invariant distribution as ϵ becomes small depends on the distribution q of tremble outcomes. This fact is not true in the discrete match model and in the other popular stochastic population models. For small positive ϵ the joint distribution of configurations and matches is ergodic, and therefore so is the projection down onto configurations. But the limit invariant distribution as ϵ becomes small depends on the distribution q of tremble outcomes. This fact is not true in the discrete match model and in the other popular stochastic population models.

The Markov process on configurations and matches has two recurrent communication classes: $C_\alpha = \{s(n) \equiv \alpha\} \times M$ and $C_\beta = \{s(n) \equiv \beta\} \times M$. Each communication class can be reached from the other by a path requiring only one "mistake", so according to Young's (1993) Theorem they both have positive probability in the limit distribution. Not surprisingly, these probabilities are proportional to the probability of the trembles needed to reach each limit set. The proof directly exploits the idea that the odds-ratios are proportional to the ratio of taboo probabilities for the passage from one class to the other without first returning to the initial class, which in turn is related to the number of mistakes along the shortest path. Thus the calculations here are closely related to Ellison's (1995) and Evans' (1993) comparisons of the radius and co-radius of basins of attraction.

Theorem 4.4: *For all strictly positive $q = (q_\alpha, q_\beta)$, as ϵ tends to 0, the probability of C_α and C_β under the invariant distribution converges to q_α and q_β, respectively.*

Proof: Consider the process of "events", that is, the cumulative process of strategy revision opportunities and matches. Events arrive at rate $\rho = N\sigma + \delta$, *independent of the history of previous events or the configuration of play.* Any given event has probability $N\sigma/\rho$ of being a strategy revision opportunity for some player, and probability δ/ρ of being a new match. Conditional on being a strategy revision opportunity, the probability of it belonging to any paticular player is $1/N$. And, conditional on it being a match, all matches are equally likely.

From the event process the process $\{s_t, \mu_t\}_{t=0}^{\infty}$ can be constucted using the strategy π, ϵ and q. Finally, suppose we construct a new event process from an old event process by permuting names. That is, there is a permutation p of the players indices such that any time a strategy revision opportunity arrives for i in the old process, it arrives for $p(i)$ in the new process, and if the pair (i, j) is in the match μ_t for the old process, the pair $(p(i), p(j))$ is in the corresponding match for the new process. Because everything in sight is equiprobable, the distribution of the new process is the same as that of the old process. It is another event process, and so it too will generate another strategy revision process.

Let τ_k denote the kth time the process $\{s_t, \mu_t\}_{t=0}^{\infty}$ has an event. Because of the strong Markov property, the process $\{s_k, \mu_k\}_{k=0}^{\infty} = \{s_{t_k}, \mu_{t_k}\}_{k=0}^{\infty}$ is a discrete time Markov chain. Because the distribution of the waiting times $\tau_k - \tau_{k-1}$ is independent of the state $(s_{\tau_k}, \mu_{\tau_k})$, the invariant distribution ν for the process $\{s_k, \mu_k\}_{k=0}^{\infty}$ is the same as that for $\{s_t, \mu_t\}_{t=0}^{\infty}$. Now choose a state $m_\alpha = (s^\alpha, \mu^\alpha) \in C_\alpha$ and another state $m_\beta = (s^\beta, \mu^\beta) \in C_\beta$. (Of course $s^\alpha(n) \equiv \alpha$ and $s^\beta(n) \equiv \beta$.) Again let k_i denote the ith time (s_k, μ_k) is one of m_α or m_β, and consider the process $\{s_i, \mu_i\}_{i=0}^{\infty} = \{s_{k_i}, \mu_{k_i}\}_{i=0}^{\infty}$. This process too is Markov, and it is a standard exercise to show that its invariant distribution ν' has the property that $\nu'(m_\alpha)/\nu'(m_\beta) = \nu(m_\alpha)/\nu(m_\beta)$.

The transition matrix for the $\{s_i, \mu_i\}_{i=0}^{\infty}$ process is

$$\begin{pmatrix} 1 - \kappa(\mu_\alpha, \mu_\beta)\epsilon q_\beta + O(\epsilon^2) & \kappa(\mu_\beta, \mu_\alpha)\epsilon q_\alpha + O(\epsilon^2) \\ \kappa(\mu_\alpha, \mu_\beta)\epsilon q_\beta + O(\epsilon^2) & 1 - \kappa(\mu_\beta, \mu_\alpha)\epsilon q_\alpha + O(\epsilon^2) \end{pmatrix}.$$

To see this, observe that the probability of moving from m_α to m_α without hitting m_β is

$$\kappa_(\mu_\alpha, \mu_\beta)\epsilon q_\beta + \sum_{k,l\,:\,k+l>1} \kappa_{k,l}\epsilon^{k+l} q_\beta^k q_\alpha^l$$

where $\kappa(\mu_\alpha, \mu_\beta)$ is the probability of observing a path that requires only one tremble, and $\kappa_{k,l}$ is the probability of drawing a path that requires k trembles to β qnd l trembles to α. These are probabilities of observing particular sequences of matches and strategy revision opportunities, and therefore do not depend upon the parameters ϵ, q_α and q_β. The κ

coefficients are all non-negative and less than 1, so this expression is $\kappa(\mu_\alpha, \mu_\beta)\epsilon q_\beta + O(\epsilon^2)$. There is no ϵq_α term because a path that reaches m_β from m_α with only one tremble must have a tremble to β.

Solving for the invariant probabilities shows that

$$\frac{\nu(m_\alpha)}{\nu(m_\beta)} = \frac{\kappa(\mu_\beta, \mu_\alpha)q_\alpha}{\kappa(\mu_\alpha, \mu_\beta)q_\beta} + O(\epsilon).$$

Thus

$$\frac{\mu(C_\beta)}{\mu(C_\alpha)} = \frac{q_\beta}{q_\alpha} \sum_{\mu'} \frac{1}{\sum_\mu r(\mu, \mu')} + O(\epsilon)$$

where $r(\mu, \mu') = \kappa(\mu', \mu)/\kappa(\mu, \mu')$. In general the $O(\epsilon)$ term will depend upon q_α and q_β.

Now let $q_\alpha = q_\beta = 1/2$. Symmetry arguments show that the ratio on the left must equal 1, so $\sum_{\mu'} \big(1/\sum_\mu r(\mu, \mu')\big) = 1$. $\square$

4.3. Lock-in in the Continuous Flow Model

In the discrete match model the large r and small σ equilibrium behavior are identical. This is not the case in the continuous flow model. Although the large r equilibrium behaviors of the continuous flow model and discrete match model are different, the small σ equilibrium behavior of the two models are the same.

As σ becomes small in the continuous flow model the Bellman equation converges to

$$rV(s, \mu) = \frac{r}{\delta + r} G\big(s(n), s(\mu_n)\big) + \frac{\delta}{\delta + r} \frac{1}{C} \sum_{\nu \in M} rV(s, \nu)$$

This equation has a unique solution, so it is the limit of the solution for the Bellman equations as σ tends to 0. Suppose the solution is of the form $rV(s, \mu) = \alpha G\big(s(n), s(\mu_n)\big)$ with $\alpha > 0$. Then

$$G\big(s(n), s(\mu_n)\big) = \frac{r}{\delta + r} G\big(s(n), s(\mu_n)\big) + \frac{\delta}{\delta + r} \frac{1}{C} \sum_{\nu \in M} G\big(s(n), s(\nu_n)\big)$$

and it follows that

$$G\big(s(n), s(\mu_n)\big) = \frac{1}{C} \sum_{\nu \in M} G\big(s(n), s(\nu_n)\big)$$

which happens only when payoffs do not depend upon the opponent's action.

The intuition is straightforward. If a player will be locked into her choice for a very long time and if she is somewhat patient, she will be concerned both about her current match and the matches she will have in the future. In fact, the solution to the limit value equation can be found from a simple computation.

Theorem 4.5: *As the revision rate σ becomes small,*

$$\lim_{\sigma \to 0} V(s, \mu) = \frac{r}{\delta + r} G\big(s(n), s(\mu_n)\big) + \frac{\delta}{\delta + r} \frac{1}{C} \sum_{\nu \in C} G\big(s(n), s(\mu_n)\big).$$

When players are locked in to their current strategy choices for long times, their optimal strategy will depend upon how patient they are. Impatient players will best-respond to their current opponent while patient players will best-respond to the play of the entire population. The intuition is straightforward. If a player will be locked into her choice for a very long time and if she is somewhat patient, she will be concerned both about her current match and the matches she will have in the future. In fact, the solution to the limit value equation can be found from a simple computation.

In two-by-two coordination games Theorem 4.5 implies risk dominant equilibrium selection when σ is small andeither r is small or δ is large. Let p^* denote the fraction of the population playing strategy α such that, against p^*, the expected payoff return in G to α and β are identical. If $p^* < 1/2$ then α is risk dominant. Even in the non-generic case where fraction p^* could arise in the population (Np^* is an integer), this affects the behavior of the model in only one state, and for large enough populations this will have no effect on the equilibrium selection result.

Corollary 4.1: *In any two-by-two coordination game with a risk-dominant strategy, for small enough σ and either small enough r or large enough δ, as ϵ becomes small, the unique equilibrium is to choose strategy α (β) precisely in those states where α (β) is a myopic best response to the current distribution of play. The invariant distribution of the equilibrium strategy revision process puts weight tending to 1 on the state in which all players choose the risk-dominant strategy. For small σ, myopic play emerges when r is large or δ is small.*

Thus the conventional risk-dominant selection result emerges in the continuous flow model, but only as σ becomes small. Risk-dominant selection also occurs in the discrete match model for small σ, but in the continuous flow model and not in the discrete match model there is the additional requirement that the discount factor r be sufficiently small relative to the matching rate δ.

5. Patient Players in 2×2 Coordination Games

This section characterizes a class of equilibria in two-by-two coordination games when players are patient. For such games the state space of the opponent process can be taken to be $\{0, \ldots, N-1\}$, which records the number of opponents playing α. Then $V(\alpha, m)$ is the value to playing α when m opponents are playing α. The domain of a policy is the probability of playing α.

Definition 5.1: An equilibrium is *monotonic* if $V(\alpha, m) - V(\beta, m)$ is strictly increasing or strictly decreasing in m.

In a monotonic equilibrium, $\pi(m)$ is non-decreasing in m. Furthermore, if $\pi(m) > 0$, then $\pi(m+1) = 1$.

Although the discrete match and continuous flow models exhibit distinct myopic behavior, they behave similarly, at least for monotonic equilibria, when players are patient.

Theorem 5.1: *In the discrete-match model, for all δ and σ and for r and ϵ sufficiently small, the following conclusions are true:*

i. There is a monotonic equilibrium in which all players choose the risk-dominant strategy in every state.

ii. Choosing the risk-dominated strategy in every state is not an equilibrium.

iii. If there is a unique payoff dominant equilibrium, then always choosing the risk-dominant strategy is the unique monotonic equilibrium.

In the continuous flow model, for all δ and σ and for all positive ϵ sufficiently small there is an $\bar{r}$ such that for all $r < \bar{r}$ the same conclusions hold.

This theorem makes a claim which initially may appear to be counter-intuitive. When players are sufficiently patient and the stochastic perturbations are sufficiently small, coordination on the risk-dominated strategy disappears, and the basin of attraction for coordination on the risk-dominant strategy is the whole state space. This result proves false the conjecture that patient players could achieve equilibrium coordination on the payoff-dominant strategy when it is not also risk-dominant.

The intuition behind the result is mostly straightforward. First, the intuition about payoff dominance is partly correct. When players care about the time-average of payoffs, "always play α" and "always play β" are both rational population equilibria. So the theorem really demonstrates a failure of lower-semicontinuity of the equilibrium correspondence at $r = 0$, and the fact that some discounting occurs is crucial to the result.

A computation shows that always playing the risk-dominant strategy is an equilibrium. Another possibility is an equilibrium with a threshold state m^* above which all players play α and at or below which all players choose β. This is ruled out because at the threshold — either the last α state or the first β state, some player will have the option to move the process across it — an α player in the first case, and a β player in the second. If ϵ is sufficiently small, the process will move very, very quickly to the sink of the basin of attraction in which the pivotal player puts it, and it will stay there a very, very long time. In this circumstance the patient pivotal player will always want to choose the payoff dominant outcome. Consequently m^* cannot be the threshold.

The remaining possibility is all players always choosing the risk-dominated equilibrium. Here the proof comes down to a calculation. It is easily seen that the benefit of deviating to the risk-dominant strategy in any state is small, and shrinks to 0 as r becomes small, but is positive.

The remainder of this section is devoted to the proof of Theorem 5.1. The proof calculations are the same in both models. The continuous flow model requires an additional step because it must be shown that as r becomes small the influence of the current match on the value of a choice disappears. We begin with the discrete match model. There are two strategies, α and β. Let M denote the number of opponents playing α. Suppose that $\pi(\,\cdot\,)$ is a rational population equilibrium ($\pi(M)$ is the probability that a player tries to play α when M opponents are playing α), and let $\rho(M)$ denote the probaility that α is played when M opponents are playing α. That is,

$$\rho(M) = (1-\epsilon)\pi(M) + \epsilon q_\alpha.$$

Let $\rho'(M) = 1 - \rho(M)$ denote the probability that β is played The value function is

$$\begin{aligned}
V(\alpha, M) =& \frac{\delta}{r+\sigma N} v(\alpha, m) + \\
& \frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho(M+1)V(\alpha, M+1) + \rho'(M+1)V(\alpha, M)\Big) + \\
& \frac{\sigma M}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\alpha, M-1)\Big) + \\
& \frac{\sigma}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\beta, M)\Big) \\
V(\beta, M) =& \frac{\delta}{r+\sigma N} v(\beta, m) + \\
& \frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho(M)V(\beta, M+1) + \rho'(M)V(\beta, M)\Big) + \\
& \frac{\sigma M}{r+\sigma N}\Big(\rho(M-1)V(\beta, M) + \rho'(M-1)V(\beta, M-1)\Big) + \\
& \frac{\sigma}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\beta, M)\Big)
\end{aligned}$$

The proof technique is to specify a ρ, take appropriate limits of parameters, and check the value function to see if the optimal policy generates the ρ that began the calculation.

The first calculation shows that for small enough r and ϵ, $\pi(M) \equiv 1$ (always choosing α) is an equilibrium if α is risk dominant, and is not if α is risk-dominated. Let $\Delta(M) = V(\alpha, M) - V(\beta, M)$, and let $\Delta_v(M) = v(\alpha, M) - v(\beta, M)$. The optimal policy π assigns positive probability to α only if $\Delta(M) \geq 0$. Computing,

$$\Delta(M) = \frac{\delta}{r+\sigma N}\Delta_v(M) +$$

$$\frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho(M+1)V(\alpha,M+1)-\rho(M)V(\beta,M+1)+$$
$$\rho'(M+1)V(\alpha,M)-\rho'(M)V(\beta,M)\Big)+$$
$$\frac{\sigma M}{r+\sigma N}\Big(\rho(M)V(\alpha,M)-\rho(M-1)V(\beta,M)+$$
$$\rho'(M)V(\alpha,M-1)-\rho'(M-1)V(\beta,M-1)\Big)$$

The policy $\pi(M)\equiv 1$ is an equilibrium if and only if, when $\rho(M)\equiv 1-\epsilon$, $\Delta(M)\geq 0$ for all M. When $\rho(M)\equiv\rho$,

$$\Delta(M)=\frac{\delta}{r+\sigma N}\Delta_v(M)+$$
$$\frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho\Delta(M+1)+(1-\rho)\Delta(M)\Big)+$$
$$\frac{\sigma M}{r+\sigma N}\Big(\rho\Delta(M)+(1-\rho)\Delta(M-1)\Big)$$

First, suppose $\epsilon=0$. Then $\rho=1$ and the difference equation becomes

$$\Delta(M)=\frac{\sigma(N-M-1)}{r+\sigma(N-M)}\Delta(M+1)+\frac{\delta}{r+\sigma(N-M)}\Delta_v(M).$$

Then

$$\lim_{r\to 0}\Delta(M)=\frac{\delta}{\sigma(N-M)}\Big(\Delta_v(M)+\cdots+\Delta_v(N)\Big).$$

It is easy to see that the right-hand side is positive for all $M\geq\max\{0,(2p^*-1)(N-1)\}$ and negative otherwise if N is large enough. The difference function $\Delta_v(M)$ is linear in M and equals 0 at $p^*(N-1)$. Therefore the claim is true so long as N is large enough that $(2p^*-1)(N-1)\geq 1$ if $p^*>1/2$. Thus $\pi(M)\equiv 1$ is optimal for all r sufficiently close to 0 if $p^*<1/2$, and is not optimal if $p^*>1/2$. Finally, the solution to the difference equation is continuous in ϵ at $\epsilon=0$, so for given r sufficiently small there is an $\hat{\epsilon}$ such that for all $\epsilon<\hat{\epsilon}$, if α is risk-dominant then $\pi(M)\equiv 1$ is optimal, and if α is risk-dominated, then $\pi(M)=1$ is not optimal.

The only other possibility for a monotonic equilibrium is the existence of an M^* between 0 and $N-1$ such that $\pi(M)=0$ for $M<M^*$ and $\pi(M)=1$ for $M\geq M^*$. For this the normalized value functions are solved. Suppose $\epsilon=0$, so that $\rho(M)=1$ for $M\geq M^*$ and $\rho(M)=0$ for $M<M^*$. Calculations show

$$\lim_{r\to 0} rV(\alpha,M)=\begin{cases}\delta v(\alpha,N-1) & \text{if } M\geq M^*,\\ \frac{N-M^*+1}{N}\delta v(\alpha,N-1)+\frac{M^*-1}{N}\delta v(\beta,0) & \text{if } M=M^*-1,\\ \delta v(\beta,0) & \text{if } M<M^*-1.\end{cases}$$

and

$$\lim_{r\to 0} rV(\beta,M) = \begin{cases} \delta v(\alpha,N-1) & \text{if } M > M^*, \\ \frac{N-M^*}{N}\delta v(\alpha,N-1) + \frac{M^*}{N}\delta v(\beta,0) & \text{if } M = M^*, \\ \delta v(\beta,0) & \text{if } M \le M^* - 1. \end{cases}$$

Therefore

$$\lim_{r\to 0} r\Delta(M^*) = \frac{M^*}{N}\delta\Big(v(\alpha,N-1) - v(\beta,0)\Big)$$

and

$$\lim_{r\to 0} r\Delta(M^*-1) = \frac{N-M^*}{N}\delta\Big(v(\alpha,N-1) - v(\beta,0)\Big)$$

Both of these have the same sign, so it is not possible that $\pi(M^*) = 1$ and $\pi(M^*-1) = 0$ unless $v(\alpha,0) = v(\beta,N)$. Again the conclusions remain valid for r sufficiently small and for ϵ sufficiently small given r. This completes the proof for the discrete match model.

For the continuous flow model there is the added complication that payoffs depend on the distribution of play and on the current match. Let $V(x,y,M)$ denote the return to playing x when M opponents play α and the current opponent plays y. Examining the value functions shows that the normalized payoffs $rV(x,y,M)$ and the payoff differences $V(\alpha,y,M) - V(\beta,y,M)$ are uniformly bounded for all M, ϵ and r.

For all $\epsilon > 0$ the strategy revision process for any given policies is strongly ergodic, and therefore the conditional payoffs at date t_0+h given the choice of action at date t_0 becomes independent of that choice for large h. Consequently for r sufficiently small the date t_0 normalized payoffs and the payoff differences are approximately independent of the date t_0 action of the current opponent. Thus for r small enough given ϵ the optimal policy can be identified (if there are no ties) by studying the expected value functions $\rho(M)V(x,\alpha,M) + \rho'(M)V(x,\beta,M)$. With these expected value functions, the same calculations carried out for the discrete match model also give the same results for the continuous flow model in exactly the same manner. □

Notice that $\lim_{r\to 0} r\Delta(M) = 0$. This is the source of the failure of lower-hemicontinuity of the equilibrium correspondence at $r = 0$. The question of uniqueness of the monotonic equilibrium when $v(\alpha,N) = v(\beta,0)$ is open. Further calculations are required to see if the same strict inequalities hold at any conjectured M^*.

The difference in hypotheses between the discrete match and continuous flow models is due to the approximation argument. The time h at which the date $t_0 + h$ conditional payoffs become independent of y to a given degree depends upon ϵ. Then r must be made sufficiently small that the payoffs are approximately determined by the behavior of the process out beyond date $t_0 + h$.

6. Conclusion

This paper studies the evolution of play in a population of players continually interacting with one another. Two models of player interaction are compared. In both models players are repeatedly and randomly matched against opponents. Players revise their strategic choices only at discrete random moments, each player independent of the others. The two models are built from Poisson processes in continuous time and differ only in the exact technology of player interaction: Discrete random encounters — the billiard ball model — versus continuous interaction randomly interrupted by the making of new matches.

Many people have argued that the risk-dominant selection results in the population-theoretic approach to coordination games are artifacts of impatience and irrelevant because, at least in the mutation models of Kandori, Mailath and Robb (1993) (although not necessarily in Young's (1993) model) the time of transition between the two fully-coordinated states are too long to make the long run analysis of interest.

In this paper we have seen two versions of risk dominant selection. In its weak form, both coordination outcomes have basins of attraction, but that of the risk-dominant outcome is larger, and so it is selected for as the stochastic element in choice is made small. In its strong form, coordinating on the risk-dominant outcome is the only state with a basin of attraction, and that basin of attraction is the entire state space.

The analysis of this paper shows that myopic risk-dominant equilibrium selection is indeed an artifact. It is present in its weak form in the discrete match model, and not at all as a consequence of myopia in the continuous flow model. It also appears in its weak form in both models as a consequence of lock-in. But in both models risk-dominant selection occurs in its strong form as a consequence of patient play. When risk dominant selection is present in its weak form, the waiting time to move from coordination on the non-dominant action to coordination on the risk-dominant action can be long for small ϵ (although how long is "too long" is not a question that can be answered outside of a concrete application.) But strong risk-dominant selection implies that the waiting times will be small for small ϵ. More telling is the fact that the waiting times are increasing in ϵ for weak selection and decreasing in ϵ for strong selection. In every sense the natural setting for the emergence of risk-dominant equilibrium selection is not in myopic models but in models with forward-looking players.

The typical myopic formulations of equilbrium selection separate the analysis of choice from the dynamics of the process. Obviously this separation cannot be maintained under more realistic assumptions of forward-looking play. But the intertwining of choice and dynamics need not lead to intractable models. This paper shows that introducing forward-looking play can lead to tractable models and interesting results.

Finally, this paper shows that the details of interaction matter. The continuous flow and discrete match models lead to distinct behavior for myopic players. Yet these two models differ only in the technology of player interaction. One promise of the population games research program is to identify just how the way in which players interact with one another matter. The two models considered here do not exhaust the possibilities. For instance, consider the following variation of the continuous-flow model: Players form matches, and each pair stays together for a random amount of time. Two players can only enter into a new match if they are both free. Strategy revision opportunities come to players both when they are matched up and when they are free. In one extreme version, players can only utilize strategy revision opportunities when matched up. This gives continuous-flow dynamics. In the other extreme, players can only use strategy revision opportunities when unmatched. This leads to discrete-match dynamics. One wonders if in the case where all players can use all strategy revision opportunities, the drift caused by the unmatched players leads to discrete-match dynamics (albeit very slowly) or whether some mixed regime arises. This and other questions will be left for future research.

REFERENCES

Binmore, K., L. Samuelson and R. Vaughan (1993) "Muddling through: Noisy equilibrium selection," *Games and Economic Behavior,* forthcoming.

Blume, L. (1993a) "The statistical mechanics of strategic interaction," *Games and Economic Behavior,* 4, 387–424.

———— (1993b) "The statistical mechanics of best response strategy revision," *Games and Economic Behavior,* forthcoming.

———— (1994) "How noise matters," unpublished, Cornell University.

Canning, D. (1990) "Social Equilibrium," Economic Theory Discussion Paper No. 150, Cambridge University.

Diamond, P. (1982) "Aggregate demand management in search equilibrium," *Journal of Political Economy,* 90, 881–94.

Ellison, G. (1995) "Basins of attraction and long run equilibria," unpublished, MIT.

Evans, R. (1993) "Observability, Imitation and Cooperation in the Repeated Prisoners' Dilemma," unpublished, University of Cambridge.

Foster, D. and H. P. Young (1990) "Stochastic evolutionary game dynamics," *Theoretical Population Biology,* 38, 219–32.

Fudenberg, D. and C. Harris (1992) "Evolutionary dynamics in games with aggregate shocks," *Journal of Economic Theory,* 57, 420–41.

Kandori, M., G. Mailath and R. Rob (1993) "Learning, mutation, and long-run equilibria in games," *Econometrica,* 61, 29–56.

Kiyotaki, N. and R. Wright (1989) "On money as a medium of exchange," *Journal of Political Economy,* 92, 927–54.

Menger, C. (1892) "On the origin of money," *Economic Journal,* 2, 239–55.

Samuelson, L. (1994) "Stochastic stability in games with alternative best replies," *Journal of Economic Theory,* 64, 35–65.

Young, H. P. (1993) "The evolution of choice behavior," *Econometrica,* 61, 57–84.