Evolutionary Equilibrium with Forward-Looking Players[†]

Lawrence E. Blume
Department of Economics, Cornell University
and
The Santa Fe Institute

January 2004

[†] This research was supported by the National Science Foundation, the John D. and Catherine T. MacArthur Foundation and The Pew Charitable Trusts and the Santa Fe Institute. It was begun during visits to the Institute for Advanced Studies at The Hebrew University and the Department of Economics at Tel Aviv University.

I. Social Norms and Coordination

Conventions and norms are the glue that holds society together. When social outcomes appear to be coordinated or predictable, the cause is most often not a remarkable correlation of taste or concerted collective action, but individual adherence to well-understood and shared guides to normative behavior. This is not to say that social order is explained by norms and conventions; instead, an explanation of social order is rooted in an understanding of their origins and consequences.

In this paper I summarize the case for evolutionary game theory as a framework for the investigation of the origins and persistence of norms and conventions. One problem with evolutionary models is that individuals are not forward-looking, a property which lies at the heart of the arguments over the application of rational choice theory to the problem of social order. I demonstrate here one way to address this problem, and demonstrate some of the consequences of taking the future seriously.

A. Homo Economicus vs. Homo Sociologicus

Homo economicus and *homo sociologicus* are the two straw men of social science.[1] For *homo economicus*, each social act is a considered choice, an exercise in naked self-interest. For *homo sociologicus* there is no choice. Man is a boat without rudder, drifting at the mercy of the powerful tides of social forces.

The substantive question underlying these caricatures is the scope of instrumental behavior. Each of us at different times pursue material gain, consciously follow social custom, act reflexively from habit or tradition and respond to emotions. Weber (1947) developed a typology of action which distinguished affectual choice, seated entirely in emotional responses; traditional actions, rooted in social custom; value-rational or axiological choice, which flows from assertions of belief, regardless of consequences; and instrumental choice, behavior whose purpose is to produce particular outcomes in order to achieve specific goals. The domains of these types are fluid. What was once rational becomes customary; and what was customary or emotionally determined can be reexamined rationally.

The answer to the question of when and why individuals feel and act upon social obligations, or act in altruistic ways, is contested by these two explanatory traditions. Some rational choice theorists argue that obedience to norms is a direct expression of self-interest. Norm obedience directly generates utility. For example, Riker and Ordeshook (1968) and Tullock (1967) eliminate the paradox of voting by having the act of voting itself generate utility beyond its (negligible) instrumental effect on the outcome. Others, such as Ellison (1994) and Kandori (1992), view social norms as implicit long-term contracts. When social interaction is recurrent, reciprocity is rational. These explanations are not without problems. Making norm obedience an argument of a utility function simply pushes back the explanatory problem. How did norms get there? Why do some social prescriptions become normative while others do not?[2] The long-term contracting explanations are based on what some might regard as a weakness of dynamic game theory. When social interaction is recurrent, reciprocity is rational if individuals are sufficiently patient. But there are many distinct self-consistent social formations. Nearly every outcome is rational if individuals are sufficiently patient. This is the content of the so-called 'folk theorems' for repeated games. Now the problem is pushed back to equilibrium selection. Why (or when) do we observe those equilibria which are norm-respecting and not some others?

Homo sociologicus offers a different set of answers. Much effort has been expended demonstrating that commitment to norms is complementary to, rather than consequent of, instrumental behavior.[3] As put succinctly by Durkheim, 'It is therefore in the nature of society itself that we must seek the explanation of social life.'[4] I will not attempt to summarize this vast and varied literature here, but I claim that any such survey would conclude that what is being rejected by this tradition is not rational action per se, but methodological individualism.[5] Thus this tradition is not only at odds with rational choice theory; it is also in opposition to exchange theorists such as Homans, whose individualist views were developed from behaviorist rather than utilitarian principles. The methodological holist position does not preclude understanding the functionality of social norms, and thus, to some degree, their persistence, for any truly dysfunctional norm would adversely affect the future of any society that took it up. But it is less helpful in understanding the origins of norms, and functionality alone cannot explain persistence.

Having given up on explanations of adherence to norms that are based solely on individual incentives, some social scientists who hold to the holist tradition nonetheless recognize that the active agents in society are individuals, and that norms emerge as an aggregate of their activities. Durkheim writes:[6]

> ... society is not the mere sum of individuals, but the system formed by their association represents a specific reality which has its own characteristics. Undoubtedly no collective entity can be produced if there are no individual consciousnesses: this is a necessary but not a sufficient condition. In addition, these consciousnesses must be associated and combined, but combined in a certain way. It is from this combination that social life arises and consequently it is this combination which explains it. By aggregating together, by interpenetrating, by fusing together, individuals give birth to a being, psychical if you will, but one which constitutes a psychical individuality of a new kind.

For Durkheim, social life, and in particular, the 'collective consciousness', is an emergent property of a system of interacting individuals.[7] Evolutionary game theory offers one approach to the study of emergent properties in systems of interacting agents. The evolutionary game theory program is to replace or supplement highly detailed explorations of individual rationality in a social environment with an examination of social interactions and rough descriptions of adaptive behavior. The outcome of an evolutionary game theory analysis is a description of equilibrium as a population state, that is to say, as an emergent property of the system whose interactions are captured in the evolutionary model.[8] As a method for modelling emergence, evolutionary game theory is one bridge across the individualism/holism gap.

B. Emergence and Equilibrium Selection

In the context of evolutionary game theory, social norms are often understood as solutions to coordination problems which arise in a population of individuals large enough that most members never directly interact with one another, thus precluding the possibility of effective collective action. The process of norm emergence has been modeled in evolutionary game theory as a population of individuals engaging in multiple strategic interactions in which some actions are thought of as 'norm adhering' and others as 'norm defecting'. Generally speaking, the more a norm is adhered to, the greater the cost which is borne by any individual who breaks the norm (although individual choice behavior is loosely specified and idiosyncratic). Thus the strategic interaction is a coordination game. Evolutionary dynamics then describes the pattern of norm adherence through time.[9] The simplest coordination games to study are two-by-two games, where two players each choose from among two actions.

Elementary evolutionary game theory, which is to say, static equilibrium concepts such as evolutionarily stable strategies and deterministic dynamic models such as replicator dynamics, has little to say about coordination games beyond the observation that polymorphic (mixed) equilibria may not be robust. In particular, these models cannot distinguish the different pure coordination outcomes. In the last decade, however, game theorists have come to understand the importance of agent heterogeneity and idiosyncratic behavior for selecting among equilibria. Blume (1993), Kandori, Mailath and Rob (1993) and Young (1993), among others, have shown how, with different microdynamics and different kinds of strategic interaction, small amounts of randomness in individuals' decision processes have a large impact on the determination of *which* equilibrium is most likely to appear.

Stochastic adjustment models like these have four distinguishing characteristics: A *matching technology*, which describes how individuals in the population interact with one another; *myopic agents*, who do not value the future at all but consider only the present in making a choice; *inertia in choice*, which is to say that agents cannot continually revise their choice but are to some degree locked in to their current action; and *noisy choice* which is a consequence either of random idiosyncratic utility or limited rationality. Inertia and noisy choice in particular speak to some of the issues raised by critics of rational choice theory. Inertia is the force of habit and tradition. Only occasionally are choices rationally examined for their instrumental value, and payoff perturbations or simple noise in choice are a first pass at accounting for a non-instrumental component in choice.

The cost of focusing on interactions in evolutionary game theory models has been the oversimplification of the individual. Indeed utility maximization is a component of choice, but the optimization problem is so spare that it is hard to say that the decision model is in any sense rational. In particular, the contrast between *homo economicus* and *homo sociologicus* is often posed as a contrast between forward-looking and backward-looking behavior. *Homo economicus* is concerned with beliefs about the effects of present intentions on future consequences, while *homo sociologicus* looks backwards to pre-existing social structure. But the actors in most evolutionary game theory models have no concept of a future. Beliefs about where the social process may be headed have no impact on current choice. The purpose of this paper is to bring consideration of the future into the evolutionary analysis of two-by-two games, and to examine its implications for emergence.

The surprising finding of this paper is the effect of patience on equilibrium selection. The natural conjecture is that, while the distribution of play is centered on the risk-dominant action when individuals are myopic, payoff dominance would emerge from the actions of patient individuals. This is not the case for a reasonable class of equilibria. I investigate *monotonic equilibria*, those with the property that individuals follow suit: As more individuals play a given choice, the more likely a given decisionmaker will make the same choice. I show that when there is a unique risk-dominant outcome in the static game, the dynamic game with sufficiently patient individuals has a unique equilibrium in which all individuals *always* choose the risk-dominant choice. In games where risk- and payoff-dominance differ, I conclude that the unique monotonic equilibrium has all individuals always coordinating on the low-payoff outcome. This finding defeats the intuitive analysis that suggests coordination failure is a consequence of myopia.

II. The Model

The model envisions a population of individuals who are matched at random moments for an instantaneous strategic interaction, just as two billiard balls might collide while both are in motion. The payoffs from the match are determined by the actions the two individuals hold at the instant of the match. At other distinct random moments, individuals can revise their strategic choice. Although the model as described is quite abstract, applications of this idea to economics are manifold: For instance, the Menger-Kiyotaki-Wright model of monetary exchange[10] and some versions of Diamond's (1982) search equilibrium model.

Models in the population game literature have been implemented in three ways. Canning (1992), Kandori et al. (1993), Samuelson (1994) and Young (1993) 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 (1993, 1994 and 2003) implements this model as a continuous-time jump process. In particular, when all individuals are ex-ante identical, this formulation is a continuous-time birth-death process. Binmore, Samuelson and Vaughn (1995) 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 individuals 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 continuous time construction with discrete events and random matching as developed in Blume (2003).

A *stochastic strategy revision process* is a population process which describes the distribution among a set of strategies of a population of individuals. The population has N individuals. Given too is a payoff matrix for a 2×2 symmetric coordination game G with strategy set $\{\alpha, \beta\}$. Each individual is labelled with a strategy. The *state* of the population process at time t is the number M_t of individuals choosing α.

From time to time each individual has an opportunity to revise her strategy. An individual's *policy* is a map that assigns to each state a probability of choosing α if given the opportunity to revise her strategy at an instant when the process is in that state (β will be chosen with the complementary probability). Formally, a policy is a map $\pi : \{0, \dots, N-1\} \to [0,1]$ with the interpretation that $\pi(M)$ is the probability that the decision-maker chooses α when M of the $N-1$ other individuals n the population are already choosing α. The constraint that the optimal policy depends only upon the current state, that policies are *Markovian*, will be justified in equilibrium.

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 individual 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 individual n is $\sum_{l=1}^{m} x_{nl}$.

When an individual's expected utility calculation indicates that a particular action is best, that choice may not in fact be implemented. Fix an ϵ strictly between 0 and 1. With probability $1-\epsilon$ the nominal expected-utility-maximizing choice is realized, and with probability ϵ the new choice is chosen by a draw from the distribution which assigns probability q to α. The probability that a decision-maker using policy π will choose α is

$$\rho_\pi(M) \equiv (1-\epsilon)\pi(M) + \epsilon q \, .$$

The probability that a β-chooser will choose to remain with β is $\rho'_\pi(M+1) = 1 - \rho_\pi(M+1)$. One can think of the deviation from notional expected utility maximization as either the effect of an idiosyncratic random utility term or as the consequence of boundedly rational choice. In the latter case, ϵ measures the deviation from rationality.

Matches are treated similarly. For each pair of individuals (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 individuals 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 individual n with anybody is exponentially distributed with rate parameter δ. Notice that matching cannot be independent across individuals (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 individuals.

When individuals are matched, each receives a utility determined by the actions each of the pair employs at the moment of the match, according to the payoff matrix G.

Each individual discounts the future at rate r. Given the policies of other individuals, the expected present discounted value of the payoff stream from a policy for any individual can be computed. Each individual chooses a policy to maximize that expected present discounted value. A stochastic strategy revision process is an *equilibrium strategy revision process* if each individual's policy is the ϵ-mixture of the expected present discounted value maximizing policy π^* and the random policy q.

In summary, the decision problem dynamics are described by three parameters which are common to all individuals in the population: 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 individuals become impatient, as r becomes large. More surprising is that myopic behavior emerges for any discount rate when the arrival rate of revision opportunities σ, is sufficiently small. This includes situations wherein all individuals are extremely patient.

Now come the details. Suppose the population contains N individuals. Choose an individual, say, individual 1, and suppose she is at a *strategy revision opportunity* — an instant of time at which she may revise her choice. Her policy, or strategy, is denoted by π_1. The argument of π_1 is the number of individual 1's $N-1$ fellow individuals who are currently using strategy α. The value of π_1 is the probability she chooses strategy α. The stochastic process which describes the evolution of the opponents' behaviors is called the *opponent process*. Typically the current choice of individual 1 will affect the evolution of opponents' play, and so the evolution of the opponent process will be contingent upon the current choice of individual 1.

I will describe the individual choice problem through its value function. I construct the value for individual 1 in state M and currently using α by computing the value of the next event. Many events can happen to individual 1. Strategy revision opportunities arrive to her at rate σ, and to everyone else at rate $(N-1)\sigma$. She receives matches at rate δ. (The matches of others are not payoff relevant to her.) So the aggregate arrival rate of events is $N\sigma + \delta$. When an event arrives, with probability $\sigma/(N\sigma+\delta)$ it is a strategy revision opportunity. With probability $\delta/(N\sigma+\delta)$ it is a match. With probability $(N-M-1)\sigma/(N\sigma+\delta)$ it is an updating opportunity for another individual currently choosing β. In this case the state of the opponent process increases by one with probability $\rho_\pi(M+1)$ and remains the same with probability $\rho'_\pi(M+1)$. With probability $M\sigma/(N\sigma+\delta)$ the event is an updating opportunity for another individual currently choosing α. In this case the state of the opponent process decreases by one with probability $\rho'_\pi(M)$ and remains the same with probability $\rho_\pi(M)$.

When individual 1 is matched, the probability that she meets an opponent playing strategy l is proportional to the number of individuals choosing strategy l in the current state a. Thus her expected return from a match when the opponent process is in state m and she uses action k is

$$v(k,a) = \frac{m}{N-1}G(k,\alpha) + \frac{N-1-m}{N-1}G(k,\beta). \tag{1}$$

To construct the value function, let τ denote the time to the next event. The random variable τ is exponentially distributed with parameter $N\sigma+\delta$. The value function is given by the Bellman equation. For the individual currently choosing α in state M of the opponent process,

$$\begin{aligned} V(\alpha, M) = E\Big\{ e^{-r\tilde{\tau}} \Big(&\frac{\delta}{N\sigma+\delta}\big(v(\alpha, M) + V\alpha, M\big) + \\ &\frac{\sigma(N-M-1)}{N\sigma+\delta}\big(\rho_\pi(M+1)V(\alpha, M+1) + \rho'_\pi(M+1)V(\alpha, M)\big) + \\ &\frac{\sigma M}{N\sigma+\delta}\big(\rho'_\pi(M)V(\alpha, M) + \rho_\pi(M)V(\alpha, M-1)\big) + \\ &\frac{\sigma}{N\sigma+\delta}\max_\pi\big(\rho_\pi(M)V(\alpha, M) + \rho'_\pi(M)V(\beta, M)\big)\Big)\Big\}. \end{aligned} \tag{2}$$

A computation shows that

$$\begin{aligned}
V(\alpha, M) = {} & \frac{\delta}{r+\sigma N}\, v(\alpha, m) + \\
& \frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho(M+1)V(\alpha, M+1) + \rho'(M+1)V(\alpha, M)\Big) + \\
& \frac{\sigma M}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\alpha, M-1)\Big) + \\
& \frac{\sigma}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\beta, M)\Big) \qquad \text{(3a)}
\end{aligned}$$

and

$$\begin{aligned}
V(\beta, M) = {} & \frac{\delta}{r+\sigma N}\, v(\beta, m) + \\
& \frac{\sigma(N-M-1)}{r+\sigma N}\Big(\rho(M)V(\beta, M+1) + \rho'(M)V(\beta, M)\Big) + \\
& \frac{\sigma M}{r+\sigma N}\Big(\rho(M-1)V(\beta, M) + \rho'(M-1)V(\beta, M-1)\Big) + \\
& \frac{\sigma}{r+\sigma N}\Big(\rho(M)V(\alpha, M) + \rho'(M)V(\beta, M)\Big)\,. \qquad \text{(3b)}
\end{aligned}$$

Standard contraction mapping arguments show that the Bellman equation has a unique solution which is the value function for individual 1's dynamic programming problem. The optimal strategy for individual 1 requires that when a strategy revision opportunity arises in state m she chooses the probability π to maximize $\rho_\pi(M)V(\alpha, M) + \rho'_\pi(M)V(\beta, M)$. Notice that individual 1's Bellman equation accounts for the possibilty of a noisy perturbation in choice.

III. Equilibrium

The equilibrium concept introduced here is closely related to Nash equilibrium. It differs only in the presence of the perturbations. Essentially, a rational population equilibrium as defined below is a Nash equilibrium of a perturbed game wherein all strategies are required to be minimally mixed.[11]

Definition 1. *A* rational population equilibrium *is a strategy π such that for all states a of the opponent process and policy functions π',*

$$E_{\pi(a)}V(\tilde{k},a) \geq E_{\pi'(a)}V(\tilde{k},a).$$

An equilibrium strategy revision process *is the stochastic strategy revision process which results from a rational population equilibrium.*

The existence of equilibrium is straightforward because the dynamic programming problems are well-behaved.

Theorem 1. *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. □

The equilibrium strategy revision process is constructed from the equilibrium policy function π in just the way the opponent process was constructed. The state space is $\{0, N\}$, and the state variable counts the number of α-choosers. The process is a birth-death process, and the birth and death rates in state M are, respectively,

$$\begin{aligned}\lambda(M) &= (N-M)\sigma\rho_\pi(M)\\ \mu(M) &= M\sigma\rho'_\pi(M)\end{aligned}$$

The possibility of idiosyncratic choice implies that all states are reachable from one another regardless of what the equilibrium strategy actually is. The Markov process of population states is irreducible, and so the process is ergodic. The invariant distribution is easily computed from the birth and death rates:[12]

Corollary 1. *For any equilibrium π, the corresponding equilibrium strategy revision process is ergodic. The invariant distribution is characterized by the relations*

$$\frac{\nu(M)}{\nu(0)} = \binom{N}{M}\prod_{m=1}^{M}\frac{\rho_\pi(m-1)}{\rho'_\pi(m)}\,.$$

In Blume (1993) 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 (1993) and (2003) 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.

One interesting class of equilibria are the *monotonic equilibria*, which are defined by the property that π is non-decreasing in its argument. Monotonic equilibria are the most natural equilibria of the games discussed in this paper, since they reflect the basic intuition of coordination games: The more people who choose α, the bigger is the payoff advantage to choosing α over β, and therefore the more likely β-choosers are to switch to α; and vice versa. In the next section it will become apparent that all equilibria are monotonic when r is large or σ sufficiently small. In the following section I will show that monotonic equilibria exist for large enough N and small enough r and ϵ. Likely a proof of the existence of monotonic equilibria for all parameter values can be constructed along the lines of the proof of Theorem 1 in Blume (2002), but this is not pursued here.

Finally, standard arguments also prove:

Corollary 2. *The correspondences from parameters r, δ and σ to equilibrium strategies and to the invariant distributions for the equilibrium strategy revision process are upper hemi-continuous.*

IV. Myopic Behavior

Myopic behavior emerges when the value function becomes proportional to the payoff function in every state. When such proportionality occurs, the solution to the intertemporal optimization problem is identical to that of maximizing immediate expected returns on the supposition that a match will take place before the opponent process changes. Not surprisingly, myopic behavior emerges as the discount factor r gets large. In addition, myopic behavior emerges as the arrival rate of strategy revision opportunities becomes small. In this case, even when looking far into the future, one expects the state of the opponent process to remain unchanged with high probability, so the future looks just like the present.

Theorem 2. *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.*

Proof. Manipulate the Bellman equations (3). Let

$$D = r + \sigma(N-1-M)\rho_\pi(M+1) + \sigma M \rho'_\pi(M) + \sigma \, .$$

One can derive from (3) that

$$\begin{aligned} &\frac{rV(\alpha,M)}{v(\alpha,M)} = \\ &\quad \frac{\sigma(N-1-M)\rho_\pi(M+1)}{D}\frac{rV(\alpha,M+1)}{v(\alpha,M)} + \frac{\sigma M\rho'_\pi(M)}{D}\frac{rV(\alpha,M-1)}{v(\alpha,M)} + \\ &\qquad\qquad \frac{\sigma}{D}\frac{r\max_{\pi(M)}\{\rho_\pi V(\alpha,M) + \rho'_\pi(M)V(\beta,M)\}}{v(\alpha,M)} + \frac{\delta r}{D} \, . \end{aligned} \tag{4}$$

The value function is uniformly bounded across states. Let v^* and v_* denote upper and lower bounds, respectively, for the payoff function v (which in turn are the largest and smallest entries in the matrix G). The value function is bounded above by the value of a reward process which pays out at every match v^*. This upper bound is

$$V^* = E\big\{e^{-r\tau}(v^* + V^*)\big\} \, ,$$

where τ is the waiting time until the next match. Computing,

$$\begin{aligned} V^* &= \sigma \int_0^\infty e^{-(r+\sigma)t}(v^* + V^*)\,dt \\ &= \frac{\sigma}{r+\sigma}(v^* + V^*) \\ &= \frac{\sigma}{r}v^*\,. \end{aligned}$$

Similarly, a lower bound is $V_* = (\sigma/r)v_*$. It follows that each of the rV/v terms in (4), including the max term, is bounded above by $\sigma v^*/v_*$ and below by $\sigma v_*/v^*$.

Taking limits as $r \to \infty$ and making use of the bounds, it is easily seen that $rV(\alpha, M)/v(\alpha, M) \to \delta$. A similar argument gives the same conclusion as $\sigma \to 0$. The same arguments apply when α is replaced with β in the arguments of V and v. $\square$

Slowing down the arrival rate of strategy revision opportunities is not the same as speeding up the matching rate. As δ grows, $\delta^{-1}V(\alpha, M)/v(\alpha, M)$ converges to r/D, which is not independent of M.

The consequences of myopic play in this model have been developed in a number of papers. Two natural questions to investigate are the model's large population behavior and its small noise behavior. The small noise results discussed below parallel the earlier results of Blume (1993), Kandori et al. (1993) and Young (1993).

Coordination games exhibit strategic complementarities, that is, the more likely it is that a player's opponent will play a particular strategy, the higher the payoff for the player to choosing that strategy. Operationally, this implies the existence of a probability p^* such that if the percentage of the population playing α exceeds p^*, α will be the unique best response (and β will be the unique best response if the fraction playing α is below p^*). The range $(p^*, 1]$ is the basin of attraction for the action α and the range $[0, p^*)$ is the basin of attraction for β. The following corollary is an immediate consequence of Theorem 2 and Corollary 2:

Corollary 3. *For fixed r and δ there is an $\sigma(r, \delta)$ and for fixed δ and σ there is an $r(\delta, \sigma)$ such that for all (r, δ, σ) such that $r > r(\delta, \sigma)$ or $\sigma < \sigma(r, \delta)$,* *the only rational population equilibrium is the* myopic equilibrium *in which* $\pi(M) = 1$ *if* $M/(N-1) > p^*$ *and* $\pi(M) = 0$ *for* $M/(N-1) < p^*$*.*

The key concept for understanding equibrium emergence in the long run is risk dominance. Risk dominance of an action can be stated in several different ways. Most convenient for our purposes is the following:

Definition 2. *Action* α *in the two-by-two coordination game with payoff matrix* G *is* risk dominant *iff* $p^* < 1/2$*.*

That is, α is risk dominant if it is a best response when some fraction less than half of the population is using it. For instance, in the game depicted in figure 1, the strategy α is payoff dominant because among all the Nash equilibria of the game it gives the highest payoff, but the strategy β is risk dominant because if α and β are equally likely, β is the best response. Intuitively, the higher payoff from coordinating on α is offset by the higher cost of playing α in the event that coordination fails.

	α	β
α	$3, 3$	$-2, 0$
β	$0, -2$	$2, 2$

Figure 1: A coordination game.

The consequences of equilibrium selection for emergence are striking because they demonstrate how social interaction extends and transforms the logic of individual decision-making. Suppose that α is risk dominant, so $p^* < 1/2$. Rescale the population states so that, instead of measuring absolute numbers they measure population fractions (that is, divide by N). Thus for all processes the state space can be taken to be $[0, 1]$.

Theorem 3. *For fixed* (r, δ, σ) *such that* $r > r(\delta, \sigma)$ *or* $\sigma < \sigma(r, \delta)$*,*

1. *For fixed* $\epsilon < p^*$*, as* $N \to \infty$ *the invariant distribution converges to point mass at* $1 - \epsilon$*.*
2. *For fixed* N *large enough, as* $\epsilon \to 0$*, the invariant distribution converges to 1.*

The first result, a strong law of large numbers for social interaction models, is proved in Blume and Durlauf (2003), and the second is proved in Blume (2003). The answer the two natural questions raised earlier in this section are simply put. Coordination on the risk-dominant action emerges in the myopic limit. This interpretation is clearly true for the second result. For the first, it can be shown that the short run dynamics of the model can be characterized by a differential equation which has two stable states, one at ϵ and one at $1-\epsilon$. These are the two candidate states where the invariant distribution can pile up as N becomes large. The Theorem shows that the invariant distribution will concentrate around the limit point in the basin of attraction of the risk-dominant equilibrium.

V. Patient Players in 2-by-2 Coordination Games

This section characterizes monotonic equilibria in two-by-two coordination games when players are patient. When players are myopic, the only rational population equilibria have invariant distributions which pile up on those states in which nearly all individuals choose the risk-dominant choice, even when it is not payoff dominant. The inability to reach the payoff-dominant outcome is sometimes regarded as a consequence of myopia. If individuals were sufficiently patient, it is argued, they would pay the penalty of high short-run deviations from equilibrium in order to achieve higher long-run gains from payoff-dominance. This argument is false. The striking result is that for small r and ϵ, the only monotonic equilibrium is one in which all individuals *always* choose the risk-dominant choice. Patience strengthens the barriers against the risk-dominated equilibrium, even when it is payoff-dominant.

Some of the intuition behind the result is straightforward. Payoff dominance does have a distinguished role in the $\beta = 1$ limit: When players care about the time-average of payoffs, 'always play α' and 'always play β' are both rational population equilibria. The subtle feature of the result is that the intuition about payoff dominance fails for β near 1. The theorem really demonstrates a failure of lower-semicontinuity of the equilibrium correspondence at $r = 0$, and the fact that some small amount of 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 the benefit is strictly positive.

Theorem 4. *Suppose that the game G has a unique risk-dominant action. For all δ and σ, for r and ϵ sufficiently small and N sufficiently large, choosing the risk-dominant strategy in every state is the unique monotonic equilibrium.*

In particular, choosing the non-risk-dominant action in every state, even when it is payoff-dominant, is not a rational population equilibrium. A simple calculation shows the following:

Corollary 4. *Under the conditions of Theorem 4, the invariant distribution for the equilibrium strategy revision process is the binomial distribution $b(N, 1-\epsilon)$.*

Proof. The proof technique is to specify a ρ, take appropriate limits of parameters, and check the value function given in equations (3) 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,

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

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 $\rho(M) \equiv \rho$,

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

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

$$\lim_{r\to 0}\Delta(M) = \frac{\delta}{\sigma(N-M)}\Big(\Delta_v(M) + \cdots + \Delta_v(N-1)\Big).$$

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 $(2p^* - 1)(N-1) < 1$ if

$p^* < 1/2$. (Strict inequality rules out indifference at the boundary state.) 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

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

and

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

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. $\square$

VI. Conclusion

This paper studies the evolution of play in a population of players continually interacting with one another. 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 model is built from Poisson processes in continuous time, but individuals interact at discrete random moments. Individual rationality alone would suggest that some kind of coordination would occur. But the dynamics of social interaction limit what kind of coordination can occur. For the pairwise coordination problems studied here, the only emergent state is risk-dominant coordination.

This paper has extended the stochastic evolutionary paradigm by considering forward-looking agents. Extensions in other directions are required to develop a full-fledged theory of the emergence of social norms. One direction has to do with the topology of social interaction. Here we have considered global random matching. Everyone is equally likely to interact with anyone else. At the other extreme, myopic random matching with a restricted subset of neighbors has been examined by Blume (1993), Blume (1995) and Morris (2000). Other matching models have not been considered.

More generally, different kinds of interaction may generate different incentives for coordination. Stigma, for instance, depends upon individuals being labelled as different and then being punished for wearing that label. Labelling and the ability to react to labels is a complexity not treated here.[13] Other interaction models have yet to be explored in depth, but there is no reason to believe that risk-dominant selection is universal. Mailath, Samuelson and Shaked (2001), for instance, have shown that in at least some models where matching is endogenous, payoff-dominant selection can obtain. The challenge for evolutionary game theory is to identify emergent properties for a rich class of interaction mechanisms and to allow for more complex individual behaviors. The promise of this work is an account of the emergence of social order, of Durkheim's 'collective entity', that satisfies the constraint of methodological individualism but which is not trivially reductive.

Notes

[1] This is not to say that each discipline is characterized by diametrically opposed views of individual behavior. Modern versions of exchange theory invoke rational choice models to explain social life. And while rational choice models have dominated economics for the last fifty years, the discipline certainly has a rich and internally respected tradition which thinks of individuals as embedded in social structures, going back to Adam Smith.

[2] There is a tendency in the literature to see all social norms as productive social capital, as 'goods'. But normative behaviors such as norms of racial or ethnic discrimination are clearly not productive social capital. So explanations based on the efficiency-inducing aspects of social prescriptions fail. See Durlauf (2002) and Portes (1998).

[3] Parsons (1937) is a prominant example. Also Elster (1989).

[4] Durkheim (1982 [1895], p. 128).

[5] In a footnote to the quotation cited below, Durkheim (1982 [1895], n. 17, p. 145) writes, 'In this sense and for these reasons we can and must speak of a collective consciousness distinct from individual consciousnesses.'

[6] Durkheim (1982 [1895], p. 129).

[7] See supra n. 4.

[8] I use the phrase 'emergent property' quite precisely to mean a behavior of the system that exists on a temporal or spatial scale different than the scale at which the equations defining the behaviors of the individual agents who comprise the system are defined. I thank Jim Crutchfield for this succinct definition.

[9] See Skyrms (1996) for a discussion of distributive justice using elementary evolutionary game theory arguments.

[10] Menger (1892) and Kiyotaki and Wright (1989).

[11] Trembling hand perfect equilibria in finite games are the limit of such

equilibria as the minimal mixing bound goes to 0, an exercise that will be carried out below.

[12] See Blume (2003) for details.

[13] See Blume (2002) for a model of this process in the spirit of evolutionary game theory.

References

Binmore, Ken L., Larry Samuelson, and Richard Vaughn. "Musical Chairs: Modelling Noisy Evolution." *Games and Economic Behavior* 11-1 (October 1995): 1–35.

Blume, L. E. "The Statistical Mechanics of Strategic Interaction." *Games and Economic Behavior* 4 (July 1993): 387–424.

——. "How Noise Matters." *Games and Economic Behavior* 44 (August 2003): 251–271.

Blume, Lawrence E. "The Statistical Mechanics of Best Response Strategy Revision." *Games and Economic Behavior* 11 (November 1995): 111–45.

——. "Stigma and Social Control." (2002). Institut für Hohere Studien, Wien, also online at http://econwpa.wustl.edu/eps/game/papers/0312/0312002.pdf.

—— and Steven Durlauf. "Equilibrium Concepts for Social Interaction Models." *International Game Theory Review* 5-3 (September 2003): 193–210.

Canning, D. “Average Behavior in Learning Models.” *Journal of Economic Theory* 57 (August 1992): 442–472.

Diamond, Peter. “Aggregate Demand Management in Search Equilibrium.” *Journal of Political Economy* 90 (October 1982): 881–94.

Durkheim, Emile. *The Rules of Sociological Method*, New York: The Free Press, (1982 [1895]).

Durlauf, Steven N. “Bowling Alone: A Review Essay.” *Journal of Economic Behavior and Organization* 47-3 (March 2002): 259–73.

Ellison, G. “Cooperation in the Prisoner’s Dilemma with Anonymous Random Matching.” *The Review of Economic Studies* 61 (July 1994): 567–88.

Elster, Jon. *The cement of society*, Cambridge: Cambridge University Press, (1989).

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

Fudenberg, D. and C. Harris. “Evolutionary Dynamics with Aggregate Shocks.” *Journal of Economic Theory* 57-2 (August 1992): 420–41.

Kandori, M., G. Mailath, and R. Rob. “Learning, Mutation and Long Run Equilibrium in Games.” *Econometrica* 61 (January 1993): 29–56.

Kandori, Michihiro. “Social Norms and Community Enforcement.” *The Review of Economic Studies* 59-1 (January 1992): 63–80.

Kiyotaki, N. and R. Wright. “On Money as a Medium of Exchange.” *Journal of Political Economy* 92 (August 1989): 927–54.

Mailath, George, Larry Samuelson, and Avner Shaked. “Endogenous Interactions.” in Antonio Nicita, Nicita, and Ugo Pagano, eds., *The evolution of economic diversity*, New York: Routledge, (2001): pp. 300–324.

Menger, C. “On the Origin of Money.” *Economic Journal* 2 (June 1892): 239–55.

Morris, Stephen. “Contagion.” *Review of Economic Studies* 67-1 (January 2000): 57–78.

Parsons, Talcott. *The Structure of Social Action*, New York: McGraw-Hill, (1937).

Portes, Alejandro. “Social Capital: Its Origins and Applications in Modern Sociology.” *Annual Review of Sociology* 24 (1998): 1–24.

Riker, William H. and Peter C. Ordeshook. “A Theory of the Calculus of Voting.” *American Political Science Review* 62-1 (January 1968): 25–42.

Samuelson, Larry. “Stochastic Stability in Games with Alternative Best Replies.” *Journal of Economic Theory* 64 (October 1994): 35–65.

Skyrms, Brian. *Evolution of the Social Contract*, Cambridge: Cambridge University Press, (1996).

Tullock, Gordon. *Toward a Mathematics of Politics*, Ann Arbor: University of Michigan Press, (1967).

Weber, Max. *The Theory of Social and Economic Organization;* being Part I of *Wirtschaft und Gesellschaft,* translated from the German by A. R. Henderson and Talcott Parsons, London: W. Hodge, (1947).

Young, H. Peyton. "The Evolution of Conventions." *Econometrica* 61-1 (January 1993): 57–84.