
Learning, Mutation, and Long Run Equilibria in Games

Author(s): Michihiko Kandori, George J. Mailath and Rafael Rob

Source: *Econometrica*, Vol. 61, No. 1 (Jan., 1993), pp. 29-56

Published by: [The Econometric Society](#)

Stable URL: <http://www.jstor.org/stable/2951777>

Accessed: 22/06/2014 18:53

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*.

<http://www.jstor.org>

LEARNING, MUTATION, AND LONG RUN EQUILIBRIA IN GAMES

BY MICHIO KANDORI, GEORGE J. MAILATH, AND RAFAEL ROB¹

We analyze an evolutionary model with a finite number of players and with noise or mutations. The expansion and contraction of strategies is linked—as usual—to their current relative success, but mutations—which perturb the system away from its deterministic evolution—are present as well. Mutations can occur in every period, so the focus is on the implications of ongoing mutations, not a one-shot mutation. The effect of these mutations is to drastically reduce the set of equilibria to what we term “long-run equilibria.” For 2×2 symmetric games with two symmetric strict Nash equilibria the equilibrium selected satisfies (for large populations) Harsanyi and Selten’s (1988) criterion of risk-dominance. In particular, if both strategies have equal security levels, the Pareto dominant Nash equilibrium is selected, even though there is another strict Nash equilibrium.

KEYWORDS: Evolutionary game theory, evolution, bounded rationality, learning, Markov chains, strict equilibria, risk dominance, equilibrium selection.

1. INTRODUCTION

WHILE THE NASH EQUILIBRIUM CONCEPT has been used extensively in many diverse contexts, game theory has been unsuccessful in explaining *how* players know that a Nash equilibrium will be played. Moreover, the traditional theory is silent on how players know *which* Nash equilibrium is to be played if a game has multiple equally plausible Nash equilibria. Introspective (eductive) theories that attempt to explain equilibrium play “directly” at the individual decision-making level impose very strong informational assumptions and so are widely recognized as having serious deficiencies (see, for example, Binmore (1987, 1988)).

As a consequence, attention has shifted to “evolutionary” explanations of equilibrium, motivated by the work of biologists in evolutionary game theory (in particular, the seminal work of Maynard Smith and Price (1973), see also Maynard Smith (1982)).² Two features of this approach distinguish it from the introspective approach. First, players are not assumed to be so “rational” or “knowledgeable” as to correctly guess (anticipate) the other players’ choices. Second (and instead), an explicit dynamic process is specified describing how players adjust their choices over time as they learn (from experience) about the

¹This paper has benefited from the comments of In-Koo Cho, Dean Foster, Drew Fudenberg, Jacob Glazer, Christopher Harris, Michael Woodford, Peyton Young, and many seminar audiences, as well as from an editor and three anonymous referees. Financial support from the NSF is gratefully acknowledged by the first and second authors (SES-9108351 and SES-8908451).

²Maynard Smith and Price (1973) introduced the notion of an evolutionary stable strategy (ESS), a notion of stability against mutations. Subsequent work (such as Taylor and Jonker (1978)) has studied when an ESS is an attractor of the replicator dynamic (in which the rate of growth in the fraction of the population playing a particular strategy is equal to the deviation of that strategy’s fitness or payoff from the average fitness or payoff). The biological literature is surveyed in Hines (1987).

other players' choices. Therefore, this approach tries to explain how an equilibrium emerges, based on trial-and-error *learning* (instead of introspective-type arguments). The evolutionary approach is surveyed by van Damme (1987) and Mailath (1992). Many papers have explored this approach, including Friedman (1991), Fudenberg and Maskin (1990), Nachbar (1990), and Binmore and Samuelson (1992). The work of Nelson and Winter (which culminated in their book (1982)) should also be mentioned for its emphasis on the importance of evolutionary ideas in explaining economic change.

The purpose of the present paper is to extend these ideas and—more importantly—to add a new perspective on the problem of equilibrium *selection*. Specifically, we follow the pioneering work of Foster and Young (1990), who were the first to argue that in games with multiple *strict* Nash equilibria,³ some equilibria are more likely to emerge than others in the presence of continual small stochastic shocks. In the present paper, we introduce a discrete framework to address this issue,⁴ and develop a general technique to determine the most likely, or *long run*, equilibrium (this concept is essentially the *stochastically stable equilibrium* of Foster and Young (1990)). We then apply this technique to the class of symmetric 2×2 games, and show that, for coordination games, the long run equilibrium coincides with the *risk dominant* equilibrium (see Harsanyi and Selten (1988)).⁵ We show this to be the case independent of all but the most crude features of the dynamics. Thus, our framework provides a link between the evolutionary approach and the risk-dominance criterion.⁶

Our specification of dynamics draws heavily on the biological literature. In that literature, animals are viewed as being genetically coded with a strategy and selection pressure favors animals which are fitter (i.e., whose strategy yields a higher reproductive fitness—or payoff—against the population). The focus there has been on the concept of evolutionary stable strategy (ESS) due to Maynard Smith and Price (1973) and its relationship to various dynamics. Selten (1991) provides a nontechnical discussion of the degree to which this literature is suited to the study of social phenomena.

While our model can be interpreted in like manner, we intend it more as a contribution to the growing literature on *bounded rationality* and *learning*. Accordingly, the hypotheses we employ here reflect limited ability (on the players' part) to receive, decode, and act upon information they get in the course of playing games.

In particular, we consider the situation where a group of players is repeatedly matched to play a game. The following three hypotheses form the basis of our analysis. (i) Not all agents need react instantaneously to their environment (the

³And so multiple evolutionary stable strategies.

⁴Moreover, the stochastic shocks are based on the micro-structure of the game.

⁵See page 46 for a definition. If the game is a coordination game with zeroes off-the-diagonal (or more generally, if the two strategies have identical security levels, i.e., payoffs from miscoordination), then the risk dominant equilibrium coincides with the Pareto dominant equilibrium.

⁶This is only shown for the symmetric 2×2 case. It is not clear that a similar link exists for the more general case.

inertia hypothesis); (ii) when agents react, they react myopically (the *myopia* hypothesis); and (iii) there is a small probability that agents change their strategies at random (the *mutation*, or experimentation hypothesis). We regard these as descriptions of boundedly rational behavior; we argue below that they can also be justified as rational behavior under some circumstances. We first lay out the general motivations of the three hypotheses, and then discuss some specific interpretations.

The logic underlying these hypotheses is as follows. A strategy in a game is, in general, a complicated object, specifying what actions one should take given various contingencies one is able to observe. However, players' observations are imperfect, their knowledge of how payoffs depend on strategy choices may be tenuous,⁷ and changing one's strategy may be costly. The presence of inertia is then suggested by the existence of such uncertainties and adjustment costs.⁸ In turn, to the extent that there is substantial inertia present, only a small fraction of agents are changing their strategies simultaneously. In this case, those who do move are justified in acting myopically: they know that only a small segment of the population changes its behavior at any given point in time and, hence, strategies that proved to be effective today are likely to remain effective for some time in the future. Thus, taking myopic best responses is justified as being fully rational when players have a high discount rate compared to the speed of adjustment.

The *myopia* hypothesis also captures a second aspect of learning which we feel is important; namely, imitation or emulation. Here the idea is that the world is a complicated place and agents cannot calculate best responses to their stage environment. People learn what are good strategies by observing what has worked well for other people. The *myopia* assumption amounts to saying that at the same time that players are learning, they are not taking into account the long run implications of their strategy choices. Thus, agents act as if each stage game is the last.⁹

Mutation plays a central role in our analysis. With some small probability, each agent plays an arbitrary strategy.¹⁰ One economic interpretation is that a player exits with some probability and is replaced with a new player who knows nothing about the game and so chooses a strategy at random (as in Canning (1989)).

The model admits a variety of interpretations, with differing types of bounded rationality. The first interpretation is that players gradually learn the strategy

⁷ Lipman (1991) argues that bounded rationality can be usefully modeled through the observation that knowing a fact is distinct from knowing the logical implications of that fact.

⁸ Not changing a strategy when a player is uncertain may, of course, be even costlier. However, it seems plausible that uncertainty leads to inertia (for example, the status quo plays a distinguished role in the Knightian decision theory of Bewley (1986)).

⁹ At the risk of repeating ourselves, we are interested in agents learning their optimal one-shot action, not their optimal repeated game strategy.

¹⁰ Fudenberg and Kreps (1989) also use repeated experimentation in their model of learning. However, the role experimentation plays in that model is very different from that in Foster and Young (1990) and here.

distribution in the society. At any given moment, a small fraction of the population is exogenously given opportunities to observe the exact distribution in the society, and take the best response against it. An important aspect of bounded rationality in this interpretation is the assumption that uninformed players do not change their strategy choice (i.e., these players remain at the status quo), even though they may receive partial knowledge of the current strategy distribution through random matching. Although such “cautious” behavior might be inconsistent with Bayesian rationality, it seems to capture a certain aspect of observed behavior (see footnote 8).

The second interpretation is that players are completely naive and do not perform optimization calculations. Rather, players sometimes observe the current performance of other players, and simply mimic the most successful strategy. Note that in the first interpretation, players are able to calculate best replies and learn the strategy distribution of play in society. In the second interpretation, players are less sophisticated in that they do not know how to calculate best replies and are using other players’ successful strategies as guides for their own choices. While the first aspect of learning (i.e., learning other players’ strategy choices) is certainly an important aspect of bounded rationality (and the focus of much recent work, e.g., Fudenberg and Kreps (1989)), the second aspect (learning the optimal choice) is also important. For example, the central hypothesis in economics that agents optimize has been justified by an appeal to natural selection (or, rather, market selection; see Alchian (1950) and Friedman (1953)).¹¹

In the third interpretation, players are rational with perfect foresight, but there is significant inertia. For example, a strategy may be a choice of technology (capital equipment) and, if the cost of new equipment is large enough, a player will change his/her strategy only when the existing equipment needs replacing. If the opportunity to change a strategy occurs infrequently, a myopic best reply to the current strategy distribution may be justified as fully rational. Section 2 discusses an example which can be interpreted along these lines.

We formalize these ideas in a model with a finite population of players in which players are repeatedly matched within a period to play a stage game. We only impose a weak monotonicity condition reflecting the inertia and myopia hypotheses on the dynamics, which describe the intertemporal changes in the number of players playing each strategy. Individual behavior is then perturbed by independent shocks. This yields “Darwinian” dynamics with a stochastic component. The novelty of the approach comes from the focus on the behavior of the system in the long run, where a non-negligible number of mutations can occur. In contrast, the early literature on evolutionary games (Maynard Smith and Price (1973), Taylor and Jonker (1978)) concerns the stability of equilibria

¹¹Of course, maximizing utility and maximizing wealth are not the same thing. Blume and Easley (1991) show that, in the context of a financial asset market, the market may not select for agents who optimize.

against mutations which are *one-shot* small perturbations of equilibria. For a given rate of mutation, we first note that there is a unique stationary distribution on the number of players playing each strategy, and that this distribution is achieved in the long run irrespective of the initial condition. We then characterize the limit of this distribution as the rate of mutation goes to zero for 2×2 symmetric games (we call this limit the *limit distribution* and its support, the set of *long run equilibria*). When the game has two symmetric strict Nash equilibria, the limit distribution places probability one on the equilibrium that satisfies (for large populations) Harsanyi and Selten's (1988) criterion of risk-dominance.¹² In particular, if both strategies have equal security levels, the limit distribution places probability one on the Pareto dominant Nash equilibrium, even though there is another strict Nash equilibrium. This result requires only the weak monotonicity condition, and so is independent of the precise specification of the Darwinian dynamic.

This is to be contrasted with Foster and Young (1990), where the stochastically stable equilibrium depends on additional details of the dynamics, such as the speed of adjustment. The main difference in the formulations of the two models is in the source of the randomness (i.e., the introduction of mutations). We introduce randomness at the individual player level by independent mutations. Foster and Young (1990), on the other hand, applying the techniques of perturbed dynamical systems in Freidlin and Wentzell (1984), consider a model with a continuum population in which a Brownian motion reflects *aggregate* randomness in the population (without generating this randomness at the individual level).¹³ Independent individual mutations yield a stochastic process that is qualitatively different from the one in Foster and Young (1990), and one where the long run equilibria depend only on the likelihood of "large jumps" across the basins of attraction of the deterministic dynamics. Since the likelihoods depend only on the payoff structure of the game (in particular, on the relative sizes of the basins of attraction), our predictions are independent of finer details of the deterministic dynamic. In contrast, gradual small movements play a crucial role in Foster and Young (1990), and therefore their results are not independent of the speed of adjustment.¹⁴

The balance of the paper is organized as follows. In the next section, we give a brief heuristic discussion. The general model is laid out in Section 3, and in Section 4 the limit distribution (as the rate of mutation becomes small), is characterized (Theorem 1), and applied to the case of dominant strategies. Section 5 considers 2×2 coordination games, demonstrates the independence

¹²In general games, the limit distribution may imply a limit cycle rather than an equilibrium. See Section 6 for an example.

¹³There are some technical difficulties in Foster and Young (1990) (see Fudenberg and Harris (1992)). Fudenberg and Harris (1992) maintain the Brownian motion formulation of Foster and Young (1990), but consider a different kind of perturbation; namely, they perturb the populations playing each strategy, rather than perturbing the population share (as in Foster and Young).

¹⁴For more detail, see Section 8.

of the limit distribution to all but the coarsest properties of the deterministic dynamics, and connects our theory to the risk-dominance criterion. We analyze games with a unique symmetric Nash equilibrium in mixed strategies (i.e., the game has no pure symmetric Nash equilibria) in Section 6. For these games, a little more structure is needed on the deterministic dynamics to obtain a limit distribution which mimics the symmetric Nash equilibrium. Section 7 discusses the behavior of the expected time until the predicted state is observed. The robustness of our results are examined in Sections 8 and 9.

2. AN INFORMAL DISCUSSION OF THE MODEL

We illustrate the nature of our analysis by means of an example, which concerns the choice of computer systems in a small community. Consider a dormitory of ten graduate students, each of whom is using one of the two computer systems, s_1 or s_2 . The students are randomly meeting each other, and whenever two students meet, they can collaborate by writing a paper or exchanging software. So the matching is fruitful if and only if they are using the same computer. If we assume computer s_1 is superior to s_2 , this induces the following 2×2 game with common interest:

		II	
		s_1	s_2
I	s_1	2, 2	0, 0
	s_2	0, 0	1, 1

.

The game has two pure strategy equilibria, $E_1 \equiv (s_1, s_1)$ and $E_2 \equiv (s_2, s_2)$. Note that the former is the unique Pareto efficient outcome in the game. There is also a third equilibrium in mixed strategies, which puts probability $1/3$ on strategy s_1 . This implies that if more than one third of the population (4 or more in this case) are using s_1 , the best response in the random matching game is also s_1 . We assume that the students occasionally have opportunities to change their computers, when, for instance, they finish writing papers or their machines break down. Assume also that their main concern is the payoffs in the near future, so that they choose the best response against the current strategy distribution whenever they have opportunities to adjust their strategy. This induces “Darwinian” adjustments, in which the student population is gradually moving towards a situation in which all play a best response to the population. In this setting, it is clear that the final outcome crucially depends on the initial condition: If at least four students are initially using computer s_1 , then all students eventually end up using s_1 (equilibrium E_1 is realized). Otherwise, the final outcome is E_2 . The dependence of the final outcome on the initial condition has attracted much attention of economic historians, and is sometimes

called “path dependence” (Arthur (1984) and David (1988)). The main point of the present paper, however, is to show that, in some situations, this indeterminacy is resolved in a very particular direction if stochastic shocks are present.

Suppose that each student leaves the dormitory with probability ε in each period. If a student leaves, a newcomer enters, and the newcomer is an s_1 -computer user with probability $m \in (0, 1)$. The number m can be interpreted as the fraction of s_1 -users in the outside world; our prediction does not depend on its specific value, as long as that value is strictly between zero and one. This adds a stochastic element or “mutation” to the basic adjustment process described above. With mutations, it is clear that the system perpetually fluctuates between the two equilibria E_1 and E_2 . For example, even though all the players are initially using s_2 , in the long run it is possible that at least four students mutate to s_1 , and whenever this happens the society moves from E_2 to E_1 by means of the Darwinian adjustments. Likewise, E_1 is also upset in the long run, when there are at least seven mutations towards s_2 . The interesting thing to note here is that *upsetting the good (or efficient) equilibrium E_1 is more difficult than upsetting the bad equilibrium E_2 , because the former requires more mutations than the latter (seven versus four)*. As we will see below, this implies that *the system spends most of the time at the good equilibrium, when the mutation rate is small*.

Suppose, for the sake of intuition, that we can think of the choice between two equilibria as being determined by a two-state Markov chain, with two states given by E_1 and E_2 , ignoring all intermediate states.¹⁵ As a result of mutation, there is nonzero probability of transition from E_1 to E_2 , denoted p , and from E_2 to E_1 , denoted p' . The argument in the previous paragraph suggests that the orders of p and p' are ε^7 and ε^4 respectively, so $p/p' \rightarrow 0$ as $\varepsilon \rightarrow 0$. Simple computation shows that the unique stationary distribution over E_1 and E_2 is $(p'/(p + p'), p/(p + p'))$. Therefore, as $\varepsilon \rightarrow 0$, the stationary distribution tends to $(1, 0)$, i.e., the Pareto dominant equilibrium receives probability 1 in the limit. Since the stationary distribution can be viewed as the vector of relative proportions of time spent on each state,¹⁶ this shows that the system spends most of the time on E_1 in the long run, if the mutation rate is small. In fact, when the mutation rate is .1, our numerical analysis in Section 7 shows that on average it takes 78 periods to upset the bad equilibrium E_2 , while upsetting the good equilibrium E_1 takes about 100,000 periods. Therefore, in this situation, we think it reasonable to expect that the students are using the superior computer s_1 most of the time.¹⁷

¹⁵This assumption is employed in this section to provide a rough intuition for the reader's convenience and is *not* assumed in the main body of the paper. However, there is a deterministic adjustment rule which allows us to use a two state Markov chain: the best response dynamic in which *all* players play the best response to last period's population. See footnote 23.

¹⁶That is, the stochastic process is ergodic. For more detail, see Section 3.

¹⁷This argument hinges on the “irrationality” of mutants in picking initial strategies. However, as Matsui and Rob (1991) show, this result is actually strengthened when (a certain type of) rationality is incorporated into the mutants' decision making.

On the other hand, the above illustration indicates the limits of our approach. It should be clear from the above argument that upsetting each equilibrium takes an extremely long time if the population size is sufficiently large and the mutation rate is sufficiently small. For example, while many people have argued that the current design of computer keyboards, known as QWERTY, is inefficient, it has been widely used for a long time.¹⁸ Given the large number of typewriter/computer users, we cannot expect that QWERTY will be replaced with a more efficient design by independent mutations of individual users within any reasonable amount of time. Therefore, in terms of expected time to upset the “bad” equilibrium, we believe that our analysis is most relevant to the case of a small number of players.¹⁹ On the other hand, in terms of the proportion spent in each equilibrium, our analysis applies to any number of players.

The above argument based on the two state Markov chain is of course only suggestive. The state space is much larger, and the probabilities of transition are functions of the myopic adjustment rule and mutation rates at the individual level. The result however does carry over. For 2×2 games of common interest in which the costs of miscoordination are the same in both equilibria, the limit distribution puts probability one on the Pareto dominant equilibrium. In general, the limit distribution puts probability one on the risk dominant equilibrium (which need not be Pareto dominant when the costs of miscoordination in the two equilibria differ—see page 46).

This result should be contrasted to the recent literature (e.g., Aumann and Sorin (1989), Anderlini (1989), and Matsui (1991)) on eliciting coordination on the Pareto dominant equilibrium. The need for more complicated devices such as a communication stage before the stage game is played is explained by their focus on the Pareto dominant equilibrium, rather than the risk dominant one.

To conclude this section, we briefly relate our work to simulated annealing (see Kirkpatrick, Gelatt, Vecchi (1983) for an introduction). Simulated annealing is an optimization algorithm which introduces randomness into the calculations in order to prevent the system from becoming stuck in a local optimum that is not a global optimum and then evaluates solutions as the randomness (“temperature”) is reduced. While there is a similarity between our work and simulated annealing (in that both introduce randomness and then examine the results from taking limits as the degree of randomness goes to zero), the role of the randomness is very different. Loosely, randomness in simulated annealing introduces regions of the “parameter space” which have not yet been searched.

¹⁸ See Dixit and Nalebuff (1991, Ch. 9) for more detail and other interesting examples of path dependence.

¹⁹ For more discussion of this subject, see Section 7. A related issue is the rate of convergence to the stationary distribution from arbitrary initial distributions. Ellison (1991) observes that, as the number of players increases, this rate of convergence decreases. Ellison (1991) compares the rates of convergence in a neighborhood model, in which each player is matched only with his neighbors, with the model analyzed here (under the best reply dynamic). He argues that when the population is large, a neighborhood model need not require “unreasonable” lengths of time for the stationary distribution to be viewed as a plausible description.

In contrast, in the model studied here, agents are behaving optimally in any equilibrium and randomness facilitates *coordination* on a particular equilibrium, which need not be Pareto dominant.

3. THE MODEL

We consider a situation where a finite number of players are repeatedly matched to play a stage game and adjust their behavior over time. In this paper, we focus on 2×2 symmetric games, although the analysis naturally extends to more general cases. Consider a group of N players (N an even number) who are going to play the 2×2 symmetric game,

		II	
	s_1	s_1	s_2
I	s_1	a, a	b, c
	s_2	c, b	d, d

where $S = \{s_1, s_2\}$ is the set of pure strategies of the game and a, b, c , and $d \in \mathbb{R}$. For our purposes, there are three types of generic 2×2 symmetric games: games with a dominant strategy equilibrium ($(a - c)(d - b) < 0$), coordination games ($a > c, d > b$), and games with a unique symmetric equilibrium which is in mixed strategies ($a < c, d < b$).

We assume that actions are taken in discrete time, $t = 1, 2, \dots$. At the beginning of period t , each player chooses his pure strategy for the period (this is part of the inertia hypothesis). Let z_t be the number of players adopting strategy s_1 at time t . This defines the state of the dynamical system, and the state space (the range of z_t) is $Z = \{0, 1, \dots, N\}$. We assume that $\pi_i(z_t)$, the average payoff of a player with strategy s_i , is given by

$$\begin{aligned}\pi_1(z) &= \frac{(z-1)}{(N-1)}a + \frac{(N-z)}{(N-1)}b, \quad \text{and} \\ \pi_2(z) &= \frac{z}{(N-1)}c + \frac{(N-z-1)}{(N-1)}d.\end{aligned}$$

There are two models of matching which generate the above payoffs. The first specifies that each player is matched with each of the remaining players exactly once (as in a tournament). The second postulates that there are an infinite number of random matches within period t , so that each player's average payoff in that period is equal to the expected payoff. The only feature of the within period matching that we use are the values for $\pi_1(z)$ and $\pi_2(z)$. Players' actions are assumed to be fixed within a period.

We assume that better strategies are better represented in the population in the next period. That is, we assume the underlying *deterministic dynamic*,

$$(1) \quad z_{t+1} = b(z_t),$$

has the following “Darwinian” property:

$$(D) \quad \text{sign}(b(z) - z) = \text{sign}(\pi_1(z) - \pi_2(z)) \quad \text{for } z \neq 0, N.^{20}$$

Note that (D) does not restrict b for $z = 0$ or N . Biologists typically assume $b(0) = 0$ and $b(N) = N$, i.e., extinct strategies stay extinct (apart from mutation). An alternative assumption, more plausible in our case, is $b(0) > 0$ if $\pi_1(0) > \pi_2(0)$, and $b(N) < N$ if $\pi_1(N) < \pi_2(N)$. Our qualitative results do not depend upon which we assume, and we will, in general, assume $b(0) > 0$ if $\pi_1(0) > \pi_2(0)$ (and $b(0) = 0$ if $\pi_1(0) \leq \pi_2(0)$), and $b(N) < N$ if $\pi_1(N) < \pi_2(N)$ (and $b(N) = N$ if $\pi_1(N) \geq \pi_2(N)$).

Assumption D is very weak and any reasonable dynamics specified on Z will satisfy it.²¹ The Introduction described three scenarios for which (D) holds. The replicator dynamics studied in evolutionary game theory also satisfy (D).

One particular deterministic dynamic that we will use is the *best reply dynamic*, denoted B . This dynamic is given by the rule:

$$B(z) = \begin{cases} N, & \text{if } \pi_1(z) > \pi_2(z), \\ z, & \text{if } \pi_1(z) = \pi_2(z), \\ 0, & \text{if } \pi_1(z) < \pi_2(z). \end{cases}$$

If the stage game is one of coordination (i.e., $a > c, d > b$), then the dynamical system (1) has either two or three steady states (e.g., for the game in Section 2, the steady states are $0, N$, and, if $N/3$ is an integer, $N/3$), and the asymptotic behavior of the system may depend on the initial condition (z_0). Such indeterminacy is resolved if we introduce some noise (“mutations”) into the system. Now assume that each player’s strategy “flips” with probability ε in each period (i.i.d. across players and over time). What we have in mind are the following stories. With probability 2ε , each player dies and is replaced with a successor at each moment of time (so that the size of the population remains fixed). The newcomer does not know anything about the game, and he simply takes each strategy with equal probability (our results do not change as long as the ratio of the initial adoption probabilities is bounded in the limit). Alternatively, we may assume that each player “experiments” every once in a while with

²⁰This form of the Darwinian property assumes that players, after observing the behavior of their $N - 1$ opponents, include their own behavior to calculate the overall distribution. An alternative would be to specify that players who chose s_2 in period t switch only if $\pi_1(z_t + 1) > \pi_2(z_t)$ (i.e., an s_2 player realizes that when choosing s_1 , the number playing s_1 goes up), and that players who chose s_1 in period t switch only if $\pi_1(z_t) < \pi_2(z_t - 1)$. This formulation is a little more complicated since $\pi_1(z_t + 1) > \pi_2(z_t)$ and $\pi_1(z_t) < \pi_2(z_t - 1)$ can hold simultaneously. Assuming $b(z) > z$ if $\pi_1(z + 1) > \pi_2(z)$ and $\pi_1(z) > \pi_2(z - 1)$ and $b(z) < z$ if $\pi_1(z + 1) < \pi_2(z)$ and $\pi_1(z) < \pi_2(z - 1)$ only results in minor changes (see footnote 29).

²¹The innocuous assumption for games with more than two strategies is to require that only the *best* strategy be better represented. For the case of two strategies, this of course coincides with (D).

exogenously fixed probability.²² This yields the nonlinear stochastic difference equation:

$$(2) \quad z_{t+1} = b(z_t) + x_t - y_t,$$

where x_t and y_t have the binomial distributions:

$$x_t \sim \text{Bin}(N - b(z_t), \varepsilon) \quad \text{and} \quad y_t \sim \text{Bin}(b(z_t), \varepsilon).$$

The dynamical system (2) defines a Markov chain on the finite state space $Z = \{0, 1, \dots, N\}$. The transition probabilities are given by

$$p_{ij} = \text{Prob}(z_{t+1} = j | z_t = i),$$

and $P = [p_{ij}]$ is the Markov matrix. Note that, under our assumptions, all elements in the matrix P are strictly positive.²³ It is well known, then, that the Markov chain has a unique stationary distribution. Let $\Delta_N \equiv \{q \in \mathbb{R}^{N+1} | q_i \geq 0 \text{ for } i = 0, 1, \dots, N \text{ and } \sum_i q_i = 1\}$ be the N -dimensional simplex. A *stationary* (or *invariant*) *distribution* is a row vector $\mu = (\mu_0, \mu_1, \dots, \mu_N) \in \Delta_N$ satisfying

$$(3) \quad \mu P = \mu.$$

We sometimes write $\mu(\varepsilon)$ to emphasize the dependence of μ on ε . When P is strictly positive, not only does there exist a unique stationary distribution, but it also has the following nice properties:

STABILITY: For any $q \in \Delta_N$, $qP^t \rightarrow \mu$ as $t \rightarrow \infty$.

ERGODICITY: For all initial states z_0 ,

$$T^{-1} \sum_{t=1}^T X_i(z_t) \rightarrow \mu_i \quad \text{almost surely as } T \rightarrow \infty,$$

where

$$X_i(z_t) = \begin{cases} 1, & \text{if } z_t = i, \\ 0, & \text{otherwise.} \end{cases}$$

²²In the model studied here, it may be required that the experimenter not change the strategy chosen, even if it yields a lower payoff than the previous strategy choice. A more satisfactory specification would allow any experimenter to immediately switch back to that player's previous strategy if that generated a higher payoff. Note that this is true of the best response dynamic $B(\cdot)$ in which *all* players choose the best response to last period's population. We do not see any difficulty in extending the results of this paper to cover the more complicated specification (now a state is the number playing the first strategy as well as the number of mutations in each direction).

²³If $b(\cdot)$ is given by the best reply dynamic and the game is a coordination game, then we can model this as a two-state Markov chain: Let S_i denote the basin of attraction of s_i . If $z_t \in S_1$, then $b(z_t) = N$, so that the probability that the system moves from z_t to *any* state in S_2 is independent of z_t . Similarly for movements from $z_t \in S_2$ to any state in S_1 . Letting E_i denote the state in which all players are choosing s_i , we thus have a two-state Markov chain, with states $\{E_1, E_2\}$ and transition probability $\Pr(E_i \rightarrow E_j)$ given by the probability that the system moves from E_i to S_j . This observation underlies one of the examples in Canning (1992).

Stability is a standard fact about aperiodic irreducible Markov chains. Asymptotically, independent of the initial condition, the strategy distribution in the society is given by μ . By introducing stochastic shocks to the dynamical system, we have obtained uniqueness and global stability. Ergodicity is an immediate implication of $\mu_i > 0$ and the Strong Ergodic Theorem.²⁴ The stationary distribution can be interpreted as the proportion of time that the society spends on each state. A standard reference for this material is Karlin and Taylor (1975).

We examine the long run behavior of the system when the probability of mutation is small. To this end, we introduce the concept of the limit distribution.

DEFINITION 1: The *limit distribution* μ^* is defined by

$$\mu^* = \lim_{\varepsilon \rightarrow 0} \mu(\varepsilon), \quad \text{if it exists.}$$

Existence and uniqueness of the limit distribution will be shown in the next section. If the limit distribution places positive probability on the strategy configuration $z \in Z$, z is called a long run equilibrium. Formally, let $C(q)$ be the carrier of probability distribution $q \in \Delta_N$; $C(q) = \{i \in Z | q_i > 0\}$.

DEFINITION 2: The set of *long run equilibria* is $C(\mu^*)$.²⁵

4. CHARACTERIZATIONS OF THE LIMIT DISTRIBUTION

In principle, we can calculate the limit distribution, for each ε , by explicitly writing down the closed form for p_{ij} and then solving the equation $\mu P = \mu$ for μ , but the procedure is rather complicated. The following characterization of stationary distributions due to Freidlin and Wentzell (1984), however, enables us to skip explicit calculations and drastically simplifies the analysis.²⁶ The characterization utilizes trees, or directed graphs, connecting the states in the state space Z . A *z-tree* h on a finite set Z is a collection of ordered pairs of (or arrows between) elements of Z , denoted $(i \rightarrow j)$, such that every state in $Z \setminus \{z\}$ is the initial point of exactly one arrow and from any state in $Z \setminus \{z\}$ there is a sequence of arrows leading to z . We will say that state j is the immediate

²⁴If the initial distribution is μ , then the Markov chain is a stationary stochastic process and the Strong Ergodic Theorem implies the time averages converge almost surely to μ_i . To prove that the same is true for other initial distributions it is enough to do this for all initial distributions δ_z , where δ_z puts probability one on state $z \in Z$. Let $A \subset Z^\infty$ be the set of sequences $\{z_i\}$ such that $T^{-1}\sum_i X_i(z_i) \rightarrow \mu_i$. Let \mathcal{P} and \mathcal{P}_z be the probability distributions induced on Z^∞ by the transition matrix P and the initial distributions μ and δ_z , respectively. Then $\mathcal{P}_z(A) = \mathcal{P}(A|z) = \mathcal{P}(A)/\mu_z = 0$, since $\mathcal{P}(A) = 0$.

²⁵Since we are dealing with a finite Markov chain with a unique stationary distribution for every positive mutation level, this is equivalent to Foster and Young's (1990, p. 228) notion of stochastically stable: They define a state s to be *stochastically stable* if the asymptotic distribution exists and assigns positive probability to every neighborhood of s for arbitrarily small noise levels.

²⁶The characterizations of limit distributions are extensively studied by Freidlin and Wentzell (1984) in continuous time-continuous state space models. What follows is basically a discrete version of their analysis.

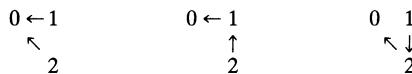


FIGURE 1

successor of state i in tree h if $(i \rightarrow j) \in h$. Thus, a z -tree is a directed graph on Z such that each state except z has a unique successor and there are no closed loops. Denote the set of z -trees by H_z . Figure 1 shows all 0-trees when the state space is $Z = \{0, 1, 2\}$.

To obtain the stationary distribution μ , we first construct a vector q which is proportional to μ . For each z -tree, calculate the product of transition probabilities along the tree, and then take the summation of the products for all z -trees. This defines a number q_z :

$$(4) \quad q_z = \sum_{h \in H_z} \prod_{(i \rightarrow j) \in h} p_{ij}.$$

For the example in Figure 1, $q_0 = p_{20}p_{10} + p_{21}p_{10} + p_{12}p_{20}$. Let q be this constructed vector, $q = (q_0, \dots, q_N)$. The following result (Freidlin and Wentzell (1984, Chapter 6, Lemma 3.1)) is key. Since its proof is simple, we present it for the readers' convenience.

LEMMA 1: *The vector q is proportional to μ .*

PROOF: For given state z , consider the set of directed graphs on Z , denoted G_z , such that (i) each state i has a unique successor, j , $i \neq j$, and (ii) there is a unique closed loop, and it contains state z . For each $g \in G_z$, calculate the product of transition probabilities along g and then take the summation of all the products over $g \in G_z$. A moment's reflection shows that the resulting number is expressed as

$$\sum_{k \neq z} q_k p_{kz} = \sum_{l \neq z} q_z p_{zl},$$

and the right hand side is equal to $(1 - p_{zz})q_z$. This yields $\sum_k q_k p_{kz} = q_z$, which shows $qP = q$. *Q.E.D.*

Since $\mu_z^* = \lim_{\varepsilon \rightarrow 0} q_z(\varepsilon)/\sum_i q_i(\varepsilon)$, the states which receive positive weight under μ^* are those states for which $q_z(\varepsilon)$ converges to zero at the lowest rate.²⁷ If such a state is unique, the limit distribution assigns probability one to that state. Let v_z be the rate at which q_z converges to zero, i.e., $q_z = O(\varepsilon^{v_z})$. From (4), the rate at which q_i converges to zero is determined by the rate at which the probability along the slowest z -tree goes to zero. Let c_{ij} be the speed of

²⁷If there is a state z such that $q_z(\varepsilon)$ does not converge to zero, then the rate of convergence of $q_z(\varepsilon)$ to zero is the lowest possible, namely 0.

convergence of p_{ij} to zero as ε tends to zero, i.e., $p_{ij} = O(\varepsilon^{c_{ij}})$. Thus, $v_z = \min_{h \in H_z} \sum_{(i \rightarrow j) \in h} c_{ij}$.

For an arrow $(i \rightarrow j)$, we can interpret c_{ij} as the *cost* of a transition from i to j . Any transition of the form $(i \rightarrow b(i))$ is costless ($c_{i,b(i)} = 0$), because $p_{i,b(i)}$ is bounded away from 0 as $\varepsilon \rightarrow 0$; in fact, $p_{i,b(i)} \rightarrow 1$ as $\varepsilon \rightarrow 0$. The cost of the transition $(i \rightarrow j)$ can be interpreted as the degree of mutation needed to transit from i to j . Since, in the absence of mutation, $b(i)$ is the state reached from i , at least $|b(i) - j|$ mutations are needed to reach j . While more mutations will also work, as long they are offsetting, the number of mutations most likely to reach j is the minimal number, $|b(i) - j|$. This motivates the next lemma.

LEMMA 2: $c_{ij} = |b(i) - j|$.

PROOF: Follows from considering the leading term of

$$p_{ij} = \sum_{k=\max\{j-b(i), 0\}}^{\min\{j, N-b(i)\}} \binom{b(i)}{b(i)+k-j} \binom{N-b(i)}{k} \times \varepsilon^{b(i)+2k-j} (1-\varepsilon)^{N+j-b(i)-2k}. \quad Q.E.D.$$

A z -tree is a “map,” indicating how to reach z from every other state without passing through any state more than once. The total cost of transition from all states to z according to the z -tree h is then $\sum_{(i \rightarrow j) \in h} c_{ij} \equiv c(h)$. Lemma 3 states that v_z is the cost of transition along that z -tree which minimizes costs, and we call it *the cost of transition to state z*. Given Lemma 2, v_z is determined as follows:

LEMMA 3: $v_z = \min_{h \in H_z} \sum_{(i \rightarrow j) \in h} |b(i) - j|$.

We can now determine the limit distribution μ^* . Each element of q is a polynomial of ε ,

$$q_z = \sum_{v=0}^N a_z(v) \varepsilon^v.$$

Note that some of the coefficients $a_z(v)$ may be equal to 0. Clearly, $v_z = \min\{v | a_z(v) \neq 0\}$. (If q_z does not converge to 0, then $a_z(0) \neq 0$ so that $v_z = 0$.) Let $v^* = \min_z v_z$ and define $a^* = (a_0(v^*), \dots, a_N(v^*))$. Then we have the following theorem.

THEOREM 1: *The limit distribution μ^* exists and is unique. In particular,*

$$(5) \quad \mu_z^* = a_z^* \Big/ \left(\sum_i a_i^* \right),$$

and the set of long run equilibria is given by

$$(6) \quad C(\mu^*) = \text{Argmin}_{z \in Z} v_z.$$

PROOF: First observe that $a_i^* \geq 0$. Straightforward calculation of $\lim_{\epsilon \rightarrow 0} q_z(\epsilon) / (\sum_i q_i(\epsilon))$ verifies (5) and (6). The uniqueness follows from the uniqueness of $\mu(\epsilon)$ for $\epsilon > 0$. *Q.E.D.*

A straightforward application of these results is provided by considering games with a dominant strategy, $(a - c)(d - b) < 0$. Suppose that s_1 is a dominant strategy. Then, $v_N = 0$, since every state must eventually transit to N under b satisfying (D). Consider $z \neq N$. Note that every z -tree must have a path from N to z , and that the “cheapest” way to do this is in one step ($N \rightarrow z$), since the cost of the transition ($N \rightarrow z'$) and then $(z' \rightarrow z)$ is $N - z' + b(z') - z > N - z$ (by (D), $z' < b(z')$). Then the lowest cost z -tree is $\{(z' \rightarrow b(z')) : z' \neq z, N\} \cup \{(N \rightarrow z)\}$, since the only costly transition is $(N \rightarrow z)$, and so $v_z = N - z > 0$. Thus N is the only state with lowest rate of convergence to zero, and so the limit distribution puts probability 1 on that state.²⁸ We have proved the following theorem.

THEOREM 2: *Suppose the stage game has a dominant strategy. For any population size $N \geq 2$ and any adjustment process satisfying (D), the limit distribution puts probability one on N if s_1 is the dominant strategy and on 0 if s_2 is the dominant strategy.*

5. COORDINATION GAMES

In this section, we study coordination games, i.e., $a > c, d > b$, and, as a normalization, $a \geq d$. In this case, $E_1 \equiv (s_1, s_1)$ and $E_2 \equiv (s_2, s_2)$ are strict Nash equilibria. The game is described as being one of *common interest* when $a > d$, in which case E_1 Pareto dominates E_2 . In such games, if players could coordinate, they would presumably coordinate on E_1 . In the absence of explicit coordination, however, it is not obvious that E_1 will be played. The riskiness of s_1 relative to s_2 is also then relevant.

Suppose that $a > d$ and the two strategies have identical security levels (i.e., $b = c$). It seems natural that players will play the best equilibrium E_1 . Players have a common interest in playing E_1 and in case of possible confusion on which strategy to play, the first strategy is no riskier than the second. This conjecture turns out to be true in our model of the dynamic evolutionary process.

Let $z^* \in \mathfrak{N}$ be the critical level of population for which the following is true (z^* need not be an integer):

$$\text{sign}(\pi_1(z) - \pi_2(z)) = \text{sign}(z - z^*).$$

The two states 0 and N have basins of attraction under b given by $\{z < z^*\}$ and

²⁸The argument still applies if we require the extinction hypothesis of $b(0) = 0$ and $b(N) = N$. In this case the cost of the N -tree is 1, since one mutation is required to shift the state from the unstable steady state into the basin of attraction of the stable steady state. It is still true that the cost of any 0-tree is at least $N \geq 2 > 1$.

$\{z > z^*\}$ respectively. The relative sizes of these basins of attraction is a crucial factor in the determination of the limit distribution. Calculation shows that $z^* = [N(d - b) + a - d]/(a - c + d - b)$. Essentially, z^* corresponds to the mixed strategy equilibrium, which puts probability $\gamma = (d - b)/(a - c + d - b)$ on strategy s_1 . However, z^* is not exactly equal to γN because players with strategies s_1 and s_2 face slightly different strategy distributions due to the finiteness of the population. As the reader can easily verify, the difference between z^* and γN vanishes as the population size becomes large. The following theorem shows that the equilibrium with the larger basin of attraction is selected in the long run.

THEOREM 3: *Suppose the stage game is a coordination game and $z^* \neq N/2$. For any population size $N \geq 2$ and any adjustment process satisfying (D), the limit distribution puts probability one on N if $z^* < N/2$, and on 0 if $z^* > N/2$.*²⁹

Before we prove Theorem 3, it might be helpful to provide some intuition. Consider a case with six players (Figure 2). The law of motion according to the Darwinian dynamic $b(\cdot)$ is indicated by the arrows. In this example, the critical level z^* is between 2 and 3, and state 6 has the larger basin of attraction $\{3, 4, 5, 6\}$ than state 0 ($\{0, 1, 2\}$). We claim that *the easiest way to upset equilibrium state 0 is to have just enough mutations to escape from its basin of attraction*. More precisely, the following 6-tree (Figure 3), which indicates a particular way to achieve the other equilibrium state 6 by an immediate jump $0 \rightarrow 3$, turns out to be of minimum cost among all 6-trees. The numbers in parentheses indicate the costs of transition c_{ij} (the number of required mutations), and the total cost of this tree is 3, which is exactly equal to the size of the basin of attraction of state 0. The fact that this is the least cost 6-tree is explained as follows. Consider another 6-tree shown in Figure 4, which represents a gradual movement towards state 6, rather than the immediate jump depicted in Figure 3. The cost of this tree is 4 rather than 3, which indicates that the gradual transition is less likely than the immediate jump. Note that after achieving state 2, the selection pressure pushes the state back to 1, so we need two mutations rather than one to achieve state 3. This is the source of the inefficiency of the tree in Figure 4, and the same argument shows that any gradual transition is more costly than the immediate jump described in Figure 3. Therefore we conclude that the cost of transition to state 6, v_6 , is 3. By the same token, the minimum cost 0-tree is given by immediate jump from 6 to 2, which requires four mutations. So the cost of transition to state 0, v_0 , is 4. It is easy to see that any other nonequilibrium states have larger costs of transitions than one of those, so the conclusion is that state 6, which has the larger basin of attraction, achieves

²⁹For coordination games, the condition in footnote 20 can be simplified to $b(z) > z$ if $z > z^\dagger + 1$ and $b(z) < z$ if $z < z^\dagger$, where $\pi_1(z^\dagger + 1) = \pi_2(z^\dagger)$ (note that $(N - 1)^{-1}z^\dagger = \gamma$). Note also that $N^{-1}|z^* - z^\dagger| \rightarrow 0$ as $N \rightarrow \infty$. If the deterministic dynamic satisfies the condition in footnote 20 (rather than (D)), then the limit distribution puts probability one on N if $z^\dagger - 1 < N/2$, and on 0 if $z^\dagger > N/2$. The proof is the same as that of Theorem 3.

$$0 \leftarrow 1 \leftarrow 2 \quad 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$$

FIGURE 2

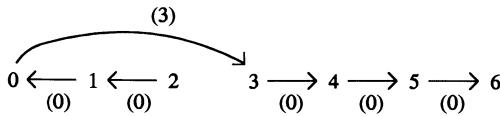


FIGURE 2

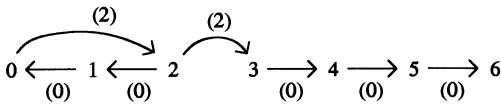


FIGURE 3

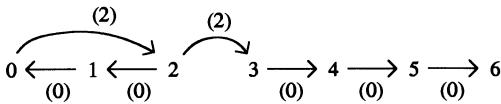


FIGURE 4

the minimum cost of transition among all states. According to our analysis in Section 4 (equation (6) in Theorem 1), this is the unique long run equilibrium.

PROOF OF THEOREM 3: Since the case of $z^* > N/2$ is just a relabelling of $z^* < N/2$, we only consider $z^* < N/2$. We will show that $\min_{z \in Z} v_z$ is uniquely achieved by $z = N$. Since we have a finite population, we need the following integers around z^* :

$$\alpha = \min \{z \in Z | \pi_1(z) > \pi_2(z)\}, \text{ and}$$

$$\beta = \max \{z \in Z | \pi_1(z) < \pi_2(z)\}.$$

Note that $\beta < z^* < \alpha \leq N/2$.

We first determine the value of v_N . Recall that for a tree h , the cost of transition along the tree is given by

$$c(h) = \sum_{(i \rightarrow j) \in h} c_{ij}.$$

For any N -tree h , define the path connecting states 0 and N ,

$$h' = \{(i \rightarrow j) \in h | i = 0 \text{ or } i \text{ is a successor of } 0\},$$

and let $h'' = h \setminus h'$. Then $c(h) = c(h') + c(h'')$. By the Lemma in the Appendix, the minimum of $c(h')$ over all paths in Z from 0 to N is achieved by $h^* = \{(z \rightarrow \alpha) \cup \{(z' \rightarrow b(z')) | z' = b^m(\alpha) \text{ for some } m \geq 0\}\}$, where $b^m(\cdot)$ is the m -fold iteration of $b(\cdot)$ and $b^0(z) = z$. Recall that, by (D), for any $z \geq \alpha$, $b^m(z) = N$ for sufficiently large m .

Thus, α is a lower bound of $c(h)$ for any N -tree h . This lower bound, however, is exactly achieved by the N -tree

$$\{(0 \rightarrow \alpha)\} \cup \{(z \rightarrow b(z)) | 0 < z < N\},$$

if z^* is not an integer, and by

$$\{(0 \rightarrow z^*), (z^* \rightarrow \alpha)\} \cup \{(z \rightarrow b(z)) | z \neq 0, z^*, N\},$$

if z^* is an integer (and so equals $\alpha - 1$), so that $v_N = \alpha$. In either case, the most efficient N -tree involves just enough mutations to upset E_2 .

Next, consider the value of v_z for $\alpha \leq z < N$. The same type of argument shows that a path connecting states 0 and z provides a lower bound α for v_z . However, there is no way to achieve this bound because there is an additional cost of transition from any state $z' > z$ to z . Hence we have $v_z > \alpha$ for $\alpha \leq z < N$.

By a symmetric argument, we have that $v_z \geq N - \beta$ for $0 \leq z \leq \beta$, and when z^* is an integer (so that it is a possible state), $v_{z^*}^* \geq N - z^*$. Since $\beta, z^* < N/2$ and $\alpha \leq N/2$, we conclude $v_z > v_N = \alpha$ for all $z \neq N$. *Q.E.D.*

The sign of $z^* - N/2$ is a function of the risk characteristics of the coordination stage game. If c is large relative to b (so that s_1 is riskier than s_2), then $z^* > N/2$ and E_2 is the long run equilibrium. It is easy to check that $z^* \leq N/2$ if and only if $a(1 - 2N^{-1}) - c \geq d(1 - 2N^{-1}) - b$. Note that this reduces to $a \geq d$ when $b = c$ (and in particular that the population size becomes irrelevant). Harsanyi and Selten (1989) have proposed a notion of *risk dominance* to capture the risk characteristics of a game. For the coordination game, E_1 risk dominates E_2 if $a - c > d - b$ and E_2 risk dominates E_1 if $a - c < d - b$. This discussion then yields the following:

COROLLARY 1: *Suppose the stage game is a coordination game and $z^* \neq N/2$. If $N \geq 2(a - d)/(a - c - d + b)$, the unique long run equilibrium is the risk dominant equilibrium.*

COROLLARY 2: *Suppose the stage game is a coordination game and its two equilibria have identical security levels ($b = c$) and $a > d$. For all $N \geq 2$, the unique long run equilibrium is the Pareto efficient equilibrium.*

We now turn our attention to the case $z^* = N/2$, which includes games of pure coordination such as:

		II	
	s_1	s_1	s_2
I	s_1	1, 1	0, 0
	s_2	0, 0	1, 1

In this game, there are two pure strategy equilibria, $E_1 = (s_1, s_1)$ and $E_2 = (s_2, s_2)$, which are Pareto efficient.

THEOREM 4: *Suppose the stage game is a coordination game and $z^* = N/2$. For any $N \geq 4$ and any adjustment process satisfying (D), the long run equilibria are E_1 and E_2 . Furthermore, the limit distribution places probability 1/2 on each of them.*

PROOF: Arguments identical to those in the proof of Theorem 3 show that $v_0 = v_N = N/2 + 1$ and $v_z > N/2 + 1$ for all $z \neq 0, N/2, N$. The cost minimizing z^* -tree is $\{(0 \rightarrow z^*), (N \rightarrow z^*)\} \cup \{(z \rightarrow b(z)) | z \neq 0, z^*, N\}$, with cost N . Thus, the long run equilibria are E_1 and E_2 .

Since the cost minimizing 0- and N -trees are unique and have the same structure, it is easy to see that the limit distribution must put equal probabilities on 0 and N . *Q.E.D.*

It is worth emphasizing that the limit distribution is symmetric (puts *equal* probability) on states E_1 and E_2 , irrespective of the precise nature of $b(\cdot)$. In particular, $b(\cdot)$ can be highly asymmetric in how it behaves on the basin of attractions of E_1 and E_2 .

The interpretation of the above result is as follows. The society stays in the neighborhood of, say, E_1 for a long time if the initial configuration of strategies is close to it. Every once in a while, E_1 is upset by a large number of deviations and society moves to the other equilibrium E_2 , and so on. Asymptotically, half the time will be spent in each equilibrium.

The limit equilibrium has the flavor of a correlated equilibrium in which agents are using history to coordinate their actions. An outside observer should predict the state in any period by looking at the history of play. If all (or most) agents had played s_1 last period, then this is a good indication that the same will occur in the current period. While this prediction is sensible from an outsiders' standpoint, players are not carrying out this calculation. On the contrary, players are simply following their learning (behavioral) rules.

We conclude this section with a comment about the behavior of the stationary distribution as N becomes large for fixed ε . For simplicity, restrict attention to the best reply deterministic dynamic.³⁰ As mentioned in footnote 23, we can analyze the system as a two state Markov chain. Let p be the transition probability from E_1 to E_2 and p' be the transition probability from E_2 to E_1 . Suppose $z^* < N/2$. Intuitively (and as suggested by the numerical example in Section 7), we might expect similar results, that is, that the stationary distribution will asymptotically place probability one on E_1 . The following line of argument, while not a proof, is suggestive: Observe that $p = \Pr(j \geq N - z^*)$ and $p' = \Pr(j \geq z^*)$, where $j \sim \text{Bin}(N, \varepsilon)$ is the number of mutations. Now argue that p/p' goes to zero as N goes to infinity by first approximating p and p' by their respective cumulative probabilities from the standard normal distribution, and then applying l'Hôpital's rule.

6. GAMES WITH NO SYMMETRIC PURE STRATEGY EQUILIBRIUM

In this section we explore the behavior of the stochastic dynamics for the remaining case, $a < c$ and $d < b$. In this case, there is a unique symmetric Nash

³⁰This is an interesting question in general, but significantly more difficult, since the state space itself is changing with N .

equilibrium, in which s_1 is played with probability $\gamma \equiv (b - d)/(c - a + b - d)$ and s_2 with the complementary probability $1 - \gamma$.

Consider first the case of $\gamma = 1/2$ and suppose that the speed of adjustment is fast—that is, the adjustment rule is given by the best reply dynamic $B(\cdot)$. In this situation, the long run equilibrium does not correspond to the mixed strategy equilibrium: the limiting distribution puts probability 1/2 on state 0 and probability 1/2 on state N (this is easily confirmed using our earlier arguments). Suppose the process is originally in state 1. Under $B(\cdot)$, the state next period is N , at which point the system oscillates between 0 and N (because $B(0) = N$ and $B(N) = 0$), until a mutation moves the process to state $N - 1$ (with highest probability), at which point the process jumps to state 0 and so on.³¹ Since $B(N/2) = N/2$, once in state $N/2$ the process remains there until a mutation occurs, after which either 0 or N will result.

However, if the speed of adjustment is slow, the long run equilibria correspond to the mixed strategy equilibrium. Let $[x]_+$ denote the smallest integer greater than or equal to x and $[x]_-$ the largest integer less than or equal to x , for any $x \in \mathbb{R}_+$. Given a mixed strategy equilibrium with $\text{Prob}(s_1) = \gamma$, say that $b(\cdot)$ is a *contraction relative to γN* if

$$[|b(z) - \gamma N|]_+ \leq [|z - \gamma N|]_+ \quad \text{for all } z,\text{³²}$$

with the inequality holding strictly if $\gamma N \in \mathbb{Z}$ or $[|z - \gamma N|]_+ \geq 2$. This assumption is best viewed as an assumption on the length of a period: periods are sufficiently short that only a small fraction of the population can adjust. For example, it is satisfied if one player can adjust at a time.

If $b(\cdot)$ is a contraction relative to γN , and if N is sufficiently large that $1 \leq \gamma N \leq N - 1$, then the limit distribution mimics the mixed strategy equilibrium γ .

THEOREM 5: *Suppose $N \geq 4$ and $b(\cdot)$ is a contraction relative to γN in the game where $a < c$ and $d < b$. If $1 \leq \gamma N \leq N - 1$, the limit distribution puts probability 1/2 on $[\gamma N]_-$ and probability 1/2 on $[\gamma N]_+$. If $\gamma N < 1$, the limit distribution puts probability one on 0 and if $\gamma N > N - 1$, the limit distribution puts probability one on N .*

PROOF: Suppose first that γN is an integer. In this case, γN is the global attractor for any $b(\cdot)$ which is a contraction relative to γN , and so trivially $v_{\gamma N} = 0$ (this is achieved by $\{(z \rightarrow b(z)): z \neq \gamma N\}$).³³ For all other z , $v_z \geq 1$.

³¹If instead we assumed extinction $B(0) = 0$ and $B(N) = N$, then after state 1, the system moves to state N and remains in N until a mutation moves the process to state $N - 1$ (with highest probability), at which point the process jumps to state 0 and so on.

³²If $b(0) = 0$ and $b(N) = N$, then this is only required for $z \neq 0$ or N .

³³If $b(0) = 0$ and $b(N) = N$, the argument needs to be adjusted as follows. If γN is an integer, $v_{\gamma N} = 2$, which is achieved by $\{(0 \rightarrow 1), (N \rightarrow N - 1)\} \cup \{(z \rightarrow b(z)): z \neq 0, \gamma N, N\}$. For all other z , $v_z \geq 3$. If γN is not an integer, then both $[\gamma N]_-$ and $[\gamma N]_+$ have identical minimum transition costs: $v_{[\gamma N]_-} = v_{[\gamma N]_+} = 2$, and the only $[\gamma N]_-$ -tree to achieve $v_{[\gamma N]_-}$ is $\{(0 \rightarrow 1), (N \rightarrow N - 1)\} \cup \{(z \rightarrow b(z)): z \neq [\gamma N]_-, 0, N\}$. The remainder of the argument is identical to the case $b(0) = N$, $b(N) = 0$.

If γN is not an integer, then both $[\gamma N]_-$ and $[\gamma N]_+$ have identical minimum transition costs: $v_{[\gamma N]_-} = v_{[\gamma N]_+} = 0$. All other states have higher transition costs. Furthermore, the only $[\gamma N]_-$ -tree to achieve $v_{[\gamma N]_-}$ is $\{(z \rightarrow b(z)): z \neq [\gamma N]_-\}$, since $b([\gamma N]_+) = [\gamma N]_-$. Since essentially the same tree achieves $v_{[\gamma N]_+}$, both states receive equal probability in the limit distribution.

The cases of $\gamma N < 1$ and $\gamma N > N - 1$ are handled similarly.

Q.E.D.

7. EXPECTED WAITING TIMES: HOW LONG IS THE LONG RUN?

In this section we return to the class of coordination games and investigate an alternative measure of how likely a particular state is to be observed. The ergodic property is a strong statement about the behavior of time averages. Thus, if the limit distribution places probability one on state N in the coordination game, for small rates of mutation the proportion of periods spent in that state is close to one. Note, however, that the state 0 is also an attractor of $b(\cdot)$ and so, *conditional* on being in that state in period t , for low rates of mutation we should expect to remain there for a long period of time.

In order to make this precise, we will restrict attention to the best reply deterministic dynamic, $b(\cdot) = B(\cdot)$. This particular deterministic dynamic makes it hardest for the process to leave the 0 state, since it requires a simultaneous mutation from 0 to leave its basin of attraction under $B(\cdot)$.

With this particular dynamic, we can view the Markov chain as having two states, E_1 and E_2 , with transition matrix (see footnote 23):

	E_1	E_2
E_1	$1-p$	p
E_2	p'	$1-p'$

Note that as $\varepsilon \rightarrow 0$, p , p' , and $p/p' \rightarrow 0$.

Suppose it is period $t = 0$. Let \mathfrak{F}_1 be the number of periods before the state E_1 is first observed. The argument in the first paragraph of this section concerns $E(\mathfrak{F}_1 | z_0 = E_2) = E(\mathfrak{F}_1 | \mathfrak{F}_1 \geq 1)$. It is easy to evaluate this expression: $E(\mathfrak{F}_1 | \mathfrak{F}_1 \geq 1) = \sum_{t \geq 1} tp'(1-p')^{t-1} = p' \sum_{t \geq 1} t(-\ln(1-p'))/dp' = 1/p'$. Thus we see that as $\varepsilon \rightarrow 0$, the expected duration in E_2 , *conditional* on starting in E_2 , goes to infinity.

However, this is the “worst case” scenario and it may not properly capture the *average* behavior of the process. Consider the following thought experiment: the process has been ongoing for a large period of time (so that the invariant distribution is a good approximation of the likelihood of being in any state) and an outside observer will observe the state at period τ . How many periods after τ will the observer have to wait before observing state E_1 ? That is, what is the *unconditional* expected value of \mathfrak{F}_1 when the original state is chosen according to the invariant distribution $\mu = (p'/(p+p'), p/(p+p'))$? Now, $E\mathfrak{F}_1 = E(\mathfrak{F}_1 | \mathfrak{F}_1 = 0) \Pr(\mathfrak{F}_1 = 0) + E(\mathfrak{F}_1 | \mathfrak{F}_1 \geq 1) \Pr(\mathfrak{F}_1 \geq 1) = (1/p') \times \Pr(z_\tau = E_2) = p/[p'(p+p')]$. Thus, the behavior of $E\mathfrak{F}_1$ is determined by the rate at

which p/p' goes to zero, relative to the rate at which $p + p'$ goes to zero.

If $E\mathfrak{F}_1$ converges to zero as the rate of mutation goes to zero, then E_1 clearly captures the long run behavior of the system in our thought experiment. Not only does this state occur with probability one in the limiting distribution, but in the presence of low rates of mutation, the expected time until E_1 is observed is close to zero, i.e., an outside observer expects to see E_1 immediately. On the other hand, if $E\mathfrak{F}_1$ goes to infinity as the rate of mutation goes to zero, then the situation is more complicated. While the state occurs with probability one in the limiting distribution, in the presence of low rates of mutation, the expected time until it is observed is large. That is, while there is a small probability of being in E_2 , the expected time in E_2 overwhelms that small probability. Thus we are interested in the behavior of $E\mathfrak{F}_1$.

THEOREM 6: Suppose the stage game is a coordination game and that $b(\cdot)$ is given by $B(\cdot)$, the best response dynamic. Then,

- (i) if $z^* < (N - 2)/3$, then $\lim_{\varepsilon \rightarrow 0} E\mathfrak{F}_1 = 0$; and
- (ii) if $z^* > (N + 1)/3$, then $\lim_{\varepsilon \rightarrow 0} E\mathfrak{F}_1 = \infty$.

PROOF: First observe that under $B(\cdot)$ states other than 0 and N are only achieved through mutation. Let x be the smallest integer strictly larger than $N - z^*$ and let y be the smallest integer strictly larger than z^* . The probability that in any period, the chain transits from any state $z > z^*$ to any state $z < z^*$ is of order ε^x (since a minimum of x mutations are required). Similarly, the probability that in any period, the chain transits from any state $z < z^*$ to any state $z > z^*$ is of order ε^y . The discussion above then implies that the asymptotic behavior of $E\mathfrak{F}_1$ is determined by the behavior of $\xi \equiv \varepsilon^x / [\varepsilon^y(\varepsilon^x + \varepsilon^y)]$ as $\varepsilon \rightarrow 0$.

Now, $\xi = (\varepsilon^y + \varepsilon^{2y-x})^{-1}$, and note that $2y - x < y$ because $y < x$. Then the asymptotic behavior of ξ is the same as ε^{x-2y} . In particular, if $x > 2y$, then $\xi \rightarrow 0$, i.e., $E\mathfrak{F}_1 \rightarrow 0$; and if $x < 2y$, then $\xi \rightarrow \infty$, i.e., $E\mathfrak{F}_1 \rightarrow \infty$, as $\varepsilon \rightarrow 0$.

Finally, observe that $x + y = N + 1$ (or $N + 2$ if z^* is an integer), so that $x > 2y$ is implied by $z^* < (N - 2)/3$ and $x < 2y$ is implied by $z^* > (N + 1)/3$.

Q.E.D.

NUMERICAL EXAMPLE: Consider the following special case of a coordination game, in which $a > 1$:

		II	
		s_1	s_2
I		s_1	a, a
		s_2	$0, 0$
		$0, 0$	$1, 1$

For this game, $z^* = (N + a - 1)/(a + 1)$, so from Theorem 5, the expected waiting time to see the long run equilibrium $E_1 = (s_1, s_1)$ tends to zero if $a > (2N - 1)/(N - 5)$, and tends to infinity if $a < (2N - 4)/(N - 2)$. Table I

TABLE I

a = THE PAYOFF FOR E_1 , ε = MUTATION RATE, N = POPULATION SIZE,
 t_1 = EXPECTED DURATION OF $E_1 = 1/p$, t_2 = EXPECTED DURATION OF $E_2 = 1/p'$,
 μ_2 = PROBABILITY OF E_2 IN STATIONARY DISTRIBUTION.

a	ε	N	t_1	t_2	μ_2
2	0.1	10	1.1×10^5	78	7.1×10^{-4}
2	0.1	20	4.6×10^9	420	9.0×10^{-8}
2	0.1	40	3.1×10^{17}	54,000	1.7×10^{-13}
3	0.1	10	2.7×10^6	14	5.3×10^{-6}
3	0.1	20	3.1×10^{12}	89	2.9×10^{-11}
3	0.1	40	9.1×10^{22}	680	7.4×10^{-21}
4	0.1	100	5.4×10^{61}	1,200	2.3×10^{-59}

TABLE II

$a = 2$, $\varepsilon = 0.1$, $N = 10$, $(\mu_1(t), \mu_2(t)) = (0, 1)P^t$

t	1	20	40	60	80	100
$\mu_2(t)$	0.93	0.23	0.054	0.013	0.0030	0.00070

shows the stationary distributions and the expected durations of two equilibria for some range of parameters. Note that for those parameter values, the stationary distributions closely approximate the limit distribution $\mu^* = (1, 0)$, and still the “wrong” equilibrium $E_2 = (s_2, s_2)$ is upset in relatively short time. For example, if $a = 2$ and the mutation rate and the population size are 0.1 and 20 respectively, upsetting the wrong equilibrium takes about 420 periods, while upsetting the long run equilibrium takes about five billion periods.

Table II illustrates global stability. It provides the probability of observing the equilibrium E_2 after t periods, given that the initial position is E_2 .

8. ROBUSTNESS OF THE MODEL

The most surprising feature of our model, perhaps, is that the equilibrium selected in the long-run does not depend on the exact details of the adjustment process; it only depends on the payoff structure of the underlying game. Our results show (Theorem 3) that as long as the monotonicity property (D) is satisfied by a *one-dimensional* dynamic,³⁴ the equilibrium with the largest basin of attraction is always selected in the long run. In turn, the sizes of the basins of attraction depend only on the payoff structure of the game. This is not true in Foster and Young (1990) or in Fudenberg and Harris (1992). In Foster and Young (1990) the selection of equilibria depends on a certain “potential function,” which reflects not only information on payoffs but also details of the adjustment process and mutations, in their case replicator dynamics and Gaussian disturbances. Fudenberg and Harris (1992) sometimes find no long-run equilibria. Rather, the dynamics are nonergodic, where the population eventually locks into one state, and never departs it thereafter. In this section, we

³⁴We discuss in Section 9 why the result can only be true for one-dimensional dynamics.

make a few remarks concerning the substantive and the modelling differences between the two formulations.

As stressed earlier our construction is based on certain behavioral postulates at the individual player level. These postulates, the discreteness of population and the independence of mutations enable us to *generate* aggregate randomness. In contrast, Foster and Young (1990) and Fudenberg and Harris (1992) consider a continuous-time, continuous state-space formulation, and *assume* aggregate randomness.³⁵ It is not obvious, however, how to generate the particular random process (Brownian motion) they use from assumptions on individual behavior. In particular, in a suitable discretization of their model, how should individual shocks be correlated across players and how should they be synchronized with the timing of strategy adjustments?

The most important difference, however, between our model and the models with Brownian motion is how a given equilibrium is upset by stochastic shocks. In our model of finite population and independent mutations, an equilibrium is upset by large “jumps” (from the equilibrium to the basin of attraction of the other equilibrium), while in the Brownian motion models have continuous sample paths and so gradual local movements are the only route out of the equilibrium. This is why details of the model like the speed of adjustment are unimportant in our formulation (for the symmetric 2×2 case), in contrast to the Brownian motion models. We now explore the difference in more detail.

In a related paper (Kandori (1991)) one of us has shown that the central result of this paper holds in a continuous time, discrete state space formulation where the opportunity to adjust one's behavior arrives according to a Poisson process. Furthermore, that model is constructed for a generally specified speed of adjustment (i.e., arrival rate) and the results can be shown to hold under the further restriction that mutations occur only among players who are actually adjusting. These specifications make simultaneous mutations—at any given point in time—impossible, and they minimize the number of mutations within any positive time interval. In Kandori (1991), the most likely way to upset an equilibrium is to have a series of single mutations in a short time interval, which is quite similar to simultaneous mutations. As a result, the selection of equilibria still only depends on the relative sizes of the basins of attraction. This suggests to us that upsetting equilibria by large jumps is a natural consequence of independent mutations, and that this result does not hinge on the fine modelling details. As long as mutations are independent, n simultaneous mutations are no less likely than a series of n mutations spread over time (loosely speaking, their probabilities both have order ε^n).

³⁵Aggregate randomness in a discrete model is the key to selecting among multiple equilibria since it ensures that it is always possible to transit from one equilibrium to the other. Strictly speaking, what is at issue here is the ergodicity of the stochastic process. While weaker conditions will guarantee that the process is ergodic, these assumptions do not have plausible interpretations. Aggregate randomness in a continuous state model, on the other hand, is not enough for ergodicity (see Fudenberg and Harris (1992)).

In contrast, the speed of adjustment does matter when random shocks that are spread over time are much more likely than simultaneous shocks. In that case, how the system behaves between random shocks (which is affected by the speed of adjustment) becomes a critical factor. On the basis of the above, we believe that this is generated by the correlation of shocks and we suspect this correlation to be the driving force behind the predictions in Foster and Young (1990) and Fudenberg and Harris (1992).³⁶

9. ANALYSIS OF MORE GENERAL GAMES

In this section we illustrate a case where the long-run equilibrium depends on the details of the Darwinian adjustment process. Consider again the 2×2 game of pure coordination in which $a > d = 1$ and $b = c = 0$. Until this point we have assumed that there is a single population of players, from which players are randomly drawn (in particular, there is no role identification). In this section, we suppose that there are two separate populations of players: row players and column players. There are M players of each type. The state of the system at each date is now described by a pair of numbers, (z_r, z_c) , where $0 \leq z_i \leq M$; the state-space itself is a grid of points over a two-dimensional square.

Consider now the Darwinian adjustment of row players (the same analysis applies to column players). See Figure 5. There is a critical value, $z^* = M/(a + 1)$ (assume, for simplicity, that z^* is not an integer), so that $\pi_1(z_c) > \pi_2(z_c)$ if and only if $z_c > z^*$. Therefore, according to our previous specification of Darwinian dynamics (and applied to each population separately), the state-space partitions into four subregions. In two subregions, namely 1 and 3, the dynamics are unambiguous. That is, they point towards one of the game's two Nash equilibria. On the other hand, in the remaining regions, the Darwinian dynamics point in conflicting directions. Hence, depending on where one starts and on the *relative speeds of adjustments* (i.e., the speed of adjustment of row players vs. column players), we may end up with an equilibrium payoff of $(1, 1)$ or (a, a) .

To illustrate the effect of this, consider the extreme case where in region 2, column players adjust infinitely faster than row players, whereas in region 4, it is the row players that adjust infinitely faster. In this case the basin of attraction of the $(1, 1)$ equilibrium consists of the region 1, 2, and 4, whereas the basin of attraction of the (a, a) equilibrium, consists of 3 alone. According to our general procedure we now have to compute costs of transitions between the two equilibria. To escape the (a, a) equilibrium one must mutate $[Ma/(a + 1)]$ players (where $[x]$ refers to the first integer exceeding x). This will bring the state into region 2, after which one follows the costless Darwinian dynamics. On the other hand, to escape the $(1, 1)$ equilibrium one must have $2[M/(a + 1)]$

³⁶One way to see the existence of correlation in their models is that the sum of independent individual mutations is not normally distributed (under independent individual mutations, the sum of mutations is only approximated by a normal distribution once the sum is normalized by \sqrt{N}). On the other hand, if the adjustment process follows a Brownian motion then the number of mutations in the society is normally distributed.

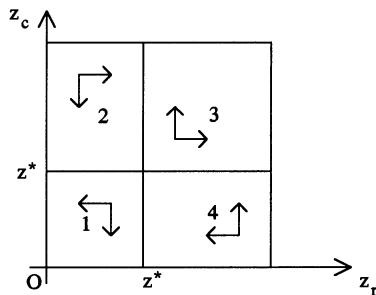


FIGURE 5

mutations. This will bring the state into region 3 (by $[M/(a + 1)]$ mutations in the row players and in the column players respectively). Therefore we must compare $[Ma/(a + 1)]$ and $2[M/(a + 1)]$. Thus, the long run equilibrium is $(1, 1)$ if $a < 2$ and (a, a) if $a > 2$.

This result is to be contrasted with the case where the Darwinian adjustment process exhibits equal speeds of adjustment in all regions. In that case, the risk dominant equilibrium (a, a) is the unique long run equilibrium. And this is independent of how a compares with 2. So this serves to illustrate that the selection of a long run equilibrium may depend on the specification of the Darwinian adjustment process, not just on the payoff structure of the game. This discussion also suggests that if a is large enough, then the results will be independent of the deterministic dynamics.

REMARK: If the Darwinian adjustment process was stochastic, i.e., if each player was independently allowed to pick a best-response with certain probability per-unit time then the choice of long-run equilibrium is again independent of the Darwinian speed of adjustment. This holds since in that case regions 2 and 4 both belong to the basin of attraction of the $(1, 1)$ equilibrium *and* the (a, a) equilibrium (so there is an overlap in the state-space segmentation into basins of attraction). Thus, the most efficient escape routes are along the horizontal (or vertical) edges, generating a comparison between $[M/(a + 1)]$ and $[aM/(a + 1)]$ (as before).

Faculty of Economics, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113, Japan, and Dept. of Economics, Princeton University, Princeton, NJ 08544, U.S.A.

Dept. of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, U.S.A.,

and

Dept. of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, U.S.A.

Manuscript received January, 1991; final revision received July, 1992.

APPENDIX

In the claim, $b^m(\cdot)$ is the m -fold iteration of $b(\cdot)$ and $b^0(z) = z$. Recall that, by (D), for any z in the basin of attraction of N , $b^m(z) = N$ for sufficiently large m .

LEMMA: *Let $A = \{0, 1, \dots, \alpha - 1\}$. For all $z \in A$, the minimum of $c(h')$ over all paths h' in Z from z to N is achieved by $h^* = \{(z \rightarrow \alpha)\} \cup \{(z' \rightarrow b(z'))|z' = b^m(\alpha) \text{ for some } m \geq 0\}$.*

PROOF: For any path h' from z to N , define $k(h') \equiv \#\{z' \in A : \exists z'' \in Z, (z' \rightarrow z'') \in h'\}$, i.e., $k(h')$ is the number of states in A which are the initial state of an arrow in h' . Define $C(z; \kappa) \equiv \min\{c(h') : h' \text{ is a path in } Z \text{ from } z \text{ to } N \text{ and } k(h') \leq \kappa\}$. The claim asserts that $c(h^*) = C(z; \kappa)$, for all κ . The claim is proved by induction on κ .

Consider first the case of $\kappa = 1$. For this case, z is the only state in A which can be the initial state of an arrow in h' , so the immediate successor of z , denoted z^0 , lies in $Z \setminus A$ (i.e., $z^0 \geq \alpha$). The cost of a path from z to N is at least $z^0 - b(z)$, and this is greater than or equal to $c(h^*) = \alpha - b(z)$. Hence, $c(h^*) = C(z; 1)$.

Assume, now, that for some κ , $2 \leq \kappa \leq \alpha$, and for all $z \in A$, $c(h^*) = C(z; \kappa - 1)$. Fix a starting state $z \in A$. Let z^0 be the immediate successor of z in some path h^0 which is cost minimizing over all paths h' in Z from z to N satisfying $k(h') \leq \kappa$, i.e., $c(h^0) = C(z; \kappa)$. It is clear that $z^0 \leq \alpha$ (otherwise $c(h^0) \geq z^0 - b(z) > \alpha - b(z) = c(h^*)$). If $z^0 = \alpha$, then h^* clearly satisfies $c(h^*) = C(z; \kappa)$. Suppose, therefore, $z^0 < \alpha$. By the induction hypothesis, we have $c(\{(z^0 \rightarrow \alpha)\} \cup \{(z' \rightarrow b(z'))|z' = b^m(\alpha) \text{ for some } m \geq 1\}) = C(z^0, \kappa - 1)$. This implies that $c(h^*) = c(h^0)$, where

$$h^* = \{(z \rightarrow z^0)\} \cup \{(z^0 \rightarrow \alpha)\} \cup \{(z' \rightarrow b(z'))|z' = b^m(\alpha) \text{ for some } m \geq 1\}.$$

We now show that $c(h^*) \geq c(h^*)$, which will prove the claim. Note that condition (D) implies $b(z^0) \leq z^0$. First, suppose $z^0 \leq b(z)$, which implies $b(z^0) \leq b(z)$. Then $c(h^*) = b(z) - z^0 + \alpha - b(z^0) \geq \alpha - b(z^0) \geq \alpha - b(z) = c(h^*)$. Next, if $b(z) < z^0$, then $c(h^*) = z^0 - b(z) + \alpha - b(z^0) = c(h^*) + z^0 - b(z^0) \geq c(h^*)$, and the claim is proved. *Q.E.D.*

REFERENCES

- ALCHIAN, A. (1950): "Uncertainty, Evolution and Economic Theory," *Journal of Political Economy*, 58, 211–221.
- ANDERLINI, L. (1989): "Communication, Computability and Common Interest Games," Mimeo, Cambridge University.
- AUMANN, R., AND S. SORIN (1989): "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1, 5–39.
- ARTHUR, W. B. (1984): "Competing Technologies and Economic Prediction," *Options*. Laxenburg, Austria: International Institute for Applied Systems Analysis.
- BEWLEY, T. (1986): "Knightian Decision Theory, Part 1," Cowles Foundation Paper No. 807, Yale University.
- BINMORE, K. (1987): "Modelling Rational Players, Part I," *Economics and Philosophy*, 3, 179–214.
- (1988): "Modelling Rational Players, Part II," *Economics and Philosophy*, 4, 9–55.
- BINMORE, K., AND L. SAMUELSON (1992): "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory*, 57, 278–305.
- BLUME, L., AND D. EASLEY (1991): "Wealth Dynamics and the Market Selection Hypothesis," mimeo, Cornell University.
- CANNING, D. (1989): "Convergence to Equilibrium in a Sequence of Games with Learning," LSE Discussion Paper TE/89/190.
- (1992): "Average Behavior in Learning Models," *Journal of Economic Theory*, 57, 442–472.
- DAVID, P. (1988): "Path-Dependence: Putting the Past into the Future of Economics," IMSSS Technical Report No. 533, Stanford University.
- DIXIT, A., AND B. NALEBUFF (1991): *Thinking Strategically*. New York: W. W. Norton.
- ELLISON, G. (1991): "Learning, Local Interaction, and Coordination," mimeo, MIT.
- FOSTER, D., AND P. YOUNG (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219–232.
- FREIDLIN, M. I., AND A. D. WENTZELL (1984): *Random Perturbations of Dynamical Systems*. New York: Springer Verlag.

- FRIEDMAN, D. (1991): "Evolutionary Games in Economics," *Econometrica*, 59, 637–666.
- FRIEDMAN, M. (1953): *Essays in Positive Economics*. Chicago: University of Chicago Press.
- FUDENBERG, D., AND C. HARRIS (1992): "Evolutionary Dynamics with Aggregate Shocks," *Journal of Economic Theory*, 57, 420–441.
- FUDENBERG, D., AND D. KREPS (1989): "A Theory of Learning, Experimentation, and Equilibrium in Games," mimeo, MIT and Stanford.
- FUDENBERG, D., AND E. MASKIN (1990): "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80, 274–279.
- HARSANYI, J. C., AND R. SELTEN (1988): *A General Theory of Equilibrium in Games*. Cambridge: MIT Press.
- HINES, W. G. S. (1987): "Evolutionary Stable Strategies: A Review of Basic Theory," *Theoretical Population Biology*, 31, 195–272.
- KANDORI, M. (1991): "Independent Mutations and Long Run Equilibria in Games: A Continuous Time Formulation," mimeo, Princeton University.
- KARLIN, S., AND H. M. TAYLOR (1975): *A First Course in Stochastic Processes*, Second Edition. San Diego: Academic Press.
- KIRKPATRICK, S., C. D. GELATT, JR., AND M. P. VECCHI (1983): "Optimization by Simulated Annealing," *Science*, 220, 671–680.
- LIPMAN, B. (1991): "How to Decide How to Decide to...: Modeling Limited Rationality," *Econometrica*, 59, 1105–1125.
- MAILATH, G. J. (1992): "Introduction: Symposium on Evolutionary Game Theory," *Journal of Economic Theory*, 57, 259–277.
- MATSUI, A. (1991): "Cheap Talk and Cooperation in the Society," *Journal of Economic Theory*, 54, 245–258.
- MATSUI, A., AND R. ROB (1991): "Evolution, Rationality and Equilibrium Selection in Societal Games," mimeo, University of Pennsylvania.
- MAYNARD SMITH, J. (1982): *Evolution and the Theory of Games*. Cambridge, United Kingdom: Cambridge University Press.
- MAYNARD SMITH, J., AND G. R. PRICE (1973): "The Logic of Animal Conflicts," *Nature*, 246, 15–18.
- NACHBAR, J. H. (1990): "Evolutionary Selection Dynamics in Games: Convergence and Limit Properties," *International Journal of Game Theory*, 19, 59–89.
- NELSON, R. R., AND S. G. WINTER (1982): *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press of Harvard University Press.
- SELTEN, R. (1991): "Evolution, Learning, and Economic Behavior," *Games and Economic Behavior*, 3, 3–24.
- TAYLOR, P. D., AND L. B. JONKER (1978): "Evolutionarily Stable Strategies and Game Dynamics," *Mathematical Biosciences*, 40, 145–156.
- VAN DAMME, E. (1987): *Stability and Perfection of Nash Equilibria*. Berlin: Springer Verlag.