JOURNAL OF ECONOMIC THEORY **65**, 383–414 (1995)

Evolution of Equilibria in the Long Run: A General Theory and Applications*

MICHIHIRO KANDORI†

Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

AND

RAFAEL ROB‡

Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, Pennsylvania 19104-6297

Received October 2, 1992; revised August 25, 1994

We extend the evolutionary process studied in Kandori *et al.*, Econometrica **61** (1993), 29–56, to $n \times n$ games. The evolutionary process is driven by two forces: players switching to the best response against the present strategy configuration, and players experimenting with new strategies. We show that a unique behavior pattern, called the long-run equilibrium, arises even if the underlying game has multiple (static) equilibria. The paper gives a general algorithm for computing the LRE, and then applies it to two classes of economic games. For games of pure coordination, the LRE is the Pareto-efficient equilibrium. For games with strategic complementarities, the geometry of the best-response correspondence helps identify the LRE. *Journal of Economic Literature* Classification Numbers: C63, C72, D43.

I. INTRODUCTION

In this paper we consider an evolutive process of learning in which players are randomly and repeatedly matched to play a two-person stage game. The purpose of this is to generate a positive theory which shows how a Nash equilibrium may be reached in the realistic situation where players possess neither full rationality nor a congruent set of expectations. A second purpose is to show that players will most frequently play a

* We thank the referees and many seminar audiences for various suggestions. Responsibility for remaining errors lies with us.

† Financial support from the NSF (SES-9108351) is gratefully acknowledged.

‡ This project was started while Rafael Rob was visiting Princeton University, 1990/1991.

0022-0531/95 $6.00

particular equilibrium, even though the stage game possesses multiple strict Nash equilibria.

The *evolutionary* approach deals with these issues by analyzing the behavior of boundedly rational players who are not capable of deducing their opponents' actions at a point in time, but who play the game repeatedly and are thereby able to *observe* (or "learn") those actions. After observing others' actions players adjust their behavior, i.e., they abandon losing strategies in favor of winning strategies. In addition to this, players experiment with (or mutate toward) new strategies independently of their payoff experience. This changes the strategy profile, which triggers further adjustments. A trial and error process is thereby generated which, in the long run, singles out a stable configuration. If such configuration corresponds to a Nash equilibrium of the stage game we can interpret this equilibrium as the eventual outcome of a learning and adjustment process. This paper shows how such process works for economic games with multiple equilibria, and how the equilibrium it selects can be identified from the structure of the underlying game.

The study of adjustment processes perturbed by "mutations" has attracted much attention lately and has been labeled *Evolutionary Game Theory*.[1] The characterization of *locally* stable equilibrium in such a system —also known as *evolutionary stable strategy* (ESS)—was initiated by the seminal work of Maynard Smith and Price [25], followed by the dynamic models of Taylor and Jonker [39]. More recently, Foster and Young [10] introduced an evolutive process which is *repeatedly* perturbed by random mutations and have shown that it possesses much stronger selection properties. Namely, the repeatedly perturbed process can select among *strict* Nash equilibria, which the ESS cannot. The basic idea behind the selection is that any locally stable configuration (equilibrium or limit cycle) is bound to be upset by a series of mutations. Some configurations, however, are more difficult to upset and are, therefore, likely to appear *more frequently* over long time horizons. Furthermore, the identity of these configurations is independent of where the adjustment process is started.[2] Therefore, this approach pins down a particular configuration, which we term the *long-run equilibrium* (LRE hereafter). Subsequent work by Kandori *et al.* [16] (KMR hereafter) reformulated this to an economic

[1] van Damme [41] and Mailath [20, 21] provide surveys of this field. Recent modifications and economic applications include Matsui [23] (for pre-play communications) and Binmore and Samuelson [1] (for cooperation in repeated games).

[2] This is in sharp contrast to the traditional literature on adjustment processes without random shocks (the Cournot tatônnement literature: Cournot [5], Seade [35], and Moulin [33]; or the fictitious-play literature: Miyasawa [29], Shapley [37], Krishna [19], and Monderer and Shapley [30]), where the prediction depends, in general, on the initial condition.

model of behavior where the population of players is finite and where perturbations come from experimentation at the *individual-player level*. Other papers based on the finite formulation include Ellison [9], Kandori [15], Noldeke and Samuelson [33], Samuelson [34], and Young [43]. Fudenberg and Harris [12] provide a useful discussion of the continuous model.

Relative to previous papers, this paper will cover the following topics. First of all, it considers a broader class of myopic adjustment processes where the identity of players who adjust and the speed of adjustment may depend on the configuration of strategies they face. The idea here is that the players whose payoff would be most increased by switching to a best response have the strongest incentive to adjust.[3] Also, configurations which are far away from a Nash equilibrium may give players stronger incentives to adjust than configurations that are close by. Our framework accommodates these possibilities, while showing that the long-run behavior is independent of further specifications of the adjustment process.[4] Therefore, the predictions we generate are robust to various specifications of the adjustment process.

The second contribution here is that we provide further details on how the algorithm to compute the LRE works. The basic idea of the algorithm was first given in Freidlin and Wentzell [11] for the continuous case. The discrete version was then introduced by KMR [16] and was generalized by Young [43]. Basically, it works as follows. First, it identifies a collection of *limit sets*. Those are the sets toward which the process tends under the best-response adjustment alone, i.e., without mutations. Second, it computes *costs of transition* between the various limit sets. Third, it computes a minimum cost *spanning-tree* among the limit sets. The root(s) of this tree is (are) the LRE. The present paper provides systematic analysis of the algorithm for various economic models.

In particular, we identify two classes of games in which a Nash equilibrium emerges in the long run (instead of a limit cycle). The first class is games with pure coordination: two players receive positive payoff if they choose the same strategy; otherwise, they get zero. One scenario which fits this description is when players choose among incompatible computers. For this game we show that the Pareto-efficient equilibrium, i.e., the best computer is the unique LRE.

The second application is to games with strategic complementarities. In those games each player's marginal payoff is increasing in the rival's strategy. A prime example is the differentiated-product oligopoly game with

[3] A slight modification of the model can also accommodate the possibility that some players are more perceptive than others and, therefore, can adjust faster.

[4] This can be contrasted with Young's [43] independent contribution. His model utilizes a specialized adjustment rule (see Footnote 11 below). On the other hand, Young considers more general classes of games than ours.

prices as strategic variables. We analyze such games under one further continuity assumption and show how the geometry of their best-response correspondence helps determine the LRE. For example, when the game possesses an equilibrium with a *uniformly deepest* basin of attraction (see below for precise definitions) this equilibrium is the unique LRE. Such results may be of interest in the macroeconomic context as well since many models of "coordination failure" (see Cooper and John [4]) have the structure of strategic complementarity and are known to have multiple static equilibria, although previous literatures did not address the question *which* equilibrium is most likely to arise. The results we obtain here are potentially useful in this context because they help characterize the equilibrium which is singled out by the dynamics.

Both applications—games with pure coordination and games with strategic complementarities—show how the stochastic evolutionary framework applies to economic problems. In particular, they show the connection between the game's payoff functions and their LRE, which is the first step towards the comparative statics analysis of such games.[5]

The remainder of the paper is organized as follows. The next section introduces the underlying game, the societal game, and the adjustment process. Section III introduces the equilibrium concept and defines long-run states and limit sets. In Section IV we analyze the model and show how to compute the long-run states. Section V applies the general theory to games with pure coordination and to those with strategic complementarities. We first prove global convergence results for them and then show how their LRE are determined.

II. Formulation

We consider a symmetric two-person game[6] with n strategies. The strategy set is $\{1, 2, ..., n\}$ with generic elements i, j. When a player and her opponent choose strategies i and j respectively, the player's payoff is u_{ij}. A mixed strategy is represented by a point in the $(n-1)$-dimensional simplex, Δ, and the payoff of strategy i against mixed strategy $\alpha \in \Delta$ is denoted by

$$u(i, \alpha) \equiv \sum_{j=1}^{n} \alpha_j u_{ij}.$$

[5] A recent contribution along these line is Milgrom and Roberts [28] who analyze the comparative statics analysis of the equilibrium *correspondence*.

[6] Extension to asymmetric N-player case is straightforward. We can employ N-populations, one for each role in the stage game, and assume that all players are randomly matched to form N-tuples. Although this formulation makes the state space a product set, all the definitions and results in Sections II and III apply. The Matching Pennies example (see working paper version) illustrates such a formulation.

The set of *pure*-strategy best responses against strategy $\alpha \in \Delta$ is denoted by $BR(\alpha)$.

Given a two-person game, the following repeated societal game is considered. Society consists of M individuals (or "members"). At the beginning of each period each individual chooses a pure strategy and sticks to it for the duration of the period. The configuration of strategy choices in the society is summarized by the *state vector* z, whose ith element, z_i, represents the number of players with strategy i. The *state space* is a finite set (a collection of grid points in the simplex):

$$Z \equiv \left\{ (z_1, ..., z_n) \mid Z_i \in \{0, 1, ..., M\},\ \sum_{i=1}^{n} z_i = M \right\}. \tag{2.1}$$

For each $z \in Z$, we define the set of existing strategies by $C(z) \equiv \{ i \mid z_i > 0 \}$.

Once individuals have chosen strategies they are randomly matched and each pair of matched players plays the above (two-person) game, implementing their preselected strategies. At the end of the period, an individual who chose strategy i collects an average payoff of

$$\pi_i(z) \equiv \frac{1}{M-1} \left[\sum_{j \neq i} z_j u_{ij} + (z_i - 1) u_{ij} \right] = \frac{1}{M-1} \left[\sum_{j=1}^{n} z_j u_{ij} - u_{ij} \right]$$

(assuming that M is even[7]). $(z_i - 1)$ in the above formula reflects the impossibility of being matched with oneself. To capture this, it is convenient to consider the strategy distribution faced by player with strategy i; denote it by $\alpha(z, i)$. Formally, for $i \in C(z)$, we define

$$\alpha_j(z, i) = z_j/(M-1) \qquad \text{for} \quad j \neq i \text{ and } \alpha_i(z, i) = (z_i - 1)/(M-1). \tag{2.2}$$

The set of pure-strategy best responses against this distribution is $BR(\alpha(z, i))$, and denoted $\beta_i(z)$. Simply put, $\beta_i(z)$ represents the best responses for player with strategy i when the state is z.

Now we introduce a dynamic process where (1) the population of players gradually adjusts toward a configuration of best responses and (2) non-best responses are adopted with small probabilities. To illustrate this process, consider the choice of personal computers in an Economics department as a leading example. Assume there are M faculty members, each of whom is using one of n types of computers. Assume also that the opportunity to buy a new computer arrives stochastically. For example,

[7] If M is odd, there is one unmatched player, and we assume that players get zero payoffs if they are not matched. In that case, the number $M-1$ in the above expression must be replaced with M. For concreteness, we assume that the population size, M, is even but all the results also hold if M is odd.

one may switch to a new computer when the current one is broken, research fund is granted, or the current project is finished. Formally we assume that the opportunity to adjust arrives *independently*—across players and time—and with *strictly positive* probability, η. This probability may depend on the current strategy distribution and the strategy one is using; thus we write $\eta = \eta(z, i)$. A player with strategy i under state z can adjust with this probability. When he can adjust, we assume that he chooses a myopic best response, an element[8] of $\beta_i(z)$. One reason for allowing η to depend on z and i is that players may have different "urgencies" to adjust.[9] For instance, it would be natural to expect that $\pi_i(z) > \pi_j(z)$ implies that $\eta(z, j) > \eta(z, i)$.

This defines a law of motion toward myopic best responses, which is similar to Gilboa and Matsui [13]. Following their terminology, we call it *best-response dynamic*. There are, however, important differences between our formulation and theirs. They consider a *continuum* of players and assume that a *deterministic* fraction, say $\eta\%$, of players always adjust. Thus the state moves along a straight line toward the best response as in Fig. 1a. In our finite population model, however, this is not always feasible, since the state space is a collection of *grid points* in the simplex. Instead, we assume that each player adjusts with *probability* η, and as a result, *adjustment toward any point in the shaded region in Fig.* 1b *is possible*. In this figure strategy 2 is the best response against z for all players, but the exact direction of adjustment depends on who adjusts. For example, adjustment toward z' realizes when a large fraction of players with strategy 3 happens to adjust to strategy 2, whereas point z'' realizes when a large fraction of players with strategy 1 adjust. Similarly, the *speed* of adjustment depends on *how many* players are adjusting.

[8] When there are multiple best responses, we assume that the player randomly chooses one. Let $\gamma_i(z)$ be the probability distribution over $\beta_i(z)$, which represents the random choice of a best response. We assume that $\gamma_i(z)$ puts probability one on i if $i \in \beta_i(z)$, and has full support otherwise. Those assumptions are introduced for concreteness and do not play major roles. The first assumption is reasonable if there is always some switching cost and guarantees that the state cannot drift away from a mixed strategy equilibrium until "mutations" (which we will introduce shortly) come in. One of our assertions in the paper is the instability of mixed strategy equilibria for some classes of games, and this assumption provides the hardest case for proving such an assertion. The full support assumption plays a role only in the analysis of supermodular games with linear payoffs.

[9] This can potentially be justified by stochastic adjustment cost. Let c be random cost of adjustment, which has distribution function $F(c)$, and let $\Delta v_i(z)$ be the increment of the discounted payoff when a player switches from her current strategy i to the myopic best response under state z. Assuming heavy discounting, the optimal switching is toward the myopic best response, and this happens if and only if $c \leqslant \Delta v_i(z)$. Hence the adjustment rate is *derived* by $\eta(z, i) = F(\Delta v_i(z))$.

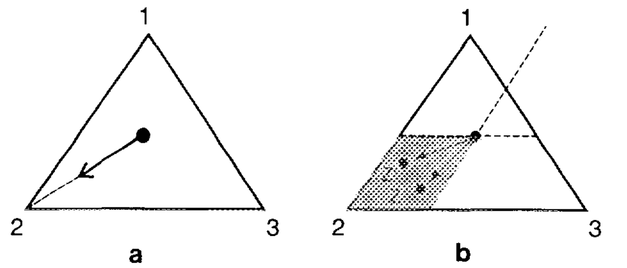

FIGURE 1

Figure 1 illustrates a feature of our formulation, *heterogeneity of adjustments*. Any specification of $\eta(z, i)$ generates a particular probability distribution over the shaded region which indicates the likely directions and speeds of adjustment. Our formulation allows *any* such distribution, as long as it has full support (i.e., as long as $\eta(z, i) > 0$), and the main results do not hinge on the exact values that $\eta(z, i)$'s assume.[10] This is shown in Theorem 1. In addition to this property (*robustness*), our formulation has two further advantages. As we will see the full support assumption drastically simplifies the calculation of the long-run equilibria. Furthermore, the adjustment rule can be interpreted as rational behavior in certain circumstances. In particular, if the adjustment is slow (η being small) and if players discount the future heavily (which may be restrictive for some applications), the present state is expected to persist for a while; thus taking a myopic best response is in fact a dynamically rational choice.[11]

In addition to the best-response dynamic, we postulate that new strategies enter into the population with small probabilities. More specifically, we assume that an individual who is expected to play strategy k "mutates" to strategy j with probability $m_j\varepsilon > 0$, where $\sum_j m_j = 1$, and m_j, $\varepsilon \in (0, 1)$. These mutations are independent across players and over time and occur (if they occur) after the best-response adjustment. One way to justify this assumption in economic contexts is to assume turnover of the

[10] This is a feature of the discrete (i.e., finite population) formulation and is in sharp constrast with the continuous models of Foster and Young [10] or Fudenberg and Harris [12], where the predictions depend on details of the adjustment process—*even for games with two strategies*. If the game has more than two strategies there is yet more leeway in specifying the direction of adjustment and, therefore, more scope for sensitivity of predictions. Kandori [15] examines a possible source of the difference between the continuous and discrete models.

[11] The last two points can be contrasted with the independent contribution of Young [43]. In his model, one pair is randomly drawn from the population at each moment. Each player randomly draws a sample of size k out of her most recent m matches and takes a best reply against the resulting empirical distribution. While Young derives results for this particular adjustment rule, we show that similar results hold for a wide class of adjustments, and that it is possible to interpret the adjustments as rational behavior in certain circumstances.

population. In our story of faculty members choosing computers, each member may resign from the department with probability ε and then his position will be replaced with a newcomer (so that the population size stays constant). The newcomer may not know which computer is a good choice in the department; thus she takes each strategy with a positive probability, reflecting her prior beliefs.[12] Alternatively, we can assume that the newcomer owns already a computer and regard $m \equiv (m_1, ..., m_n)$ as the distribution of computers in the outside world. One of the striking features of our analysis is that once mutations become frequent (i.e., as $\varepsilon \downarrow 0$) the specification of m is immaterial: as long as $m = (m_1, ..., m_n)$ is fixed[13] and all m_i's are positive we get the same long-run behavior. To emphasize this fact we state our maintained assumption:

Assumption A. $m_i, \eta(z, i) > 0$, for all i and all $z \in Z$.

The composition of myopic best responses and mutations generates a Markov chain over the finite state space, Z, whose transition matrix is denoted $P(\varepsilon) = (p_{zz'}(\varepsilon))$. The $p_{zz'}(\varepsilon)$ element of this matrix represents $\text{Prob}(z(t+1) = z' \mid z(t) = z)$. The best-response dynamic corresponds to $P(0)$ where the mutation rate is zero. A positive mutation rate transforms this into a "perturbed" system, $P(\varepsilon)$. Since mutations are independent, it can be seen that $p_{zz'}(\varepsilon)$ is a polynomial in ε.

To analyze the behavior of our system, we first define the *distance* between two states z and z' by $d(z, z') \equiv (1/2) \sum_i |z_i - z'_i|$. This represents the minimum number of strategy changes which are necessary to achieve state z' from state z. Given this definition of distance, we introduce the *cost of transition* between two states, which plays a crucial role in what follows.

DEFINITION 1. The cost of transition between z and z' is defined by

$$c(z, z') = \min_{z'' \in b(z)} d(z', z''), \tag{2.3}$$

where $b(z)$ is the set of possible intended strategy configurations at $t+1$ when strategy configuration at time t is given by z, i.e., $b(z) \equiv \{z' \in Z \mid p_{zz'}(0) > 0\}$. This corresponds to the shaded region in Fig. 1b.

The number $c(z, z')$ defined above measures *the minimum number of mutations to achieve state z' from state z in one period.* By the independence

[12] In the case of a rational player who knows the payoff function, it would be better to assume that the newcomer picks only *rationalizable* strategies with positive probabilities. We will come back to this issue later when we analyze supermodular games (Section V.C.2.).

[13] When players hold forward-looking expectations, it may be sensible to consider a mutation process where m_i is allowed to depend on the state, z, and thus where mutations may eventually cease; see Matsui and Rob [24].

of mutations, $p_{zz'}(\varepsilon)$ is a polynomial: $a_0\varepsilon^k + a_1\varepsilon^{k+1} + \cdots + a_r\varepsilon^{k+r}$. Since the transition from z to z' requires at least $c(z, z')$ mutations, the order of this polynomial, k, must be $c(z, z')$. In other words, we have

$$p_{zz'}(\varepsilon) = O(\varepsilon^{c(z, z')}). \tag{2.4}$$

That is, the cost of transition $c(z, z')$ measures *how fast* the transition probability from z to z' tends to zero as the mutation rate (ε) goes to zero.

III. The Equilibrium Concept

The presence of mutations implies that the society will perpetually fluctuate among the different states, $z \in Z$. Accordingly, we consider a stochastic equilibrium concept, measuring the long run (or time-average) behavior of the system. Formally, let $\Delta^{|Z|}$ denote the $|Z| - 1$ dimensional simplex, where $|Z|$ is the number of elements in the state space. Then, we introduce the following concept.

Definition 2. $\mu(\varepsilon) \in \Delta^{|Z|}$ is a *stationary distribution* if

$$\mu(\varepsilon)\, P(\varepsilon) = \mu(\varepsilon). \tag{3.1}$$

When the mutation rate ε is positive, the system can jump from any initial state to any final state, since we have strictly positive transition probabilities, $p_{zz'}(\varepsilon) > 0$. This in particular implies that the system is irreducible and aperiodic,[14] and the standard theory of Markov chains shows the following.[15]

Proposition 1. *If the mutation rate ε is strictly positive, then we have*

Uniqueness. There exists a unique stationary distribution, $\mu(\varepsilon)$.

Global stability. For any initial distribution, q, the future distribution converges to $\mu(\varepsilon)$: $\lim_{t \to \infty} qP(\varepsilon)^t = \mu(\varepsilon)$.

Time average property. $\mu(\varepsilon)$ represents the average time spent on each state. Let $1_z(z')$ be 1 if $z = z'$ and 0 otherwise. Then the random variable

[14] A Markov chain is irreducible if $\text{Prob}(z(T) = z' \mid z(0) = z) > 0$ for all z and z' for some $T \in \mathbf{N}$ ($\mathbf{N}$ is the set of natural numbers). It is aperiodic if the greatest common divisor of $\{T \in \mathbf{N} \mid \text{Prob}(z(T) = z' \mid z(0) = z) > 0\}$ is 1 for all z and z'.

[15] These results can be found in Hoel *et al.* [14]. The uniqueness and the time average property follows from the finiteness of the state space and the irreducibility only and *do not require aperiodicity* (Hoel *et al.* [14, Corollaries 4–6, pp. 66–67]). The global stability requires the aperiodicity as well as the finiteness and irreducibility [14, Corollaries 4, 5, and Theorem 7, pp. 66–73]).

$[1_z(z(1)) + \cdots + 1_z(z(t))]/t$ *tends to the number* $\mu_z(\varepsilon)$ *almost surely as* $t \to \infty$.

To illustrate the meaning of these properties, suppose $\mu_z(\varepsilon) = 0.9$ for some state z. The global stability means that if we were to "sample" the state of the system after a long enough time, the state z should come up with probability 0.9, irrespective of where we started. The time average property means that over a long time horizon, the system spends 90% of the time at state z. Hence the system is well behaved when the mutation rate is positive, and its long-run behavior is nicely summarized by the stationary distribution $\mu(\varepsilon)$.

This is in sharp contrast to the case of $\varepsilon = 0$, where the linear system of equations (3.1) may very well possess multiple solutions. This will occur whenever the underlying game possesses multiple pure strategy Nash equilibria, because the distribution which puts probability one to the configuration where all players choose the same pure strategy Nash equilibrium is always a stationary distribution of $P(0)$. Moreover, if the best-response dynamic, $P(0)$, possesses a limit cycle, there is another stationary distribution which puts all probability to the states on the cycle. Unlike the case with positive mutation rate, which distribution is eventually chosen is crucially dependent on the initial condition. The purpose of introducing (a small rate of) mutations is to resolve this indeterminancy and select a *particular* stationary distribution. This idea is reflected in the following definition (see also KMR [16]) and proposition.

DEFINITION 3. (i) The *limit distribution* $\mu^* \in \Delta^{|z|}$ is defined by $\mu^* = \lim_{\varepsilon \to 0} \mu(\varepsilon)$. (ii) The set of *long-run states* is the support of $\mu^*: \{z \in Z \mid \mu^*_z > 0\}$.

Note that neither $P(0)$ nor $P(\varepsilon)$ converges over time to μ^*. Instead, for a small (but fixed) ε the process $P(\varepsilon)$ converges over time to $\mu(\varepsilon)$ which *approximately* equals μ^*. Thus, μ^* is the approximate time average of different states as the mutation rate is made arbitrarily small ($\varepsilon \to 0$).

PROPOSITION 2. *There exists a unique limit distribution, and it is a stationary distribution of the best—response dynamic* $P(0)$: $\mu^* P(0) = \mu^*$.

Proof. The existence of $\lim_{\varepsilon \to 0} P(\varepsilon)$ and $\lim_{\varepsilon \to 0} \mu(\varepsilon)$ comes from the fact that (i) the elements of $P(\varepsilon)$ are polynomials in ε, and (ii) the elements of $\mu(\varepsilon)$ lie in $[0, 1]$ and are the ratios of polynomials in ε (by Cramer's rule). Thus we can take the limits of both sides of $\mu(\varepsilon) P(\varepsilon) = \mu(\varepsilon)$. ∎

The limit distribution captures the long-run behavior of the system (in the sense of Proposition 1) when the mutation rate is low. It should be noted, however, that the lower the mutation rate, the longer we must wait to see the long-run effects. In particular, the length of time needed to reach the limit distribution is of the order $1/\varepsilon^K$, where K is the number of mutations needed to disrupt (see below) a non-limit-distribution and K is proportional to the population size; see KMR [16]. Thus, our analysis is most relevant for a small population. On the other hand, Ellison's [9] analysis shows that this time can be considerably shorter when the players interact *locally* and the long-run equilibirum has large enough basin of attraction. When those conditions are met, the long-run effects show within a reasonably short time, even for a large population.

Having provided the basic idea and the definitions, we will examine our equilibrium selection procedure in more detail. For the dynamic without mutations (i.e., $P(0)$), we have seen that the equilibrium points and the limit cycles provide different stationary distributions. Those objects are limit sets which are formally defined as follows.

Definition 4. A set $A \subset Z$ is a limit set if, under $P(0)$,

(i) $\Pr(z(t+1) \in A \mid z(t) \in A) = 1$, and

(ii) for all $z, z' \in A$, $\Pr(z(t+k) = z' \mid z(t) = z) > 0$ for some $k > 0$.

The collection of all limit sets is denoted Ω.

A limit set is closed in a probabilistic sense (Definition 4(i)) and its elements are mutually reachable (4(ii)).[16] By the definition, it is easy to see that *the state under the best-response dynamic will eventually converge to one of the limit sets with probability one*. It is well known that there is a unique stationary distribution, μ_A, for *each* limit set A, and the set of all stationary distributions of $P(0)$ is the convex hull of those, $\mathrm{Co}\{\mu_A, A \in \Omega\}$.[17] As indicated above, out of those multiple distributions we select a particular distribution, by first introducing a positive mutation rate ε and then let it tend to zero: $\mu^* = \lim_{\varepsilon \to 0} \mu(\varepsilon)$. This selection procedure is continuous, which is shown in Proposition 2 above.

Given the above argument, we have the following proposition which gives the classification of the long-run states.

Proposition 3. *The limit distribution μ^* is uniquely decomposed as*

$$\mu^* = \sum_{A \in \Omega^*} r_A \mu_A,$$

[16] A limit set is a special case of what is known as a *communication class*, i.e., it is a non-transient communication class.

[17] This is shown as follows. By Definition 4, $P(0)$ induces a finite, irreducible Markov chain on each limit set. Then the result cited in Footnote 15 shows that the unique stationary distribution exists for the induced chain.

where Ω^* is a subset of Ω, μ_A is the unique stationary distribution on the limit set A, and $r_A \in (0, 1]$ represents the likelihood of the limit set A. An element of Ω^* is called a long-run equilibrium.

In what follows we will show that for some classes of games, a particular Nash equilibrium is the unique long-run equilibrium. In general, however, long-run equilibrium can be a limit cycle,[18] rather than a steady state of the best-response dynamic, $P(0)$.

IV. Algorithm

We start with the definition of the cost of transition between two limit sets. For $z, z' \in Z$, let $G(z, z')$ be the set of *directed paths* from z to z'. A directed path form z to z' is a sequence of states $(z^1, z^2, ..., z^T)$, where $z^1 = z$ and $z^T = z'$. Given such path, $g = (z^1, z^2, ..., z^T)$, we can count the total number of mutations on it as

$$N(g) \equiv \sum_{t=1}^{T-1} c(z^t, z^{t+1}).$$

Recall that $c(z^t, z^{t+1})$ is the cost of transition *between two states*, and it represents the minimum number of mutations to achieve state z^{t+1} from z^t in *one* period (Definition 1). Given this, we define the cost of transition *between two limit sets* $A, A' \in \Omega$, as

$$C(A, A') = \min_{z \in A,\, z' \in A'} \min_{g \in G'(z, z')} N(g),$$

where for $z \in A$ and $z' \in A'$, $G'(z, z')$ represents the collection of paths from z to z' *which do not intersect with any other limit set* $A'' \neq A, A'$. This represents the required number of mutations to achieve A' from A *over time*.

Next, we consider a reduced Markov chain defined on the limits sets. Note that if we restrict our attention to the subset of periods when the state is in the limit sets, the resulting process is again a Markov chain. Formally, let X be the set of states in the limit sets, and let $\tau(t)$ be the tth data when the state z lies in X. Then $x(t) = z(\tau(t))$ can be regarded as a Markov chain (sometimes called an imbedded process) defined on

[18] The working paper version shows an example of a limit cycle which is a long-run equilibrium.

the reduced state space, X. Then we can apply the Freidlin–Wentzell's [11] graph theoretic technique to this reduced Markov chain; for a self-contained explanation of this approach, see KMR [16]. Let P' be the transition matrix of the reduced chain, and let $c'(x, x')$ be the order of smallness of the transition probability ($p'_{xx'} = O(\varepsilon^{c'(x, x')})$). Then a moment's reflection shows that $c'(x, x') = C(A, A')$ for $x \in A$ and $x' \in A'$. From this observation it can be seen that the Freidlin–Wentzell (1984) approach reduces to a program defined on the *collection of limit sets*[19]:

$$\operatorname*{Min}_{A \in \Omega} \operatorname*{Min}_{h \in H_A} \sum_{(A', A'') \in h} C(A', A''). \tag{4.1}$$

H_A in the above expression refers to the set of A-trees *whose vertices are the collection of limit sets*, Ω. An A-tree is a collection of directed branches (A^0, A^1) (A^1 being the successor of A^0), where (1) except for A, each limit set has a unique successor, and (2) there are no closed loops. In other words, it is a tree directed into root A. Formally we have:

PROPOSITION 4. *The set of long-run equilibria is the solutions to program* (4.1).

This is a discrete version of Freidlin and Wentzell's analysis, which was worked out for the continuous time and continuous state space case [16, Chap. 6]. A direct proof of Proposition 4 was first given by Young [43],[20] using a "cutting and pasting" technique (Lemma 2 in his Appendix). The above argument, based on the reduced chain, shows an alternative proof.

Now it is easy to see why the set of long-run equilibria is independent of modeling details. Program (4.1) involves the costs of transition, $C(A, A')$, and those depend on the basic nature of the dynamic process which, in turn, depends only on the payoff function. More specifically, the cost of transition depends only on the *support* of the intended state distribution ($b(z)$ in Definition 1) and not on the distribution itself. Also, what matters is the rate of convergence of the probability of mutations for each strategy, while the exact distribution is immaterial. Thus, we have:

[19] Formally, the Chu–Liu/Edmonds algorithm (see below), when applied to the minimum-cost tree problem defined on the state space Z, retracts each limit set to a single state. Therefore, the reduced problem (4.1) is equivalent to the minimum-cost tree problem defined on X.

[20] Young considers *all* paths from z to z' in the definition of $C(A, A')$, while our approach considers a *subset*, $G'(z, z')$. This difference becomes inconsequential when we take the minimum in the final program, (4.1).

THEOREM 1. *The set of long-run equilibria*[21] *is independent of the distribution of mutations* $(m_1, \ldots, m_n)$, *and the speed and directions of adjustment*, $\eta(z, i)$, *as long as* $m_i, \eta(z, i) > 0$ *for all* z *and* i.

What remains is to solve the reduced program (4.1). This can be broken into two subprograms: (1) the determination of transition costs among the limit sets, $C(A, A')$; and (2) an algorithm to solve program (4.1) *given* the transition costs. Efficient algorithms for handling both problems are available in the combinatorial optimization literature. A general algorithm for solving the first is the Dijkstra [7] algorithm, while the second, which is known as the *optimum branching problem*, is solvable by the Chu–Liu [3]/Edmonds algorithm [8]. In a working paper version we provide a complete description of the latter, and in the next section we illustrate how it works for particular classes of games. As to the costs of transition, we will determine them *directly*, using the special structure of games analyzed here. To this end, let $e_i = (0, \ldots, M, \ldots, 0)$ be the state where all players take strategy i and define the *best-response region* by $\mathrm{BR} = \{z \in Z \mid i \in \beta_j(z) \text{ for all } j \in C(z)\}$. Given this, we have:

PROPOSITION 5 (THE TRIANGLE INEQUALITY). *Suppose strategy* i *constitutes pure strategy Nash equilibrium. Then, for all* $x \in Z$ *and* $y \in \mathrm{BR}_i$, *we have*

$$c(e_i, x) \leqslant c(e_i, y) + c(y, x).$$

Note that the triangle inequality says that an immediate jump from a *Nash state*, e_i, to another state x is less costly than any gradual transition *through the best-response region of the starting point* (BR_i). The proof of Proposition 5 is available in a working paper version.

V. APPLICATIONS: SUPERMODULAR AND PURE COORDINATION GAMES

In this section we show how the above algorithm applies to two classes of games with multiple (static) equilibria. For these games the "relevant" equilibria are in pure strategies, and the best-response dynamic always converges to one of them; hence, limit cycles are ruled out. We start by defining the games and listing a few of their properties.

[21] The *distribution* over the long-run equilibria, on the other hand, may well depend on m and $\eta(z, i)$. In the class of games we consider in this paper, however, the unique long-run equilibrium is a singleton set (a Nash equilibrium); thus this problem does not arise. We conjecture that generically Ω consists of a unique limit set for large M, but we have been unable to confirm this for supermodular games. Kandori and Rob [17] show this for the class of "Bandwagon games."

V.A.1. *Supermodular Games*

Supermodularity is defined by the requirement that for any pair $1 \leqslant i < k \leqslant n$ the payoff differences $u_{kj} - u_{ij}$ are strictly increasing in j. A maintained assumption here is that the strategies are completely and linearly ordered.

Three leading examples of supermodular games are: (i) Certain differentiated-product oligopoly games in which price is the strategic variable. In particular, a discretized Hotelling model in which firms' locations are *fixed* and are sufficiently apart. (ii) The two-firm Cournot model with linear demand and linear costs. Here the strategic variable is quantity, but the natural ordering over strategies is reversed as to satisfy supermodularity (this trick will not work with more than two firms). (iii) A macroeconomic coordination game where the strategic variable is search effort, in particular, the Diamond [6]–Mortensen [31] matching model. A variety of other examples is given in Milgrom and Roberts [26].

The fact that a player's payoff differences are increasing in the opponent's strategy often leads to multiple, Pareto-ranked equilibria in supermodular games (acquiring special significance in the macroeconomic literature). A necessary and sufficient condition for this to occur is that the game is not dominance solvable. In other words, after iteratively eliminating all strictly dominated strategies, the game still possesses more than one strategy. This result is stated in the next proposition (for proof see the Appendix).

PROPOSITION 6. (i) *Suppose all strictly dominated strategies have been iteratively removed from the game. Then both* $(1, 1)$ *and* (n, n) *are Nash equilibria (NE, for short).* (ii) *No asymmetric NE in pure strategies exist:* (i, j) *is a NE only if* $i = j$. (iii) *If* l *is a best response to* i *and* k *is a BR to* j, *where* $j > i$, *then* $k \geqslant l$. (iv) *For a generically selected supermodular game all pure strategy NE are strict.*

According to Proposition 6(ii), the set of pure strategy NE is a subset of the "main diagonal," $\{(i, i)\}_{i=1}^{n}$. Let us denote this set by $N \equiv \{1 \leqslant i \leqslant n \mid (i, i) \text{ is a pure-strategy NE}\}$. The next proposition extends part (iii) of Proposition 6 to the case where we compare across *mixed* strategies (as opposed to comparing across pure strategies). In this case a similar monotonicity property obtains when mixed strategies are *partially* ordered according to *first-order stochastic-dominance*.[22] This property will play an important role in the rest of the paper.[23]

[22] The definition is as follows. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\alpha' = (\alpha'_1, \ldots, \alpha'_n)$ be two probability distributions over $\{1, \ldots, n\}$. We say that α' stochastically dominates α if the cumulative distributions of α and α', denoted $F(i) = \sum_{k \leqslant i} \alpha_k$ and $G(i) = \sum_{k \leqslant i} \alpha'_k$, are such that $F(i) \geqslant G(i)$ for all i with strict inequality for at least one i. In this case, the expected value of any increasing function under α' is no smaller than the expected value of the same function under α.

[23] Independent work by Krishna [19] contains a similar result.

PROPOSITION 7. (MONOTONICITY OF BEST-RESPONSES OVER Δ). *Assume $\alpha' \succ \alpha$, where $\succ$ refers to first-order stochastic-dominance. Then we have*

$$\text{Min BR}(\alpha') \geqslant \text{Max BR}(\alpha).$$

Proof. Let $j = \text{Max BR}(\alpha)$. For any $i < j$, we have

$$u(j, \alpha') - u(i, \alpha') = \sum_{k=1}^{n} \alpha'_k [u_{jk} - u_{ik}] > \sum_{k=1}^{n} \alpha_k [u_{jk} - u_{ik}]$$
$$= u(j, \alpha) - u(i, \alpha) \geqslant 0,$$

where the first inequality follows from the presumed stochastic dominance of α' over α (see Footnote 22) and from the strict supermodularity assumption; the second inequality follows from the optimality of j relative to α. Hence, any element in $\text{BR}(\alpha')$ must be no smaller than j, establishing the claim.

V.A.2. *Pure Coordination Games*

Pure coordination games have positive payoffs along the main diagonal and zero payoffs elsewhere, $u_{ij} = 0$ for $i \neq j$. For convenience, let us order the strategies so that $u_{jj} > u_{ii} > 0$ for $j > i$. Clearly, each strategy constitutes a pure strategy Nash equilibrium, and the strategies are Pareto-ranked, where strategy 1 generates the least efficient equilibrium and strategy n the most efficient. It is easy to verify that these games are *not* supermodular unless $n = 2$. For example, the marginal payoff $u_{1j} = u_{2j}$ increases when j changes from 1 to 2, but it *decreases* when j is further increased from 2 to 3.

Examples of pure coordination games arise in the network externality/product compatibility literature. Consider, for instance, the situation where each individual is to choose one of n different computers (or software packages). Two individuals can "collaborate" if and only if they have the same computer, and the inherent quality of the computers is different. In such a game, there are as many pure-strategy equilibria as the number of strategies. In addition to those, for any subset of strategies there exists a unique mixed-strategy equilibrium which puts positive weights on all strategies in this subset.

V.B. *Global Convergence*

We now provide global convergence results for the best-response dynamic, $P(0)$, i.e., we show that starting from an arbitrary initial condition, the system will converge to one of the game's pure-strategy Nash equilibria and stay there thereafter. Which equilibrium is attained as the limit point depends, of course, on the initial condition, but the fact that the system converges does not. These results are of interest by their own right, since they counter examples of cyclical behavior. See, for instance, Shapley's example [37]. This situation does not occur in the games we consider here.

V.B.1. *Supermodular Games*

First we show that in symmetric strict supermodular games, the states which mimic mixed strategy equilibria are unstable in a strong sense: they are not even stationary points of the best-response dynamic.

PROPOSITION 8. *Suppose all players are taking best responses under z, that is, $i \in \beta_i(z)$ for all $i \in C(z)$. In symmetric strict supermodular games this implies that only one strategy is played under z.*

Proof. Suppose to the contrary that we have $i < j$, $i, j \in C(z)$ and $i \in \beta_i(z)$, $j \in \beta_j(z)$. Since $i < j$, the strategy distribution facing a player with strategy i stochastically dominates that for a player with strategy j: $\alpha(z, j) \prec \alpha(z, i)$. Then Proposition 7 implies that $\text{Min}\, \beta_i(z) \geqslant \text{Max}\, \beta_j(z)$, a contradiction. ∎

Now we are ready to show the global convergence result.

THEOREM 2. *For symmetric strict supermodular games, the set of limit sets is in one-to-one correspondence with the collection of all pure-strategy Nash equilibria. That is, for any initial state $z^0 \in Z$, the system under the best-response dynamic converges to one of the pure strategy Nash states with probability one.*

A note on related results. Similar global convergence results for supermodular games are found in Milgrom and Roberts [26] and Krishna [19]. The case of 2×2 games was treated earlier by Miyasawa [29]. Milgrom and Roberts consider supermodular games with a *unique* Nash equilibrium, while we do *not* place such a restriction. When a supermodular game has a unique Nash equilibrium, the game is dominance-solvable (see Proposition 6(i)); thus a wide class of adjustment processes will naturally converge to the unique Nash equilibrium. Krishna, on the other hand, obtains a global convergence result for supermodular games with *multiple* equilibria, but considers a different adjustment process, namely, the *fictitious play* dynamics. His basic assumption is that each player takes a best response against the *empirical distribution* of strategies in the *entire* past history of play. Since his model is more complicated, Krishna employs one further assumption, that a player's payoff function is *concave* with respect to her own strategy. We do not need such an assumption in Theorem 2. In a similar vein, Monderer and Shapley [30] prove a global convergence result for potential games with fictitious-play dynamics. Their results do not apply here because there are examples (see Sela [36]) showing that supermodular games are not a special case of potential games, or vice versa.[24]

Proof of Theorem 2. By the nature of the best-response dynamic, a singleton set is a limit set if and only if all players are taking best responses. Proposition 8 shows that those states correspond to pure-strategy Nash equiibria. Next we will show that starting from any other state, the process converges to a pure-strategy Nash equilibrium with positive probability. By the definition of limit sets, this is sufficient to show that there are no other limit sets.

Take a non-Nash state z^0 and consider a particular trajectory which emanates from it. To define the trajectory, let us introduce the following notation. For state z, define the set of sub-optimal strategies by $S(z) = \{i \in C(z) \mid i \notin \beta_i(z)\}$. Now consider a trajectory $\{z^k, k = 0, 1, ...\}$ such that as long as $S(z^k)$ is non-empty, only one player adjusts at a time. Furthermore, we require the same adjustment to occur whenever a state is revisited. Such an adjustment realizes with a positive probability because of the independence of stochastic adjustments across players.

Since the state space is finite, this trajectory forms a cycle (which might be a singleton set). Let R be the set of strategies whose populations change on this cycle, and let $x = \text{Max } R$ (recall that we have complete ordering over strategies). Suppose a player with strategy i is adjusting to strategy x at point z^T on this cycle. Then we have

$$\alpha(z^T, i) = \alpha(z^{T+1}, x) \prec \alpha(z^{T+1}, j) \qquad \text{for all} \quad j < x.$$

Recall that under the population configuration z, a player with strategy $k \in C(z)$ faces strategy distribution $\alpha(z, k)$. The equality follows from the fact that only *one* player is adjusting from i to x, and the stochastic dominance follows from $j < x$. Given this stochastic dominance and Proposition 7, at z^{T+1} any player with strategy $j < x$ has best responses which are no smaller than x. On the other hand, a player with strategy x is already taking a best response, because a best response to $\alpha(z^T, i) = \alpha(z^{T+1}, x)$ is x. By the definition of x, it must be that the player who moves at $T+1$ adjusts to x. Proceeding inductively, we conclude that the cycle is actually a *singleton* $\{z^*\}$ with $z^*_x = M$ and $S(z^*) = \varnothing$, which is a pure-strategy Nash state with all players choosing x. ∎

[24] Even for the games that belong to the join, one has to establish convergence for all *better-response* paths, not just the *best-response* ones. Note also that their model pertains to games played by the *same* players, not to the random matching scenario. Therefore, their model formalizes a learning story rather than an evolution story.

As an immediate corollary, we have the following.

THEOREM 2′. *Each long-run equilibrium of a supermodular game corresponds to a pure-strategy Nash equilibrium.*

V.B.2. *Pure Coordination Games*

Consider now the games where $u_{jj} > u_{ii} > 0$ for $j > i$ and $u_{ij} = 0$ for $i \neq j$. Then we have the following analogues of Proposition 8 and Theorem 2.

PROPOSITION 9. (1) *Any mixed strategy configuration is unstable, i.e., if* $|C(z)| > 1$ *then at least for one* $i \in C(z)$, $i \notin \beta_i(z)$. (2) *The collection of limit sets is* $\{\{e_1\}, ..., \{e_n\}\}$. *Therefore, the best-response dynamic for a pure coordination game converges to a pure-strategy Nash equilibrium with probability one.*

Proof. (1) Assume $i, j \in C(z)$ with $i < j$, and assume $i \in \beta_i(z)$, $j \in \beta_j(z)$. Then we have

$$\alpha_j(z, j)\, u_{jj} \geqslant \alpha_i(z, j)\, u_{ii} > \alpha_i(z, i)\, u_{ii} \geqslant \alpha_j(z, i)\, u_{jj} > \alpha_j(z, j)\, u_{jj},$$

where the strict inequalities follow from the definition of $\alpha(z, k)$, (2.2); and the weak inequalities follow from the presumed optimality of i and j relative to z. Thus this contradicton establishes that for at least one strategy, $i \in C(z)$, $i \notin \beta_i(z)$.

(2) Note that each pure-strategy configuration, e_i, satisfies (as a singleton) the definition of a limit set. Therefore it remains to show that starting from any other state there is a positive probability path leading into one of these singleton states. Let $z \in Z$. By (1) there exists an $i \in C(z)$ so that $i \notin \beta_i(z)$. Let $k \in \beta_i(z)$, and let a player taking strategy i switch over by herself to strategy k, and let that occur before any other player adjusts its strategy. This event realizes with positive probability. After that, k becomes a BR for all individuals. This is shown as follows. First, given that one player switched optimally from strategy i to strategy k, the players with strategy k must be taking best responses. Next consider a player who is not taking strategy k. Compared to the players with strategy k, this player faces a strategy distribution which puts *more* probability to k and less (or equal) probability on each other strategy. By the structure of pure coordination games, this implies that strategy k is even more desirable for this player than for those who are already taking strategy k. Therefore, we conclude that k should be the best response for all players. Given this, we let the continuation path be one where everyone switches to k. Since this path realizes with positive probability, the proof is complete. ∎

V.C. *Costs of Transition between Limit Sets; Optimum Branching Algorithm*

We now show how to compute long-run equilibria for the above classes of games. We start with pure coordination games since the computations are illustrated more straightforwardly for them.

V.C.1. *Pure Coordination Games*

The first thing is to compute costs of transition $C_{ij} \equiv C(\{e_i\}, \{e_j\})$ between the limit sets $\{e_i\}$ and $\{e_j\}$. Assume the population is initially clustered at e_i, $i \neq n$. Since n is Pareto efficient, the easiest way to escape from the basin of attraction of e_i is to enter the basin of attraction of e_n by m mutations, where m is the smallest number for which

$$\frac{mu_{nn}}{M-1} \geqslant \frac{(M-m-1)\,u_{ii}}{M-1}.$$

This represents an immediate jump to escape the best-response region of i, and the triangle inequality (Proposition 5) guarantees that no gradual transition is less costly than this jump. Thus, for $i \neq n$, we have

$$C_{in} = m = \left\lceil \frac{(M-1)\,u_{ii}}{u_{ii}+u_{nn}} \right\rceil, \qquad (5.1)$$

where $\lceil x \rceil$ stands for the smallest integer weakly exceeding x. Since this is the minimum number of mutations to escape the basin of attraction of e_i, we must have $C_{in} \leqslant C_{ij}$ for all $j \neq i$. A similar argument applies to C_{nj}: we have $C_{n,n-1} \leqslant C_{n,j}$ for all $j \neq n$ and thus $C_{n,n-1} = \lceil (M-1)\,u_{nn}/(u_{nn}+u_{n-1,n-1}) \rceil$.

Next observe from (5.1) that for large M,

$$C_{in} < C_{ij}, \text{ for } i, j \neq n \qquad \text{and} \qquad C_{n,n-1} < C_{n,j}, \text{ for } j < n-1. \qquad (5.2)$$

We need "large M" because of an integer problem. More precisely, it can be seen that (5.2) holds if $u_{nn}/(u_{nn}+u_{n-1,n-1}) - u_{n-1,n-1}/(u_{n-1,n-1}+u_{nn}) \geqslant 1/(M-1)$.

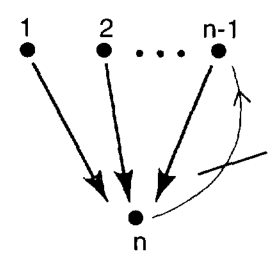

FIGURE 2

Therefore, the first step of the optimum branching algorithm which establishes the most likely transitions from each state results in the system of branches depicted in Fig. 2.

According to (5.2) the longest branch among those is of length $C_{n,n-1}$. Therefore we drop it and are left with an n-tree. This terminates the algorithm. The conclusion in this case is that the Pareto-superior equilibrium is a unique long-run state. This result is summarized as follows:

THEOREM 3. *In a pure coordination game, the unique long-run equilibrium is the Pareto-efficient Nash equilibrium, if* $u_{nn}/(u_{nn}+u_{n-1,n-1}) - u_{n-1,n-1}/(u_{n-1,n-1}+u_{nn}) \geqslant 1/(M-1)$.

V.C.2. *Supermodular Games*

The next result shows how the costs of transition, C_{ij}, are computed for the class of supermodular games. The proof of the result is constructive and shows that the transition from i to j (for $j > i$) involves an initial step where k individuals are mutated from strategy i to strategy n (the *largest* strategy), which is then followed by (costless) best-response adjustments. An analogous construction applies to the transition from j to i. To validate this result we need one further restriction which we call the "continuity" assumption.

Assumption B. If $\mathrm{BR}(\alpha) = i$ and $\mathrm{BR}(\alpha') = j$ and $i < k < j$, there exists a $\lambda \in (0, 1)$ such that $\mathrm{BR}(\lambda\alpha + (1-\lambda)\alpha') = \{k\}$.

The idea behind this is that the best-response changes "gradually", i.e., as we change the strategy distribution continuously the best response changes continuously as well (or, in other words, it does not skip intermediate strategies). We note that Assumption B is satisfied for a broad class of games, including the examples we present below and many supermodular games with continuous strategy sets. On the other hand, it is not satisfied by *all* supermodular games.

THEOREM 4 (MUTATIONS TO EXTREME STRATEGIES). *Assume we have a supermodular game which satisfies Assumption B. Let i and j be two Nash strategies where $j > i$. If the population size is sufficiently large, the cost of transition between e_i and e_j is realized as follows.*

(i) *C_{ij} is realized by initially mutating k individuals from i to n. The number k represents the critical mass of mutants beyond which best-response adjustment to higher numbered strategies between i and j becomes feasible (Fig. 4 illustrates how k is computed).*

(ii) *C_{ji} is realized by mutating k' individuals from j to 1. The number k' is determined analogously.*

Proof. See the Appendix.

Remark. Theorem 4 shows that the cost of transition depends crucially on the smallest and largest strategies. In some games those are naturally specified, but in other games there is no obvious way of specifying them. For example, it is not obvious how to choose a minimal and a maximal price in oligopoly games. One way to overcome this ambiguity is to assume that mutants know the payoff function, that they form "rationalizable" beliefs, and that they choose an optimal strategy against these beliefs. In this case we should eliminate all non-rationalizable strategies from the game before applying the algorithm. After doing that, Proposition 6 shows that the minimal and maximal strategies in supermodular games are pure strategy Nash equilibria. Therefore, strategies 1 and n in Theorem 4 are identified as the smallest and largest *Nash equilibria*.

When we have adjacent equilibria $i<j<k$, where j is the only equilibrium strategy between i and k, we denote $i=j-$ and $k=j+$. Given this notation, the following result is an immediate consequence of Theorem 4.

THEOREM 5. *Under the same assumptions as in Theorem* 4, $C_{ij} = \mathrm{Max}_{i<k\leq j} C_{k-,k}$ *for all* $i, j \in N$.

Therefore, Theorem 5 simplifies the problem of computing costs of transition between the various equilibria: we need only compute costs between *adjacent* Nash equilibria.

The determination of transition costs can be further simplified in the case of supermodular games with *linear* payoff functions:

$$u_{ij} = f(s_i) + g(s_i)\, s_j, \tag{5.3}$$

where strategies are denoted here by s_i and s_j, and $s_i < s_j$ for $i<j$. The basic assumption here is that the payoffs are linear in the opponent's strategy. One way of generating such games is in the context of differentiated product oligopoly. Assume two firms produce differentiated products, and let the market demand for the product of firm i be

$$Q^i = X(p^i) + Y(p^j), \tag{5.4}$$

where Q^i is the quantity sold by firm i when it charges price p^i, and its rival charges p^j. $X(p^i)$ is a strictly decreasing function, and $Y(p^j)$ is strictly increasing.

Firms' strategic variable is price, which varies discretely over the set[25]

$$S \equiv \{p_1, p_2, \ldots, p_n\}.$$

[25] If the strategy set is not finite, the limit distribution may fail to exist, and our analysis requires additional regularity conditions.

Note that superscripts denote individual *firms*, whereas subscripts denote individual *strategies*. Let c denote the (constant) per-unit cost of production. Then, when firms interact duopolistically, this leads to a payoff of

$$u_{ij} = (p_i - c)[X(p_i) + Y(p_j)],$$

and to an *expected* payoff of

$$\pi_i = (p_i - c)[X(p_i) + \bar{y}], \tag{5.5}$$

where

$$\bar{y} \equiv \frac{1}{M-1} \sum_k Y(p^k),$$

with k ranging over all remaining $M-1$ firms in the industry. It is straightforward to verify from (5.4) that the game is supermodular. Also, Eq. (5.5) makes it clear that the expected payoff to a firm posting price p_i depends only on the *average*, $\bar{y}$. This simplifies the analysis considerably.

This embeds the differentiated-product oligopoly model into the framework of randomly matched firms. Alternatively, we can consider a framework without random matching and where firms interact in one *global* market. In other words, all M firms sell their (differentiated) product simultaneously in one market. The demand curve that firm i faces in this market is given by

$$Q^i(p^i, ..., p^M) = X(p^i) + \frac{1}{M-1} \sum_{j \neq i} Y(p^j), \tag{5.6}$$

reflecting the idea that the quantity sold by a firm depends on the entire configuration of other firms' prices. This gives rise to the same payoff function (5.5). Hence, while the interpretation is different the formal analysis applies to either formulation.

This example shows, therefore, that random matching is not an essential part of our model. The most general formulation, which includes random matching as a special case, is as follows. When M players are adopting a strategy profile $s = (s^1, ..., s^M)$, player m's payoff is $u^m(s)$. The state space is $Z = S^1 x, ..., x S^m$, where S^m is player m's strategy set. The random matching story, as well as the above market interaction story with symmetric additive payoffs, simplifies the state space, because the identities of players do not matter. The analysis in Sections II and III apply equally to the general case described above, with a higher dimensional state space.

Next, since the function $Y(\cdot)$ is strictly increasing, we can transform the strategy set and consider

$$S' = \{y_1, ..., y_n\}, \qquad \text{where} \quad y_i = Y(p_i).$$

Each player is envisioned as choosing a $y_i \in S'$, which is equivalent to choosing $p_i = Y^{-1}(y_i)$.

The crucial property of this example is that a firm's best response only depends on the average of the opponents' strategies. Thus let $\text{BR}(y) \subset S'$ be the set of best responses when y is the average of the opponents' strategies, and let $N \subset S'$ be the set of Nash equilibria. We now show how to apply Theorem 4 to this game. Let y' and y'' be two equilibria, where $y' < y''$. Define a parametrized family of correspondences,

$$B(y, a, i) = \text{BR}(ay_i + (1-a)y), \qquad a \in [0, 1],$$

where i is either 1 or n. We then have, for large M,

$$C(y', y'')/M \approx \min\{a \mid B(y, a, n) \geq y, y' \leq y \leq y''\}$$

$$C(y'', y')/M \approx \min\{a \mid B(y, a, 1) \leq y, y' \leq y \leq y''\}.$$

The idea behind this is illustrated in Fig. 3. For the sake of this illustration, we suppose a continuous strategy set $[y_1, y_n]$ and a smooth best-response function (rather than a step function). The figure shows that the best-response function moves to the left as we mutate aM individuals from y' to y_n. For a sufficiently large a—say, a^*—the function becomes *just tangent* to the 45° line between y' and y''. At this point a best-response path connecting y' and y'' emerges (which is nothing but the path specified in the proof of Theorem 4). Since this is the first time that such a path emerges, the cost of transition between y' and y'' is a^*M; the computation for $C(y'', y')$ is analogous, except that individuals are mutated to strategy 1 which perturbes the best response function to the *right*, not to the *left*. Figure 4 shows a convenient way to find a^*. We can measure the *depth* of the

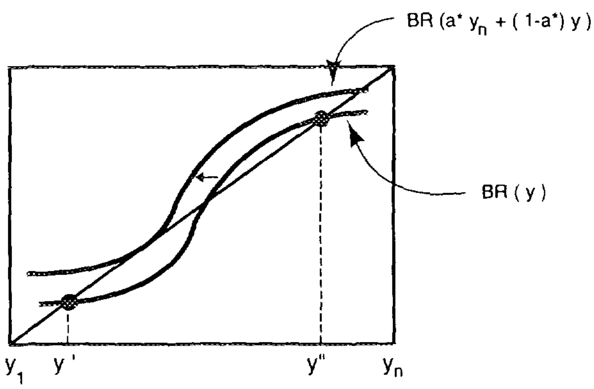

FIGURE 3

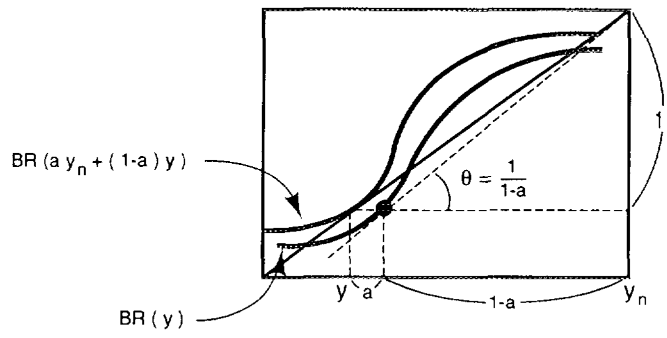

FIGURE 4

best-response function over the interval $[y', y'']$ by the slope θ. Then, by definition, $C(y', y'')$ is an increasing function of this slope, $\theta = 1/(1-a)$. From this characterization we can see that $C(y', y'') > C(y'', y')$ if and only if $\theta > \theta'$.

THEOREM 6. *Suppose the assumptions in Theorem 4 are satisfied. If there is an equilibrium i such that* $C_{i,i-}, C_{i,i+} > C_{j,j-}, C_{j,j+}$ *for all* $j \neq i$, *then i is the unique long-run equilibrium.*[26]

Proof. Suppose $k \neq i$ is a long-run equilibrium. Let h^* be the optimal k-tree. Delete the outgoing branch at i from this tree and add the branch from k to i. The resulting graph is an $i-$tree, denoted h^{**}. By Theorem 5 and the inequality in Theorem 6, we have $C(h^{**}) < C(h^*)$, a contradiction. ∎

V.D. *Examples*

Let us provide a few examples, showing how to identify the long-run equilibrium. First, consider the supermodular game with linear payoffs, (5.3), whose best-response function is depicted in Fig. 5. For the ease of exposition, we draw a smooth function rather than a step function. There are two equilibria, s_1 and s_n.[27] Equilibrium s_n has a larger *basin of attraction* than s_1, i.e., $s_n - s' > s' - s_1$. But this is not relevant for the long-run stability. What matters is the depth of basin of attraction, as measured by C_{1n} and C_{n1}. Since $C_{1n} > C_{n1}$, s_1 is the long-run equilibrium.

[26] For $i = 1$ delete $C_{i,i-}$ from the left-hand side. For $i = n$ delete $C_{i,i+}$.

[27] Assume, for simplicity, that the "unstable" equilibrium s' does not correspond to a pure strategy and therefore it cannot be a long-run equilibrium (Proposition 7).

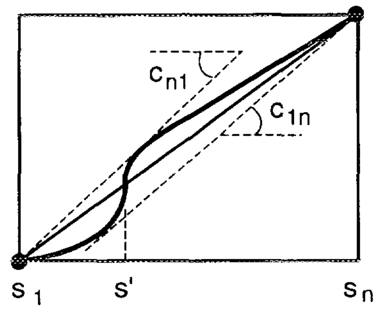

FIGURE 5

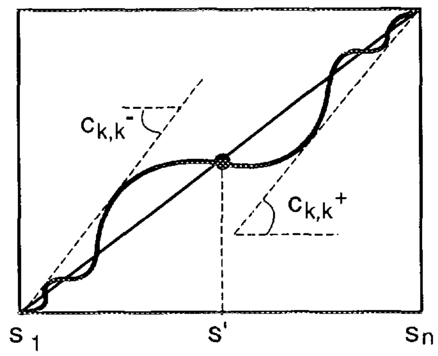

FIGURE 6

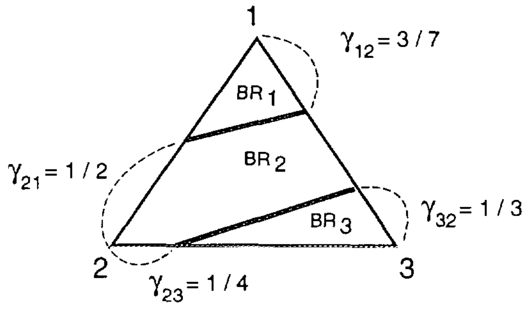

FIGURE 7

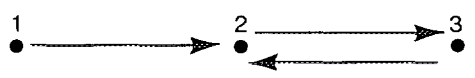

FIGURE 8

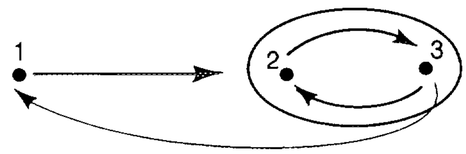

FIGURE 9

The next example (see Fig. 6) illustrates Theorem 6. Equilibrium s_k is surrounded by the largest "bumps": $C_{k,k+}$ and $C_{k,k-}$. Thus, Theorem 6 shows that s_k is the long-run equilibrium.

In general, however, it may not be possible to find such an equilibrium. Then we can utilize the optimal branching algorithm. The next example illustrates how it works. Consider the following payoff matrix (u_{ij}):

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 12 \\ -6 & 8 & 15 \end{pmatrix}.$$

It can be easily checked that this is a supermodular game. The best response regions are depicted in Fig. 7.[28] From this we can see that Assumption B is satisfied (although the game is not linear); thus we can use Theorem 4 to determine C_{ij}: for adjacent pairs (i, j) we have $C_{ij}/M \approx \gamma_{ij}$. By Theorem 5 we can calculate C_{ij} for non-adjacent pairs: $C_{13}/M \approx \text{Max}\{\gamma_{12}, \gamma_{23}\} = \gamma_{12}$ and similarly $C_{31}/M \approx \gamma_{21}$. Thus, we have $C_{31} = C_{21} > C_{13} = C_{12} > C_{32} > C_{23}$. The optimum branching algorithm (see also working paper version) applies to this data as follows.

Step 1. Establish the shortest branch leaving each node. State 1 has two outgoing branches with the same cost; thus we can choose either one. Let us choose (1, 2) to better illustrates the algorithm. This results in Fig. 8.

Step 2. Drop the longest branch from the set of branches established in Step 1: $\text{Max}(C_{12}, C_{23}, C_{32}) = C_{12}$

Step 3. This generates a cycle ((2, 3), (3, 2)). Determine the longest branch along the cycle: $\text{Max}(C_{23}, C_{32}) = C_{32}$. Adjust the cost of outgoing

[28] γ_{ij} in the figure is calculated as

$$\gamma_{12} = \frac{u_{11} - u_{21}}{u_{11} - u_{21} + u_{23} - u_{13}} \qquad \gamma_{32} = \frac{u_{33} - u_{23}}{u_{33} - u_{23} + u_{21} - u_{31}}$$

$$\gamma_{21} = \frac{u_{22} - u_{12}}{u_{22} - u_{12} + u_{11} - u_{12}} \qquad \gamma_{23} = \frac{u_{22} - u_{32}}{u_{22} - u_{32} + u_{33} - u_{23}}.$$

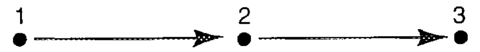

FIGURE 10

branches from this cycle whose origin is different from the origin of the longest branch. This results in

$$\tilde{C}_{21} = C_{21} + C_{32} - C_{23}.$$

(The adjustment rule is to add the cost of the longest branch along the cycle (C_{32}) and to subtract the cost of the outgoing branch from the same origin (C_{23}); this leaves the cost of branches starting at 3 unchanged).

The states $\{2, 3\}$ are now retracted to a single state. This gives us a two-state set of nodes as in Fig. 9. This suggests the nature of the dynamic when the mutation rate in non-negligible: the state usually fluctuates between 2 and 3 and only occasionally visits 1. We now return to Step 1.

Step 1. The branch $\tilde{C}_{31}$ is picked over $\tilde{C}_{21} = C_{21} + C_{32} - C_{23}$ since $C_{31} = C_{21}$ and $C_{32} > C_{23}$.

Step 2. Drop $\max(C_{12}, C_{31}) = C_{31}$
This terminates the algorithm's first phase. Now we have to open loops. The loop $\{2, 3\}$ is a root component so we can open it according to optimality. Thus, we drop $\max(C_{23}, C_{32}) = C_{32}$, which gives us the final result as seen in Fig. 10.

Net conclusion. e_3 is the unique long-run state.

For games with more than three strategies these computations are, of course, more involved. But the computational complexity of the optimum branching algorithm is of order n^2 (see Tarjan [38]). Therefore, as far as numerical implementation is concernced the computation of LRE is quite tractable.

APPENDIX

Proof of Proposition 6. With the exception of (ii) and (iv) these properties are well known (see Milgrom and Roberts [26] or Vives [42]). Thus here we prove (ii) and (iv) only.

(ii) Assume that (i, j) is a NE for $i \neq j$. Then

$$u_{ij} \geqslant u_{jj}, \tag{N.1}$$

and

$$u_{ji} \geqslant u_{ii}. \tag{N.2}$$

Assume, without loss of generality, that $j > i$. Then $u_{jk} - u_{ik}$ increases in k by the supermodularity. In particular,

$$u_{ji} - u_{ii} < u_{jj} - u_{ij}.$$

But this contradicts (N.1) and (N.2) above.

(iv) Assume (i, i) is a NE which is not strict. Then we can modify the game by letting $\tilde{u}_{ij} = u_{ij} + \varepsilon$ for $j = 1, \ldots, n$ and for sufficiently small ε. The modified game is still supermodular, it contains no dominated strategies, and every NE in the u game is also an NE in the $\tilde{u}$ game. Also, i is now a unique best response against i. This modification can be performed for every weak NE, resulting in a game which has the same set of NE's but where each NE is strict. ▮

Proof of Theorem 4. We will show (i); the proof for (ii) is identical. The proof is in two steps. The first step shows a feasible path. The second proves its optimality.

Step 1. Consider the state where all players adopt i. Call this state $z(0) = (0, \ldots, M, \ldots, 0)$, and let $y(0) = i$. Fix a number x, $0 < x \leqslant M$, and let x individuals mutate from i to n. Call these individuals "mutants" and the remaining individuals "non-mutants". Call the resulting state $z(1) = (0, \ldots, M - x, 0, \ldots, x)$. Define an adjustment path $(y(t), z(t))$ as

$$y(t) = \operatorname{Max} \beta_{y(t-1)}(z(t))$$

$$z_r(t+1) = \begin{cases} x & r = n \\ M - x & r = y(t) \\ 0 & \text{otherwise.} \end{cases}$$

(Along this path mutants stick to strategy n, while non-mutants are adjusting to their best responses). By Proposition 7 the sequences $y(t)$ and $z(t)$ are increasing in x and t. Choose the smallest x for which $y(t) \geqslant j$ for some t and call it k. We show that e_j is achieved from e_i with k mutations. Let T be the first time along the path at which $y(T) \geqslant j$. We distinguish between three cases.

Case 1a. $y(T) = j$. At stage $T - 1$, $y(T)$ is the best response of all non-mutants. If it is also the best response of all mutants we let all players adjust to j and the process is terminated.

Case 1b. $y(T) = j$ but $y(T)$ is *not* the best response of mutants. In that case we let a particular fraction of the non-mutants adjust to j, creating an "intermediate" state $\tilde{z}(T+1)$. This fraction is such that j becomes the largest best response of mutants. The distribution faced by the remaining non-mutants (i.e., those with strategy $y(T-1)$) is slightly larger than that faced by mutants. However, j is a *strict* NE and j is the *largest* best response of mutants. Therefore, if M is sufficiently large we have $j \in \beta_{y(T-1)}(\tilde{z}(T+1))$. Therefore, we can let all individuals adjust to j, and the process is terminated.

Case 2. $y(T-1) < j < y(T)$. Here we let again a particular fraction of the non-mutants adjust to $y(T)$, creating an intermediate state $\tilde{z}(T+1)$ in which j is the largest best response of mutants. This can be done because of Assumption B. Furthermore, by the same argument as in Case 1b we can ensure that j is the best response of all non-mutants. Then we let all players adjust to j.

Step 2. On the other hand, if the total number of mutations is less than k, then we will prove by contradiction that e_j cannot be achieved from e_i by the best-response adjustment. Suppose in the transition path we have $k' < k$ mutations. Then since $k' < M$, at some stage in the transition, at least one player should be adjusting to j. Consider the following modification of the transition path. Instead of having k' mutations over time, let k' players simultaneously mutate into strategy n *at the beginning*. As before, those players are called mutants, and the other players are non-mutants. In the modified adjustment, the mutants stick to strategy n. In contrast, the non-mutants simultaneously adjust to their largest best responses. This generates a process $\tilde{z}(t)$, and $k' < k$ implies that the non-mutants are eventually stuck with some strategy $s' < j$.

However, by Proposition 7, the strategy distribution at each step, t, on the new path stochastically dominates that on the original path. This implies that the best response of any non-mutant at any stage on the original path is no larger than s', which implies that the non-mutants can never adjust to j, a contradicton. ∎

References

1. K. Binmore and L. Samuelson, Evolutionary stability in repeated games played by finite automata, *J. Econ. Theory* **57** (1992), 278–305.
2. D. Canning, Average behavior in learning models, *J. Econ. Theory* **57** (1992), 442–472.
3. Y. J. Chu and T. H. Liu, On the shortest arborescence of a directed graph, *Sci. Sinica* **14** (1965), 1396–1400.
4. R. Cooper and A. John, Coordinating coordination failures in Kaynesian models, *Quart. J. Econ.* **103** (1990), 441–463.

5. A. Cournot, "Recherches sur les Principes Mathématiques de la Théorie des Richesses" (1838); N. T. Bacon, Trans. "Researches into the Mathematical Principles of the Theory of Wealth," Hafner, London, 1960.
6. P. Diamond, Aggregate demand management in search equilibrium, *J. Polit. Econ.* **90** (1982), 881–894.
7. E. W. Dijkstra, A note on two problems in connexion with graphs, *Numerische Math.* **1** (1959), 269–271.
8. J. Edmonds, Optimum branchings, *J. Res. Nat. Bur. Stand.* **71B** (1967), 233–240.
9. G. Ellison, Learning, local interaction, and coordination, *Econometrica* **61** (1993), 1047–1071.
10. D. Foster and P. Young, Stochastic evolutionary game dynamics, *Theor. Population Biol.* **38** (1990), 219–232.
11. M. I. Freidlin and A. D. Wentzell, "Random Perturbations of Dynamical Systems," Springer-Verlag, New York, 1984.
12. D. Fudenberg and C. Harris, Evolutionary dynamics with aggregate shocks, *J. Econ. Theory* **57** (1992), 420–441.
13. Y. Gilboa and A. Matsui, Social stability and equilibrium, *Econometrica* **59** (1991), 859–867.
14. P. G. Hoel, S. C. Port, and C. J. Stone, "Introduction to Stochastic Processes," Houghton Mifflin, Boston, 1972.
15. M. Kandori, Independent mutations and long run equilibria in games: A continuous time formulation, mimeo, Princeton University, 1991.
16. M. Kandori, G. Mailath, and R. Rob, Learning, mutation and long equilibria in games, *Econometrica* **61** (1993), 29–56.
17. M. Kandori and R. Rob, Bandwagon games and long-run technology choice, CARESS Working Paper 93-05, University of Pennsylvania, 1993.
18. E. Kohlberg and J. F. Mertens, On the strategic stability of equilibria, *Econometrica* **54** (1986), 1003–1038.
19. V. J. Krishna, Learning in games with strategic complementarities, mimeo, Harvard Business School, 1991.
20. G. Mailath, Introduction: Symposium on evolutionary game theory, *J. Econ. Theory* **57** (1992), 259–277.
21. G. Mailath, Perpetual randomness in evolutionary economics, *Econ. Lett.* **42** (1993), 291–299.
22. A. Matsui, Cheap-talk and cooperation in the society, *J. Econ. Theory* **54** (1991), 245–258.
23. A. Matsui and K. Matsuyama, An approach to equilibrium selection, CARESS Working Paper 91-36, University of Pennsylvania, 1991.
24. A. Matsui and R. Rob, Evolution with forward-looking players, mimeo, University of Pennsylvania, 1993.
25. J. Maynard Smith and G. R. Price, The logic of animal conflict, *Nature (London)* **246** (1973), 15–18.
26. P. Milgrom and J. Roberts, Rationalizability, learning and equilibrium in games with strategic complementarities, *Econometrica* **58** (1990), 1255–1278.
27. P. Milgrom and J. Roberts, Adaptive and sophisticated learning in repeated normal form games, *Games Econ. Behav.* **3** (1991), 82–100.
28. P. Milgrom and J. Roberts, Comparative statics of games with strategic complementarities, mimeo, Stanford University, 1993.
29. K. Miyasawa, On the convergence of the learning process in a 2×2 non-zero-sum game, Economic Research Program, Research Memorandum 33, Princeton University, 1961.
30. D. Monderer and L. Shapley, Potential games, mimeo, UCLA and Technion, Haifa, 1991.

31. D. T. MORTENSEN, Property rights and efficiency in mating, racing and related games, *Amer. Econ. Rev.* **72** (1982), 968–979.
32. H. MOULIN, Dominance solvability and cournot stability, *Math. Soc. Sci.* (1984), 83–102.
33. G. NOLDEKE AND L. SAMUELSON, The evolutionary foundations of backward and forward induction, *Games Econ. Behav.* **5** (1993), 425–454.
34. L. SAMUELSON, Stochastic stability in games with alternative best-replies, *J. Econ. Theory* **64** (1993), 35–65.
35. J. SEADE, The stability of cournot revisited, *J. Econ. Theory* **23** (1980), 15–27.
36. A. SELA, Learning processes in games, M.Sc. thesis, The Technion, Haifa, Israel, 1992.
37. L. SHAPLEY, Some topics in two-person games, *Adv. Game Theory Ann. Math. Stud.* **5** (1964), 1–28.
38. R. E. TARJAN, Finding optimum branching, *Networks* **7** (1977), 25–35.
39. P. D. TAYLOR AND L. B. JONKER, Evolutionary stable strategies and game dynamics, *Math. Biosci.* **40** (1978), 145–156.
40. D. M. TOPKIS, Equilibrium points in nonzero-sum n-person submodular games, *SIAM J. Control and Optimization* **17** (1979), 773–787.
41. E. VAN DAMME, "Stability and Perfection of Nash Equilibria," Springer-Verlag, Berlin, 1987.
42. X. VIVES, Nash equilibrium with strategic complementarities, *J. Math. Econ.* **19** (1990), 305–321.
43. P. YOUNG, The evolution of conventions, *Econometrica* **61** (1993), 57–84.