CONVERGENCE TO RATIONAL EXPECTATIONS
IN A STATIONARY LINEAR GAME

by

James S. Jordan

Discussion Paper No. 258, July 1990

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minn 55455

Convergence to Rational Expectations in a

Stationary Linear Game

by

J.S. Jordan*

Department of Economics
University of Minnesota
271 19th Avenue S.
Minneapolis, MN 55455

July 1990

Abstract

This paper describes several learning processes which converge, with probability one, to the rational expectations (Bayesian-Nash) equilibrium of a stationary linear game. The learning processes include a test for convergence to equilibrium, and a method for changing the parameters of the process when nonconvergence is indicated. This self-stabilization property eliminates the need to impose stability conditions on the economic environment. Convergence to equilibrium is proved for two types of self-stabilizing learning mechanisms: a centralized forecasting mechanism and a decentralized strategy adjustment process. For a version of the decentralized process it is also shown that the publicly observable information, on which learning is based, does not necessarily reveal anything about the economic environment except the equilibrium.

1. Introduction

The rational expectations hypothesis asserts that economic agents anticipate the future according to the true probability distribution of future events. A probability distribution which is consistent with this hypothesis is called a rational expectations equilibrium. Since beliefs influence decisions which, in turn, affect the probabilities of future events, a rational expectations equilibrium is a fixed-point of a mapping from believed expectations to correct expectations. In most economic models, individual agents cannot observe directly all of the parameters of the economic environment. In particular, they cannot observe the characteristics of other agents. Hence, individuals do not have enough knowledge to deduce the correct expectations by solving for a fixed-point. Therefore, the plausibility of the rational expectations hypothesis depends on the ability of economic agents to learn the correct expectations from repeated observations of publicly available data. For this reason, the solution of the learning problem has become a principal objective of current research in rational expectations theory.

A learning process is a method by which agents compute estimated expectations from previously observed data. The asymptotic behavior of the estimates is a statistical problem greatly complicated by the influence of the estimates on the expectations being estimated. To make the problem tractable, most authors have confined their analysis to highly simplified economic models. In particular, linearity is a prominent feature of most models. However, even within the simplified models, two sources of difficulty have confined most studies in this area to partial solutions:

Identification: A rational expectations equilibrium is sensitive to virtually all parameters of the economic environment, yet few of them can be identified by publicly observable data. Frydman (1982, p. 658) emphasizes the difficulties posed by this problem.

Stability: Learning mechanisms typically proceed by adjusting agents' beliefs, or expectations, so the casual influence of beliefs has the potential to create unstable feedbacks.

1

In many recent studies (see Marcet and Sargent (1989) and references therein) the casual influence of beliefs is formalized as a mapping, say f, from believed parameters, to the correct parameters. A rational expectations equilibrium is a fixed-point of f. A powerful analytical technique due to Ljung (1977) reduces the convergence of a learning process to the stability of solutions to a deterministic differential equation system based on f. Therefore, the convergence of learning processes is confined to economic environments which satisfy this stability requirement. Marcet and Sargent (1989) give an extensive discussion of Ljung's method and the resulting stability requirement.

The object of the present paper is to confront directly and overcome both the identification and stability problems. The learning processes constructed below involve a combination of two extremely simple-minded mechanisms. One mechanism adjusts expectations, or decisions, in response to past errors. The second mechanism selects the adjustment coefficients for the first. At each iteration the second mechanism performs a crude test to determine whether the first mechanism is converging. If nonconvergence is indicated, new adjustment coefficients are selected at random and maintained until the next failure of the convergence test. These learning processes are shown to converge to rational expectations equilibrium with probability one. Moreover, convergence does not depend on the revelation of all environmental parameters. It is shown that nothing need be revealed with asymptotic certainty except the equilibrium itself.

The model described below is very simple, so this paper does not pretend to offer a general solution to the learning problem. Nor are the crudely constructed learning mechanisms of this paper intended to represent plausible economic behavior. What is intended is a clear demonstration that the identification and stability problems are not impassable barriers to learning.

The learning problem will be studied here in an N-person noncooperative game model, so the agents will be termed players. Each player's strategy is a real number

which must be chosen before the other players' strategies are known. Each player's utility maximizing strategy is a linear function of the expectation of the other player's strategies, plus an i.i.d. noise term. The parameters of this linear function are determined by the parameters of the player's utility function, which are not directly observable by other players.

The set of environments is described and the rational expectations equilibrium, which is also a Bayesian-Nash equilibrium, is derived in Section 2. As the game is repeated over time, with fixed values of the nonrandom parameters and i.i.d. drawings of the noise terms, players adjust their strategies in response to the observation of the past strategies of all players. Section 3 is devoted to a forecasting process by which a central agency computes, prior to each play, a forecast of all N strategies based on observed past strategies. Players then choose their current strategies as their optimal response to the forecast. Section 4 is devoted to a decentralized strategy adjustment process in which each player adjusts his strategy in response to the average discrepancy between his past strategies and his best response to past strategies of the other players. Theorems 3.2 and 4.1 state that for almost all environments, the respective process will eventually stabilize itself and converge to equilibrium with probability one. The identification problem is discussed in Section 5. It is shown there that the class of learning processes studied in Section 4 can be enlarged, in a simple way, to maintain the convergence result (Theorem 5.1) but prevent time series of strategies from revealing with asymptotic certainty any property of the environment other than the equilibrium (Theorem 5.2). Some implications of this result for the strategic manipulation of learning processes are discussed in that section. Some open problems are mentioned in the concluding section.

The role of the centralized forecasting model in Section 3 is primarily expositional. In this model, only one agent, namely the forecaster, is actually engaged in learning. This, together with the linearity of the model, makes it easy to derive the deterministic differential equation system associated with the learning process. The stabilizing role of the

3

adjustment coefficients is then fairly transparent, as is the fact that stabilizing coefficients can be selected through a sequence of random draws. The convergence test, which is based on Lemma 3.1, is a bit less obvious. The decentralized learning process in Section 4 is more complicated than the centralized forecasting model, but involves the same ideas. All of the proofs in this paper are essentially technical and, with the possible exception of Lemma 3.1, rather straightforward exercises in probability theory, so they have been relegated to an appendix.

2. Equilibrium

This section describes an N-person game and its equilibrium. Convergence to the equilibrium is the object of the learning processes described in subsequent sections. Learning occurs as the game is repeated over time, but in this section, to define the equilibrium, we consider a one-shot version of the game.

There are $N \geq 2$ players, and each player's strategy space is the real line, R. Player i's payoff function is defined as follows. Let a^i, $b^i_j \in R$, $j \neq i$, and let ε^i be a random variable with zero mean and finite variance. The random variables ε^i, $1 \leq i \leq N$, are mutually independent. We will use the symbol ε^i to denote both the random variable and its realization. Given a^i and b^i_j, $j \neq i$, and a realization ε^i, player i's payoff is

$$u^i(s^i, s^{)i(}, \varepsilon^i) = -(s^i - [a^i + \Sigma_{j=1}^N b^i_j s^j + \varepsilon^i])^2,$$

where $s^{)i(} = (s^1, \ldots, s^{i-1}, s^{i+1}, \ldots, s^N)$ and $b^i_i = 0$. At the time s^i is chosen, player i knows his nonrandom characteristics a^i and $b^i_j, j \neq i$, and the realization ε^i, but does not know s^j, $j \neq i$. Player i's expected payoff maximizing strategy is thus

$$(2.1) \qquad s^i = a^i + \Sigma_{j=1}^N b^i_j E^i s^j + \varepsilon^i$$

where $E^i s^j$ is player i's expectation of s^j, assuming that player i regards the strategies s^j, $j \neq i$, as independent random variables with finite mean and variance. It follows from (2.1) that

4

(2.2) $Es^i = a^i + b^i E^i s$

for each i, where $b^i E^i s = \Sigma_j b^i_j E^i s^j$. By setting $E^i s = Es$ for each i, j, we can use (2.2) to obtain the "rational expectations"

(2.3) $\sigma^* = -(B - I)^{-1} a,$

where B is the matrix $(b^i_j)_{\substack{1 \le i \le N \\ 1 \le j \le N}}$, I is the $N \times N$ identity matrix, and a is the column vector $(a^1, \ldots, a^N)$. For each player i, the strategy

(2.4) $s^i = a^i + b^i \sigma^* + \varepsilon^i = \sigma^{*i} + \varepsilon^i$

constitutes a rational expectations equilibrium.

Suppose that the nonrandom characteristics $(a^i, b^i)_i$ and the distributions of the random characteristics $(\varepsilon^i)_i$ are common knowledge. Then (2.4) is also a Bayesian-Nash equilibrium. If we impose the restriction that each s^i have finite mean and variance, it is the unique Bayesian-Nash equilibrium.

It may be useful at this point to consider a more concrete example. Suppose that the players are producers of differentiated products in a Cournot oligopoly. Since the products are differentiated, the price functions facing the different produces need not be identical. Suppose that the profit function of producer i takes the form

(2.5) $\pi^i(s, \varepsilon'^i) = (a'^i - b'^i s + \varepsilon'^i) s_i - c^i s_i^2$

for each i, where $a'^i > 0$, b'^i is a nonnegative N-vector with $b'^i_i > 0$, ε'^i satisfies the same assumptions as ε^i, above, and $c^i > 0$. In this example, $c^i s_i^2$ is interpreted as the cost of producing the output quantity s_i, and $a'^i - b'^i s + \varepsilon^i$ is interpreted as the price. Each producer i chooses the output quantity s^i after observing ε'^i but before observing the quantities chosen by the other producers. Since the profit function π^i is quadratic, a Taylor series expansion shows that π^i can be written as

(2.6) $\pi^i(s, \varepsilon'^i) = -(c^i + b'^i_i)(s^i - [a^i + \Sigma^N_{j=1} b^i_j s^j + \varepsilon^i])^2 + g^i(s)^{i(}, \varepsilon'^i)$

5

where $a^i = a'^i/2(b_i'^i + c^i)$, $b_j^i = -b_j'^i/2(b_i'^i + c^i)$ if $j \neq i$ and $b_i^i = 0$, and $\varepsilon^i = \varepsilon'^i/2(b_i'^i + c^i)$, and $g^i(s)^{i(}, \varepsilon'^i)$ equals $\pi^i(s, \varepsilon'^i)$ evaluated at $s^i = a^i + \Sigma_{j=1}^N b_j^i s^j + \varepsilon^i$. Since g^i does not depend on s^i, (2.6) shows that π^i is an affine transformation of a payoff function u^i as defined above. Hence this Cournot oligopoly model is an example of the general linear model described above.

For the sake of analytical simplicity we will not impose any nonnegativity or other boundary constraints on the strategies s^i. Such constraints, especially if they were binding at equilibrium, would introduce nonlinearities which are beyond the scope of the linear analysis described in this paper.

The severity of the stability problem is indicated by some results which have been obtained elsewhere for the deterministic version of this model (Jordan (1984, §4) and Williams (1985)). Suppose that the stochastic disturbance terms are deleted, and consider a dynamic strategy adjustment process of the form

$$\dot{s}^i = f^i(a^i, b^i; s) \qquad i \leq i \leq N,$$

for some C^1 functions $f^1, \ldots, f^N$. That is, player i adjusts his strategy as a function of his own characteristics and the current strategy N-tuple, but with no direct knowledge of the other players' characteristics. The stability question is whether there exist adjustment functions $(f^i)_i$ which make the equilibrium $\sigma^* = -(B-I)^{-1}a$ locally stable. If $N = 2$, for example, the obvious adjustment functions

$$\dot{s}^1 = a^1 + b_2^1 s^2 - s^1$$

$$\dot{s}^2 = a^2 + b_1^2 s^1 - s^2$$

are stable if $b_2^1 b_1^2 < 1$ (since $N = 2$ and $b_i^i = 0$ for all i, the trace of $B - I$ is -2 and $\det(B-I) = 1 - b_2^1 b_1^2$). However, if $N > 2$, there do not exist adjustment functions which are stable for all environments satisfying $(-1)^N \det(B - I) > 0$. Impossibility theorems obtained by Jordan (1984) and Williams (1985) indicate that strong restrictions on the space of environments are necessary for the existence of stable adjustment functions.

3. A Convergent Forecasting Process

The ability of each player to compute his equilibrium strategy depends on the assumption that the characteristics $(a^i, (b^i_j)_j)_i$ are common knowledge, or on the rational expectations assumption that each player knows the equilibrium distribution of the other players' strategies. In most situations of economic interest, it is more natural to assume that the characteristics of the participants are private information which is revealed, if at all, only through observed behavior. The rational expectations assumption also begs the question of how the assumed knowledge could be acquired. Thus, the justification of either assumption requires an explicit demonstration that the assumed knowledge can be learned, at least asymptotically, from the repeated observation of publicly available data.

If the game described above is played only once, there is no possibility of learning, so suppose instead that the game is played repeatedly over time. Assume that the deterministic characteristics $(a^i, b^i)_i$ are fixed throughout, and that the random characteristics $(\varepsilon^i)_i$ are drawn independently each period from a fixed distribution satisfying the assumption stated in Section 2. The publicly available data are the strategies, that is, the realized N-tuple of numbers, which are revealed after each play of the game. Thus, at the time player i chooses the number s^i_t, his information set consists of $(a^i, b^i; \varepsilon^i_1, \ldots, \varepsilon^i_t; s_1, \ldots, s_{t-1})$. We are interested in whether σ^i_t, defined as $s^i_t - \varepsilon^i_t$, converges to σ^{*i}.

There are at least three different types of learning processes:

A) players estimate the characteristics $(a^i, b^i)_i$ and compute the implied indirect estimates of the equilibrium strategies;

B) players forecast the mean of s_t and each player i chooses s^i_t as the optimal response to the forecast, via (2.1) above; and

C) player i adjusts σ^i_t in response to past errors, that is, in response to the differences $\sigma^i_\tau - (a^i + b^i s_\tau)$, $\tau = 1, \ldots, t-1$.

Convergence is established for a type B process in Theorem 3.2 and for a type C process

7

in Theorem 4.1. Section 5 shows that type A learning is unnecessarily demanding. In particular, Theorem 5.2 establishes that a certain family of type C processes converges without revealing player characteristics.

Since player i must choose s_t^i before observing $s_t^{)i(}$, the learning problem is naturally modelled as the problem of learning how to forecast $s_t^{)i(}$. Furthermore, since forecasting is, in practice, a specialized activity undertaken by academic and government economists, it is natural to model the forecast as a public announcement made by a separate forecasting agency. Accordingly, we will assume that in each period t, the forecasting agency makes a prediction, p_t, of s_t based on the data $s_1, \ldots, s_{t-1}$. Player i then chooses s_t^i as the optimal response to the prediction:

$$(3.1) \qquad s_t^i = a^i + b^i p_t + \varepsilon_t^i,$$

where $b^i p_t = \Sigma_{j \neq i} b_j^i p_t^j$. Strategies will converge to equilibrium if p_t converges to σ^*.

It is useful to express (3.1) in terms of deviations from equilibrium. Define σ_t as $s_t - \varepsilon_t$, and recall that $\sigma^* = a + B\sigma^*$. Then we have

$$(3.2) \qquad \sigma_t - \sigma^* = B(p_t - \sigma^*); \text{ or}$$

$$(3.3) \qquad \sigma_t - p_t = (B - I)(p_t - \sigma^*).$$

The difference equations (3.2) and (3.3) indicate that the forecasting problem is less a statistical problem than a dynamic adjustment problem. For example, since the game is stationary over time, the forecasting agency might use the prediction rule $p_t = \bar{s}_{t-1}$, where $\bar{s}_{t-1} = (t-1)^{-1}\Sigma_{\tau=1}^{t-1} s_\tau$, the empirical mean of past strategy N-tuples (throughout the sequel, $\bar{x}_t$ denotes $t^{-1}\Sigma_{\tau=1}^{t} x_\tau$, for any time indexed variable x). Since (3.1) is linear, we can take empirical means on both sides to obtain $\bar{s}_t = a + B\bar{p}_t + \bar{\varepsilon}_t$, so

$$(3.4) \qquad p_{t+1} = a + B\bar{p}_t + \bar{\varepsilon}_t.$$

By (3.4) and the fact that $\sigma^* = a + B\sigma^*$,

$$(3.5) \qquad p_{t+1} - \bar{p}_t = (B - I)(\bar{p}_t - \sigma^*) + \bar{\varepsilon}_t.$$

8

By the definition of the empirical mean, $\bar{p}_{t+1} - \bar{p}_t = (t+1)^{-1}(p_{t+1} - \bar{p}_t)$, so by (3.5)

(3.6) $\qquad \bar{p}_{t+1} - \bar{p}_t = (t+1)^{-1}(B - I)(\bar{p}_t - \sigma^*) + (t+1)^{-1}\bar{\varepsilon}_t.$

By the strong law of large numbers, $\bar{\varepsilon}_t \to 0$ almost surely, so (3.5) implies that $p_t \to \sigma^*$ if $\bar{p}_t \to \sigma^*$. Since $\bar{\varepsilon}_t \to 0$ almost surely, and since the factor $(t+1)^{-1}$ drives the "step size" to zero, equation (3.6), speaking very loosely, is asymptotically similar to the linear differential equation:

(3.7) $\qquad \dot{\bar{p}} = (B - I)(\bar{p} - \sigma^*),$

which suggests that convergence to equilibrium depends on the characteristic roots of the matrix $B - I$.

The analogy between discrete-time stochastic adjustment procedures and continuous-time deterministic differential equations has been exploited in the mathematical theory of stochastic adjustment processes (see Ljung (1977) and Arthur, et. al. (1987)) and applied to the study of economic learning models (Marcet and Sargent (1989) and Woodford (1990)). Unfortunately, this analogy suggests that the possibility of learning is confined to environments which satisfy quite restrictive stability conditions. In the present example, Ljung (1977, Theorem 2.2) implies that if some characteristic root of $B - I$ has a positive real part, then the forecasting process described by (3.4) has zero probability of converging to σ^*. However, the main results of the present paper show that learning processes can be designed which, as they proceed, modify themselves to produce the desired stability conditions for virtually any environment.

The forecasting rule $p_{t+1} = \bar{s}_t$ can be written as $p_{t+1} = \bar{p}_t + (\bar{s}_t - \bar{p}_t)$, which suggests the modification $p_{t+1} = \bar{p}_t + A(\bar{s}_t - \bar{p}_t)$ for some nonsingular matrix A. Working through the previous algebra for the modified forecasting rule leads to the equations

(3.5') $\qquad p_{t+1} - \bar{p}_t = A(B - I)(\bar{p}_t - \sigma^*) + A\bar{\varepsilon}_t;$ and

(3.7') $\qquad \dot{\bar{p}} = A(B - I)(\bar{p} - \sigma^*).$

Equation (3.7′) suggests that stability is ensured if the matrix A is chosen appropriately; e.g., near $-(B-I)^{-1}$. However, the forecasting agency has no direct knowledge of $B-I$, and must infer the appropriate choice of A from past strategy N-tuples. It may be possible to estimate $B-I$ and adjust the choice of A as time proceeds. Fortunately, however, there is a much simpler method available. If A is chosen near enough to $-(B-I)^{-1}$, for example, one can derive from the law of the iterated logarithm an almost sure rate of convergence for $\bar{p}_t$. Via the equation

$$(3.8) \quad \bar{s}_t - \bar{p}_t = (B-I)(\bar{p}_t - \sigma^*) + \bar{\varepsilon}_t,$$

the rate of convergence of p_t to σ^* determines a rate at which $\bar{s}_t - \bar{p}_t$ converges to zero almost surely. More precisely, there is a nonrandom sequence $\{c_t\}$ with $c_t \to 0$ such that, with probability one, $\|\bar{s}_t - \bar{p}_t\| < c_t$ for all t sufficiently large. Moreover, the numbers c_t are independent of any environmental characteristics. Therefore, the forecasting agency can test its choice of A by checking whether $\|\bar{s}_t - \bar{p}_t\| < c_t$. If not, a new choice can be made and maintained until the next failure of this convergence test. Each time a new choice is made, the numbers c_t must be rescaled upward to prevent the test from failing again immediately, but this is easily done. If each choice of A is an independent drawing from a probability distribution with full support, then with probability one, eventually a stabilizing choice will be made and maintained, and p_t will converge to σ^*. We will now give a precise definition of this forecasting process.

Centralized Forecasting: Fix $p_1 \in R^N$ and $k_1 > 0$ arbitrarily. Let μ be a probability distribution on R^{N^2}, the space of $N \times N$ matrices, with the property that for any open set $U \subset R^{N^2}$, $\mu(U) > 0$. Fix the $N \times N$ matrix A_1 arbitrarily, and for each $t \geq 1$, let

$$(3.9) \quad s_t^i = a^i + b^i p_t + \varepsilon_t^i, \ 1 \leq i \leq N; \ \text{and}$$

$$(3.10) \quad p_{t+1} = \bar{p}_t + A_{t+1}(\bar{s}_t - \bar{p}_t); \ \text{where}$$

$$(3.11) \quad A_{t+1} = A_t \ \text{and} \ k_{t+1} = k_t \ \text{if} \ \|\bar{s}_t - \bar{p}_t\| < k_t t^{-1/3}; \ \text{and}$$

10

(3.12)　$A^{t+1} = A'$ and $k_{t+1} = t(\|\bar{s}_t - \bar{p}_t\| + 1)$ otherwise;

where A' is a matrix drawn from R^{N^2} according to μ independently of previous drawings.

The equation $k_{t+1} = t(\|\bar{s}_t - \bar{p}_t\| + 1)$ represents the rescaling mentioned in the previous paragraph. Lemma 3.1, below, establishes the basis for the convergence test (3.11). The convergence result for Centralized Forecasting, Theorem 3.2, follows the Lemma. Both are proved in the appendix, but some informal explanatory comments follow the statement of Theorem 3.2

Lemma 3.1: Let $\{\delta_t\}_{t=1}^{\infty}$ be a sequence in R^m such that for some positive constants K_1, T_1, γ,

(3.13)　$\|\delta_t\| < K_1 t^{-\gamma}$ for all $t \geq T_1$.

Let $\alpha, \beta > 0$ and let $U = \{D \in R^{N^2} : \|D\| < \beta, \ \max\{x'Dx : \|x\| = 1\} < -\alpha\}.$[1] Assume that $\gamma < \alpha$, and that for some $T_3 > T_2 \geq T_1$,

(3.14)　$x_{t+1} = \bar{x}_t + D_{t+1}\bar{x}_t + \delta_t$

and $D_{t+1} \in U$ for each $T_2 \leq t < T_3$. If T_2 is large enough (more precisely, if $1 - (\alpha + \gamma)T_2^{-1} < (1 - T_2^{-1})^{2\gamma}$ and $T_2 \geq 2(\beta + 1)^2(\min\{\alpha - \gamma, 1\})^{-1})$, then there is some K_2 large enough (more precisely, $K_2 = \max\{T_2^{\gamma}(\|\bar{x}_{T_2}\| + 1), \ 4K_1(\min\{\alpha - \gamma, 1\})^{-1}\}$) that

(3.15)　$\|\bar{x}_t\| < K_2 t^{-\gamma}$ for all $T_2 < t \leq T_3$.

Theorem 3.2: Suppose that $B - I$ is nonsingular. Then for the Centralized Forecasting process defined above, $p_t \to \sigma^*$ and $\sigma_t \to \sigma^*$ with probability one.

The role of Lemma 3.1 can be described very informally as follows. Equation (3.14) can be interpreted as (3.5′) by interpreting x_t as $p_t - \sigma^*$ and interpreting δ_t as a typical

---

[1]　For an $N \times N$ matrix D, define the norm $\|D\| = \max\{\|Dx\| : \|x\| = 1\}$.

11

drawing of $\bar{\varepsilon}_t$. The law of the iterated logarithm implies that the hypothesis (3.13) is satisfied with $\gamma = 1/3$. Therefore the Lemma implies that if A_{t+1} is chosen from a bounded open set of matrices A such that $\max\{x'A(B-I)x : \|x\| = 1\} < -1/2$, for example, then $\bar{p}_t \to \sigma^*$ at the asymptotic rate $t^{-1/3}$. It follows from (3.8) that $\|\bar{s}_t - \bar{p}_t\|$ converges to zero at the same rate as $\bar{p}_t \to \sigma^*$. Since the probability measure μ gives positive probability to every open set, and A_{t+1} is chosen independently each time the convergence test fails, we can conclude that with probability one, an appropriate matrix A_{t+1} will eventually be chosen and the convergence test will eventually always be satisfied.

4. A Convergent Strategy Adjustment Process

This section is devoted to a learning process in which each player, instead of responding to a public forecast, simply adjusts his strategy according to the performance of his previous choices. Specifically, let $s_t^i = \sigma_t^i + \varepsilon_t^i$, and suppose that player i adjusts σ_t^i according to the equation

$$(4.1) \quad \sigma_{t+1}^i = \bar{\sigma}_t^i + \alpha^i(a^i + b^i\bar{s}_t - \bar{\sigma}_t^i),$$

where α^i is a real number. Ideally, player i would like to set $\sigma_t^i = a^i + b^i s_t$, so the quantity $(a^i + b^i\bar{s}_t - \bar{\sigma}_t^i)$ is the average of player i's past mistakes. Thus player i chooses σ_{t+1}^i by adjusting the average of his previous choices by a multiple of the average of his past mistakes. Since $\bar{s}_t^i = \bar{\sigma}_t^i + \bar{\varepsilon}_t^i$, we have

$$(4.2) \quad \sigma_{t+1} = \bar{\sigma}_t + A(a + (B-I)\bar{\sigma}_t) + AB\bar{\varepsilon}_t,$$

where A is the diagonal matrix of adjustment coefficients $\alpha^1, \ldots, \alpha^N$. Then since $\sigma^* = a + B\sigma^*$, we can rewrite (4.2) as

$$(4.3) \quad \sigma_{t+1} - \bar{\sigma}_t = A(B-I)(\bar{\sigma}_t - \sigma^*) + AB\bar{\varepsilon}_t,$$

which is virtually identical to the equation

$$(3.5') \quad p_{t+1} - \bar{p}_t = A(B-I)(\bar{p}_t - \sigma^*) + A\bar{\varepsilon}_t$$

12

of Section 3.

Therefore, the decentralized strategy adjustment process will exhibit virtually the same asymptotic behavior as the centralized forecasting process. There are, however, two important distinctions. First, the matrix A must be a diagonal matrix, so to ensure that $B - I$ can be stabilized by a diagonal matrix, we will assume that all principal minors of $B - I$ are nonzero. Second, a new convergence test must be specified. If the adjustment coefficients are chosen appropriately, Lemma 3.1 again implies that $\bar{s}_t \to \sigma^*$ at the rate $t^{-1/3}$. The convergence of a sequence at a specified rate can be tested without knowledge of the limit, so we can specify such a test for $\bar{s}_t$.

We now give a precise definition of this process.

Decentralized Strategy Adjustment:

For each i, let μ^i be a probability distribution on the real line, with $\mu^i(\{0\}) = 0$ and $\mu^i(U) > 0$ for every open interval U. For any $x \in R$ and $\delta > 0$, let $K(x, \delta)$ denote the open cube centered at x with radius δ. Fix $(\sigma_1^i)_i$, $(\alpha_1^i)_i$, and k_1 arbitrarily, and for each i and each $t \geq 1$, let

$$(4.4) \quad s_t^i = \sigma_t^i + \varepsilon_t^i; \text{ and}$$

$$(4.5) \quad \sigma_{t+1}^i = \bar{\sigma}_t^i + \alpha_{t+1}^i (a^i + b^i \bar{s}_t - \bar{\sigma}_t^i);$$

where, if $C_t \neq \emptyset$,

$$(4.6) \quad \alpha_{t+1}^i = \alpha_t^i;$$

$$(4.7) \quad k_{t+1} = k_t; \text{ and}$$

$$(4.8) \quad C_{t+1} = C_t \cap K(\bar{s}_{t+1}, k_{t+1}(t+1)^{-1/3}); \text{ and}$$

if $C_t = \emptyset$

$$(4.6') \quad \alpha_{t+1}^i = \alpha'^i;$$

$$(4.7') \quad k_{t+1} = t(\|\bar{s}_t\| + 1); \text{ and}$$

13

$$(4.8') \quad C_{t+1} = K(\bar{s}_{t+1}, \ k_{t+1}(t+1)^{-1/3});$$

where α'^i is drawn from the distribution μ^i independently of previous drawings.

Theorem 4.1: Suppose that all principal minors of $B - I$ are nonsingular. Then for the decentralized strategy adjustment process, $\sigma_t \to \sigma^*$ with probability one.

The players choose adjustment coefficients and adjust their strategies independently, but all players use the same convergence test. It might appear that player i should only be concerned with the convergence of σ_t^i, and therefore should test only the convergence of σ_t^i or perhaps $\bar{s}_t^i$. Given player i's preferences, as defined in Section 2, above, it would be more accurate to assert that player i's goal is the convergence to zero of the deviation $|a^i + b^i \sigma_t - \sigma_t^i|$. The first paragraph of the proof of Theorem 4.1 shows that the convergence of $\bar{s}_t$ implies convergence to equilibrium, and, therefore, the convergence to zero of $|a^i + b^i \sigma_t - \sigma_t^i|$ for all i. Hence the convergence of $\bar{s}_t$ is of natural interest to all players.

Nonetheless, it is possible to replace the common convergence test with independent tests of the convergence of $\bar{s}_t^i$ or the convergence to zero of $\|a^i + b_t^i \bar{s}_t - \bar{\sigma}_t^i\|$. Independent tests create several complications in the proof of convergence. If player i's convergence test fails, causing a new choice of α^i, the resulting discrete change in player i's behavior may cause other players to experience nonconvergence of their own strategies; and so on. The simplest way to mitigate this instability appears to be to replace the terms $k_t = t(\|\bar{s}_t\|+1)$, which are used after a test failure, with terms which diverge much more rapidly. Also, to avoid the possibility that some players' strategies may converge while others may not, it is convenient to strengthen the nondegeneracy assumption on $B - I$ to require that the principal minors are nonzero under any permutation of player indices. Subject to these changes, and a much more intricate proof, almost sure convergence to equilibrium can be obtained with independent convergence tests. (This result is stated formally and proved

14

in an earlier version of this paper, Jordan (1985, Theorem 4.1)).

5. A Nonrevealing Adjustment Process

The strategy adjustment process formulated in the preceding section appears to use very little information about the environment $(a^i, b^i)_i$. A natural question is whether the observed strategy N-tuples $\{s_t\}_{t=1}^{\infty}$ reveal much information. In particular, does the time series $\{s_t\}_{t=1}^{\infty}$ identify the environment $(a^i, b^i)_i$? This seems a difficult question, although the fact that the adjustment coefficients are eventually constant suggests a positive answer. However, this property of the adjustment coefficients can be dropped without affecting convergence. For example, suppose that player i has a countable collection of open intervals $I^i = \{U_j^i\}_{j=1}^{\infty}$ which is a base for the usual topology on R. Instead of drawing a coefficient α^i from the distribution π^i, player i draws an interval U_j^i from a strictly positive probability distribution on the countable set I^i, and then chooses $\alpha_t^i \in U_j^i$ according to any measurable function $f_{jt}^i : (s_1, \ldots, s_{t-1}) \to \alpha_t^i$ which is bounded away from zero over time, that is, $\lim \inf |f_{jt}^i(s_1, \ldots, s_t)| > 0$ for every sequence $\{s_t\}_{t=1}^{\infty}$. Then the proof of Theorem 4.1 would be unaffected, but the indeterminacy of the functions f_{jt}^i would seem to prevent identification.

There is an even simpler modification of the adjustment process which directly prevents identification. Suppose that player i selects false characteristics $(\hat{a}^i, \hat{b}^i)$ and uses the strategy adjustment process as though these were his true characteristics. Of course, this charade can be maintained only as long as σ_t is converging to the true equilibrium, so player i must monitor the error $\|a^i + b^i \bar{s}_t - \bar{s}_t^i\|$ in his true equilibrium condition. If this error does not converge to zero sufficiently fast, he reverts to his true characteristics. Convergence might be prevented by repeated switches between true and false characteristics, so we require that player i choose, at the beginning, to revert permanently to his true characteristics after a specified number of failures of the convergence test. If both the true and false characteristics have the same equilibrium, there is a positive probability

15

that the payers never revert to their true characteristics. Therefore, under this convergent modification of the strategy adjustment process, the time series $\{s_t\}_{t=1}^{\infty}$ cannot reveal with certainty anything about the environment other than its equilibrium.

The new strategy adjustment process is defined formally below. Theorem 5.1 asserts that convergence is preserved, and Theorem 5.2 is the nonrevelation result.

The Mimicing Adjustment Process:

Fix $(\sigma_1^i)_i$, $(\alpha_1^i)_i$, $(\hat{a}^i, \hat{b}^i)_{i=1}^N$, and k_1 arbitrarily, and let $(F^i)_i$ be an arbitrary N-tuple of positive numbers. For each i and each $t \geq 1$, let

(4.4) $\quad s_t^i = \sigma_t^i + \varepsilon_t^i;$ and

(5.1) $\quad \sigma_{t+1}^i = \bar{\sigma}_t^i + \alpha_{t+1}^i(\hat{a}^i + \hat{b}^i \bar{s}_t - \bar{\sigma}_t^i)$ if

$\qquad \# \{\tau \leq t : C_\tau = \emptyset\} < F^i$ and

$\qquad \|a^i + b^i \bar{s}_t - \bar{s}_t^i\| / F^i < \sup\{\|\hat{a}^i + \hat{b}^i x - x^i\| : \|x - \bar{s}_t\| < k_t t^{-1/3}\};$ and

(4.5) $\quad \sigma_{t+1}^i = \bar{\sigma}_t^i + \alpha_{t+1}^i(a^i + b^i \bar{s}_t - \bar{\sigma}_t^i)$ otherwise,

where C_t, α_{t+1}^i, and k_t are given by $(4.6 - 4.8)$ and $(4.6' - 4.8')$ in Section 4 above.

Theorem 5.1: Suppose that all principal minors of $B - I$ are nonsingular. Then for the Mimicing Adjustment Process, $\sigma_t \to \sigma^*$ with probability one.

We will formulate the nonrevelation property as the assertion that any function of the environment $(a^i, b^i)_i$ which is independent of the arbitrary initial conditions of the strategy adjustment process, and which is revealed with certainty, must be a function only of the equilibrium $\sigma^* = -(B - I)^{-1}a$. More precisely, let Y be a Hausdorff topological space, and let $f : E \to Y$. For each $t \geq 1$, let S^t denote the t-fold cartesian product of R^N and let $f_t : S^t \to Y$.

Theorem 5.2: Let $E = \{(b_i)_i: \text{all principal minors of } B - I \text{ are nonsingular}\}$. Suppose that for the Mimicing Adjustment Process $f_t(s_1, \ldots, s_t) \to f((a^i, b^i)_i)$ with probability

one for all $(a^i, b^i)_i$ and $(\hat{a}^i, \hat{b}^i)_i$ such that $(b^i)_i$, $(\hat{b}^i)_i \in E$. Then there is a function $g : R^N \to Y$ such that for all $(a^i, b^i)_i$ with $(b^i)_i \in E$, $f((a^i, b^i)_i) = g(\sigma^*)$, where $\sigma^* = -(B - I)^{-1} a$.

Theorem 5.2 demonstrates that it is not necessary to solve the identification problem in order to construct convergent learning processes. The Theorem also implies that the Mimicing Adjustment Process does not necessarily reveal any information which could be used by a strategically sophisticated player to manipulate the strategy adjustments of other players in order to achieve a higher asymptotic payoff. However, this latter implication is rather weak. Since each player i does not know the characteristics of the other players, it seems a priori unlikely that player i would select the pseudo characteristics $(\hat{a}^i, \hat{b}^i)$ from the set of characteristics which yield the same equilibrium σ^i, since the set has Lebesgue measure zero. Thus, from a Bayesian viewpoint, one would expect all players to revert eventually to their true characteristics. These brief remarks indicate that a careful study of the incentive compatibility of learning processes is beyond the scope of this paper.

6. Open Problems

Several interesting aspects of the learning problem have been ignored or treated inadequately in the present paper and in the learning literature in general.

Efficiency: In disequilibrium, players make suboptimal decisions which result in payoff losses. A convergent learning process eliminates such losses asymptotically, but it would be interesting to investigate whether a given learning process economizes on the losses along the path. The literature to date does not appear to contain any general definition of efficiency for learning processes in multi-agent models. Some notion of Pareto efficiency subject to informational limitations appears to be needed. However, in models such as the Cournot oligopoly example in Section 2, even the equilibrium which is the limit of the learning process may not be Pareto efficient. As in the case of the convergence prob-

lem, the interactions among the various decision makers introduce significant economic complications into an otherwise standard statistical problem.

Rate of Convergence: The rate of convergence, though not itself an economic quantity, may have some bearing on the efficiency issue. The convergence test in the learning processes described above ensures convergence at the rate $0(t^{-1/3})$ with probability one. The proof of Theorem 3.2 indicates that a sharper convergence test could improve this to $0(t^{-((1/2)-\varepsilon)})$ for any $\varepsilon > 0$. It seems likely that further improvement would be possible with a more careful probabilistic analysis. Of course the asymptotic rate of convergence ignores the possibly large but finite number of failures of the convergence test. The expected number of periods before the distance from equilibrium is reduced to any given magnitude might be more relevant, but would be quite difficult to compute.

Incentive Compatibility: As the game is repeated over time, players may have an incentive to learn more than just the equilibrium of the one-shot game. In the Cournot oligopoly example, a player who learns the reaction functions of the other players may be able to obtain a higher long-run payoff by acting as a "Stackelberg leader" rather than playing a Nash strategy. Theorem 5.2 states that for the Mimicing Process, the equilibrium is all that can be learned in a classical statistical sense, but even for the Mimicing Process, players who behave as Bayesians may have an incentive to depart from the decisions prescribed by that learning process. Moreover, since the game is repeated indefinitely, sophisticated players might seek a repeated game equilibrium other than the Nash equilibrium of the one-shot game. Of course, if there are a large number of players, or more particularly, if each "player" represents a large number of players of the same type, the incentive to depart from the prescribed asymptotic behavior may be negligible. In the small numbers case, however, the strategic aspects of learning appear difficult to model. For two-person repeated games, Kalai and Lehrer (1990) have formulated a Bayesian learning process which converges to a repeated game Nash equilibrium.

Nonlinear Models: The role of linearity in the above model is two-fold. First, adjustment coefficients can be found which stabilize the process globally, not just locally. Second, a global upper bound on the rate of convergence can be computed given the appropriate adjustment coefficients. Both of these properties are lost if linearity is significantly weakened. Global stability is especially important in models with stochastic disturbances since variables typically cannot be confined, with probability one, to any small neighborhood. For example, the convergence result of Ljung (1977) requires, as a separate hypothesis, that the adjusted variables remain within the domain of attraction of the deterministic differential equation associated with the stochastic adjustment processes. In economic applications, this hypothesis may be ad hoc if the domain of attraction is local rather than global. However, if a nonlinear version of the above model has at least one "regular" equilibrium, in the usual sense of that term, it may possible to obtain general convergence with the following modification of the Centralized Forecasting Process.

Suppose that whenever the convergence test fails, in addition to replacing the adjustment coefficients, the vector $\bar{p}_t$ is replaced by a randomly selected $\tilde{p}_t$. That is $\bar{p}_{t+1} = (t/(t+1)) \tilde{p}_t + (t+1)^{-1} p_{t+1}$; etc. Then, with probability one, $\tilde{p}$ will eventually be selected close enough to a regular equilibrium that the model is sufficiently linear to ensure convergence as in the proof of Theorem 3.2. Unfortunately, in the more interesting Decentralized Strategy Adjustment Process, it would be extremely artificial to suppose that the players make a common selection of $\tilde{s}_t$ to replace $\bar{s}_t$ at each failure of the convergence test.

There is an important class of nonlinear models, namely normal form games, for which the author has constructed generally convergent learning processes (Jordan (1990)). However, these learning processes are based on Bayesian inference and are much more complex than those in the present paper. How much additional complexity is needed in nonlinear models is an interesting question.

Appendix

This appendix contains the proofs of all of the formal results in this paper.

Lemma 3.1: Let $\{\delta_t\}_{t=1}^{\infty}$ be a sequence in R^m such that for some positive constants K_1, T_1, γ,

$$(3.13) \quad \|\delta_t\| < K_1 t^{-\gamma} \text{ for all } t \geq T_1.$$

Let $\alpha, \beta > 0$ and let $U = \{D \in R^{N^2} : \|D\| < \beta, \max\{x'Dx : \|x\| = 1\} < -\alpha\}$. Assume that $\gamma < \alpha$, and that for some $T_3 > T_2 \geq T_1$,

$$(3.14) \quad x_{t+1} = \bar{x}_t + D_{t+1}\bar{x}_t + \delta_t$$

and $D_{t+1} \in U$ for each $T_2 \leq t < T_3$. If T_2 is large enough (more precisely, if $1 - (\alpha + \gamma)T_2^{-1} < (1 - T_2^{-1})^{2\gamma}$ and $T_2 \geq 2(\beta + 1)^2(\min\{\alpha - \gamma, 1\})^{-1})$, then there is some K_2 large enough (more precisely, $K_2 = \max\{T_2^{\gamma}(\|\bar{x}_{T_2}\| + 1), 4K_1(\min\{\alpha - \gamma, 1\})^{-1}\}$) that

$$(3.15) \quad \|\bar{x}_t\| < K_2 t^{-\gamma} \text{ for all } T_2 < t \leq T_3.$$

Proof: For each t

$$(A.1) \quad \|\bar{x}_{t+1}\|^2 = \|\bar{x}_t\|^2 + 2\bar{x}_t(\bar{x}_{t+1} - \bar{x}_t) + (\bar{x}_{t+1} - \bar{x}_t)(\bar{x}_{t+1} - \bar{x}_t).$$

Since $\bar{x}_{t+1} - \bar{x}_t = (t+1)^{-1}(x_{t+1} - \bar{x}_t)$, (3.14) and $A.1$ imply

$$(A.2) \quad \|\bar{x}_{t+1}\|^2 = \|\bar{x}_t\|^2 + 2(t+1)^{-1}\bar{x}_t(C_{t+1}\bar{x}_t + \delta_t) + (t+1)^{-2}(C_{t+1}\bar{x}_t + \delta_t)(C_{t+1}\bar{x}_t + \delta_t).$$

Let $T_2 \leq t < T_3$ and suppose $\|\bar{x}_t\| < K_2 t^{-\gamma}$. Then $(A.2)$ implies that

$$(A.3) \quad \|\bar{x}_{t+1}\|^2 < K_2^2 t^{-2\gamma}\{1 - (t+1)^{-1}2(\alpha - (K_1/K_2)) + (t+1)^{-2}[\beta^2 + 2\beta(K_1/K_2) + (K_1/K_2)^2]\}.$$

We need to show that $\|\bar{x}_{t+1}\|^2 < K_2^2(t+1)^{-2\gamma}$, which will follow if $t^{2\gamma}(t+1)^{-2\gamma} > \{\cdot\}$. Since $t + 1 > T_2$, we have $t^{2\gamma}(t+1)^{-2\gamma} = (1 - (t+1)^{-1})^{2\gamma} > 1 - (\alpha + \gamma)(t+1)^{-1}$.

20

Hence, it suffices to show that $\{\cdot\} < 1 - (\alpha + \gamma)(t+1)^{-1}$, that is $2(\alpha - (K_1/K_2)) > (\alpha + \gamma) + (t+1)^{-1}[\cdot]$, or

$(A.4) \quad a - \gamma > 2(K_1/K_2) + (t+1)^{-1}[\cdot]$.

Since $[\cdot] < (\beta + 1)^2$, the specification of T_2 and K_2 in the hypothesis of the Lemma imply that the terms $2(K_1/K_2)$, and $(t+1)^{-1}[\cdot]$ are each less than $(\alpha - \gamma)/2$, which proves $(A.4)$. Hence, for any $T_2 \leq t < T_3$, if $\|\bar{x}_t\| < K_2 t^{-\gamma}$ then $\|\bar{x}_{t+1}\| < K_2(t+1)^{-\gamma}$. Since $K_2 \geq (\|\bar{x}_{T_2}\| + 1)T_2^{\gamma}$, $\|\bar{x}_{T_2}\| < K_2 T_2^{-\gamma}$, so the Lemma is proved by induction on t. ∎

Theorem 3.2: Suppose that $B - I$ is nonsingular. Then for the Centralized Forecasting process defined above, $p_t \to \sigma^*$ and $\sigma_t \to \sigma^*$ with probability one.

Proof: Let $\{p_t, s_t, \varepsilon_t, A_t\}_{t=1}^{\infty}$ be a realized sequence of predictions, strategies, random characteristics, and prediction adjustment matrices. Since we are proving convergence with probability one, we can invoke the law of the iterated logarithm (Breiman (1968), p. 291)), and assume that for some $K > 0$, $\|\bar{\varepsilon}_t\| < K[(\log \log t)/t]^{1/2}$ for all t sufficiently large.

We first show that if $\|\bar{s}_t - \bar{p}_t\| \geq k_t t^{-1/3}$ only finitely often, then $p_t \to \sigma^*$. Suppose there is some T such that $\|\bar{s}_t - \bar{p}_t\| < k_t t^{-1/3}$ for all $t \geq T$. Then $k_t = k_T$ for all $t > T$, so $\|\bar{s}_t - \bar{p}_t\| \to 0$. Recall

$(3.8) \quad \bar{s}_t - \bar{p}_t = (B - I)(\bar{p}_t - \sigma^*) + \bar{\varepsilon}_t$.

Since $\bar{\varepsilon}_t \to 0$ and $B - I$ is nonsingular, $\bar{p}_t \to \sigma^*$. Since $A_t = A_T$ for all $t > T$, $p_{t+1} = \bar{p}_t + A_T(\bar{s}_t - \bar{p}_t)$ for all $t > T$, so $p_t \to \sigma^*$.

Now suppose, by way of contradiction, that $\|\bar{s}_t - \bar{p}_t\| \geq k_t t^{-1/3}$ infinitely often. Let U be a bounded open subset of R^{N^2} such that if $A \in U$, $\max\{x'A(B-I)x : \|x\| = 1\} < -1/2$. Since U is open, $\pi(U) > 0$. Therefore we can assume that for infinitely many times T, $\|\bar{s}_T - \bar{p}_T\| \geq k_T T^{-1/3}$ and $A_{T+1} \in U$. Equations (3.8) and (3.10) imply

that $p_{t+1} = \bar{p}_t + A_{t+1}(B - I)(\bar{p}_t - \sigma^*) + A_{t+1}\bar{\varepsilon}_t$, so that

$$(A.5) \quad p_{t+1} - \sigma^* = \bar{p}_t - \sigma^* + A_{t+1}(B - I)(\bar{p}_t - \sigma^*) + A_{t+1}\bar{\varepsilon}_t.$$

Equation $(A.5)$ is in the form of equation (3.14) of Lemma 3.1, with $x = p - \sigma^*$. Since U is bounded, there is some $c > 0$ with $\|A\| < c$ for all $A \in U$, where $\|A\|$ is defined as $\max\{\|Ax\| : \|x\| = 1\}$. Hence, for T sufficiently large if $A_{t+1} \in U$, then $\|A_{t+1}\bar{\varepsilon}_t\| < cK[(\log \log t)/t]^{1/2} < t^{-1/3}$ for all $t > T$. Hence, if $A_{t+1} \in U$ for large t, the hypothesis of Lemma 3.1 is satisfied, with $K_1 = 1$, $\gamma = 1/3$, $\alpha = 1/2$, and $\beta = c\|B - I\|$. Hence, for any large T with $\|\bar{s}_T - \bar{p}_T\| > k_T T^{-1/3}$ and $A_{T+1} \in U$, we can conclude from Lemma 3.1 that for $t > T$, as long as $A_t = A_{T+1}$,

$$(A.6) \quad \|\bar{p}_t - \sigma^*\| < T^{1/3}(\|\bar{p}_T - \sigma^*\| + 1)t^{-1/3}.$$

Let $c_{T+1} = \|B - I\| [T^{1/3}(\|\bar{p}_T - \sigma^*\| + 1)] + 1$. Then (3.8) and $(A.6)$ imply that for $t > T$, as long as $A_t = A_{T+1}$,

$$(A.7) \quad \|\bar{s}_t - \bar{p}_t\| < c_{T+1}t^{-1/3}.$$

Recall that $k_{T+1} = T(\|\bar{s}_T - \bar{p}_T\| + 1)$. By (3.8), $\|\bar{s}_T - \bar{p}_T\| > m\|\bar{p}_T - \sigma^*\| - T^{-1/3}$, where $m = \min\{\|(B - I)x\| : \|x\| = 1\}$, so $k_{T+1} > c_{T+1}$ for large T. Hence, taking T large enough, as long as $A_t = A_{T+1}$,

$$(A.8) \quad \|\bar{s}_t - \bar{p}_t\| < k_{T+1}t^{-1/3} = k_t t^{-1/3},$$

so $A_{t+1} = A_{T+1}$, and, by induction, $(A.8)$ holds for all $t > T$. Hence, $\|\bar{s}_t - \bar{p}_t\| \geq k_t t^{-1/3}$ only finitely often. ∎

Theorem 4.1: Suppose that all principal minors of $B - I$ are nonsingular. Then for the decentralized strategy adjustment process, $\sigma_t \to \sigma^*$ with probability one.

Proof: Let $\{(s_t^i, \varepsilon_t^i, \alpha_t^i)_{i=1}^N\}_{t=1}^\infty$ be a realized sequence of strategies, random characteristics, and adjustment coefficients. Again, we can assume that for some $K > 0$, $\|\bar{\varepsilon}_t\| < K[(\log \log t)/t]^{1/2}$ for large t.

It is clear from the definition of C_t that if $C_t = \emptyset$ only finitely many times, then $\bar{s}_t \to s^*$ for some $s^* \in R^N$. Since $\bar{s}_t = \bar{\sigma}_t + \bar{\varepsilon}_t$ and $\bar{\varepsilon}_t \to 0$, it follows that $\bar{\sigma}_t \to s^*$. Also, since $C_t \neq \emptyset$ for all large t, α_t^i is constant for large t, so for each i, (4.5) implies that σ_t^i converges. Hence $\sigma_t^i \to s^{*i}$. From (4.5), $s^{*i} = s^{*i} + \alpha^i(a^i + b^i s^* - s^{*i})$ for some $\alpha^i \in R$. Since $\pi^i(\{0\}) = 0$, we can assume $\alpha^i \neq 0$, so $a^i + b^i s^* = s^{*i}$ for each i. Therefore $s^* = \sigma^*$, so $\sigma_t \to \sigma^*$.

Now suppose, by way of contradiction, that $C_t = \emptyset$ infinitely often. For each t, let A_t be the diagonal matrix of adjustment coefficients $(\alpha_t^i, \ldots, \alpha_t^N)$. Then

$$(A.9) \qquad \sigma_{t+1} - \bar{\sigma}_t = A_{t+1}(B - I)(\bar{\sigma}_t - \sigma^*) + A_{t+1} B \bar{\varepsilon}_t,$$

which is simply (4.3), with the appropriate time subscript on A. Since all principal minors of $B - I$ are nonzero, the Fisher-Fuller Theorem (Fisher and Fuller (1958)) implies the existence of a diagonal matrix A° such that the characteristic roots of $A^\circ(B - I)$ are real, distinct, and negative. Therefore, there is a nonsingular diagonalizing matrix P such that $\max\{y' P A^\circ(B - I) P^{-1} y : \|y\| = 1\} < 0$. Hence, for each i, there is a bounded open interval $U^i \subset R$ such that if $\alpha^i \in U^i$ for each i, then $\max\{y' P A(B - I) P^{-1} y : \|y\| = 1\} < -1/2$. Let U denote the set of diagonal matrices A with $\alpha^i \in U^i$ for each i. Since $\pi^i(U^i) > 0$ for each i, we can assume that $C_t = \emptyset$ and $A_{t+1} \in U$ infinitely often.

As in the proof of Theorem 3.2, the remainder of the proof relies on Lemma 3.1. In this case we apply the Lemma to equation (A.9) in the coordinates $y = Px$. Since the conclusion of the Lemma is clearly invariant under linear coordinate changes, we can ignore the coordinate change to economize on notation. Thus, as in the proof of Theorem 3.2, if T is large, $C_T = \emptyset$ and $A_{T+1} \in U$, then Lemma 3.1 implies

$$(A.10) \qquad \|\bar{\sigma}_t - \sigma^*\| < T^{1/3}(\|\bar{\sigma}_T - \sigma^*\| + 1) t^{-1/3}$$

for $t > T$ as long as $A_t = A_{T+1}$. Since $\|\bar{s}_t - \sigma^*\| < \|\bar{\sigma}_t - \sigma^*\| + \|\bar{\varepsilon}_t\|$ and $\|\bar{s}_t\| > \|\bar{\sigma}_t - \sigma^*\| - \|\bar{\varepsilon}_t\| - \|\sigma^*\|$, it follows from (A.10) that

$$(A.11) \quad \|\bar{s}_t - \sigma^*\| < T(\|\bar{s}_T\| + 1)t^{-1/3}$$

for $t > T$ as long as $A_t = A_{T+1}$. Since $C_T = \emptyset$, $C_{T+1} \neq \emptyset$ by definition. Suppose that for some $T' \geq T+1$, $C_t \neq \emptyset$ for all $T+1 \leq t \leq T'$. We will show that $C_{T'+1} \neq \emptyset$. Since $C_t \neq \emptyset$ for all $T+1 \leq t \leq T'$, $A_{T'+1} = A_{T+1}$, so $(A.11)$ holds through $t = T'+1$. Hence, for each $T+1 \leq t \leq T'+1$, the right hand side of $(A.11)$ is $k_t t^{-1/3}$ for each such t. Hence, for each such t, $\sigma^* \in K(\bar{s}_t, k_t t^{-1/3})$, so $C_{T'+1}$, which is the intersection of these sets, is nonempty. It follows by induction that $C_t \neq \emptyset$ for all $t > T$. ∎

Theorem 5.1: Suppose that all principal minors of $B - I$ are nonsingular. Then for the Mimicing Adjustment Process, $\sigma_t \to \sigma^*$ with probability one.

Proof: Let $\{(s_t^i, \varepsilon_t^i, \alpha_t^i)_{i=1}^N\}_{t=1}^\infty$ be a realized sequence of strategies, random characteristics, and adjustment coefficients. As in the proof of Theorem 4.1, we first show that $\sigma_t \to \sigma^*$ if $C_t = \emptyset$ only finitely often. In this case, as in the proof of Theorem 4.1, $\bar{s}_t \to s^*$ and $\sigma_t \to s^*$ for some s^*. Since $C_t = \emptyset$ only finitely often, α_t^i is eventually constant and, with probability open, nonzero for each i. Suppose that, for some player i, equation (4.5) holds infinitely often. Then, taking limits on both sides, we obtain $s^{*i} = a^i + b^i s^*$ exactly as in the proof of Theorem 4.1. If (4.5) holds only finitely often, then (5.1) holds for all large t and

$$(A.12) \quad \|a^i + b^i \bar{s}_t - \bar{s}_t^i\|/F^i < \sup\{\|\hat{a}^i + \hat{b}^i x - x^i\| : \|x - \bar{s}_t\| < k_t t^{-1/3}\}$$

for all large t. Since (5.1) holds for all large t, the above argument implies that $s^{*i} = \hat{a}^i + \hat{b}^i s^*$. Since $k_t t^{-1/3} \to 0$ and $\bar{s}_t \to s^*$, it follows that the right hand side of $(A.12)$ converges to zero. Therefore $s^{*i} = a^i + b^i s^*$. Hence, we have shown that if $C_t = \emptyset$ only finitely often, then $s^{*i} = a^i + b^i s^*$ for each i. Therefore $s^* = \sigma^*$ and $\sigma_t \to \sigma^*$. If $C_t = \emptyset$ infinitely often, then (4.5) holds for all i and all large t, so the second paragraph of the proof of Theorem 4.1 applies exactly to show that this event has probability zero. ∎

Theorem 5.2: Let $E = \{(b_i)_i: \text{all principal minors of } B - I \text{ are nonsingular}\}$. Suppose that for the Mimicing Adjustment Process $f_t(s_1, \ldots, s_t) \to f((a^i, b^i)_i)$ with probability one for all $(a^i, b^i)_i$ and $(\hat{a}^i, \hat{b}^i)_i$ such that $(b^i)_i$, $(\hat{b}^i)_i \in E$. Then there is a function $g : R^N \to Y$ such that for all $(a^i, b^i)_i$ with $(b^i)_i \in E$, $f((a^i, b^i)_i) = g(\sigma^*)$, where $\sigma^* = -(B - I)^{-1}a$.

Proof: It suffices to show that for any $(a^i, b^i)_i$, $(\bar{a}^i, \bar{b}^i)_i \in E$ with $-(B - I)^{-1}a = -(\bar{B} - I)^{-1}\bar{a}$, $f((a^i, b^i)_i) = f((\bar{a}^i, \bar{b}^i)_i)$. Given two such environments, let σ^* denote the common equilibrium. For each i, let $c^i = \sup\{\|a^i + b^i x - x^i\| : \|x - \sigma^*\| < 1\}$, and define $\bar{c}^i$ analogously for $(\bar{a}^i, \bar{b}^i)$. Note that for any $\delta > 0$ and any $s \in R^N$, if $\|\sigma^* - s\| < \delta$ then $\|\bar{a}^i + \bar{b}^i s - s^i\|/\bar{c}^i < \delta \leq \sup\{\|a^i + b^i x - x^i\| : \|x - s\| < \delta\}/c^i$. Given $(a^i, b^i)_i$, fix $(\sigma_1^i, \alpha_1^i)_i$ and k_1. Set $(\hat{a}^i, \hat{b}^i)_i = (a^i, b^i)_i$, so that equations (5.1) and (4.5) are identical. Then there are positive constants K and T such that, with positive probability, $\|\bar{s}_t - \sigma^*\| < K$ for all t and $C_t \neq \emptyset$ for all $t > T$. Hence, on this event,

$$(A.13) \quad \#\{\tau < t : C_\tau = \emptyset\} < T \quad \text{for all} \quad t.$$

Also on this event, $\|\sigma^* - \bar{s}_t\| < k_t t^{-1/3}$ for all $t > T$, so we have

$$(A.14) \quad \|\bar{a}^i + \bar{b}^i \bar{s}_t - \bar{s}_t^i\|/\sup\{\|a^i + b^i x - x^i\| : \|x - \bar{s}_t\| < k_t t^{-1/3}\}$$
$$< \max\{\max\{\bar{c}^i K/c^i k_t t^{-1/3} : t \leq T\}, \bar{c}^i/c^i\} \quad \text{for all } i \text{ and all } t.$$

Now, given the environment $(\bar{a}^i, \bar{b}^i)_i$ and the same initial conditions $(\sigma_1^i, \alpha_1^i)_i$ and k_1 as above, set $(\hat{a}^i, \hat{b}^i)_i = (a^i, b^i)$, and for each i, set F^i larger than both T and the right hand side of $(A.14)$. Under these specifications of the adjustment processes for the respective environments, the sequences $\{\varepsilon_t, (\alpha_t^i)_i\}_{t=1}^{\infty}$ which determine the above event result in the same $\{s_t\}_{t=1}^{\infty}$ sequences for both environments, and thus the same $\{f_t(s_1, \ldots, s_t)\}_{t=1}^{\infty}$ sequences for both environments. Since the event has positive probability and Y is a Hausdorff space, $f((a^i, b^i)_i) = f((\bar{a}^i, \bar{b}^i)_i)$.; ∎

25

References

Arthur, W.B., Y. Ermoliev, and Y. Kaniovski (1987), "Non-Linear Urn Processes: Asmptotic Behavior and Applications", Working Paper WP-87-85, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Breiman, L. (1968), *Probability* (Reading: Addison-Wesley).

Fisher, M.E., and A.T. Fuller (1958), "On the Stabilization of Matrices and the Convergence of Linear Processes", *Proceedings of the Cambridge Philosophical Society*, 54, 417-425.

Frydman, R. (1982), "Towards an Understanding of Market Processes: Individual Expectations, Learning, and Convergence to Rational Expectations Equilibrium", *American Economic Review*, 72, 652-668.

Jordan, J. (1990), "Bayesian Learning in Normal Form Games", *Journal of Games and Economic Behavior*, forthcoming.

Jordan, J. (1985), "Convergence to Bayesian Equilibrium in a Stationary Linear Game", mimeo.

Jordan, J. (1984), "Information Flows Intrinsic to the Stability of Economic Equilibrium", Preprint #75, Institute for Mathematics and its Applications, University of Minnesota, Minneapolis.

Kalai, E. and E. Lehrer (1990), "Rational Learning Leads to Nash Equilibrium", Discussion Paper No. 858, Northwestern University.

Ljung, L. (1977), "Analysis of Recursive Stochastic Algorithms", *IEEE Trans. Auto. Control*, 22, 551-575.

Marcet, A., and T. Sargent (1989), "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, 48, pp. 337-368.

Williams, S. (1985), "Necessary and Sufficient Conditions for the Existence of a Locally

Stable Message Process", *Journal of Economic Theory*, **35**, 127-154.

Woodford, M. (1990), "Learning to Believe in Sunspots", *Econometrica*, **58**, 277-307.