JOURNAL OF ECONOMIC THEORY 26, 340–351 (1982)

Learning to Be Rational*

LAWRENCE E. BLUME

Department of Economics, University of Michigan, Ann Arbor, Michigan 48109

AND

DAVID EASLEY

Department of Economics, Cornell University, Ithaca, New York 14853

Received February 25, 1981

We study the dynamical system of expectations generated by a simple general equilibrium model of an exchange economy in which each agent considers a finite collection of models, each of which specifies a relationship between payoff-relevant information and equilibrium prices. One of the models under consideration is a correct description of the rational expectations equilibrium. We find that under a Bayesian type of learning process the rational expectations equilibrium is locally stable, but that nonrational equilibria may also be locally stable. *Journal of Economic Literature* Classification Numbers: 021, 026.

One of the most important problems in the theory of equilibrium with uncertainty is the relationship between equilibrium prices and traders' payoff-relevant information. The issue is: Can traders accurately forecast information available to others by observing only the equilibrium price? Such a forecast, if it exists, is a "rational expectation."

However, in any model of equilibrium with uncertainty, non-rational expectations equilibria exist. Why then are rational expectations equilibria of particular interest? The usual answer is, "If traders held incorrect beliefs, then over the course of time they would discover their error and modify their behavior accordingly." In other words, over time traders learn the true

*The authors gratefully acknowledge the comments offered by Richard Erikson, John Ledyard, and Roy Radner. Discussions with Jim Jordan have been especially insightful. Also helpful were discussions in seminars at Stonybrook, Washington University, and at the Midwestern Conference on Mathematical Economics. This research was supported by the National Science Foundation and by the Center for Mathematical Studies in Economics and Management Science at Northwestern University.

0022-0531/82/020340–12$02.00/0

relationship between information and prices. Rational expectations are, so this argument goes, globally stable equilibria of a dynamical system whose laws of motion are given by specifying traders' learning behavior.

In this paper we test this argument. Suppose each trader considers a finite collection of economic models, each mapping price to a probability on private information. One model under consideration is assumed to be the rational model, i.e., a model that puts probability one on the actual information vector associated with an equilibrium price. Given an observation of an equilibrium price and an information vector each trader learns by updating a prior distribution over models. The learning rule takes a simple form: the prior weight on a model is increased if and only if the model predicts the observed outcome better than does the "average model." This learning rule would be a generalization of Bayesian learning if each trader had a correctly specified model of the relation between prices and information vectors that prevails when traders are learning. However, the learning process itself affects the observed relation between prices and information vectors so none of the models (including the rational model) need be correct out of equilibrium.

The learning process that we describe can best be thought of as a boundedly rational version of Bayesian learning. Fully rational learning would require each trader to take into account the effect of his learning and the learning of others on equilibrium prices. However, the necessary structure for such an analysis requires a degree of sophistication on the part of traders that seems highly implausible.

In our boundedly rational learning structure we find that for regular economies, rational expectations equilibria are locally stable, and there exist non-rational expectations equilibria that are also locally stable.

Our arguments are simply constructed. They depend upon finiteness of signals and of models, and continuity of the trader's learning rules.

The equilibrium model is described in Section 1. In Section 2, expectations are discussed. Dynamics are treated in Section 3 and examples are presented in Section 4. Conclusions are drawn in Section 5. See Blume *et al.* [1] for a discussion of the related literature.

1. Notation and Assumptions

We consider economies with K goods and N traders. Prices are vectors in the positive unit simplex Δ^{K-1}. Trader n in period t observes a signal s_{nt} drawn from a finite set S_n. Let $S = \prod_{n=1}^{N} S_n$ denote the set of joint signals. These signals are correlated with some payoff-relevant random variables; thus knowledge of other traders' signals is payoff-relevant information for each trader.

Suppose that joint signal $s_t \in S$ were to occur. Denote by $(s_t)_n$ the nth trader's component of that signal. Each trader has a guess, given his own signal, as to the value of the joint signal. This guess can be represented by a probability distribution α_{nt}; an element of $\Delta^{\#S-1}$, the positive unit simplex of dimension $\#S-1$. Denote by $\alpha_{nt}(s)$ the coordinate of α_{nt} corresponding to joint signal s. We shall impose the following rationality requirement at the outset. Suppose that joint signal $s \in S$ has occurred. Then each trader's subjective beliefs about the true joint signal must satisfy $\alpha_{nt}(s') > 0$ implies $(s')_n = (s)_n$. This is to say, trader n assigns probability 0 to all joint signals inconsistent with his own information. Let $L_n = \sum_{i \neq n} \#S_i$. Let $\Delta(s_n)$ denote the $L_n - 1$ dimensional face of $\Delta^{\#S-1}$ whose vertices correspond to those joint signals for which $(s')_n = s_n$. We require that if joint signal s_t occurs at time t, then $\alpha_{nt} \in \Delta((s_t)_n)$ for all n. Let $\Delta(s) = \prod_{n=1}^N \Delta((s)_n)$, and let $\alpha = (\alpha_n)_{n=1}^N$, $\alpha \in \Delta(s)$, represent the vector of all trader's beliefs in the event that joint signal s occurs.

With each joint signal s we identify an excess demand function Z_s. This excess demand function is assumed to satisfy *Walras Law*. If the excess demand for $K-1$ goods is known, then the excess demand for the Kth good is known as well. Thus the excess demand function has range $\mathbb{R}^{K-1}$; Z_s: $(\operatorname{Int} \Delta^{K-1}) \times \Delta(s) \to \mathbb{R}^{K-1}$. We also assume:

A.1. *Z_s is bounded from below.*

A.2. *For all $\alpha \varepsilon \Delta(s)$, if a price sequence with $p_i \subset \operatorname{Int} \Delta^{K-1}$ and $p_i \to p \in \partial \Delta^{K-1}$, then* $\liminf \|Z_s(p_i, \alpha)\| > 0$.

A.3. *Z_s is C^1 and 0 is a regular value of Z_s.*

A.4. *Excess demand for commodity K is bounded from below.*

Thus each Z_s is very well behaved. We know that for each $\alpha \in \Delta(s)$ there exists $p \in \operatorname{Int} \Delta^{K-1}$ such that $Z_s(p, \alpha) = 0$. Since $0 \in \mathbb{R}^{K-1}$ is a regular value of Z_s the graph of the equilibrium correspondence is a manifold of codimension $K-1$ in $(\operatorname{Int} \Delta^{K-1}) \times \Delta(s)$. It follows from the transversality theorem (Guilleman and Pollack [3, 68]), that for generic $\alpha \in \Delta(s)$, $Z_s(p, \alpha) = 0$ implies that $D_p Z_s(p, \alpha)$ is surjective. In this instance α is said to be a regular vector of beliefs.

DEFINITION 1. $\alpha \varepsilon \Delta(s)$ is a *regular vector of beliefs* if $Z_s(p, \alpha) = 0$ implies $D_p Z_s(p, \alpha)$ is surjective.

Denote by $\alpha_s \in \Delta(s)$ the vector of probability distributions such that each trader assigns probability 1 to joint signal s.

DEFINITION 2. A *revealing rational expectations equilibrium* (RREE) is a set of prices $\{p_s; s \in S\}$ satisfying

(i) $Z_s(p_s, \alpha_s) = 0$ for all $s \in S$,

(ii) $p_s = p_{s'}$ implies $s = s'$.

A RREE $\{p_s: s \in S\}$ is regular if α_s is a regular vector of beliefs. Thus each trader can unambiguously assign probability 1 to joint signal s upon observing p_s. Were each trader to do so and act accordingly, then excess demand at p_s would be 0.

Our last assumption about excess demand is

A.5. *A regular* RREE *exists.*

The justification of this assumption depends upon the particular model used to derive the Z_s. It is generically true, for example, in Radner's model [4]. The model that we have in mind has excess demand depending upon beliefs about some exogenous payoff-relevant events. These events are correlated with joint signals, and so if traders make use of information about joint signals, they can improve their guesses about the occurrence of the payoff-relevant events. The excess demand functions Z_s are derived from expressing excess demand as a function of prices and beliefs about events. If 0 is a regular value of this function, then for generic conditional probabilities of events given joint signals, A.5 will be true.

2. MODELS

In this section we consider how traders arrive at their beliefs. In brief, they do so by using models. Each trader observes his own signal and the equilibrium price vector. Using this information, the trader then forecasts which s has occurred. This forecast is made with a model of the relationship between (s_n, p) and s.

DEFINITION 3. A *model* for trader n is a function $\Psi_n: S_n \times \Delta^{K-1} \to \Delta$ satisfying

(i) $\Psi_n(s_n, \cdot): \Delta^{K-1} \to \Delta(s_n)$,

(ii) $\Psi_n(s_n, \cdot) \in C^1$.

In a RREE there is only one price associated with each of a finite number of joint signals, so only a finite number of equilibrium prices are possible. However, during the learning process many prices other than the RREE price may occur for any given joint signal. Hence, models must be defined for all feasible signal-price pairs if the boundedly rational version of Bayesian learning is to result in a well-defined learning rule.[1]

DEFINITION 4. Let $\{p_s; s \in S\}$ be a RREE. A *rational model* for trader n is a model Ψ_n such that $\Psi_n((s)_n, p_s)(s) = 1$ for all $s \in S$.

We denote by $\Psi(s, p)$ the vector of models $(\Psi_n((s)_n, p))_{n=1}^N$. A vector of rational models has the property that $\Psi(s, p_s) = \alpha_s$.

We are not concerned with the precise relationship between models and equilibrium prices. Our only concern is that in a neighborhood of some models of particular interest the relationship is continuous. We guarantee this by studying models regular in a neighborhood of some equilibrium. Recall that a point x in the domain of a function f is a regular point if $Df(x)$ is a surjective linear map from the tangent space of the domain at x to the tangent space of the range at $f(x)$.

DEFINITION 5. A vector of models Ψ is *regular* at $\{p_s; s \in S\}$ if, for all $s \in S$,

(i) $Z_s(p_s, \Psi(s, p_s)) = 0$,

(ii) each p_s is a regular point of $Z_s(\cdot, \Psi(s, \cdot))$.

Fortunately it is easy to find regular models. Proposition 1 shows that for any economy satisfying A.5, any rational model is regular at the RREE.

PROPOSITION 1. *Let Ψ be a vector of rational models with regular* RREE *prices $\{p_s; s \in S\}$. Then Ψ is regular at $\{p_s; s \in S\}$.*

Proof. $DZ_s(p, \Psi(s, p)) = D_p Z_s(p, \Psi(s, p)) + D_\alpha Z_s(p, \Psi(s, p)) \cdot D_p \Psi(s, p)$. Choose $s \in S$.

$$D_p \Psi(s, p) = \begin{bmatrix} D_p \Psi_1((s)_1, p) \\ \vdots \\ D_p \Psi_N((s)_N, p) \end{bmatrix}.$$

For trader n, the coordinate of $\Psi_n((s)_n, p_s)$ corresponding to joint signal s', $\Psi_n((s)_n, p_s)(s')$ is either 1 or 0 as s' does or does not equal s. Since, for all p, $0 \leqslant \Psi_n((s)_n, p)(s') \leqslant 1$ and $p_s \in \operatorname{Int} \Delta^{K-1}$, $\Psi_n((s)_n, p_s)(s')$ is either at a maximum or a minimum, and so $D_p \Psi_n((s)_n, p_s) = 0$. Hence, $DZ_s(p_s, \Psi(s, p_s)) = D_p Z_s(p_s, \Psi(s, p_s))$. But since $\{p_s; s \in S\}$ is a regular RREE, $D_p Z_s(p_s, \Psi(s, p_s))$ is surjective. Q.E.D.

[1] Although we have assumed models to be C^1, it should be clear from the proofs that models need only be locally C^1 in certain neighborhoods.

The next proposition shows that regularity is a generic property of vectors of models.

PROPOSITION 2. *Let $\beta_s \in \operatorname{Int} \Delta(s)$ for all $s \in S$ and let each β_s be a regular vector of beliefs. Let $\{p_s; s \in S\}$ be a set of prices such that $Z_s(p_s, \beta_s) = 0$ for all $s \in S$. The set of vectors of models Ψ regular at $\{p_s; s \in S\}$ with $\Psi(s, p_s) = \beta_s$ is relatively open and dense in the set of all vectors of models Ψ with $\Psi(s, p_s) = \beta_s$ in the topology of uniform C^1 convergence.*

Proof. Let $H = \{\Psi: \Psi(s, p_s) = \beta_s$ for all $s \in S\}$. We need to show for almost all $\Psi \in H, DZ_s(p_s, \Psi(s, p_s))$ is surjective. $DZ_s(p_s, \Psi(s, p_s)) = D_p Z_s(p_s, \beta_s) + D_\beta Z_s(p_s, \beta_s) \cdot D_p \Psi(s, p_s)$. Choose $s \in S$. Denote by A the $(K-1) \times (K-1)$ matrix $D_p Z_s(p_s, \beta_s)$, and by B the $(K-1) \times L$ matrix $D_\beta Z_s(p_s, \beta)$, where $L = \sum_{n=1}^{N} L_n - N$. Let C be an $L \times (K-1)$ matrix representing $D_p \Psi(s, p_s)$ for some $\Psi \in H$. Then $DZ_s(p_s, \Psi(s, p_s)) = A + BC$ which is $(K-1) \times (K-1)$. Since $\beta_s \in \operatorname{Int} \Delta(s)$ there are no restrictions on admissible C matrices. Therefore the proposition is true if the set of $L \times (K-1)$ matrices C such that $A + BC$ has full rank $K-1$ is open and dense. This set is obviously open. To see that this set is dense, choose arbitrary C and consider the function from $\mathbb{R}$ to $\mathbb{R}$ defined by $p(\lambda) = \det(A + B(\lambda C))$. The function p is evidently a polynomial, and since $p(0) = \det A \neq 0$, p is not the degree 0 polynomial identically equal to 0. Thus it has only a finite number of roots. Hence if $\det(A + BC) = 0$, then for λ arbitrarily near to 1, $p(\lambda) \neq 0$, and so there exists C' arbitrarily near to C such that $A + BC$ has full rank. Q.E.D.

We conclude from Propositions 1 and 2 that restricting attention to regular models is not a particularly binding constraint.

3. DYNAMICS

We posit dynamics in which a trader tries to learn which of several models is correct. Each trader considers models $\Psi_{n1}, \ldots, \Psi_{nM_n}$, and has initially a prior distribution representing his beliefs about which models are correct. The dynamics are as follows. Each trader begins with a vector λ_{n0} representing his prior distribution on the models. λ_{n0} is an element of the positive unit simplex $\Delta^{M_n - 1}$. A joint signal $s \in S$ is randomly drawn, and each trader learns $(s)_n$. When trader n observes price vector p and signal $(s)_n$, his beliefs about $s \in S$ are given by the probability distribution $\sum_{m=1}^{M_n} \lambda_{nom} \Psi_{nm}((s)_n, p)$. After the market clears, traders are informed of the true joint signal s. They compute their posterior distributions λ_{n1} by applying a learning rule. Then the process begins again.

We begin by considering learning rules.

DEFINITION 6. A *learning rule* for trader n is a function $l_n: S \times \Delta^{K-1} \times \Delta^{M_n-1} \to \Delta^{M_n-1}$, $\lambda_{nt} = l_n(s_{t-1}, p_{t-1}, \lambda_{n,t-1})$.

We impose only two restrictions on learning rules.

A.6. (i) The learning rules l_n, $n = 1,\ldots, N$ are continuous.

(ii) (Likelihood Property) $\lambda_{n,t+1,m} > \lambda_{ntm}$ if

$$\Psi_{nm}((s_t)_n, p_t)(s_t) > \sum_{m=1}^{M_n} \lambda_{ntm} \Psi_{nm}((s_t)_n, p_t)(s_t)$$

and $1 > \lambda_{ntm} > 0$, otherwise $\lambda_{n,t+1,m} = \lambda_{ntm}$.

The likelihood property requires that trader n's confidence in model Ψ_{nm} increase if and only if model Ψ_{nm} predicts the outcome in the previous period better than does the "average model." For a correctly specified model, Bayesian updating generates a specific learning rule that satisfies A.6.[2] In our equilibrium framework, however, the learning rule need not be statistically correct as the updating mechanism in A.6 may be misspecified. We call the learning rule boundedly rational as it applies Bayesian updating to models that ignore only the effect of learning itself on observations.

We write $\lambda_t = (\lambda_{1t},\ldots, \lambda_{Nt})$, and

$$Z_s(p, \lambda_t) = Z_s\left(p, \sum_{m=1}^{M_1} \lambda_{1tm} \Psi_{1m}((s)_1, p),\ldots, \sum_{m=1}^{M_N} \lambda_{Ntm} \Psi_{Nm}((s)_N, p)\right).$$

Also, let $\bar{\lambda} = (\bar{\lambda}_1,\ldots, \bar{\lambda}_N)$, where $\bar{\lambda}_n = (1, 0,\ldots, 0)$, and $\Psi_1 = (\Psi_{11},\ldots, \Psi_{N1})$.

In order to specify the sequence of temporary equilibria that obtain, we need to specify how the equilibrium prices vary with λ. All we require is some continuity, which comes from regularity properties of the models.

PROPOSITION 3. *Let models $\Psi_{11},\ldots, \Psi_{N1}$ be regular at prices $\{p_s; s \in S\}$. Also let $\Psi_{12},\ldots, \Psi_{NM_N}$ be any other models. Then for each $s \in S$ there exists an $\varepsilon_s > 0$ and a function $g_s: \prod_{n=1}^{N} \Delta^{M_n-1} \to \Delta^{K-1}$ such that*

(i) *g_s is C^1 on $\{\lambda: \|\lambda - \bar{\lambda}\| < \varepsilon_s\}$,*

(ii) *$Z_s(g_s(\lambda), \lambda) = 0$ if $\|\lambda - \bar{\lambda}\| < \varepsilon_s$,*

(iii) *$g_s(\bar{\lambda}) = p_s$.*

[2] Specifically, posterior beliefs about any model will be more favorable than prior beliefs if and only if the model outperforms the average and if prior beliefs were not such that the model was surely correct or surely false. The likelihood property can be modified by deleting the condition $\lambda_{ntm} > 0$ without affecting our results.

Proof. $DZ_s(p_s, \lambda) = [D_p Z_s(p_s, \lambda)\, D_\lambda Z_s(p_s, \lambda)]$. At $\lambda = \bar{\lambda}$, $D_p Z_s(p_s, \lambda) = DZ_s(p_s, \Psi_1(s, p_s))$. Since the vector of models Ψ_1 is regular at $\{p_s; s \in S\}$, $D_p Z_s(p_s, \bar{\lambda})$ is surjective. The proposition now follows from the implicit function theorem. Q.E.D.

The formal specification of the dynamical system under consideration is

$$(*)\quad \lambda_{t+1} = l(s_t, g_{s_t}(\lambda_t), \lambda_t), \qquad \text{where } \lambda_{n,t+1} = l_n(s_t, g_{s_t}(\lambda_t), \lambda_{nt});\ n = 1,..., N,$$

and the domain of definition is the intersection of the domains of definition of the g_{s_t}. We give sufficient conditions for $\bar{\lambda}$ to be a locally stable stationary point of (*), and then apply this condition to rational models.

THEOREM 1. *Let* $\{\Psi_{nm}, n = 1,..., N, m = 1,..., M_n\}$ *be a collection of models with the following properties:*

(i) *the vector of models* Ψ_1 *is regular at prices* $\{p_s; s \in S\}$,

(ii) $\Psi_{n1}((s)_n, p_s)(s) > \Psi_{nm}((s)_n, p_s)(s)$ *for all* $m = 2,..., M_n$; $n = 1,..., N$; *and* $s \in S$.

Then $\bar{\lambda}$ *is a locally stable rest point of* (*).

The intuition behind the proof is illustrated in Fig. 1. Trader n is choosing between models Ψ_{n1} and Ψ_{n2}. If λ_t is sufficiently near $\bar{\lambda}$, then $g_s(\lambda_t)$ will be near enough to p_s so that $\Psi_{n1}((s)_n, g_s(\lambda_t))(s) > \Psi_{n2}((s)_n, g_s(\lambda_t))(s)$. The likelihood property then implies that $\lambda_{n,t+1,1} > \lambda_{nt1}$. If this is the situation for all n and all $s \in S$, then $\|\bar{\lambda} - \lambda_t\| > \|\bar{\lambda} - \lambda_{t+1}\|$. Continuity of the functions l_n then imply that $\lim_{t\to\infty} \lambda_t = \bar{\lambda}$.

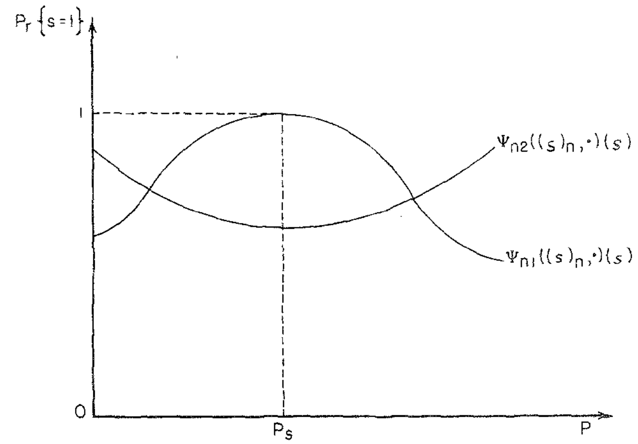

FIGURE 1

This theorem relies on the assumption that each agent has a finite signal set and considers only a finite number of models. Whether or not a similar result would hold with a continuum is an open question. Bray [2] has demonstrated stability of the rational model with a continuum in a specific partial equilibrium model with linear supply and demand functions.

Proof of Theorem 1. Take as a norm for $\prod_{n=1}^{N} \Delta^{M_n - 1}$ the l_1 norm $\|\lambda - \lambda'\| = \sum_{n=1}^{N} \sum_{m=1}^{M_n} |\lambda_{nm} - \lambda'_{nm}|$. From Proposition 3 and the continuity of the Ψ_{nm} and the hypotheses of the theorem, there exists an $\varepsilon > 0$ such that if $\|\bar{\lambda} - \lambda_t\| < \varepsilon$, then for all $s \in S g_s(\lambda_t)$ is continuous and for all $n = 1,..., N$ and $m = 2,..., M_n$, $\Psi_{n1}((s)_n, g_s(\lambda_t))(s) > \Psi_{nm}((s)_n, g_s(\lambda_t))(s)$. Now $\|\bar{\lambda} - \lambda_t\| = 2 \sum_{n=1}^{N} (1 - \lambda_{n_t 1})$. The likelihood property implies that $\lambda_{n,t+1,1} > \lambda_{n_t 1}$ for all $n = 1,..., N$ and $s \in S$. Thus $\|\bar{\lambda} - \lambda_{t+1}\| < \|\bar{\lambda} - \lambda_t\|$.

Let $\delta_t = \|\bar{\lambda} - \lambda_t\|$. Then $0 \leqslant \delta_t < \varepsilon$ and $\{\delta_t\}_{t=1}^{\infty}$ is a decreasing sequence. Hence it has a limit $\bar{\delta}$. Suppose $\bar{\delta} > 0$. Then the sequence $\{\lambda_t\}$ has a limit set A with $\lambda \in A$ implying that $\|\bar{\lambda} - \lambda\| = \bar{\delta} > 0$. But if $\|\bar{\lambda} - \lambda\| = \bar{\delta} > 0$, then $\|\bar{\lambda} - l(s, g_s(\lambda), \lambda)\| = 2 \sum_{n=1}^{N} (1 - l(s, g_s(\lambda), \lambda)_{n1}) < 2 \sum_{n=1}^{N} (1 - \lambda_{n1}) = \|\bar{\lambda} - \lambda\| = \delta$, by the likelihood property. Thus λ cannot be in A, and so $A = \varnothing$. Thus $\delta_t \to 0$, and so $\lambda_t \to \bar{\lambda}$. Q.E.D.

COROLLARY 1. *Let $\Psi_1 = (\Psi_{11},..., \Psi_{N1})$ be a vector of rational models with regular RREE prices $\{p_s; s \in S\}$. Let Ψ_{nm}, $n = 1,..., N$, $m = 2,..., M_n$ be any other models with the property that $\Psi_{nm}((s)_n, p_s)(s) < 1$. Then $\bar{\lambda}$ is locally stable.*

Proof. It is easy to see that the conditions of Theorem 1 are satisfied. Q.E.D.

It is easy to construct examples using Theorem 1 for which, even though traders may be considering rational models, a non-rational model is locally stable.

COROLLARY 2. *Let $\sigma(s)$ be a permutation of S such that $\sigma(s) \neq s$ for all $s \in S$. Suppose that $M_n = 2$ for all n and that Ψ_2 is a vector of models with RREE prices $\{p_s : s \in S\}$. If there exists a vector of models Ψ_1 regular at $\{p_{\sigma(s)}; s \in S\}$ and for which $\Psi_{n1}((s)_n, p_{\sigma(s)})(s) > \Psi_{n2}((s)_n, p_{\sigma(s)})(s)$ for all $s \in S$, then $\bar{\lambda}$ is locally stable.*

Proof. Again, the conditions of Theorem 1 are easily seen to be satisfied. Q.E.D.

4. Examples

Since the hypothesis of rationality ties down the prediction of a model only at the RREE prices it is easy to construct examples in which a rational model is locally stable, a non-rational model is locally stable, or cycles may occur. Such an example may be constructed using the utility functions and endowments from Radner's example [4] of the existence of RREE. In this example there are two types of traders; informed traders who all receive a signal $s = 1$ or 2 and uninformed traders who receive no signal. Suppose the uninformed traders are choosing between a rational model Ψ_1 and a non-rational model Ψ_2 as illustrated in Fig. 2. Whether $\bar{\lambda} = (1, 0)$ or $\hat{\lambda} = (0, 1)$ is locally stable or cycles occur depends on the relationship between β, ε and η. Let λ_t be the uninformed traders weight on Ψ_1. If:

(i) $\beta > \varepsilon > \eta$ then $\hat{\lambda}$ is locally stable.

(ii) $\eta > \varepsilon > \beta$ then $\bar{\lambda}$ is locally stable.

If $\lambda_t > 0$ and sufficiently less than 1:

(iii) $\varepsilon > \beta$ and $\varepsilon > \eta$ then if signal 1 occurs $\lambda_{t+1} > \lambda_t$ and if signal 2 occurs $\lambda_{t+1} < \lambda_t$.

(iv) $\beta > \varepsilon$ and $\eta > \varepsilon$ then if signal 1 occurs $\lambda_{t+1} < \lambda_t$ and if signal 2 occurs $\lambda_{t+1} > \lambda_t$.

Hence, we can have the rational model locally stable, the non-rational model locally stable, or cycles (by a suitable choice of a stochastic process on signals and a learning rule).

In this example, where all uninformed traders consider the same set of models (Ψ_1 and Ψ_2), if $\beta > \varepsilon > \eta$ then beliefs converge to a non-rational model. At this point all uninformed traders have "learned" Ψ_2, but Ψ_2 is

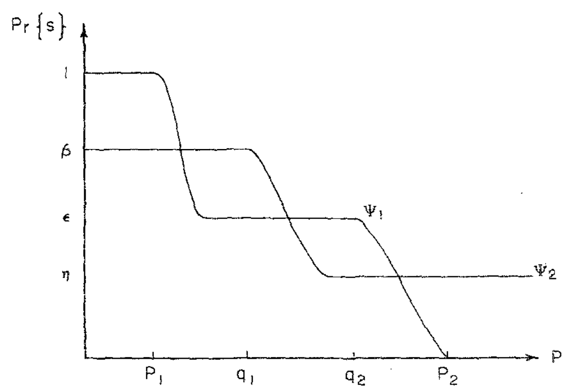

FIGURE 2

systematically wrong. For example, when $s = 1$ occurs q_1 always results but Ψ_2 puts probability $\beta < 1$ on this event (see Fig. 1). It could be argued that traders should recognize this error and modify their behavior. The obvious modification is to include a method of hypothesis testing and model revision in the learning rule. This modification is not included in our analysis as it raises numerous conceptual and technical questions. In particular, the observed relation between prices and signals will not persist if traders somehow consider a new model (i.e., one other than Ψ_1 or Ψ_2). So there seems to be no obvious way to correctly revise a model.

5. Conclusions

This paper has two results: a local stability theorem for rational models and also a global instability theorem. The learning process may get stuck at an incorrect model because all of the admissible models are incorrect away from the RREE. No model tries to forecast the expectations of other traders, and so the set of admissible models is too small to describe the full behavior of the economy. Whether enlarging the set of admissible models—in particular, by admitting richer models that estimate the expectations of other traders—will eliminate stable non-rational models is an open question.

With the finite collection of models treated in this paper there does exist a trivial learning rule which gives almost sure convergence to a RREE. Suppose that at date 1 each trader arbitrarily chooses a model and acts as if he believes this model to be correct. Each trader continues to use this model until it gives a prediction which is incorrect.[3] When a trader decides that his model is incorrect, he randomly chooses another model from his finite set of models. With probability 1 there exists $t < \infty$ such that at t each trader is using model Ψ_{n1}. Then each Ψ_{n1} predicts correctly and so no trader will abandon it.

We have two objections to this learning mechanism. First, it is not at all related to anything that statistical decision theory might suggest. The learning models of Section 3 arise from Bayesian learning with incorrectly specified conditional distributions. The learning model we have just described has no trace of rationality attached to it. For example, traders will often reject models that are only slightly incorrect in favor of models that are abominably wrong. Second, the nice behavior of this rule is a clear artifact of the finiteness of the model.

[3] In the case where a model predicts a distribution of states, each trader might do a hypothesis test based on the empirical distribution of states at a given price and switch models only when the null hypothesis (that the model is correct) is rejected. The Glivento–Cantelli Theorem and the corresponding CLT can be used to show that such models will ultimately be rejected.

REFERENCES

1. L. E. BLUME, M. M. BRAY, AND D. EASLEY, Introduction to the stability of rational expectations equilibrium, *J. Econ. Theory* **26** (1982), 313–317.
2. M. M. BRAY, Learning, estimation and the stability of rational expectations, *J. Econ. Theory* **26** (1982), 318–339.
3. V. GUILLEMIN AND A. POLLACK, "Differential Topology," Prentice-Hall, Englewood Cliffs, N.J., 1974.
4. R. RADNER, Rational expectations equilibrium: Generic existence and the information revealed by prices, *Econometrica* **47** (1979), 655–678.