

Learning, Estimation, and the Stability of Rational Expectations*

MARGARET BRAY

Faculty of Economics and Politics, Cambridge University, Cambridge CB3 9DD, England

Received September 29, 1980; revised August 4, 1981

The stability of the rational expectations equilibrium of a simple asset market model is studied in a situation where a group of traders learn about the relationship between the price and return on the asset using ordinary least squares estimation, and then use their estimates in predicting the return from the price. The model which they estimate is a well-specified model of the rational expectations equilibrium, but a misspecified model of the situation in which the traders are learning. It is shown that for appropriate values of a stability parameter the situation converges almost surely to the rational expectations equilibrium. *Journal of Economic Literature* Classification Numbers: 022, 026.

1. INTRODUCTION

This paper addresses the question of convergence to rational expectations equilibrium. It has been claimed that agents can learn how to form rational expectations. Grossman and Stiglitz [5] write that, if expectations are not rational and the stochastic process of the underlying random variables is stationary,

an individual will eventually observe that the frequency distribution of returns, conditional on the observable variables, is different from the subjective distribution, and accordingly, ought to revise his expectations.

This statement is very plausible. However, even if it is accepted, it does not imply that the revised expectations are rational, or even that multiple revisions eventually lead to rational expectations. The difficulty is that in many models with rational expectations equilibria (including that of

* This paper owes much to J. A. Mirrlees and J. E. Stiglitz, who supervised the thesis from which it is taken. I am also grateful to K. W. S. Roberts and D. Gale for the comments they gave in the course of examining the thesis. D. M. Kreps and J. M. Harrison were extremely helpful on probability theory. R. Radner's detailed comments were invaluable in revising the paper. All responsibility for errors of course rests with me.

Grossman and Stiglitz [5]) the objective distribution of variables depends upon agents' subjective beliefs about the distribution. Outside a rational expectations equilibrium the objective distribution of variables differs both from agents' subjective beliefs, and from the objective distribution which would prevail in a rational expectations equilibrium. Agents may learn about the relationship between variables, given their current beliefs. However, when they modify their beliefs in the light of what they have learned, the relationship changes. Agents may attempt to learn about the relationship, using Bayesian or classical statistical techniques which are based on a correct specification of the rational expectations equilibrium. In doing so they fail to take into account the dependency of outcomes on beliefs. Their estimation technique is based on a misspecification of the situation. For example, in the rational expectations equilibrium of the asset market model studied in this paper, the price and return on the asset at different dates form an i.i.d. sequence of normal random variables. In this situation ordinary least squares is an appropriate statistical procedure for learning about the price–return relationship. However, suppose that agents in the model use OLS (or any other statistical procedure) to estimate the price–return relationship outside the rational expectations equilibrium, and use their estimates in forecasting returns. The stochastic process of price and return is then neither stationary, nor independent. The use of OLS estimation is inappropriate.

Fully rational agents should estimate a correctly specified model, which takes into account the feedback from forecasts to outcomes. This is likely to entail a complicated learning strategy based on a considerable degree of understanding of the situation. Outside the rational expectations equilibrium, it is not usually rational to use estimation techniques which are based on a correct specification of the rational expectations equilibrium, such as OLS in the model considered here. Nevertheless the use of such techniques might be described as reasonable. The major propositions of this paper establish conditions, for the model presented here, under which OLS estimation ultimately generates rational expectations.

The model is basically an infinitely repeated version of the Grossman–Stiglitz [5, 6] model of an asset market with informed and uninformed traders. As in the Grossman–Stiglitz model, the informed agents in this model know how to form rational expectations, given the information available to them. In the Grossman–Stiglitz model the uninformed traders also form rational expectations about return, given the price. In this model, however, the uninformed traders lack the knowledge about the structure of their world needed to form rational expectations. Instead they estimate the price–return relationship from past history, using OLS regression, and use their estimates in forecasting return from price.

Two learning processes are studied. In the first, agents revise the estimates

used in forecasting return at infrequent intervals. In the second, they revise the estimates each time a new data point is observed. In both cases conditions are derived under which expectations are, in the limit, rational.

The model is described in Section 2. Section 3 contains a discussion of the relationship between forecasts and expectations. The major propositions on learning are in Section 4. The nature of the parameter which determines whether the learning processes converge is discussed in Section 5. Section 6 concludes the paper. See Blume *et al.* [1], for a discussion of the related literature.

2. THE MODEL

The model is simple and special. There are two types of traders, $N_i \geq 0$ informed traders and $N_u > 0$ uninformed traders. Both types of traders observe the market price p_t at t , and the return on the asset r_t at a date after t but before $t + 1$. In addition the informed traders observe information I_t at t . Information I_t may be a vector or a scalar random variable. The asset is supplied in quantity s_t at t . Demand is a function of the asset price and of traders' point or mean predictions of the return. Under the assumptions of the model the informed traders need use only their private information I_t in forming predictions which are rational expectations, in the sense of being the correct conditional expectations of return given all the variables, past and present, observed by the informed traders. The uninformed traders forecast return from price on the basis of their estimates of the price–return relationship derived from past history using ordinary least squares.

ASSUMPTION 1. *It is common knowledge that $\{(I_t, r_t, s_t)\}$ is a sequence of independent, identically distributed, multivariate normal, random variables which are exogenously determined.*

Assumption 1 ensures that, if traders do not revise their beliefs about the price–return relationship, (p_t, r_t) is a sequence of i.i.d. random variables.

ASSUMPTION 2. *Each informed trader demands*

$$\theta_i(E(r_t | H_{t-1}, I_t, p_t) - p_t), \quad (2.1)$$

where θ_i is a strictly positive constant, and $H_{t-1} = \{(I_1, p_1, r_1) \cdots (I_{t-1}, p_{t-1}, r_{t-1})\}$ is the history of prices, returns and information known by the informed traders at $t - 1$.

This is essentially an ad hoc specification of the demand function. However, it will be shown that the conditional distribution of r_t given the

information available to the informed traders is normal with constant variance. The demand function can thus be derived from a constant absolute risk aversion utility function, in which case

$$\theta_i = 1/[\text{risk aversion} \times \text{var}(r_i | H_{t-1}, I_t, p_t)].$$

Note that the informed traders are assumed to have rational expectations. The following assumption will be shown to imply that these expectations in fact depend only on I_t .

ASSUMPTION 3. *It is common knowledge that*

$$E(r_t | I_t, s_t) = E(r_t | I_t). \quad (2.2)$$

Assumption 3 ensures that the informed traders can learn nothing about r_t from s_t which they do not already know from I_t . The assumption certainly holds when s_t is a component of I_t , but is also true in other circumstances.

ASSUMPTION 4. *Each uninformed trader demands*

$$\theta_u(\Phi(r_t | h_{t-1}, p_t) - p_t), \quad (2.3)$$

where θ_u is a positive constant, and the uninformed traders forecast r_t by

$$\Phi(r_t | h_{t-1}, p_t) = a_{t-1} + b_{t-1}p_t. \quad (2.4)$$

The forecasting coefficients a_{t-1} and b_{t-1} are functions of $h_{t-1} = \{(p_1, r_1), \dots, (p_{t-1}, r_{t-1})\}$, the history of price and return known by the uninformed traders at $t-1$.

It will be shown that there are numbers a^* and b^* such that, if $a_{t-1} = a^*$ and $b_{t-1} = b^*$, the forecast is the rational expectation, in the sense of being the correct conditional expectation of p_t , given the information available to the uninformed traders at t . In these circumstances the conditional distribution of r_t is normal with constant variance, and the demand is that which would be generated by a constant absolute risk aversion utility function. However, in general $a_{t-1} \neq a^*$ and $b_{t-1} \neq b^*$, the forecasts using these coefficients are not in any sense rational expectations.

ASSUMPTION 5. *It is common knowledge that the market clears, and that the market clearing price p_t is determined by supply s_t , information I_t , and the past history of information, price and return H_{t-1} .*

Note that the history observed by the informed traders, H_{t-1} , includes the history observed by the uninformed traders, h_{t-1} .

3. FORECASTS AND EXPECTATIONS

The two propositions in this section clarify the nature of traders' forecasts and expectations. The market is said to be in *temporary equilibrium* at t if the asset market clears, in which case

$$N_i \theta_i (E(r_t | H_{t-1}, I_t, p_t) - p_t) + N_u \theta_u (a_{t-1} + b_{t-1} p_t - p_t) = s_t. \quad (3.1)$$

Note that in temporary equilibrium the uninformed traders do not, in general, form rational expectations.

PROPOSITION 1. *In the temporary equilibrium the informed traders' conditional expectations of r_t depend only on their current information I_t . That is,*

$$E(r_t | H_{t-1}, I_t, p_t) = E(r_t | I_t). \quad (3.2)$$

Proof. Assumption 5 implies that the informed traders realize that p_t is a function of H_{t-1} , I_t and s_t and so that $E(r_t | H_{t-1}, I_t, s_t, p_t) = E(r_t | H_{t-1}, I_t, s_t)$. Assumption 1 implies that (I_t, r_t, s_t) is independent of past history H_{t-1} , and so that $E(r_t | H_{t-1}, I_t, s_t) = E(r_t | I_t, s_t) = E(r_t | I_t)$ (using Assumption 3). Thus $E(r_t | H_{t-1}, I_t, s_t, p_t) = E(r_t | I_t)$. Taking conditional expectations over s_t given (H_{t-1}, I_t, p_t) implies that $E(r_t | H_{t-1}, I_t, p_t) = E(r_t | I_t)$. ■

An identical argument can be used to show that the conditional distribution of r_t given (H_{t-1}, I_t, p_t) is the same as the conditional distribution of r_t given I_t . As (I_t, r_t) is i.i.d. normal (Assumption 1), this conditional distribution is normal with constant variance.

Proposition 1 makes it possible to write p_t explicitly as a function using (3.1) and (3.2):

$$p_t = \frac{x_t + N_u \theta_u a_{t-1}}{(N_u \theta_u + N_i \theta_i) - N_u \theta_u b_{t-1}}, \quad (3.3)$$

where

$$x_t = N_i \theta_i E(r_t | I_t) - s_t. \quad (3.4)$$

The estimated regression coefficients a_{t-1} and b_{t-1} are functions of past history h_{t-1} . From Assumption 1 they are independent of (I_t, r_t, s_t) . The term x_t is a linear function of (I_t, r_t, s_t) . Thus (r_t, x_t) is bivariate normal and

x_t is independent of h_{t-1} . Thus $E(r_t | h_{t-1}, p_t) = E(r_t | x_t)$. Now $E(r_t | x_t) = Er_t$ if $\text{var } x_t = 0$, and if $\text{var } x_t > 0$

$$E(r_t | x_t) = Er_t + \frac{\text{cov}(r_t, x_t)}{\text{var } x_t} (x_t - Ex_t). \quad (3.5)$$

Eliminating x_t from (3.3) and (3.5) establishes Proposition 2.

PROPOSITION 2.

$$E(r_t | h_{t-1}, p_t) = E(r_t | x_t) = a'_{t-1} + b'_{t-1} p_t, \quad (3.6)$$

where

$$a'_{t-1} = Er_t - (k/N_u \theta_u) Ex_t - ka_{t-1}, \quad (3.7)$$

$$b'_{t-1} = k(N_i \theta_i + N_u \theta_u) / N_u \theta_u - kb_{t-1}, \quad (3.8)$$

$$\begin{aligned} k = 0 & \quad \text{if } \text{var } x_t = 0 \\ & = \frac{N_u \theta_u \text{cov}(r_t, x_t)}{\text{var } x_t} \quad \text{if } \text{var } x_t > 0. \end{aligned} \quad (3.9)$$

The informed traders always form rational expectations in this model. Thus the market is in a rational expectations equilibrium if the uninformed traders also form rational expectations.

DEFINITION. The market is in a *rational expectations equilibrium* if it is in temporary equilibrium and if in addition the uninformed traders' forecasts are the correct conditional expectation of r_t , given h_{t-1} and p_t . This requires that for all values of p_t

$$\Phi(r_t | h_{t-1}, p_t) = E(r_t | h_{t-1}, p_t).$$

It follows immediately from (2.4) and (3.6) that in a rational expectations equilibrium $a_{t-1} = a'_{t-1}$ and $b_{t-1} = b'_{t-1}$. It is easy to verify Proposition 3 from (3.7) and (3.8).

PROPOSITION 3. A rational expectations equilibrium exists if $k \neq -1$. The equilibrium is unique. In the rational expectations equilibrium

$$a_{t-1} = a'_{t-1} = a^* \equiv \frac{(Er_t - (k/N_u \theta_u) Ex_t)}{(1+k)}, \quad (3.10)$$

$$b_{t-1} = b'_{t-1} = b^* \equiv \frac{k(N_i \theta_i + N_u \theta_u)}{(1+k) N_u \theta_u}. \quad (3.11)$$

Note that if $k = -1$ and $b_{t-1} = b'_{t-1}$ (3.8) implies that $(N_i\theta_i + N_u\theta_u)/(N_u\theta_u) = 0$. This is impossible as it has been assumed that $N_i \geq 0$, $N_u > 0$, $\theta_i > 0$, and $\theta_u > 0$. Thus if $k = -1$ no rational expectations equilibrium exists. This is similar to other non-generic examples of the non-existence of rational expectations equilibrium (e.g., Kreps [8]).

4. ORDINARY LEAST SQUARES LEARNING

Equations (3.6), (3.7), (3.8), (3.10) and (3.11) establish that the relationship between price and return is of the form

$$r_t = a'_{t-1} + b'_{t-1}p_t + u_t, \quad (4.1)$$

where

$$u_t = r_t - E(r_t | x_t), \quad (4.2)$$

$$a'_{t-1} = (1 + k)a^* - ka_{t-1}, \quad (4.3)$$

$$b'_{t-1} = (1 + k)b^* - kb_{t-1}. \quad (4.4)$$

The variables $\{u_t\}$ form a sequence of i.i.d. normal random variables; u_t is independent of past history and in particular of a'_{t-1} and b'_{t-1} . As (r_t, x_t) is normal, (4.2) implies that u_t is independent of x_t . If a'_{t-1} and b'_{t-1} were constants over time equal to a'_0 and b'_0 , the conditions of both the Gauss–Markov theorem (Johnston [7]), and the assumptions of Zellner [8] under which OLS constitutes a Bayesian learning procedure, would be satisfied. The regression coefficients when r_t is regressed linearly on p_t would converge in probability and almost surely to a'_0 and b'_0 . If the initial forecast rule yielded rational expectations, ($a_0 = a^*$ and $b_0 = b^*$), OLS estimation would confirm that the expectations were rational, as in this case $a'_0 = a^*$ and $b'_0 = b^*$. An outside observer could appropriately use OLS to estimate the price–return relationship in repeated realizations of the rational expectations equilibrium. However, the uninformed agents who use OLS fail to take into account the fact that the relationship changes as they learn; their estimation procedure is based upon a misspecification of the situation.

I investigate two different regimes in which the uninformed agents use estimated regression coefficients in forecasting. In the first regime studied traders initially forecast r_t by $\Phi(r_t | h_{t-1}, p_t) = a_0 + b_0 p_t$. They continue to use this forecasting rule for a long period, during which they run a linear OLS regression of r_t on p_t . While they use the original forecasting rule the actual relationship between price and return is $r_t = a'_0 + b'_0 p_t + u_t$. As the length of the estimation period tends to infinity the estimated values of the regression coefficients converge almost surely to a'_0 and b'_0 . At some date all

the uninformed traders simultaneously drop the initial forecasting rule, and adopt the new rule $\Phi(r_t | h_{t-1}, p_t) = a_1 + b_1 p_t$, where $a_1 = a'_0$ and $b_1 = b'_0$. They then start to re-estimate the regression coefficients, ultimately reaching new probability limits a'_1 and b'_1 and changing the forecasting rule again.

After the m th change of forecasting rule the coefficients are (a_m, b_m) , which are defined recursively by (a_0, b_0) and the difference equations

$$a_m = (1 + k) a^* - k a_{m-1}, \quad (4.5)$$

$$b_m = (1 + k) b^* - k b_{m-1}. \quad (4.6)$$

The elementary theory of difference equations implies:

PROPOSITION 4. *If the estimates are revised periodically, and the forecasting rule after the m th revision is given by (4.5) and (4.6), expectations converge to rationality (in the sense that a_m tends to a^* and b_m tends to b^*) if and only if $|k| < 1$.*

In the second OLS procedure studied, traders revise their estimates each time a new data point is observed. The resulting, highly non-linear stochastic difference equations are hard to analyze. The problem is rendered tractable by reducing the number of coefficients estimated by one, by assuming that the traders know the means of (p_t, r_t) . If the uninformed traders believed the means to be (p, r) , they would regress $(r_t - r)$ upon $(p_t - p)$. This yields a forecasting rule of the form $\Phi(r_t | h_{t-1}, p_t) = r + b_{t-1}(p_t - p)$. Comparing this with (2.4) shows that here

$$a_{t-1} = r - b_{t-1} p. \quad (4.7)$$

These beliefs about the means of (p_t, r_t) are rational given the past history observed by the uninformed traders if and only if $E(r_t | h_{t-1}) = r$ and $E(p_t | h_{t-1}) = p$. Since r_t is independent of past history this requires that

$$r = E r_t. \quad (4.8)$$

From (3.3), (4.7) and (4.8), as x_t is independent of h_{t-1} and a_{t-1} and b_{t-1} are functions of h_{t-1} ,

$$E(p_t | h_{t-1}) = \frac{E x_t + N_u \theta_u (E r_t - b_{t-1} p)}{N_u \theta_u + N_i \theta_i - N_u \theta_u b_{t-1}}. \quad (4.9)$$

Rearranging (4.9) implies that $p = E(p_t | h_{t-1})$ if and only if

$$p = \frac{E x_t + N_u \theta_u E r_t}{N_u \theta_u + N_i \theta_i}. \quad (4.10)$$

Note from (4.8) and (4.10) that as (r_t, x_t) is i.i.d., if beliefs about means are correct, (p, r) is not in fact a function of time or past history.

If the uninformed traders have rational expectations about the means of (p_t, r_t) , (3.3) and (4.10) imply that

$$p_t - p = \frac{x_t - Ex_t}{N_u \theta_u + N_i \theta_i - N_u \theta_u b_{t-1}}. \quad (4.11)$$

In one special case, if the uninformed traders have rational expectations about the means of p_t and r_t , they also have rational expectations about the conditional mean of r_t given the information available to them. If $\text{var } x_t = 0$, $p_t = p$ for all t ; no variation in the price is observed, and it is impossible to regress $r_t - r$ upon $p_t - p$. The uninformed traders' forecast is

$$\Phi(r_t | h_{t-1}, p_t) = r - b_{t-1}(p_t - p) = Er_t,$$

which is fully rational, given that the constant price conveys no information.

Assumption 6 eliminates this exceptional case.

ASSUMPTION 6.

$$\text{var } x_t > 0. \quad (4.12)$$

Assumption 7 specifies precisely how the uninformed agents use past history to define a forecasting rule.

ASSUMPTION 7. *The uninformed agents forecast r_t by*

$$\Phi(r_t | h_{t-1}, p_t) = a_{t-1} + b_{t-1} p_t, \quad (4.13)$$

where

$$a_{t-1} = Er_t - b_{t-1} Ep_t = r - b_{t-1} p, \quad (4.14)$$

$$p = Ep_t = \frac{Ex_t + N_u \theta_u Er_t}{N_u \theta_u + N_i \theta_i}, \quad (4.15)$$

and

$$b_{t-1} = \frac{S_0 b_0 + \sum_{i=1}^{t-1} (p_i - p)(r_i - r)}{S_0 + \sum_{i=1}^{t-1} (p_i - p)^2}, \quad (4.16)$$

where $0 \leq S_0 < \infty$ and $|b_0| < \infty$ describe the initial beliefs of the uninformed traders, and are determined exogenously to the model.

If $S_0 = 0$ (4.16) is the standard ordinary least squares formula for the regressions of $r_t - r$ upon $p_t - p$ with zero intercept. The initial value b_0 can

be interpreted as the mean, and $S_0/\text{var } u_t$ as the precision of a normal prior on the coefficient in a Bayesian regression (see Zellner [9]). If $S_0 = 0$ the prior is diffuse. It should be noted that the likelihood function which would be used to give a Bayesian interpretation to (4.16) is misspecified in the situation in which all the uninformed traders are learning, just as the classical ordinary least squares model is a misspecification of the situation.

The major proposition of this paper is:

PROPOSITION 5. *Given Assumptions 1–7, if $k > -1$, where*

$$k = N_u \theta_u \text{cov}(r_t, x_t) / \text{var } x_t,$$

then a_t and b_t tend to their rational expectations equilibrium values a^ and b^* almost surely. In the limit expectations are rational.*

Proof. From (3.10), (3.11), (4.14) and (4.15) a_t is a continuous function of b_t , and if $b_t = b^*$, $a_t = a^*$. Thus it is sufficient to show that b_t tends to b^* almost surely. This is demonstrated by showing that if $k > -1$ and if S_t is defined by

$$S_t = S_0 + \sum_1^t (p_i - p)^2, \quad (4.17)$$

(b_t, S_t) is a time homogeneous Markov Process with the following properties:

PROPERTY 1. *The random variable b_t recurs to every neighbourhood of b^* . That is, for any $\delta > 0$,*

$$P(|b_t - b^*| < \delta \text{ infinitely often}) = 1.$$

In addition

$$P(S_t \rightarrow \infty) = 1.$$

PROPERTY 2. *For any π in $(0, 1)$ and any $\varepsilon > 0$, there exist δ in $(0, \varepsilon)$ and $N > 0$, such that for any (b, S) with $|b - b^*| \leq \delta$ and $S \geq N$*

$$P(|b_t - b^*| < \varepsilon \text{ for all } t \geq 0 \mid b_0 = b \text{ and } S_0 = S) \geq \pi.$$

The following argument demonstrates that these two properties imply that for any $\varepsilon > 0$, $P(|b_t - b^*| > \varepsilon \text{ infinitely often}) = 0$, and thus that b_t tends to b^* almost surely.

Let ε be any positive number, let π be any number in $(0, 1)$, and let δ and N be such that Property 2 holds. Let E be the event in which S_t exceeds N at some finite date and $|b_t - b^*|$ subsequently alternates between being less than δ and greater than ε infinitely often. Property 1 implies that, if there is a

positive probability that $|b_t - b^*| > \varepsilon$ infinitely often, then E has positive probability. It will be shown that E has probability zero.

Let E_n be the event in which this alternation happens at least n times after $S_t > N$. Formally E_n is the event that the stopping time t_n is finite, where t_n is defined by

$$t_0 = \min\{t: |b_t - b^*| < \delta, S_t > N\},$$

$$t_n = \min \left\{ t: \begin{array}{l} |b_t - b^*| < \delta, t > t_{n-1}, \\ |b_i - b^*| > \varepsilon \text{ for some } i \text{ between } t_{n-1} \text{ and } t_n \end{array} \right\}.$$

Note that $E_{n-1} \subset E_n$, $E = \bigcap_{n=0}^{\infty} E_n$, and, from Property 1, $P(E_0) = 1$.

Let A_t be the event that $|b_j - b^*| > \varepsilon$ and $|b_i - b^*| < \delta$ for some finite $i > j > t$. Thus

$$P(E_{n+1} | E_n) = P(A_{t_n} | E_n). \quad (4.18)$$

E_n is the past and A_{t_n} is in the future of the stopping time t_n . Thus from the strong Markov property

$$P(A_{t_n} | E_n) = \int P(A_{t_n} | b_{t_n} = b, S_{t_n} = S) dF(b, S | E_n), \quad (4.19)$$

where $F(b, S | E_n)$ is the distribution function of (b_{t_n}, S_{t_n}) given E_n . Note that as $t_n > 0$ and, from (4.1) S_t is increasing, $S_{t_n} \geq S_0 > N$. By definition $|b_{t_n} - b^*| \leq \delta$, so the support of F lies entirely in $\{b, S | |b - b^*| \leq \delta, S \geq N\}$. For these values of (b, S) Property 2 and the fact the (b_t, S_t) is time homogeneous Markov implies that

$$P(A_{t_n} | b_{t_n} = b, S_{t_n} = S) \leq 1 - \pi < 1. \quad (4.20)$$

Thus from (4.18)–(4.20),

$$P(E_{n+1} | E_n) = P(A_{t_n} | E_n) \leq 1 - \pi$$

and so as $E_n \subset E_{n+1}$ and $P(E_0) = 1$,

$$P(E_n) = \left\{ \prod_{i=1}^n P(E_i | E_{i-1}) \right\} P(E_0) \leq (1 - \pi)^n.$$

Thus as $\pi \in (0, 1)$,

$$P(E) \leq \lim_{n \rightarrow \infty} (1 - \pi)^n = 0.$$

This establishes that the theorem follows from Properties 1 and 2. Properties 1 and 2 are shown to hold in the Appendix.

5. THE STABILITY PARAMETER

The proof that Properties 1 and 2 hold in the Appendix makes repeated use of $k > -1$. If $k \leq -1$ it is easy to show that $(b_t - b^*)^2$ is a submartingale; that is, $E((b_t - b^*)^2 | b_{t-1}, S_{t-1}) \geq (b_{t-1} - b_t)^2$. This suggests, but does not prove, that if $k \leq -1$, b_t not only fails to converge to b^* , but is also unbounded. If this is so the boundary between the two types of behaviour occurs at precisely the point ($k = -1$) where there is no rational expectations equilibrium.

The parameter k is clearly crucial. It is defined in (3.9), but is not given any interpretation. When the supply s_t has zero variance (3.4) and (3.9) imply that k has a particularly simple form, $k = (N_u \theta_u) / (N_i \theta_i)$. In this case, given a_{t-1} and b_{t-1} , the price is a linear function of $E(r_t | I_t)$ and a sufficient statistic for the information. Thus in the rational expectations equilibrium $\Phi_t = a^* + b^* p_t = E(r_t | I_t)$. Here k is the equilibrium ratio of uninformed to informed demand and is non-negative. The period by period least squares learning rule is therefore stable. The process in which the forecasting rule is changed only when the estimates reach their probability limits is stable if the uninformed traders demand less than the informed traders in equilibrium.

If $\text{var } s_t > 0$, k can be negative if the return on the asset and its supply are positively correlated. From (3.9) the stability conditions can be reformulated as

$$b^* < \frac{(N_u \theta_u + N_i \theta_i)}{2N_u \theta_u}$$

implies that $-1 < k < 1$ so both learning processes are stable, and

$$b^* < \frac{(N_u \theta_u + N_i \theta_i)}{N_u \theta_u}$$

implies that $-1 < k$, so period ordinary least squares learning is stable. Thus learning tends to generate instability if either the ratio $N_u \theta_u / N_i \theta_i$ is large so the uninformed traders dominate the market, or the equilibrium value of the regression coefficient of price on return b^* is large.

6. CONCLUSIONS

The results of this paper suggest that a learning method can eventually yield rational expectations even if it is based upon a misspecification of the model in the situation when agents are learning. However—as one might expect—the stability properties of the system are different for different learning procedures, and instability seems to be a real possibility.

These results seem to lend weight to the rational expectations hypothesis. However, it must be borne in mind that expectations are not rational, and indeed are biased due to the misspecification of the model which is estimated, at all finite dates. Rational expectations are, if anything, a long run rather than a short run phenomenon.

APPENDIX

The Appendix contains the proof that (b_t, S_t) is a time homogeneous Markov process, and establishes Properties 1 and 2 as Lemmas 4 and 5. Lemmas 1–3 are used in the proof of Lemma 4.

It is convenient to introduce some more notation, and use it to spell out the relationship between (b_t, S_t) and (b_{t-1}, S_{t-1}) . The history of the model is determined by the price at each date (4.11), the estimation rule (4.16), the initial values (b_0, S_0) and the realizations of the exogenous i.i.d. normal random variables (r_t, x_t) . Let

$$v_t = (x_t - Ex_t)/N_u \theta_u, \quad (\text{A.1})$$

$$u_t = r_t - E(r_t | x_t), \quad (\text{A.2})$$

$$c_t = (N_u \theta_u + N_t \theta_t)/N_u \theta_u - b_t, \quad (\text{A.3})$$

$$c^* = (N_u \theta_u + N_t \theta_t)/N_u \theta_u - b^*. \quad (\text{A.4})$$

From (3.11)

$$c^* = (N_u \theta_u + N_t \theta_t)/(N_u \theta_u(1+k)) = b^*/k. \quad (\text{A.5})$$

Note that c_t is a linear function of b_t , and $|c_t - c^*| = |b_t - b^*|$. Note also from (A.5) that, if $k > -1$, c^* is strictly positive. From Assumption 1 $\{(u_t, v_t)\}$ is a sequence of i.i.d. zero mean normal random variables; u_t is independent of v_t . From Assumption 6 $\text{var } v_t > 0$. Finally as $k = N_u \theta_u \text{cov}(r_t, x_t)/\text{var } x_t$, (3.5), (4.8), (A.1) and (A.2) imply that

$$r_t - r = kv_t + u_t, \quad (\text{A.6})$$

and from (4.11), (A.1) and (A.3)

$$p_t - p = v_t/c_{t-1}. \quad (\text{A.7})$$

Equations (4.16), (A.5), (A.6) and (A.7) imply that

$$c_t - c^* = \frac{S_0(c_0 - c^*) - k \sum_1^t (v_i/c_{i-1})^2 (c_{i-1} - c^*) - \sum_1^t v_i u_i/c_{i-1}}{S_0 + \sum_1^t (v_i/c_{i-1})^2} \quad (\text{A.8})$$

Equations (4.17) and (A.8) can be used to write (c_t, S_t) as a function of c^* , c_{t-1} , S_{t-1} , u_t and v_t :

$$c_t - c^* = \frac{(c_{t-1} - c^*)\{S_{t-1} - k(v_t/c_{t-1})^2\} - v_t u_t / c_{t-1}}{S_{t-1} + (v_t/c_{t-1})^2}, \quad (\text{A.9})$$

$$S_t = S_{t-1} + (v_t/c_{t-1})^2. \quad (\text{A.10})$$

These are the equations which serve to define the stochastic process (c_t, S_t) . As c^* is constant and $\{u_t, v_t\}$ is i.i.d. it follows immediately from (A.9) and (A.10) that (c_t, S_t) is a time homogeneous Markov Process. It is obvious from (A.3) and (A.4) that (b_t, S_t) is also a time homogeneous Markov Process.

To complete that notation let (Ω, \mathcal{F}, P) be the probability space on which all the random variables are defined, and let \mathcal{F}_t be the smallest sub-sigma field of on which $(r_1, x_1), \dots, (r_t, x_t)$ are measurable. Then $\mathcal{F}_1 \subset \mathcal{F}_2 \dots$ and (c_t, S_t, u_t, v_t) is \mathcal{F}_t but not \mathcal{F}_{t-1} measurable.

LEMMA 1. *If $k \neq -1$ the event in which b_{t-1} tends to any limit apart from b^* has probability zero.*

The intuitive argument here is that if the coefficient used in forecasting b_{t-1} converged to b , the conditional expectation of $r_t - r$ would converge to $((1+k)b^* - kb)(p_t - p)$. Thus the OLS estimator would converge to $(1+k)b^* - kb$. Unless $b = b^*$ this contradicts the original assumption that b_{t-1} tends to b .

Proof. As $|c_t - c^*| = |b_{t-1} - b^*|$ the lemma can be proved by showing that c_t almost surely does not tend to any limit other than c^* .

Suppose that there is an event of positive probability on which c_t tends to c , $c \neq 0$ and $c \neq c^*$. If this is so the strong law of large numbers implies that there is an event G of positive probability on which

$$c_t \rightarrow c \quad \frac{1}{t} \sum_1^t v_i^2 \rightarrow \sigma_v^2 \quad \frac{1}{t} \sum_1^t u_i v_i \rightarrow 0.$$

Thus on G

$$\begin{aligned} \frac{1}{t} \sum_1^t (v_i/c_{i-1})^2 &\rightarrow \sigma_v^2/c^2 & \frac{1}{t} \sum_1^t u_i v_i / c_{i-1} &\rightarrow 0, \\ \frac{1}{t} \sum_1^t (v_i/c_{i-1})^2 (c_{i-1} - c^*) &\rightarrow (\sigma_v^2/c^2)(c - c^*), \end{aligned}$$

implying that the right-hand side of (A.8) tends to $-k(c - c^*)$ while the left-hand side tends to $c - c^*$. This is impossible unless $c = c^*$.

If c tends to 0 with positive probability, for any $\varepsilon > 0$ there is a date n and event G' on which, for all $t > n$, $|c_t| < \varepsilon$, and using the strong law of large numbers

$$w_t \equiv \frac{1}{t} \sum_{n+1}^t (v_t/c_{t-1})^2 \rightarrow \infty \quad \text{as } t \rightarrow \infty.$$

From the Cauchy-Schwarz inequality

$$\left| \frac{1}{t} \sum_{n+1}^t (v_t u_t / c_{t-1}) \right| \leq (w_t)^{1/2} \left(\frac{1}{t} \sum_{n+1}^t u_t^2 \right)^{1/2}.$$

Thus from (A.8)

$$|c_t - (1+k)c^*| \leq \frac{S_n |c_n - (1+k)c^*|/t + |k| w_t \varepsilon + (w_t)^{1/2} (\sum_{n+1}^t u_t^2)^{1/2}}{S_n/t + w_t}.$$

Taking limits as t tends to infinity all terms on the right-hand side of this inequality tend to zero except

$$w_t \rightarrow \infty \quad \text{and} \quad \frac{1}{t} \sum_{n+1}^t u_t^2 \rightarrow \sigma_u^2.$$

Thus the right-hand side tends to $|k| \varepsilon$. If ε is sufficiently small and $k \neq -1$ this is inconsistent with the initial supposition that $|c_t| < \varepsilon$ unless in fact c_t tends to $c^* = 0$. ■

Lemma 2 makes use of the Martingale convergence theorem possible.

LEMMA 2. *If $k > -1$ there exist real functions f defined on R , and g_1 and g_2 defined on R^{++} , with the following properties:*

(i) *f is continuous, decreasing on $x < 0$, and increasing on $x > 0$. $f(0) = 0$.* (A.11)

(ii) *For all $S > 0$, $-\infty < g_1(S) < c^* < g_2(S) < \infty$.* (A.12)

(iii) *For all $S > 0$, g_1 is non-decreasing and g_2 is non-increasing.* (A.13)

(iv) $\lim_{S \rightarrow \infty} g_1(S) = \lim_{S \rightarrow \infty} g_2(S) = c^*$. (A.14)

(v) *If c_{t-1} is not in $(g_1(S_{t-1}), g_2(S_{t-1}))$*

$$E(f(c_t - c^*) | \mathcal{F}_{t-1}) \leq f(c_{t-1} - c^*). \quad (\text{A.15})$$

The function $f(c_t - c^*)$ is a measure of the "distance" of c_t from c^* , (A.11). Note from (4.17) and Assumption 7 that $S_t > 0$ for all $t > 0$ almost

surely. Thus (A.12) and (A.15) imply that $f(c_t - c^*)$ satisfies the supermartingale property except in a neighbourhood of c^* . The critical neighbourhood becomes smaller as S_{t-1} increases (A.13), collapsing on c^* as S_{t-1} tends to infinity (A.14).

The rigorous proof of this lemma is a tedious exercise in calculus. (See Bray [2 or 3] for details.) The argument is suggested by noting that, given any $c_{t-1} \neq 0$, (A.9) implies that if S_{t-1} is sufficiently large, c_t is likely to be closer to c^* than c_{t-1} . There are apparent difficulties when c_{t-1} is close to zero. In these cases $(v_t/c_{t-1})^2$ is likely to be large, and $c_t - c^* \simeq -k(c_{t-1} - c^*) \simeq kc^*$. These difficulties are avoided by choosing f so that $f(kc^*) < f(-c^*)$.

Lemmas 1 and 2 are used in the proof of Lemma 3.

LEMMA 3. *If $k > -1$ and $S_0 > 0$, c_t almost surely visits $(g_1(S_0), g_2(S_0))$ infinitely often.*

Proof. The proof begins by showing that the first date at which c_t enters $(g_1(S_0), g_2(S_0))$, t_0 , is almost surely finite. Define c'_t by $c'_t = c_t$ if $t \leq t_0$, and $c'_t = c_{t_0}$ if $t > t_0$. Note that t_0 is a stopping time, and that c'_t is \mathcal{F}_t measurable. It follows immediately from Lemma 2, (A.11), (A.13) and (A.15), that $E(f(c'_t - c^*) | \mathcal{F}_{t-1}) \leq f(c'_{t-1} - c^*)$; $f(c'_t - c^*)$ is a positive supermartingale. Thus from the Martingale Convergence Theorem (Chung [4, corollary to Theorem 9.4.4]) $f(c'_t - c^*)$ almost surely tends to a finite limit.

If t_0 is infinite $c'_t = c_t$ for all t , in which case, if $f(c'_t - c^*)$ tends to a finite limit, $f(c_t - c^*)$ does so also. Given (A.11) this implies that c_t is bounded. Now from (A.9)

$$|c_t - c_{t-1}| = \left| \frac{(1+k)(c_{t-1} - c^*)v_t^2/t + c_{t-1}v_t u_t/t}{c_{t-1}^2 \{\sum_1^t ((v_i/c_{i-1})^2)/t + c_{t-1}^2 S_0/t\}} \right|.$$

Recall that, as (u_t, v_t) is i.i.d. normal,

$$v_t^2/t \rightarrow 0, \quad v_t u_t/t \rightarrow 0, \quad \text{and} \quad \frac{1}{t} \sum_1^t v_i^2 \rightarrow \sigma_v^2 \quad \text{almost surely.}$$

Thus if c_t is bounded, $|c_t - c_{t-1}|$ tends to zero almost surely. From (A.11) if $f(c_t - c^*)$ tends to a limit and $|c_t - c_{t-1}|$ tends to zero, c_t tends to a limit. If t_0 is infinite c_t is not in $(g_1(S_0), g_2(S_0))$ for any t . In this case (A.12) implies that, as c^* is in $(g_1(S_0), g_2(S_0))$, c_t cannot tend to c^* . Thus if there is positive probability that t_0 is infinite there is positive probability that c_t tends to a limit other than c^* . However, (Lemma 1) the probability that c_t tends to any limit other than c^* is zero, so t_0 is almost surely finite.

Now define the sequence of stopping times t_0, t_1, \dots recursively by letting t_n be the first date after t_{n-1} at which c_t is in $(g_1(S_0), g_2(S_0))$. From (4.17), S_t

is non-decreasing with time, so $S_{t_n} \geq S_0 > 0$. As (c_t, S_t) is Markov the argument used to establish that t_0 is finite establish that, if t_n is finite, t_{n+1} is almost surely finite. Thus t_n is almost surely finite for all n . Finally note from (A.13) that as $S_{t_n} \geq S_0$, $(g_1(S_{t_n}), g_2(S_{t_n}))$ is a subinterval of $(g_1(S_0), g_2(S_0))$, so c_{t_n} is in $(g_1(S_0), g_2(S_0))$ for an almost surely infinite sequence of stopping times $\{t_n\}$, which proves the lemma. ■

Property 1 is now established by proving Lemma 4.

LEMMA 4. For any $\delta > 0$

$$P(|b_t - b^*| < \delta \text{ infinitely often}) = 1. \quad (\text{A.16})$$

In addition

$$P(S_t \rightarrow \infty) = 1. \quad (\text{A.17})$$

Proof. The proof begins by showing that (A.17) holds. Note from (A.1) and (A.7) that as $\text{var } v_t > 0$, $p_1 - p \neq 0$ almost surely, and so from (4.17), even if $S_0 = 0$, $S_1 > 0$ almost surely. Thus, as the process is Markov, the dates can always be renumbered so that $S_0 > 0$.

Lemma 3 implies that there is a sequence of stopping times $\{t_n\}$, such that c_{t_n} is in $(g_1(S_0), g_2(S_0))$ for all n . Let

$$c_m = \max(|g_1(S_0)|, |g_2(S_0)|) < \infty \quad (\text{A.18})$$

Thus from (A.10), as S_t is a sum of non-negative terms

$$\lim_{t \rightarrow \infty} S_t \geq \sum_{n=0}^{\infty} (v_{t_n+1}/c_{t_n})^2 \geq \left(\sum_{n=0}^{\infty} v_{t_n+1}^2 \right) / c_m^2. \quad (\text{A.19})$$

$\{t_n\}$ is almost surely an infinite sequence of stopping times defined relative to the stationary independent stochastic process $\{(u_t, v_t)\}$. Thus from the optional stopping theorem for such processes (Chung [4, Theorem 8.2.3]) v_{t_n+1} is a sequence of i.i.d. variables with the same distribution as v_1 . As v_1 has zero mean and strictly positive variance

$$\sum_{n=1}^{\infty} v_{t_n+1}^2 \rightarrow \infty \quad \text{almost surely}$$

and so from (A.19) S_t tends to infinity almost surely.

Using (A.14) this implies that there is, almost surely, a finite date u such that

$$c^* - \delta < g_1(S_u) < c^* < g_2(S_u) < c^* + \delta.$$

Renumbering time starting at u , and using the Markov property, Lemma 3 implies that $|c_t - c^*| < \delta$ infinitely often. As $|b_t - b^*| = |c_t - c^*|$ this establishes that (A.16) holds. ■

Property 2 is shown to hold by proving Lemma 5.

LEMMA 5. *If $k > -1$ then for any π in $(0, 1)$, and any $\varepsilon > 0$, there exists $\delta > 0$ and $N > 0$, such that for any (c, S) with $|c - c^*| \leq \delta$ and $S \geq N$*

$$P(|c_t - c^*| < \varepsilon \text{ for all } t \geq 0 \mid c_0 = c^*, S_0 = S) \geq \pi.$$

As $|c_t - c^*| = |b_t - b^*|$, and the process $\{(c_t, S_t)\}$ is time homogeneous Markov, this is equivalent to Property 2.

Proof. Heuristically the argument proceeds as follows. Equation (A.8) shows that if $c_0 \simeq c^*$ and S_0 is sufficiently large, $c_t \simeq c^*$ for a long time after 0. From (A.9) if $k > -1$, c_t is in fact likely to move closer to c^* as time progresses. If $c_t \simeq c^*$ for a long time, the model has been close to its rational expectations equilibrium, the OLS estimator is likely to be close to its rational expectations equilibrium value b^* , and so c_t is likely to continue close to c^* .

The lemma is proved by defining new variables (d_t, e_t) by

$$e_0 = c_0, \quad (\text{A.20})$$

$$d_t - c^* = \frac{S_0(e_0 - c^*) - k \sum_1^t (v_i/e_{i-1})^2 (e_{i-1} - c^*) - \sum_1^t v_i u_i / e_{i-1}}{S_0 + \sum_1^t (v_i/e_{i-1})^2} \quad (\text{A.21})$$

and

$$\begin{aligned} e_t &= c^* + 2\varepsilon & \text{if } d_t \geq c^* + 2\varepsilon, \\ e_t &= d_t & \text{if } |d_t - c^*| < 2\varepsilon, \\ e_t &= c^* - 2\varepsilon & \text{if } d_t \leq c^* - 2\varepsilon. \end{aligned} \quad (\text{A.22})$$

There is no loss of generality in restricting attention to $\varepsilon < \frac{1}{2}c^*$ (recall that c^* is strictly positive). In this case

$$0 < c^* - 2\varepsilon \leq e_t \leq c^* + 2\varepsilon < 2c^*. \quad (\text{A.23})$$

These bounds will be used extensively later in the proof.

Comparing (A.8) and (A.20)–(A.22) it is clear that, if $|e_t - c^*| < \varepsilon$, $t = 1, 2, \dots, T$, then $c_t = d_t = e_t$, and so $|c_t - c^*| < \varepsilon$ for $t = 1, 2, \dots, T$. It can

be shown by an induction argument that if $|k| < 1$, sufficient conditions for $|e_t - c^*| < \varepsilon$, and thus for $|c_t - c^*| < \varepsilon$ on $t = 1, 2, \dots, T$ are

$$|c_0 - c^*| \leq \delta \equiv \frac{1}{2}\varepsilon(1 - |k|)$$

and

$$|y_t| < \varepsilon_y \equiv \frac{1}{2}\varepsilon(1 - |k|), \quad t = 1, \dots, T,$$

where

$$y_t = \frac{\sum_1^t v_i u_i / e_{i-1}}{S_0 + \sum_1^t (v_i / e_{i-1})^2}. \quad (\text{A.24})$$

If $k \geq 1$, sufficient conditions for $|e_t - c^*| < \varepsilon$, and thus for $|c_t - c^*| < \varepsilon$ on $t = 1, 2, \dots, T$ are

$$\begin{aligned} |c_0 - c^*| &\leq \delta \equiv \frac{1}{2}\varepsilon/(1 + k), \\ |y_t| &< \varepsilon_y \equiv \frac{1}{2}\varepsilon/(1 + k), \quad t = 1, \dots, T, \\ |z_t| &< \varepsilon_z \equiv 1/k, \quad t = 1, \dots, T, \end{aligned}$$

where y_t is defined by (A.24) and

$$z_t = \frac{(v_t / e_{t-1})^2}{S_0 + \sum_1^t (v_i / e_{i-1})^2}. \quad (\text{A.25})$$

See Bray [2 or 3] for details of these arguments.

The proof of the lemma is completed by showing that for any $\varepsilon_y > 0$ and $\varepsilon_z > 0$, if S_0 is sufficiently large, there are events D_1 to D_5 , such that $|z_t| < \varepsilon_z$ and $|y_t| < \varepsilon_y$ for all t on $(\bigcap_{i=1}^5 D_i)$, and $P(\bigcap_{i=1}^5 D_i) \geq \pi$. To this end let n be a positive integer and define events $D_1 - D_5$ by

$$\begin{aligned} D_1 &= \{\omega: v_t^2 < (c^* - 2\varepsilon)^2 S_0 \varepsilon_z \text{ for all } t < n\}, \\ D_2 &= \left\{ \omega: \left| \sum_1^t (v_i u_i / e_{i-1}) \right| < \frac{1}{2} S_0 \varepsilon_y \text{ for all } t < n \right\}, \\ D_3 &= \left\{ \omega: \frac{v_t^2}{t} < \frac{1}{2} \frac{(c^* - 2\varepsilon)^2}{(c^* + 2\varepsilon)^2} \sigma_v^2 \varepsilon_z \text{ for all } t \geq n \right\}, \\ D_4 &= \left\{ \omega: \frac{1}{t} \sum_1^t v_i^2 > \frac{1}{2} \sigma_v^2 \text{ for all } t \geq n \right\}, \\ D_5 &= \left\{ \omega: Q_t < \frac{\sigma_v^2 \varepsilon_y}{8(c^* + 2\varepsilon)^2} \text{ for all } t \geq n \right\}, \end{aligned}$$

where

$$Q_t = \sum_n^t \frac{v_i u_i}{i e_{i-1}}. \quad (\text{A.26})$$

Now observe from (A.23) and (A.25)

$$|z_t| < \frac{(c^* + 2\varepsilon)^2}{(c^* - 2\varepsilon)^2} \cdot \frac{v_t^2}{S_0(c^* + 2\varepsilon)^2 + \sum_1^t v_i^2}$$

so on D_1 , if $t < n$

$$|z_t| < \frac{v_t^2}{S_0(c^* - 2\varepsilon)^2} < \varepsilon_z$$

and on $D_3 \cap D_4$, if $t \geq n$

$$|z_t| < \frac{(c^* + 2\varepsilon)^2}{(c^* - 2\varepsilon)^2} \frac{v_t^2/t}{(\sum_1^t v_i^2)/t} < \varepsilon_z.$$

Thus on $D_1 \cap D_3 \cap D_4$ $|z_t| < \varepsilon_z$ for all t . On D_2 , if $t < n$, from (A.24)

$$|y_t| \leq \left| \sum_1^t v_i u_i / e_{i-1} \right| / S_0 < \frac{1}{2} \varepsilon_y < \varepsilon_y.$$

It is easy to check from (A.26), using a simple induction argument, that

$$\sum_n^t (v_i u_i / e_{i-1}) = (t+1) Q_t - \sum_n^t Q_i.$$

Thus on $D_2 \cap D_4 \cap D_5$, if $t \geq n$,

$$|y_t| \leq \frac{|\sum_1^{n-1} (v_i u_i / e_{i-1})|}{S_0} + \frac{\{(t+1)|Q_t| + \sum_n^t |Q_i|\}(c^* + 2\varepsilon)^2/t}{(\sum_1^t v_i^2)/t} < \varepsilon_y.$$

Thus on $\cap_{i=1}^5 D_i$, $|y_t| < \varepsilon_y$ and $|z_t| < \varepsilon_z$ for all t . The proof concludes by showing that there exists an n such that, if S_0 is sufficiently large, $P(\cap_{i=1}^5 D_i) \geq \pi$.

Observe that as $\{v_i\}$ is i.i.d. and $Ev_i^2 = \sigma_v^2 < \infty$,

$$v_t^2/t \rightarrow 0 \quad \text{as } t \rightarrow \infty \quad \text{a.s.} \quad (\text{A.27})$$

$$\left(\sum_1^t v_i^2 \right) / t \rightarrow \sigma_v^2 \quad \text{as } t \rightarrow \infty \quad \text{a.s.} \quad (\text{A.28})$$

Note also from (A.26), that as u_i and v_i have zero mean, and are

independent of each other and of past values of u_i , v_i and e_i , Q_t is a Martingale, and

$$EQ_t^2 = \sigma_v^2 \sigma_u^2 \sum_n^t E(1/e_{i-1})^2 < \frac{\sigma_v^2 \sigma_u^2}{(c^* - 2\varepsilon)^2} \sum_n^t (1/i)^2.$$

Thus from Kolmogorov's inequality for Martingales (Chung [4, Corollary 1 to Theorem 9.4.1]), if $\lambda > 0$

$$P\left(\max_{i=n \dots \infty} |Q_i| < \lambda\right) \geq 1 - \frac{\sigma_v^2 \sigma_u^2}{\lambda^2 (c^* - 2\varepsilon)^2} \sum_n^\infty (1/i)^2. \quad (\text{A.29})$$

Thus, from (A.27)–(A.29), if n is sufficiently large

$$P(D_i) > (4 + \pi)/5, \quad i = 3, 4, 5. \quad (\text{A.30})$$

Note also that

$$\sum_1^t (v_i u_i / e_{i-1})$$

is a Martingale, and so using Kolmogorov's inequality for Martingales, for any $\lambda > 0$

$$P\left(\max_{t=1 \dots n} \left| \sum_1^t (v_i u_i / e_{i-1}) \right| < \lambda\right) \geq 1 - \frac{n}{\lambda^2} \cdot \frac{\sigma_u^2 \sigma_v^2}{(c^* - 2\varepsilon)^2}.$$

Thus, given n , if S_0 is sufficiently large

$$P(D_2) > (4 + \pi)/5. \quad (\text{A.31})$$

As $\{v_i\}$ is i.i.d., if S_0 is sufficiently large

$$P(D_1) > (4 + \pi)/5. \quad (\text{A.32})$$

Thus from (A.30)–(A.32), there exists an n such that, if S_0 is sufficiently large

$$P\left(\bigcap_1^5 D_i\right)' = P\left(\bigcup_1^5 D_i'\right) \leq \sum_1^5 P(D_i') < 1 - \pi.$$

(Here D_i' is the complement of the event D_i in Ω .) Thus, as required,

$$P\left(\bigcap_1^5 D_i\right) \geq \pi.$$

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,” Research Paper NO. 550, Graduate School of Business, Stanford, 1980.
3. M. M. BRAY, “Expectation Formation and Information in Economic Systems,” D. Phil. thesis, Oxford, 1980.
4. K. L. CHUNG, “A Course in Probability Theory,” Second Ed., Academic Press, New York, 1974.
5. S. J. GROSSMAN AND J. E. STIGLITZ, Information and competitive price systems, *Amer. Econ. Rev.* **66** (1976), 246–253.
6. S. J. GROSSMAN AND J. E. STIGLITZ, On the impossibility of informationally efficient markets, *Amer. Econ. Rev.* **70** (1980), 393–408.
7. J. JOHNSTON, “Econometric Methods,” McGraw-Hill, New York, 1972.
8. D. M. KREPS, A note on “Fulfilled Expectations” equilibria. *J. Econ. Theory* **14** (1977), 32–43.
9. A. ZELLNER, “An Introduction to Bayesian Inference in Econometrics,” Wiley, New York, 1971.